The Sir Peter MacCallum Department of Oncology

CONFIRMATION REPORT

Title of the project: Characterizing the molecular heterogeneity of neuroendocrine prostate cancer

Student name: Rosalia Quezada Urban
Student ID: 1051094

Supervisors: Dr. David Goode, Dr. Roxanne Toivanen, Professor Gail Risbridger

Mentor: Dr. Amanda Lee

Field of Research codes:
060102 – Bioinformatics
060408 – Genomics
111203 - Cancer Genetics
111209 - Solid Tumours

Table of Contents

Abstract (approximately 100 words) that includes a concise statement of the research question

Abstract

Neuroendocrine prostate cancers (NEPCs) are aggressive tumours, with limited treatment options and poor prognosis.

A major challenge in identifying new treatments for neuroendocrine tumours is that their pathology is highly heterogeneous. The main hypothesis of this project is that to effectively treat these tumours we need to investigate intra-tumoural heterogeneity in NEPC at the molecular level.

I will analyse intra-tumoural heterogeneity in NEPC at three levels: 1) at the transcriptomic level, using single-cell RNA sequencing, 2) within the microenvironment, by assessing interactions of cancer-associated fibroblasts with neuroendocrine cells and 3) at the genomic level, by evaluating the evolutionary relationships between neuroendocrine pathologies within the same tumour.

This work will provide an understanding of the neuroendocrine tumour heterogeneity, guiding future searches for new drug targets for NEPC.

Section 1: Literature Review

1.1 Prostate cancer

Prostate cancer (PC) is a disease that predominantly affects men over 65 years old and is the second most common cancer diagnosed around the world**,** with approximately 1.6 million cases each year. Despite the recent advances in detection and therapy, PC is still one of the most common causes of death in men, with over 3000 men dying of prostate cancer in Australia in 2018 [1].

Approximately 95% of prostate tumours are adenocarcinomas, which are comprised of androgen dependant luminal cells that express the androgen receptor (AR) and excrete prostatic specific antigen (PSA). PSA is commonly used as a serum marker to detect and monitor the progression disease of adenocarcinoma. Because adenocarcinoma is androgen dependant, blocking androgen production and/or signalling (Androgen deprivation therapy; ADT) is the leading treatment for metastatic PC [2]. Initially, ADT is efficient in stopping tumour growth but rarely eradicates all malignant cells, leading to treatment resistance and progression to castration-resistant prostate cancer (CRPC). Treatment options for CRPC are second-generation androgen signalling inhibitors, taxane-based chemotherapies, or radiotherapy [2-4]. CRPC will respond to these treatments, but they will eventually fail, resulting in PC-associated death [5] [6] (Figure 1.1).

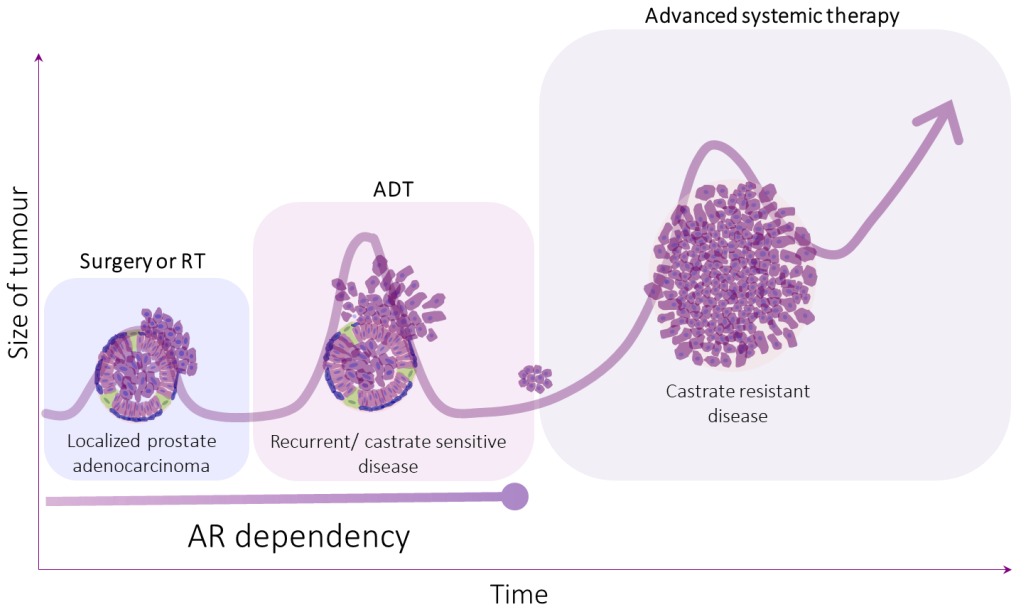


Figure 1.1 Disease progression of prostate adenocarcinoma. There are three stages to the clinical progression of prostate cancer: localised prostate cancer, advanced prostate cancer,

and castration-resistant prostate cancer (CRPC). In early localized stages PC adenocarcinoma can be treated with surgery. In later stages, Androgen deprivation therapy is the standard treatment of PC. However, due to the heterogeneity of adenocarcinoma there are still some cell populations that do not respond to therapy, and ultimately almost all patients will progress to a castrate-resistant phenotype [7].

In PC, the vast majority of tumours are adenocarcinomas, which are androgen-dependent. The disease progression of adenocarcinoma depends on the resistance to ADT, leading to CRPC, which is lethal. However, there is a class of very aggressive, androgen independent tumours with a poor prognosis; they are known as neuroendocrine tumours.

1.2 Neuroendocrine prostate tumours

Neuroendocrine prostate cancer (NEPC) is an aggressive pathology of PC that can emerge *de novo* early in localised disease [3-6]. The median survival rate of *de novo* NEPC is less than two years, making NEPC a more aggressive disease and challenging to treat [5]. Unlike adenocarcinomas that are manageable with ADT, most neuroendocrine tumours do not express AR and therefore do not respond to ADT. NEPC is currently treated with cisplatin-based chemotherapy, however, they quickly develop resistance to this treatment.

With data emerging from clinical research, studies have identified a range of distinct NEPC pathologies [8]. The World Health Organization (WHO) has classified these tumours in five categories based on their histopathology (Table 1) [8, 9]. While some of the NEPC tumours are entirely comprised of neuroendocrine cells which are AR negative, others have a mixed phenotype in which they coexist with adenocarcinoma or co-express adenocarcinoma markers. In some cases, men can develop multiple NE pathologies within the same tumour or either across different metastatic lesions [10]. Therefore, NEPC shows considerable heterogeneity at the pathological level.

Table 1. Histopathology of NE tumours in prostate cancer.

Histopathology of Neuroendocrine tumour	Immunohistochemical phenotype
Adenocarcinoma with neuroendocrine differentiation	Luminal tumour (AR+, PSA+) with focal NE marker expression (Chromogranin A+ and Synaptophysin+)
Adenocarcinoma with Paneth cell-like	Luminal tumour (AR+, PSA+) with focal Paneth cell cytology (Chromogranin A+, AMACR+, Synaptophysin+ and AR-)

neuroendocrine differentiation	
Well-differentiated neuroendocrine tumour (carcinoid tumour)	Chromogranin A+, Synaptophysin+, AR−, PSA−. Low Ki-67 labelling index.
Poorly- differentiated neuroendocrine tumours (Small or Large cell carcinoma)	Chromogranin A+, Synaptophysin+, AR−, PSA−. High Ki-67 labelling index (55%). Loss of p53 and Rb expression is common.
Amphicrine carcinoma	Diffuse expression for luminal (AR+, PSA+) and NE markers (Chromogranin A +, Synaptophysin+) in all tumour cells.

The complexity of NEPC pathology and the aggressive progression of these tumours results in a challenging task for detection [11]. Currently there are no serum markers to detect NEPC. The diagnosis of NEPC depends on a combination of pathologic and clinical features. Clinically, a rapid radiographic progression with low PSA levels is suggestive of AR signalling independence [12]. Tumour heterogeneity represents an additional challenge. If patients harbour multiple lesions, the tumour biopsy may not capture the NE component [13]. This approach, although effective for pure NE tumours, limits the detection of mixed NE pathologies with adenocarcinoma. This represents a challenge in detecting and effectively treating NEPC.

NEPC is a heterogeneous pathology and a complex disease that is difficult to detect. To improve the detection and treatment of NEPC, we need to fully understand the molecular characteristics of these tumours.

1.2.1 Molecular characteristics in neuroendocrine tumours

As there are several types of NE tumours and treatment options are limited, the molecular characteristics of these tumours have arisen as a major interest in the field, mainly to identify new treatment targets for these tumours. Studies of molecular mechanisms that drive the NE differentiation of PC have been limited. Specifically, they focus on poorly differentiated neuroendocrine tumours, which represent between 1 and 5% of all prostate tumours [8, 14]. A summary of molecular features that have been detected by bulk DNA and RNA sequencing in NEPC is shown in Table 2.

At the genomic level, there is evidence that the loss of RB and TP53 induces a NE phenotype in small cell neuroendocrine carcinoma (SCNE) [15]. RB and TP53 are tumour suppressors with a major role in cell cycle regulation and tumour progression. PTEN is another tumour suppressor gene that one of the most frequently inactivated genes in PC and play a role in the progression of the NEPC phenotype [15-17].

The Aurora serine-threonine kinases play an important role for cell cycle regulation. Aurora-A (AURKA) is involved in the regulation, function and maturation of the centrosome. This kinase has a role in the formation and function of the bipolar spindle, and in cytokinesis [18]. In cancer, the overexpression of AURKA causes aneuploidy and genomic instability, these mechanisms are well known for their role in the pathogenesis of malignancy for many types of cancers [19]. AURKA has been associated with several key targets involved in tumorigenesis; some of the key functional interactions include MYCN, AKT and p53 [20]. In NEPC, AURKA is highly expressed, suggesting a role in the development of NEPC [21, 22].

The proto-oncogene MYCN is a multifunctional transcription factor that encodes a nuclear phosphoprotein involved in cell cycle progression and apoptosis. In neuroendocrine prostate cancer, the overexpression of AURKA and MYCN induces the expression of neuroendocrine markers neuron-specific enolase (NSE) and synaptophysin (SYN) [23].

Another relevant gene involved in NEPC is RE-1 silencing transcription factor (REST). It functions as a master transcriptional repressor of approximately one thousand eight hundred neuronal genes in non-neuronal cells to limit the expression of neuronal genes [24]. REST plays a role in prostate cancer as a tumour suppressor; aberrant function of REST has been associated with NE phenotype. More importantly, when REST is supressed, NE genes (CHGA, SYN, ASCL1, SNAP25 and SRRM4) have been found to be overexpressed in CRPC metastases [25].

In NEPC, neuroendocrine transcription factors have been detected to be overexpressed, suggesting they play a role in tumour pathogenesis and progression. The ASCL1 transcription factor plays a critical role in neuronal differentiation, acting as an initiating transcription factor by accessing closed chromatin to allow other factors to bind and activate neural pathways [26].

Another relevant NE transcription factor is SOX2. This transcription factor is considered a crucial regulator in sustaining the pluripotency and self-renewal properties of embryonic stem cells [27]. Recent studies have indicated that SOX2 could potentially drive the downregulation of adenocarcinoma markers during the progression of NEPC [23, 27]

Table 2. Characteristics of molecular components in PC with NE.

Molecular components	Characteristics	Implication	REF
AURKA	Overexpression	Modulation of the neuroendocrine phenotype	[20]
MYCN	Overexpression	Modulation of the neuroendocrine phenotype	[20]
REST	Under-expression	Increased NE differentiation	[28]
Rb	Loss	Development of small cell carcinomas	[15]
ASCL1	Overexpression	Induced the expression of SOX2 and SYP	[26]
SOX2	Overexpression	Driver gene involved in the upregulation of NEPC specific genes	[27]
PTEN	Deletion at 10q23	Downregulation of AR signalling	[15]
TP53	Loss	Accelerates the onset of neuroendocrine marker expression	[15]

There has been several but limited attempts to categorize NEPC by pathology. A recent study from Labrecque et. al. [14] performed a whole transcriptome profiling by RNA sequencing (RNASeq) of 98 metastatic CRPC (mCRPC) tumours from 55 patients. The initial histology and immunohistochemistry assessment in their cohort identified adenocarcinomas, amphicrine tumours; small cell neuroendocrine pathologies and tumours lacking detectable expression of AR, PSA, CHGA, and SYP, classified as double-negative PC (DNPC).

Expression of the AR, NEURO I (REST-repressed neuronal factors) and NEURO II (NE-associated transcription factors) gene sets were used to assess the mCRPC tumours. They detected that adenocarcinomas can be segregated by AR responsive markers from the AR gene set. Amphicrine tumours expressed AR-associated genes and NEURO I, but lacked expression of NEURO II. Small Cell neuroendocrine PC tumours lacked AR expression and signalling but expressed both the NEURO I and NEURO II genes.

Previous studies have identified molecular features of NEPC, mainly in poorly differentiated carcinomas but haven't been able to identify the molecular intra and inter heterogeneity of each of the NE pathologies. Further research needs to be done to elucidate the heterogeneity of NEPC because it can lead to a more suitable therapeutic option.

1.3 Single-cell RNA sequencing

Previous approaches used to assess the genomics of NE tumours have been done using bulk RNA and DNA sequencing technology [14, 27, 29]. The complex heterogeneity of NEPC is masked when bulk tumour tissues are used to study the unique molecular characteristics of the NE tumour cells. Single-cell analyses provide insights that reveal new biological traits that previous bulk RNA analysis could not achieve. Single-cell RNA sequencing (scRNA-seq), can detect complex and rare cell populations, expose regulatory interactions between genes, and track the trajectories of distinct cell lineages in development [30].

Single-cell isolation is the first step for obtaining transcriptome information from an individual cell. One of the methods that is increasingly being used to isolate cell is micro droplet-based microfluidics [31], which allows the distribution of unique cells in aqueous droplets in a continuous oil phase. The platform from 10X Genomics offers high-throughput profiling of 3' ends of RNAs of single cells with high capture efficiency [32]. Therefore, this method allows the analysis of rare cell subtypes in a diverse biological setting.

The steps necessary for the generation of scRNA-seq libraries include cell lysis, reverse transcription into the first-strand cDNA, second-strand synthesis, and cDNA amplification [33]. Recently, there has been an incorporation of unique molecular identifiers (UMIs). With this approach, each read can be allocated to its original cell, successfully removing PCR bias and consequently improving accuracy [34].

Once the cells have been sequenced, quality control (QC) is performed. The first step of QC can be done with the FastQC tool. FastQC examines distributions of quality scores across the entire reads. The next step is the alignment of the reads to a reference genome. There are several tools available for this process, including the Burrows-Wheeler Aligner (BWA). After alignment, reads are quantified using tools like Salmon or HTseq. High mapping quality reads are used to create a gene expression matrix with the number of transcripts in each gene for each cell.

The next step is normalization. This process is essential to remove cell-specific biases, which can affect downstream analysis (e.g., differential gene expression) [35]. After normalization, we can estimate and remove the confounding sources of variation. In scRNA-seq experiments,

cells from one condition are captured and sequenced. Therefore, batch effects, unrelated systematic variances resulting from sample preparation conditions, are often noticeable. Furthermore, in addition to technical noise, biological variables (e.g., state, cell cycle, and apoptosis) may affect gene expression profiles [35]. Once we remove the different unwanted sources of variation, multiple analyses like differential gene expression, lineage trajectories, and cell type annotation can be performed.

Even though experimental methods for scRNA-seq are becoming more accessible to many laboratories, bioinformatics pipelines for processing raw data files are still limited. Some commercial companies provide software tools, but this area remains at an early stage. Although, some tools are rapidly growing and are becoming a gold-standard for single cell analysis. R packages like Seurat and Monocle offers a comprehensive set of functions that ranges from QC to differential expression and trajectory analysis.

Single-cell RNA sequencing technology has been used to assess the heterogeneity of different types of cancer [36]. In PC, scRNA sequencing studies have mainly been on cell lines and circulating tumour cells [37, 38]. Horning et.al., investigated PC cell subpopulations based on the transcriptome profiling of 144 single LNCaP cells treated or untreated with androgen after cell-cycle synchronization [38]. This study provided insight into how prostate cancer cells respond heterogeneously to androgen deprivation therapies and uncovered characteristics of subpopulations that are resistant to this treatment. Another study using single cell was done by Lambros et. al., they analysed liquid biopsies obtained by apheresis to increase circulating tumour cells yield from patients with metastatic prostate cancer. This project allowed the deconvolution of intra-patient heterogeneity and clonal evolution in adenocarcinoma in PC. Nevertheless, the amount of publications of scRNA sequencing in prostate cancer is still very limited. The field would benefit from more studies using PC tumours. To our knowledge there isn't any publication that investigates NE pathologies with this technology, mainly because of the difficulty of obtaining enough viable cells from fresh tissue.

Single-cell RNA sequencing can be used to elucidate unknown tumour features that cannot be detected trough bulk transcriptomic assay, although the success of this technology relies in the quality and quantity of the input material.

1.4 Patient-derived Xenograft (PDX) models

It is ideal to use fresh tumour material, but we can only obtain limited material from biopsies which are mainly used for patient diagnosis. This is an obstacle to obtaining enough fresh tumour material to perform single-cell RNA sequencing. Furthermore, specimens obtained are contaminated with benign tissue and tumour cells, and at the time of collection, we are blinded to that composition.

Patient-derived xenografts (PDXs) offer another approach to investigate these tumours. Several of the current models (e.g., mice, cell lines) do not represent the disease progression and the complexity of NEPC. Additionally, with PDXs it is more feasible to harvest and isolate single cells, when compared with tumour biopsies.

PDXs are models of cancer where tumour samples obtained from patients are implanted into immune-deficient mice. This model is considered the gold standard for a pre-clinical model in prostate cancer because:

a) Accurately represent main features of the patient disease: it will maintain the molecular, genetic, and histological heterogeneity of tumours of origin through serial passaging in mice.[39].

b) Readily obtainable material: It offers an infinite source of fresh tissue samples, which translates into a higher quality of the start material.

c) Intact tissue architecture: PDX maintain intercellular signalling and physiological reactions that occur in the tissue's original environment, specifically cancer-associated fibroblasts, which plays an important role in prostate cancer progression. [39].

Our group has developed a method to enable the derivation of PDX lines of PC [40, 41]. They collected 109 tumour samples from 29 patients via the Melbourne Urological Research Alliance (MURAL) platform. This generated serially transplantable PDXs, 10 of which correspond to NE pathologies, including poorly differentiated carcinomas, amphicrine tumours and mixed NE with adenocarcinoma. Genomic profiling with low-coverage whole-genome sequencing and targeted deep sequencing of 623 cancer-related genes demonstrated that the genomic features of the original tumours were maintained by the

corresponding PDXs at early and late generations. These lines thereby allow the study of NEPC intra and inter tumour heterogeneity, cellular interactions, and individual tumour responses to therapeutic compounds [42].

1.5 Project significance and hypothesis

Currently, NEPC is an aggressive pathology with limited treatment options and poor prognosis. To provide better treatment options, we need to understand the molecular heterogeneity of neuroendocrine pathologies encountered in prostate cancer. This project aims to investigate the intra and inter heterogeneity of neuroendocrine prostate cancer, using genomic and transcriptomic approaches.

This project will contribute by performing a comprehensive assessment of the transcriptional heterogeneity in NEPC that has not been possible before, exploring the different NE pathologies, their microenvironments, and spatial heterogeneity that could potentially guide future pre-clinical studies.

Hypothesis

To effectively treat NE tumours we need to fully understand the molecular heterogeneity of NE prostate cancer.

I will address this hypothesis through three aims:

Aim 1: Determine the transcriptomic intra and inter heterogeneity of neuroendocrine pathologies.

Due to the poor molecular characterization of NEPC, scRNA-seq technology will be used to determine the intra and inter tumour heterogeneity of the multiple pathologies of NEPC. A PDX model will serve as a pre-clinical model to assess the different NE pathologies. This approach will allow us to reveal the degree and nature of transcriptional intra and inter tumour heterogeneity of each NE pathology that previous bulk RNA analysis could not achieve.

Aim 2: Determine the interactions between neuroendocrine tumours and their microenvironment.

Cancer-associated fibroblasts (CAFs) have an important role in tumour microenvironment, but there is limited information about CAFs in the microenvironment of NE pathologies. We will use the CAFs from PDX tumours of the single-cell RNA sequencing data to study the interactions between different NE pathologies and their microenvironment. This approach

will provide a sight into the different CAF populations and their interactions with the different NE pathologies.

Aim 3: Characterization of the evolutionary relationship and genomic heterogeneity of neuroendocrine pathologies in prostate tumours.

In some patients with NEPC, multiple NE pathologies have been detected within the same tumour. The relationship between these pathologies at a molecular level has not been demonstrated genetically. We will sequence the exome of different regions of the same FFPE tumour from different patients with different NE pathologies to determine the evolutionary relationship and genomic heterogeneity in NEPC. This approach will provide an inferred phylogeny of each of the NE, elucidating the evolutionary trajectories of NEPC and whether or not they share a common origin.

Section 2: Aims and summary of progress

Aim 1: Determine the transcriptomic intra and inter heterogeneity of neuroendocrine pathologies.

To determine the transcriptomic heterogeneity of neuroendocrine tumours, single-cell sequencing technology, and bioinformatics toolkits will be used. This approach will allow us to unmask the distinct cell populations of NEPC. Due to the limitations in obtaining fresh specimens of NEPC we will use novel PDX tumour lines which represent the different NE pathologies found in PC. The PDX model will offer an accurate representation of the NE tumours, providing us with readily available material that is amenable to single-cell technology. This approach will allow us to explore the intra and inter tumour heterogeneity across pathologies and the interaction of the distinct cell types encountered.

This aim is comprised of the following sub aims.

1.1 Perform single-cell RNA sequencing on PDX models of NE prostate cancer

1.2 Establish filtering strategies to analyse the NE tumours.

1.3 Characterize the transcriptional profiles of NE cells within neuroendocrine pathologies (Intra-tumoral heterogeneity).

1.4 Compare transcriptional profiles of NE cells from different patients (Inter-tumoral heterogeneity)

1.5 Identify the interactions between adenocarcinoma cells and NE cells in tumours with mixed cell types

Summary of progress: Aim 1

I have completed a preliminary analysis of four out of the ten NE tumours available from our PDX collection. These PDXs include two samples of adenocarcinoma with neuroendocrine differentiation, an amphicrine tumour and a tumour with both Small Cell NE (SCNE) and Adenocarcinoma cell populations.

Results from PDX 224, the tumour with SCNE and Adeno cell populations are described below. To our knowledge, this is the first time that NE prostate tumours have been successfully analysed using single-cell RNA sequencing technology.

Aim 1.1 Perform single-cell RNA sequencing on PDX of NE prostate cancer

The tumours were harvested and an enzymatic dissociation into single cells was performed using a method developed by Dr. Roxanne Toivanen. Then we used the chromium 10X platform to separate and encapsulate the cells, single-cell sequencing was performed in the Single Cell Innovation Lab (SCIL) by Dr. Luciano Martelotto. After that, the transcriptome of each cell was sequenced in a NovaSeq from Illumina.

Once we had the cells sequenced, we have to identify the tumour cells from the stroma cells from the mouse. Briefly, the tool XenoCell [43] was used to perform the alignment for both mouse (mm10) and human (GRCh38) genome references. The aligned sequences were then separated by species of origin (Figure 2.1). Alevin generates a cell-by-gene count matrix for each species [44]. This input data is used for downstream analysis.

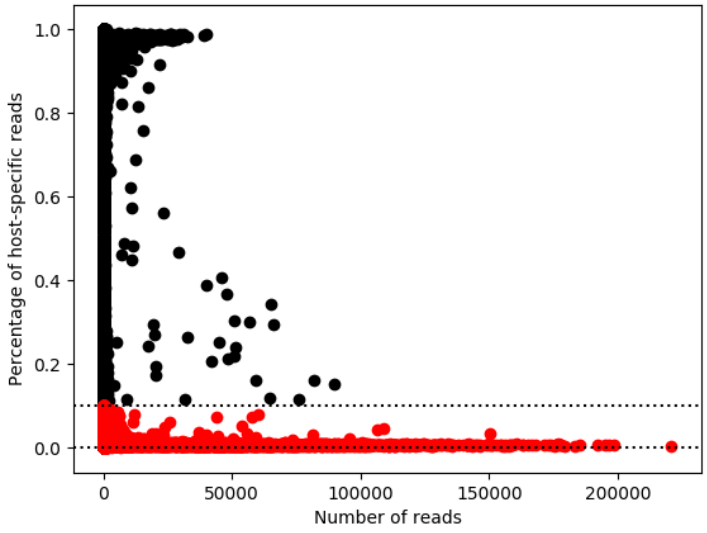


Figure 2.1 Human cells filtered by Alevin. Alevin filters all host cells by separating them base on the percentage of host specific reads, all cells below 10% have high probability of being human cells conforming the tumour. Cells highlighted in red were selected as tumour cells and carried forward for analysis

Aim 1.2 Establish filtering strategies to analyse the NE tumours

One of the key challenges is to ensure that only single and live cells are included in the downstream analyses. The inclusion of compromised cells inevitably affects data interpretation.

The existing available protocols often result in the captured cells being stressed, broken, or killed. Moreover, some droplets produced by the 10X cell capture system can be empty and some may contain multiple cells. We refer to all such cells as low-quality. These cells can lead to misinterpretation of the data and therefore need to be excluded. There is no standard method in regards to filtering single-cell data, therefore I had to develop a specific filtering strategy for the NEPC tumours (Figure 2.1).

Metrics used to filter cells:

For low-quality cells:

- Cells with a low number of gene counts. The number of unique genes detected in each cell has to be considered because low-quality cells or empty droplets will often have few genes. A threshold of < 1000 genes per cell was set for all samples (Figure 2.2 B).
- Cell doublets or multiplets may exhibit an aberrantly high gene count; hence we removed any cell that had a transcript count over 40,000. This parameter has to be adjusted in each sample. (Figure 2.2 B).
- Dying cells or low-quality cells will often exhibit extensive mitochondrial contamination; so the percentage of reads that map to the mitochondrial genome is used to filter out such cells. Here, a <25% threshold of the percentage of mitochondrial reads was considered for the SCNE with adenocarcinoma sample (Figure 2.2 A). This parameter needs to be adjusted in every sample.

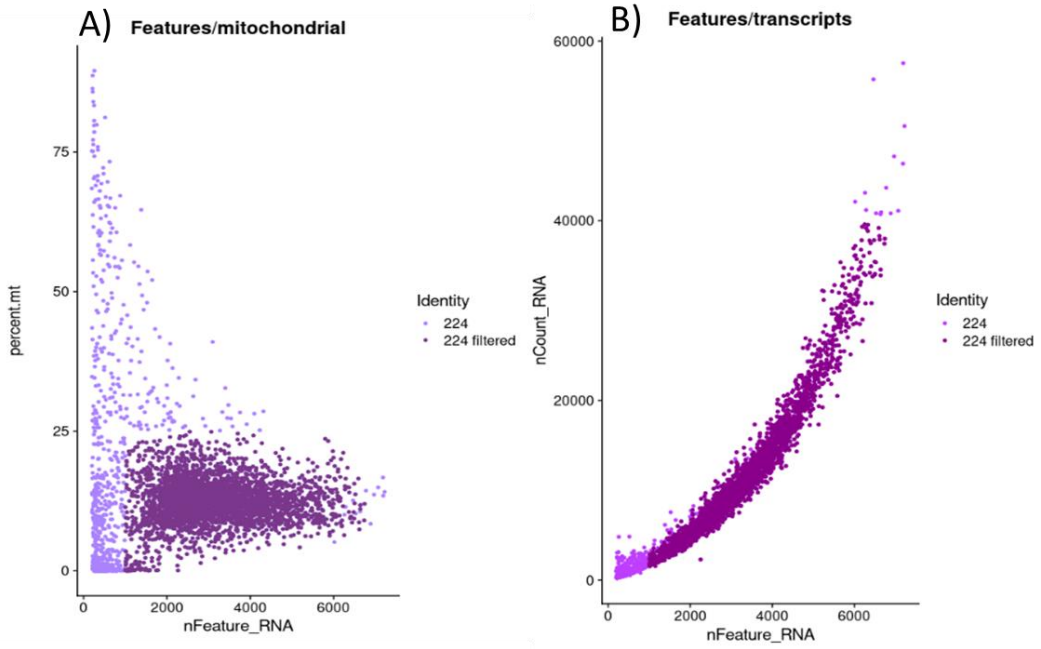


Figure 2.2 Quality control plots used to select filtering thresholds. QC parameters are shown, cells removed by filters are coloured with a lighter tone. A) The correlation between the number of features and the percentage of mitochondrial expression, B) The correlation of feature counts (genes) and transcript count (UMI's)

After removing low-quality cells, we retain 78% of our cell population (Table 2.1). In other studies, the percentage of cells retained is between 60 % and 90 %, depending on the tissue type [45]. This indicates that our cells comply with the quality required for downstream analysis.

Sample ID	Type of tumour	QC	Cells	% of cells retained	Total Features (Genes)
387	Amphicrine Carcinoma	Unfiltered	3130	74.08	15572
		Filtered	2319		14928
272	Adenocarcinoma with NED	Unfiltered	3196	78.97	15796
		Filtered	2524		15121
369	Adenocarcinoma with NED	Unfiltered	2897	71.10	15501
		Filtered	2060		14943
224	Mixed Adenocarcinoma - Small cell NE	Unfiltered	3997	78.95	15934
		Filtered	3156		15367

Table 2.1 QC of tumours analysed through single-cell RNA sequencing. The table depicts the number of cells and genes that remained after QC was done in each sample. The raw data from the scRNA sequencing was processed through our pipeline, which allowed us to identify high quality tumour cells. We obtained cells that have passed quality control and genes that can be detected in each of these cells.

The data was then normalised using the “SCTransform” function in Seurat. In this step, we subtracted (“regressed out”) sources of unwanted heterogeneity. The main sources of unwanted variation that were subtracted were the mitochondrial mapping percentage, ribosomal mapping percentage, and the effects of the cell cycle phase in each cell. The main goal of this process is to remove the influence of technical effects in the underlying molecular counts while preserving true biological variation.

A principal component analysis (PCA) was performed with the normalised counts; this method is used to compress a robust dataset into main components and identify unique populations of cells. Then, we look into each of the principal components and select the highly significant components to cluster the cells.

We then visualize the datasets using a nonlinear dimensional reduction method. The goal of these algorithms is to learn the underlying features of the data in order to place similar cells together in low-dimensional space. Cells clustered on previous steps will co-localize using the dimension reduction technique, Uniform Manifold Approximation and Projection (UMAP) that helps visualize our single cell data set. The UMAP of the SCNE with Adeno sample correlates with the PCA, two main clusters were detected, however within one cluster, three sub-clusters could be observed (Figure 2.2).

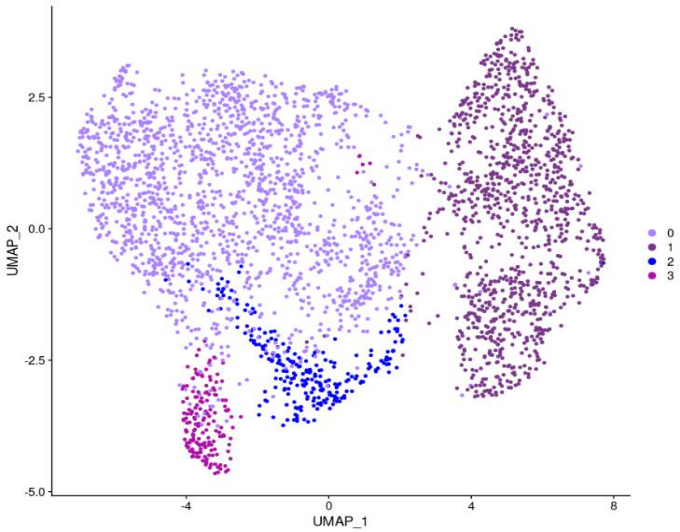


Figure 2.2 UMAP of SCNE sample. A dimensional reduction technique to visualize and explore the dataset. Small cell neuroendocrine mixed with adenocarcinoma sample is represented in the UMAP. Nine dimensions were selected as main components. Two major clusters are

formed, the cluster on the left is divided by 3 sub clusters. The UMAP suggest two major populations of cells, one of those are comprised by 3 subpopulations.

NE is a highly heterogeneous pathology and there are many biomarkers used to detect NE components, such as CHGA, SYN, NSE, and CD56. Nonetheless not all cells will necessarily express the same NE markers simultaneously. In the SCNE with Adeno sample, we use these markers to detect the landscape of the expression of markers in the cells (Figure 2.3). Although this approach can identify adenocarcinoma and NE cell populations, the distinction between NE subpopulations cannot be assed trough individual NE markers.

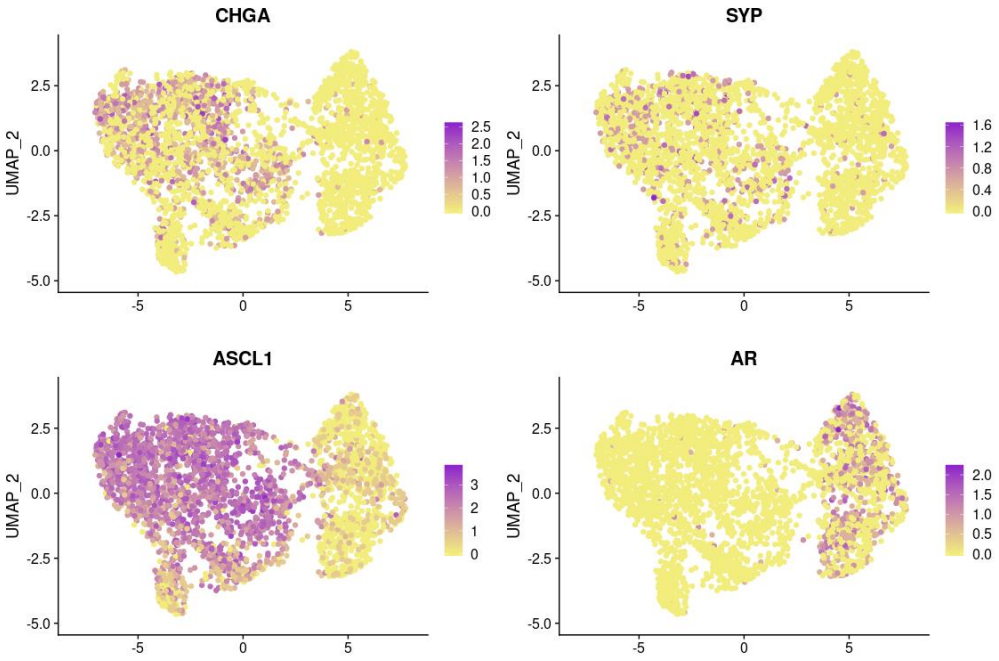


Figure 2.3 UMAP with NE and Adenocarcinoma markers in the SCNE sample. CHGA, SYP and ASCL1 were used to detect NE clusters. Regions in clusters 0, 2 and 3 were enriched for ASCL1, SYP and CHGA suggesting these are neuroendocrine clusters. Cluster 1 expressed AR, suggesting the presence of adenocarcinoma cells.

Aim 1.3 Characterize the transcriptional profiles of NE cells within neuroendocrine pathologies

Although identifying main markers per cluster allowed us to distinguish between NE and Adeno, this approach did not allow us to clearly distinguish the differences between NE clusters. The presence of sub clusters within the NE population suggested heterogeneity in the NE tumour cells. To characterize this heterogeneity we applied the previously described

Labrecque gene signatures. These signatures were used for molecular classification of bulk prostate tumours with adenocarcinoma, small cell neuroendocrine carcinoma, and amphicrine tumours [14]. We annotated the clusters based on the predominant signature in each (Figure 2.4a) and named them accordingly.

The results showed that cluster 1 expressed a signature for androgen signalling (Figure 2.4b), which was high in AR signatures, confirming that this cluster is composed of adenocarcinoma cells. The other 3 sub clusters expressed NE signatures, suggesting these are NE cells. However, each sub cluster expressed higher proportions of different NE signatures, confirming molecular heterogeneity (Figure 2.4b). Cluster 0 predominantly expressed the signature (SYP, CHGA, ASCL1; NEURO I that involves the suppression of the RE1 silencing transcription factor (REST), which normally suppresses NE expression. Cluster 2 expressed NE transcription factors (SOX2, NKX2-1; NEURO II), and cluster 3 expressed squamous markers (KRT6B and FGFBP1; SQUAM) typically used to identify double negative cells that don't express either AR or NE markers. The squamous signature clearly predominates in this cluster which explains why it is segregated from the other NE clusters.

The rest of the PDX samples were analysed in the same way, with similar results demonstrating heterogeneity was detected in each of the samples.

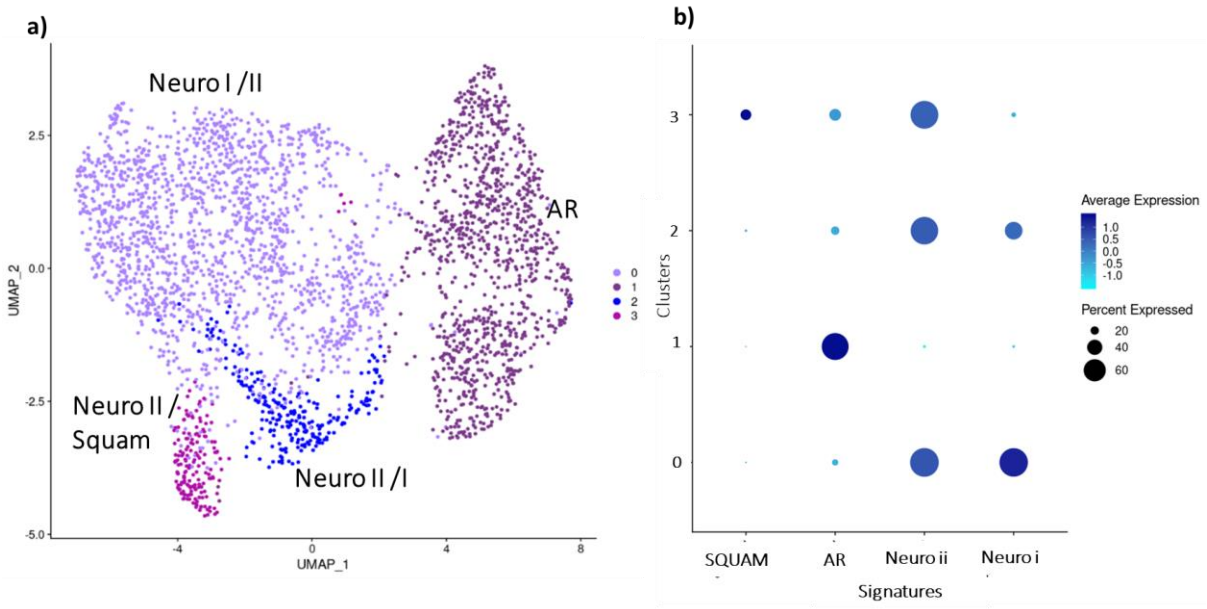


Figure 2.4 Small cell NE carcinoma signature expression. a) UMAP depicting the 4 clusters in SCNE with signatures annotation b) Dot plot representing the expression of each signature. The colour intensity represents the average expression of signatures whereas the size of the circle represents the percentage of cells that expressed the signature. NEURO I and II are predominant in the NE sub clusters, although it's the SQUAM signature segregating cluster 3

from the rest of the NE sub clusters. In the AR responsive panel, there is a presence of a low AR region, suggesting some differences between the adenocarcinoma cells.

Summary and future directions for Aim 1

To our knowledge, this is the first time that NEPC has been studied at a single-cell level. This preliminary result demonstrates that it is possible to perform single-cell RNA sequencing on these types of tumours with the quality and resolution needed to investigate the heterogeneity of NE prostate cancer. Furthermore, this result reveals that there are several subpopulations in NE tumours, implying that there is an intra tumour heterogeneity in these pathologies. To further elucidate the tumour heterogeneity, the next step would be to integrate all NEPC pathologies into one assay and perform a differential gene expression analysis through Seurat, which will allow us to determine the inter tumour heterogeneity of the NE pathologies. The interactions between NE cells and adenocarcinoma in mixed pathologies will be inferred using CellphoneDB, a tool that predicts enriched cellular interactions between two cell types from single-cell transcriptomics data [46].

The diverse molecular features and interactions of each subtype will provide valuable information to detect potential drug targets.

Aim 2: Determine the interactions between neuroendocrine tumours and their microenvironment

The second aim will look at the NE tumour microenvironment. Tumours consists not only of mutated cancer cells, but also non-mutated stroma cells, including endothelial cells, immune cells, and activated fibroblasts, also known as cancer-associated fibroblasts (CAFs) [47]. CAF populations often make up the majority of the stroma cell population and contribute significantly to cancer growth, metastasis, immune suppression, and drug resistance [48].

CAFs have an important role in the microenvironment, but there is limited information on the role of CAFs in NE PC.

I will further analyse the single-cell RNA sequencing data generated in Aim 1 to study the interactions between different NE pathologies and their microenvironment. Because we are using PDXs we can cleanly separate stromal from tumour cells by segregating cells by species of origin.

Sub aims

2.1 Establish filtering strategies to analyse the stroma from single-cell data
2.2 Characterize the stromal components in different neuroendocrine pathologies.
2.3 Identify the interactions between NE cells and cancer associated fibroblasts (CAFs)

Aim 2.1 Establish filtering strategies to analyse the stroma from single-cell data

From the previously obtained single-cell data, we can successfully separate the stroma mouse cells from the human cells using Xenocell (Figure 2.6). Cells that had over 90%, mouse-specific reads will be classified as mouse cells. These cells represent the stroma. Alevin generates a cell-by-gene count matrix for the mouse stroma cells. This input data will be used for downstream analysis. We will use Seurat to establish a filtering strategy for low-quality cells in the stroma, with the same approach previously described on aim 1.2.

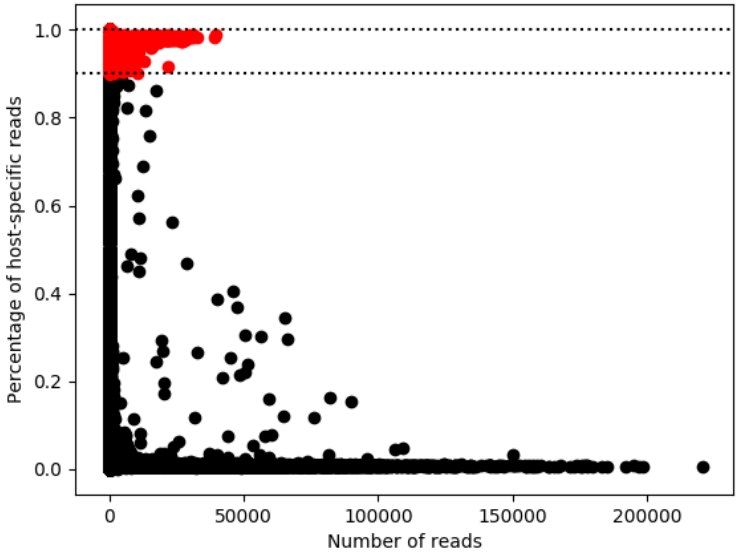


Figure 2.6 Mouse cells filtered by Alevin. Alevin filters all host cells by separating them based on the percentage of host-specific reads. All cells above 90% are highly probable of being stroma cells. Cells highlighted in red were selected as stroma cells.

Aim 2.2 Characterize the stromal components in different neuroendocrine pathologies

The pipeline to analyse the stroma cells is currently under construction. The same approach that was used for the tumour cells will be used for the stroma cells. Once we establish the pipeline, we will focus on characterizing the components of the tumour microenvironment in different NE pathologies.

Aims 2.3 Identify the interactions between NE cells and cancer associated fibroblasts (CAFs)

To identify the interactions between NE cells and cancer-associated fibroblasts (CAFs) we will compare the expression of receptor and ligand pairs between the NE cells and the CAFs. We will use CellphoneDB, a tool that combines a repository of ligand-receptor interacting pairs with a statistical framework that predicts cell-to-cell interactions. This tool will allow us to analyse the crosstalk between the CAF components of the stroma and the NEPC cells.

Aim 3: Characterize the evolutionary relationship and genomic heterogeneity of neuroendocrine pathologies in prostate tumours.

The presence of different NE pathologies mixed with adenocarcinoma in the same patient suggests that these tumours have a common genetic origin, although this remains unknown and untested. To fulfil this aim, we selected five PC patients who have multiple adenocarcinoma and NE pathologies within their prostate tumour. Starting from FFPE tissue we are going to extract the DNA to test the relationship of these pathologies. We will macro dissect the different lesions with their paired normal sample and characterize their genomic features and relationship using targeted genomic sequencing and bioinformatics tools to analyse the intra tumour heterogeneity and the relationship between them.

Sub aims

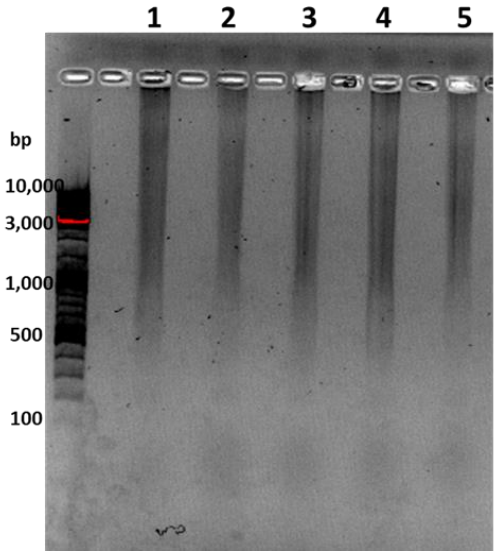
3.1 Establish a workflow to obtain with DNA for exome sequencing from FFPE samples.
3.2 Identify the genomic alterations within regions harbouring different neuroendocrine pathologies within the same patient.
3.3 Identify the evolutionary relationship between different NE pathologies in the same patient.

Aim 3.1: Establish the methodologic workflow to obtain enough material with enough quality to perform exome sequencing

To establish the methodologic workflow to obtain enough DNA to perform exome sequencing, we test the QIAamp DNA FFPE Tissue kit. We used FFPE tissue from our NEPC PDX lines. We macro dissected different regions from the same patient. None of them exceeded 10 µm thickness per section. We used the QIAamp DNA FFPE Tissue Kit and processed the samples following the manufacturer's instructions. DNA was quantified by a fluorescence-based assay (Qubit) which we obtained a sufficient amount of DNA for library preparation (Table 2.2). Nucleic acids were observed on an electrophoresis gel (Figure 2.7). Fragments of DNA observed from 500 pb to over 10,000 pb, typical from FFPE tissue. The DNA yield and fragment sizes obtained meet the requirements for library preparation for whole-exome sequencing. We intend to start processing the samples once the COVID-19 restrictions are lifted.

Sample ID	ng/uL	Total ng (50ul)
PDX 272 sample 1	30.2	1,510
PDX 27 sample 2	22	1,100
PDX 272 sample 3	16.3	815
PDX 272 sample 4	35.6	1,780
PDX 27 sample 5	48.6	1,430

Table 2.2 DNA quantification of FFPE test samples. Different methods of dissection were used for each test sample. The sample with the smallest amount of tissue was 272 sample 3. Regardless of that sufficient DNA was obtained from the sample.



Sample ID	ng/uL	Total ng (50ul)
PDX 272 sample 1	30.2	1,510
PDX 27 sample 2	22	1,100
PDX 27 sample 3	16.3	815
PDX 27 sample 4	35.6	1,780
PDX 27 sample 5	48.6	1,430

Figure 2.7 Electrophoresis gel from FFPE tissue. 1% agarose gel in TBE was used to run the assay. A 10kb ladder was used as a fragment size control. The quantity of DNA used per sample was ~80 ng. Each sample had fragments ranging in size from 500 pb to over 10, 000 bp.

Aim 3.2 Identify the genomic alterations of the different neuroendocrine pathologies

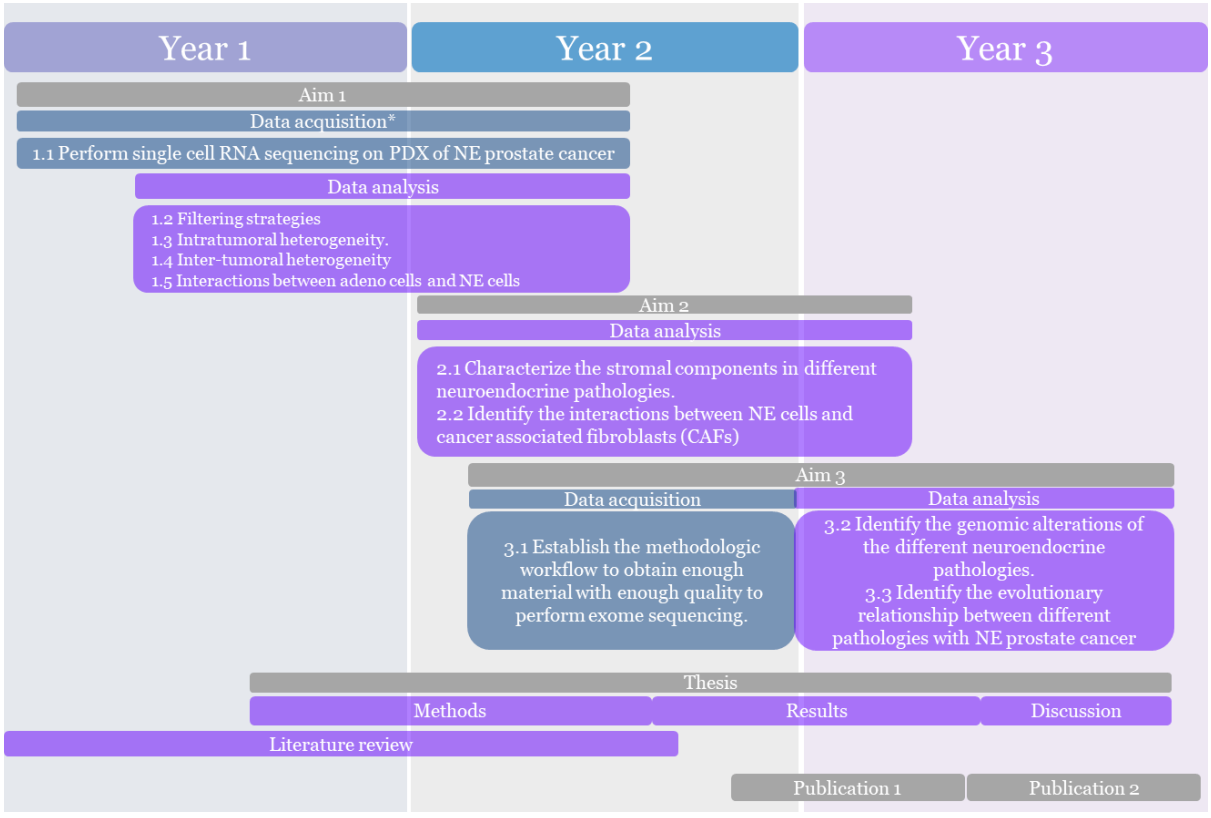
The alignment to the reference genome of exome sequences will be done by BWA, duplicates will be marked with Picard. Copy Number Alterations will be inferred using CNVkit. Genome Analysis Toolkit will be used for local realignment and base quality recalibration. Variant calling will be done with Mutect2 from GATK, designed to call somatic variants only, germline variants will be excluded. Then, we will annotate the variants to identify their biological relevance with an annotator such as Oncotator or ANNOVAR. Next, we will filter and manually

curate mutations. The first exclusion criteria will be on mutations that are synonymous, with depth <5.0× or have a mutant allele fraction <1%. Then, to capture potentially pathogenic mutations, we will filter out variants with a minor allele frequency <0.001 in gnomeAD database, 1000 Genomes (1000G) project or the Exome Sequencing Project (ESP6500). All splicing and null variants (stop-gain/loss, frameshift indels) and missense variants defined as pathogenic in ClinVar or annotated in COSMIC will be retained. These filtering criteria will allow us to identify the genetic alterations to generate a mutational profile of the different neuroendocrine pathologies.

Aim 3.3 Identify the evolutionary relationship between different pathologies with NE prostate cancer

Once we know the mutations within these tumours we can identify if there is an evolutionary relationship between different pathologies. Phylogenetic trees will be constructed using the tool PhyloWGS, which utilizes the combination of variant allele frequencies of point mutations and copy number alterations for accurate sub clonal reconstruction [49].

Section 3: Research Plan



Section 4: Research Outputs

August 2019- Comprehensive Cancer Student Symposium- Beyond FISH and ChIPs: the next generation of cancer research.

- Poster presentation
- Title: Profiling neuroendocrine differentiated cells in prostate cancer

Section 5: Attendance

- **Induction and Orientation program**

Peter Mac New Employees:

  - 100 Welcome Aboard! Orientation for new employees
  - Facility Induction
  - Emergency Procedures
  - Occupational Health & Safety (Employees)
  - Laboratory Safety
  - Cluster and Laboratory induction
  - Spill training

- **Specialist laboratory skills training**
  None
- **Hub, Lab meetings and Journal clubs**
  - Attend to the weekly Bioinformatics meeting from the Gail Risbridger group
  - Attend to the weekly Gail's Risbridger lab meetings
  - Attend to all David's Goode lab meetings
- **Seminar and workshop programs**
  - Seminar: WEHI Molecular Medicine. Long-read single-cell RNA-sequencing with 10X and nanopore sequencing
  - Research Seminar Series: 'A big data approach to microbiome-based therapeutics and diagnostics'
  - WEHI Bioinformatics Seminar: Portable Pipelines Project: Developing reproducible bioinformatics pipelines with standardised workflow languages
  - University of Melbourne: Centre for Cancer Research Seminar Series. 'Single-cell genomics - Recent advances and future directions'
  - WEHI Bioinformatics Seminar: Tools for preprocessing and benchmarking single cell RNA-sequencing data
  - WEHI Bioinformatics Seminar: Impact of gene annotation choice on the quantification of RNA-seq data by David Chisanga
  - Single Cell Research User Meeting (SCRUM): Advances, Present and Future of Single Cell Genomics
  - Grand Round VCCC Breast Tumour Stream. 'The Polygenic Risk Score for breast cancer – clinical implications and implementation'. Associate Professor Paul James: Director, Parkville Familial Cancer Centre
  - Next Generation Sequencing and its clinical applications for cancer patients related with immunotherapies presented by A/Professor Mingzhi Ye, Director, BGI Genomics R & D Center, China

- WEHI Bioinformatics Seminar. By Sepideh Foroutan from University of Melbourne Centre for Cancer Research. Using transcriptomic data to predict drug efficacy in breast cancer
- Multidimensional single cell analysis of the tumour microenvironment. Associate Professor Alex Swarbrick
- 'Integrative analysis of cancer transcriptomes'. presented by Professor Eduardo Eyras: EMBL Group Leader and Professor at Australian National University, Canberra
- PhD Completion Seminar. 'Uncovering dual roles for RYK, a WNT-binding receptor tyrosine kinase, in cancer' presented by James Roy: PhD Student
- WEHI Bioinformatics Seminar. by Anna Trigos from PeterMac. Pre-spatial transcriptomics: integrating multiplex immunohistochemistry (OPAL) with gene expression to study the immune microenvironment of prostate cancer
- PhD Completion Seminar. 'Defining functional drivers of oesophageal tumourigenesis'. presented by Jovana Gotovac: PhD Student
- WEHI Bioinformatics Seminar SVEnsemble: an algorithm for ensemble structural variant calling using re-evaluated quality scores via probabilistic random forest by Ruining Dong
- WEHI Bioinformatics Seminar by Gordon Smyth. XenoSplit: distinguishing human from mouse in RNA-seq from mouse xenograft models
- PhD Confirmation Seminar. 'Circulating tumour DNA for precision medicine in non-small cell lung cancer' presented Sebastian Hollizeck
- WEHI Bioinformatics Seminar by Luke Gandolfo. Sample weights for count based differential expression analysis
- Workshop: Single-cell RNA-seq analysis. Hosted by Melbourne Integrative Genomics (MIG)
- Workshop: RNA-seq Differential Gene Expression analysis in R
- Workshop: 10X single-cell RNA-seq analysis workshop
- Workshop: Introduction to Linux and HPC for biologists
- Workshop: Advanced Linux and Shell Scripting for biologists

- **Annual Student Retreat**
  - The Annual Student Retreat 2019
- **Topics in Cancer program**
  - *Cancer genome overview*: Professor Sean Grimmond, University of Melbourne Centre for Cancer Research
  - *Breaking bad: lessons on genome integrity from study of inherited cancer syndromes*. Associate Professor Andrew Deans, St Vincent's Institute of Medical Research
  - *Translating genetic discovery into clinical practice*. Professor Ingrid Winship, Melbourne Health and University of Melbourne
  - *Learning from pan-cancer genome studies*. Dr Ann-Marie Patch, QIMR Berghofer Medical Research Institute
  - *When the sum is so much greater than the parts*: *uncovering genetic predisposition to cancer.* Professor Georgia Chenevix-Trench, QIMR Berghofer Medical Research Institute
  - *Cancer molecular pathology: an anatomical pathologist's perspective*. Professor Stephen Fox, Peter MacCallum Cancer Centre

  - *Blood-based biomarkers: circulating tumor DNA analysis*. Professor Sarah-Jane Dawson, Peter MacCallum Cancer Centre
- **Research Seminar program**
  - CC PhD Program Grand Debate. Topic: Thankfully- basic science research in Australia is dead**.**
  - 'Single cell profiling of breast tissue refines lineage relationships and reveals rare cell subtypes in the epithelial compartment' presented by Dr Bhupinder Pal: Head, Cancer Cell Genomics Laboratory, Translational Breast Cancer Program, Olivia Newton-John Cancer Research Institute, ONJ Centre
  - 'The ecology and evolution of cancer'. Dr Beata Ujvari: Senior lecturer in Bioinformatics and Genetics at the School of Life and Environmental Sciences, Deakin University
  - Research Seminar 'Transcriptomes, single cells and cancer: How bioinformatics unlocks the power of genomic technologies' presented by Dr Alicia Oshlack
- **Research lunches attended**
  - May 16. Prof Gene Tyson - University of Queensland, Areas of Research: Molecular genetics and genomics.
  - July 25. Dr. Eduardo Eyras. Computational RNA biology. John Curtin School of Medical Research, ANU
  - August 15, Shannon McWeeney, Translational Bioinformatics, Oregon Clinical and Translational Research Institute, USA
  - September 5, Alicia Oshlack Head of Bioinformatics, Murdoch's Children's Research Institute. Cell Biology and Bioinformatics
- **Research Governance program**
  - None
- **Research reporting (Presentations, written reports submitted etc)**
  - Comprehensive Cancer Student Symposium (Poster)
  - Victorian Cancer Bioinformatics Symposium (Poster)
- **Technical seminars**
  - None

Section 6: References

1. GLOBOCAN. *Estimated number of incident cases and deaths worldwide, males, all ages*. 2018 [cited 2019 21/05/19].
2. Yap, T.A., et al., *Drug discovery in advanced prostate cancer: translating biology into therapy.* Nat Rev Drug Discov, 2016. **15**(10): p. 699-718.
3. Anassi, E. and U.A. Ndefo, *Sipuleucel-T (provenge) injection: the first immunotherapy agent (vaccine) for hormone-refractory prostate cancer.* P & T : a peer-reviewed journal for formulary management, 2011. **36**(4): p. 197-202.
4. Rice, M.A., S.V. Malhotra, and T. Stoyanova, *Second-Generation Antiandrogens: From Discovery to Standard of Care in Castration Resistant Prostate Cancer.* Frontiers in oncology, 2019. **9**: p. 801-801.
5. Nuhn, P., et al., *Update on Systemic Prostate Cancer Therapies: Management of Metastatic Castration-resistant Prostate Cancer in the Era of Precision Oncology.* Eur Urol, 2019. **75**(1): p. 88-99.

6. Nevedomskaya, E., S.J. Baumgart, and B. Haendler, *Recent Advances in Prostate Cancer Treatment and Drug Discovery.* Int J Mol Sci, 2018. **19**(5).
7. Shen, M.M. and C. Abate-Shen, *Molecular genetics of prostate cancer: new prospects for old challenges.* Genes Dev, 2010. **24**(18): p. 1967-2000.
8. Fine, S.W., *Neuroendocrine tumors of the prostate.* Modern Pathology, 2018. **31**(1): p. 122-132.
9. Priemer, D.S., et al., *Neuroendocrine Tumors of the Prostate: Emerging Insights from Molecular Data and Updates to the 2016 World Health Organization Classification.* Endocr Pathol, 2016. **27**(2): p. 123-35.
10. Bellur, S., T. Van der Kwast, and O. Mete, *Evolving concepts in prostatic neuroendocrine manifestations: from focal divergent differentiation to amphicrine carcinoma.* Human Pathology, 2019. **85**: p. 313-327.
11. Beltran, H., et al., *Aggressive Variants of Castration-Resistant Prostate Cancer.* Clinical Cancer Research, 2014. **20**(11): p. 2846-2850.
12. Beltran, H., et al., *The Initial Detection and Partial Characterization of Circulating Tumor Cells in Neuroendocrine Prostate Cancer.* Clin Cancer Res, 2016. **22**(6): p. 1510-9.
13. Hu, J., B. Han, and J. Huang, *Morphologic Spectrum of Neuroendocrine Tumors of the Prostate: An Updated Review.* Arch Pathol Lab Med, 2020. **144**(3): p. 320-325.
14. Labrecque, M.P., et al., *Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer.* J Clin Invest, 2019. **130**: p. 4492-4505.
15. Tan, H.-L., et al., *Rb Loss Is Characteristic of Prostatic Small Cell Neuroendocrine Carcinoma.* Clinical Cancer Research, 2014. **20**(4): p. 890-903.
16. Robinson, D., et al., *Integrative clinical genomics of advanced prostate cancer.* Cell, 2015. **161**(5): p. 1215-1228.
17. Poluri, R.T.K. and É. Audet-Walsh, *Genomic Deletion at 10q23 in Prostate Cancer: More Than PTEN Loss?* Frontiers in oncology, 2018. **8**: p. 246-246.
18. Tang, A., et al., *Aurora kinases: novel therapy targets in cancers.* Oncotarget, 2017. **8**(14): p. 23937-23954.
19. Mehra, R., et al., *Aurora kinases in head and neck cancer.* The Lancet. Oncology, 2013. **14**(10): p. e425-e435.
20. Beltran, H., et al., *Molecular characterization of neuroendocrine prostate cancer and identification of new drug targets.* Cancer Discov, 2011. **1**(6): p. 487-95.
21. Kivinummi, K., et al., *The expression of AURKA is androgen regulated in castration-resistant prostate cancer.* Scientific Reports, 2017. **7**(1): p. 17978.
22. Beltran, H., et al., *Molecular Characterization of Neuroendocrine Prostate Cancer and Identification of New Drug Targets.* Cancer Discovery, 2011. **1**(6): p. 487-495.
23. Li, Q., C.S. Zhang, and Y. Zhang, *Molecular aspects of prostate cancer with neuroendocrine differentiation.* Chinese journal of cancer research = Chung-kuo yen cheng yen chiu, 2016. **28**(1): p. 122-129.
24. Jin, H., et al., *Identification of RE1-Silencing Transcription Factor as a Promoter of Metastasis in Pancreatic Cancer.* Frontiers in Oncology, 2019. **9**(291).
25. Zhang, X., et al., *SRRM4 Expression and the Loss of REST Activity May Promote the Emergence of the Neuroendocrine Phenotype in Castration-Resistant Prostate Cancer.* Clinical Cancer Research, 2015. **21**(20): p. 4698-4708.

26. Sinha, S., et al., *Abstract LB-199: Role of ASCL1 in neuroendocrine prostate cancer progression.* Cancer Research, 2018. **78**(13 Supplement): p. LB-199-LB-199.
27. Li, H., et al., *SOX2 has dual functions as a regulator in the progression of neuroendocrine prostate cancer.* Laboratory Investigation, 2020. **100**(4): p. 570-582.
28. Lapuk, A.V., et al., *From sequence to molecular pathology, and a mechanism driving the neuroendocrine phenotype in prostate cancer.* The Journal of pathology, 2012. **227**(3): p. 286-297.
29. Corella, A.N., et al., *Identification of Therapeutic Vulnerabilities in Small-cell Neuroendocrine Prostate Cancer.* Clin Cancer Res, 2020. **26**(7): p. 1667-1677.
30. AlJanahi, A.A., M. Danielsen, and C.E. Dunbar, *An Introduction to the Analysis of Single-Cell RNA-Sequencing Data.* Mol Ther Methods Clin Dev, 2018. **10**: p. 189-196.
31. Marcus, J.S., W.F. Anderson, and S.R. Quake, *Microfluidic Single-Cell mRNA Isolation and Analysis.* Analytical Chemistry, 2006. **78**(9): p. 3084-3089.
32. Vento-Tormo, R., et al., *Single-cell reconstruction of the early maternal–fetal interface in humans.* Nature, 2018. **563**(7731): p. 347-353.
33. Haque, A., et al., *A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications.* Genome Med, 2017. **9**(1): p. 75.
34. Nguyen, Q.H., et al., *Experimental Considerations for Single-Cell RNA Sequencing Approaches.* Front Cell Dev Biol, 2018. **6**: p. 108.
35. Hwang, B., J.H. Lee, and D. Bang, *Single-cell RNA sequencing technologies and bioinformatics pipelines.* Experimental & Molecular Medicine, 2018. **50**(8): p. 96.
36. Stewart, C.A., et al., *Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer.* Nature Cancer, 2020. **1**(4): p. 423-436.
37. Lambros, M.B., et al., *Single-Cell Analyses of Prostate Cancer Liquid Biopsies Acquired by Apheresis.* Clinical Cancer Research, 2018. **24**(22): p. 5635-5644.
38. Horning, A.M., et al., *Single-Cell RNA-seq Reveals a Subpopulation of Prostate Cancer Cells with Enhanced Cell-Cycle–Related Transcription and Attenuated Androgen Response.* Cancer Research, 2018. **78**(4): p. 853-864.
39. Risbridger, G.P., R. Toivanen, and R.A. Taylor, *Preclinical Models of Prostate Cancer: Patient-Derived Xenografts, Organoids, and Other Explant Models.* Cold Spring Harb Perspect Med, 2018. **8**(8).
40. Lawrence, M.G., et al., *A preclinical xenograft model of prostate cancer using human tumors.* Nat Protoc, 2013. **8**(5): p. 836-48.
41. Lawrence, M.G., et al., *Patient-derived Models of Abiraterone- and Enzalutamide-resistant Prostate Cancer Reveal Sensitivity to Ribosome-directed Therapy.* Eur Urol, 2018. **74**(5): p. 562-572.
42. Davies, A.H., Y. Wang, and A. Zoubeidi, *Patient-derived xenografts: A platform for accelerating translational research in prostate cancer.* Molecular and cellular endocrinology, 2018. **462**(Pt A): p. 17-24.
43. Cheloni, S., et al., *XenoCell: classification of cellular barcodes in single cell experiments from xenograft samples.* bioRxiv, 2019: p. 679183.
44. Srivastava, A., et al., *Alevin efficiently estimates accurate gene abundances from dscRNA-seq data.* Genome Biology, 2019. **20**(1): p. 65.
45. Ilicic, T., et al., *Classification of low quality cells from single-cell RNA-seq data.* Genome biology, 2016. **17**: p. 29-29.

46. Efremova, M., et al., *CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes.* Nature Protocols, 2020. **15**(4): p. 1484-1506.
47. Hanahan, D. and L.M. Coussens, *Accessories to the crime: functions of cells recruited to the tumor microenvironment.* Cancer Cell, 2012. **21**(3): p. 309-22.
48. Cives, M., et al., *The Tumor Microenvironment in Neuroendocrine Tumors: Biology and Therapeutic Implications.* Neuroendocrinology, 2019. **109**(2): p. 83-99.
49. Deshwar, A.G., et al., *PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors.* Genome Biology, 2015. **16**(1): p. 35.