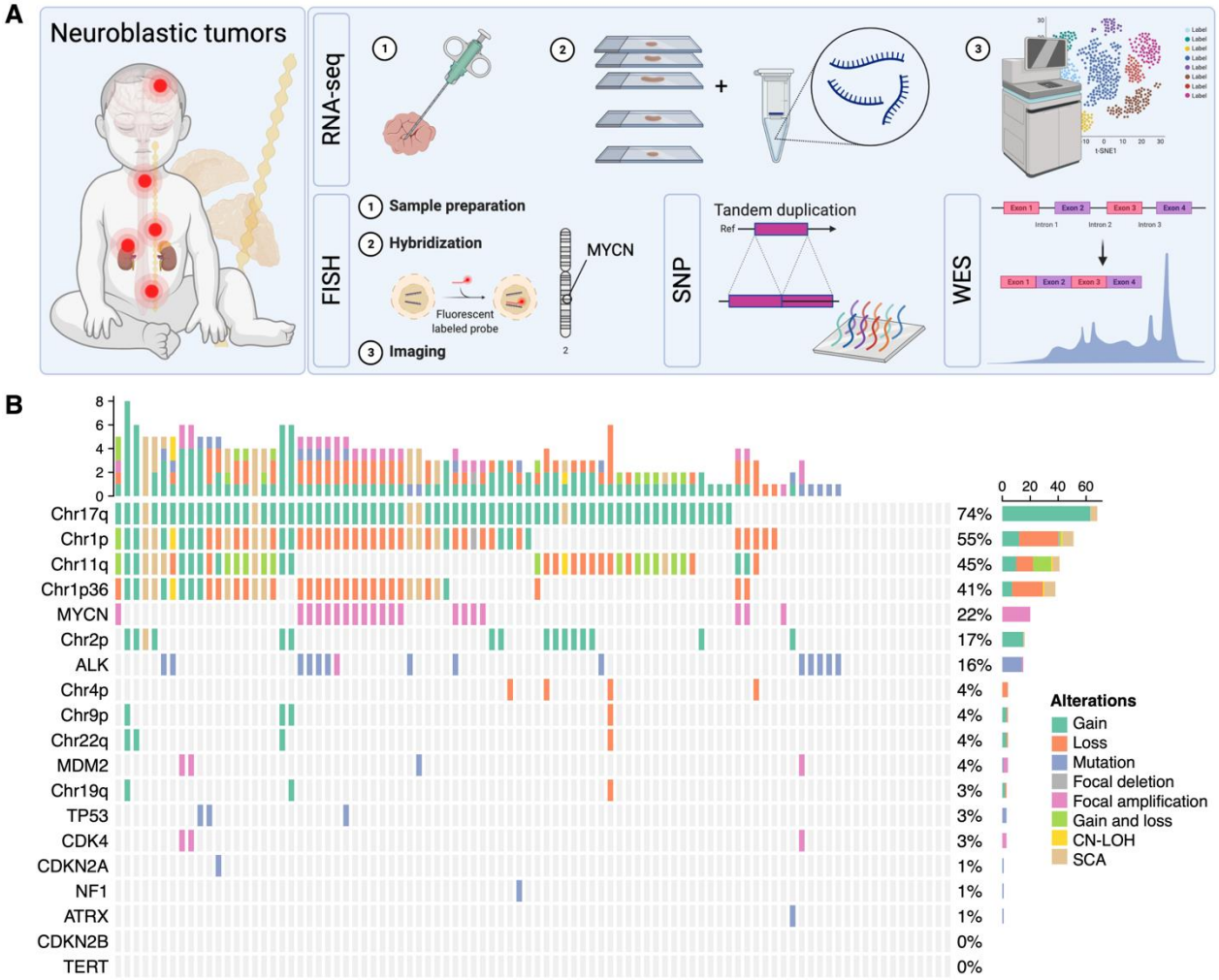
Report

Classification of Neuroblastic Tumors by Whole Transcriptome Profiling.

Louise C.M.C. Bourdery[1, 2], BSc under the supervision of Ronald R. de Krijger[1,3] , MD, PhD and Lennart A. Kester[1], PhD.
[1]Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands
[2]Utrecht University, Biology of Disease MSc programme, Utrecht University, Utrecht, The Netherlands.
[3]Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands

GRAPHICAL ABSTRACT



In Brief

RNA sequencing data of a representative, morphologically well-characterized neuroblastic tumor cohort allows tumor subgrouping and classification, opening up avenues for further gene and protein expression analysis and comparison with outcome.

KEYWORDS

Pediatric Tumors, Peripheral Neuroblastic Tumors, Neuroblastoma, Ganglioneuroblastoma, Ganglioneuroma, Oncology, Pathology, NGS, Whole Exome Sequencing, RNA Sequencing, Data Analysis, Data Visualization

HIGHLIGHTS

- The Máxima peripheral neuroblastic tumor cohort is representative with regard to INPC category and MYCN status
- Histology-based diagnoses and RNA-sequencing-based classification data show strong correlation
- Expression data of peripheral neuroblastic tumors describe a differentiation gradient
- Undifferentiated neuroblastic tumors display pathways involved in proliferation and development
- Differentiated neuroblastic tumors show enrichment in immune pathways

SUMMARY

Although primary pediatric tumors are extremely rare, they remain the leading cause of non-accidental death in children. Peripheral neuroblastic tumors (pNTs) account for about 10% of pediatric malignancies. Their incidence is approximately ten cases per million children each year. They are highly heterogeneous and are classified based on morphological characteristics. Since 2018, the Princess Máxima Center routinely uses RNA-sequencing to detect fusion genes, one of the most common genomic alterations in pediatric cancers. Through our study we aim to refine tissue-based diagnosis as well as predict the outcome of pNT patients. We reviewed the morphology of frozen tumor tissue fragments from patients with neuroblastic tumors from which RNA was extracted. Based on this morphology review we created three groups: undifferentiated, differentiating and differentiated tumors and sought to determine whether there was a correlation between the morphology and the RNA-seq data of the samples. We then investigated the RNA-seq data of pNTs, neurofibromas and Schwannomas to observe the relation between the expression profile and the histology-based diagnosis. The expression profiles described a differentiation gradient, moving from benign differentiated tumors such as schwannomas and neurofibromas to malign undifferentiated tumors such as poorly differentiated neuroblastomas. Subsequently, we explored the expression level of *MYCN* throughout the pNT cohort and found a correlation between the *MYCN* expression level and the differentiation gradient. We then analyzed differentially expressed genes (DEGs) and enriched pathways in undifferentiated versus differentiated tumors and in MYCNA versus non-MYCNA-PDNBs. Overall, based on our cohort, we showed that pNTs, neurofibromas and Schwannomas cluster together by morphological diagnosis and follow a cellular differentiation gradient. Further, we showed that these gene sets are differentially expressed between undifferentiated and differentiated tumors as well as between MYCNA and non-MYCNA PDNBs. Combining our data with clinical follow-up data, may provide new avenues for prediction of prognosis and targeted treatment for patients with peripheral neuroblastic tumors.

LAYMEN'S SUMMARY

Cancers in children are very rare, and neuroblastic tumors are the most common solid tumors in newborns and the third most common pediatric tumor in general. Neuroblastic tumors (NTs) come from cells responsible for generating neurons during early development: the so-called neural crest cells. They are very diverse tumors that can be divided into three main groups: neuroblastomas (NBs), ganglioneuroblastomas (GNBs) and ganglioneuromas (GNs), accounting for about 80, 17 and 3% respectively. NTs mainly occur in children during their first year of life and most cases are diagnosed before the age of 5 during a routine medical visit. They may appear in many places in the trunk and belly and spread to the liver, bone marrow, lymph nodes and skin. Because NTs are so diverse, they

are classified based on their morphology. In the Princess Máxima Center (PMC), a piece of tumor (a biopsy) is taken from all new patients and molecular analyses such as RNA-sequencing (RNA-seq) are performed to detect changes in their genetic material (DNA and RNA). RNA-seq is also performed on tissue samples after the whole tumor or a part of it was removed. The purpose of our study was to enable us to find a correlation between the microscopic image and the RNA-seq data of NTs to later on link these data to patient outcome data and facilitate prediction of prognosis.

We looked at the microscopic images of NT tissues used to perform RNA-seq. Based on the proportions of cell types (neuroblasts, ganglion cells) found, we categorized the tumors. NBs were categorized as undifferentiated (NBUD), poorly differentiated (PDNB) or differentiating (DNB) and we also identified GNBs and GNs. We then investigated a possible link between the microscopic image and the RNA-seq data of the samples and found that the RNA-seq data followed a trajectory based on the level of differentiation or maturity and aggressivity of the tumor samples. We then searched for differentially expressed genes (DEGs) and their corresponding enriched pathways (CEP) in the RNA-seq data of our samples. DEGs are highly expressed in one group compared to the other and vice versa. We created three groups to analyze DEGs and their CEP: undifferentiated (NBUD and PDNBs), differentiating (DNBs) and differentiated (GNBs and GNs) tumors. We performed two analyses: between undifferentiated and differentiated tumors and between *MYCN*-amplified (MYCNA) and non-MYCNA PDNBs. Finally, we looked at the expression level of *MYCN*, a crucial gene in NTs, throughout the differentiation gradient of NTs.

We showed that NTs cluster together based on their microscopic image and follow a differentiation gradient. Further, we were able to state that the morphological and the RNA-seq data were correlated and that the expression level of *MYCN* follows the same abovementioned gradient. We found that undifferentiated tumors are proliferative while differentiated ones and non-MYCNA PDNBs have enriched immune pathways, possibly due to immune cell infiltration in the tumors. Our data may provide new avenues for targeted treatment and should be linked to its corresponding outcome data and to future prediction models for NTs.

INTRODUCTION

Pediatric cancers, although very rare, constitute the leading cause of non-accidental death in children[1], among which peripheral neuroblastic tumors (pNTs) account for about 10% of all pediatric malignancies[9]. pNTs constitute a spectrum of benign to malignant embryonal neoplasms[15, 17] with clinical behaviors going from spontaneous regression to a disastrous outcome[18, 19]. This heterogeneity is reflected in the overall survival rates (OS) ranging from 85% for low-to-intermediate risk disease to 50% for high-risk disease patients[19]. pNTs are derived from neural crest cells and mainly arise in the adrenal medulla or the sympathetic ganglia of the neck, thorax, abdomen, or pelvis[6, 8, 9, 10, 11]. pNTs encompass neuroblastoma (NB) and ganglioneuroblastoma nodular (GNBn), both considered malignant, and ganglioneuroblastoma intermixed (GNBi) as well as ganglioneuroma (GN) considered benign[9].

Neuroblastic tumors (NTs) are the commonest extracranial solid malignancy in the pediatric population and the most common tumor in infants[2, 3, 4] with a slight male predominance (male-to-female ratio: 1.1-1.2)[8, 12]. Their incidence is approximately ten per million children per year[5]. They mainly occur in the first year of life[4], and most cases are diagnosed before five years of age[3, 7]. Diagnosis is usually established during a routine clinic visit, through imaging studies for other purposes, or based on a variety of symptoms, that may depend on the location of the tumor[4]. Primary NTs may arise from immature neural crest cells[4] anywhere along the sympathetic nervous system (SNS) and may metastasize to the bone marrow, lymph nodes, liver or skin[8]. Neuroblastic malignancies occur almost exclusively sporadically, however they may also arise in familial or syndromic contexts[18].

Clinically and biologically, pNTs are heterogeneous malignancies[9, 15] and are clinically organized into low-, intermediate- and high-risk groups based on multiple prognostic factors [20]. Histologically, pNTs have been classified (Figure 2A) by the International Neuroblastoma Pathology Classification (INPC) into four categories[15]. Neuroblastomas, the most common pNTs (80%), are also defined as Schwannian stroma-poor and are in turn subclassified into three subtypes: undifferentiated neuroblastoma (NBUD), poorly differentiated neuroblastoma (PDNB) and differentiating neuroblastoma (DNB) with each subtype representing about 4%, 90% and 6% of NB cases respectively[14, 15, 16]. The second and third most common pNTs are GNB intermixed (9%) and nodular (8%), also known as Schwannian stroma-

rich and composite Schwannian stroma-rich/stroma dominant and stroma poor[9, 15]. The last form of pNT is ganglioneuroma or Schwannian stroma-dominant and represents 3% of pNT cases[15] (Figure S1).

Stage of the tumor is a crucial clinical prognostic marker and is determined based on two clinical staging systems (Figure 2B). They are based on either pre- or post-surgical determination of tumor extent: the International Neuroblastoma Risk Group Staging System (INRGSS) and the International Neuroblastoma Staging System (INSS) respectively[15]. The pre-surgical staging system is used to determine the extent of the pNT and its stage based on the presence or absence of image-defined risk factors (IDRFs). IDRFs are important to help predict surgical outcomes. Their presence is associated with high-risk histopathologic and molecular features of neuroblastic tumors[39]. Moreover, the INRGSS distinguishes distant metastases (M) from those confined to specific areas (MS)[15]. The post-surgical staging system is also used for the evaluation of the disease extent and enables the distinction between localized (stages 1, 2 and 3), metastatic disease (stage 4) as well as special metastatic disease (stage 4S) correlated to a favorable outcome for the patient.

Genetic and molecular alterations are of particular interest to predict prognosis in NT patients. About 20% of pNTs and 25% of NB cases carry a *MYCN* amplification[18, 20]. *MYCN* is an oncogene that induces neoplastic transformation, codes for a transcription factor that regulates target gene expression, and promotes cancer hallmarks such as cell proliferation and growth among others[14, 18, 19]. In 80% to 90% of cases, *MYCN* amplification results in an increased amount of MYCN protein, making it a driver of aggressive NTs and a strong predictor of poor prognosis[14, 17, 19, 20]. This amplification is suspected to be an early and driving event in the development of NTs[19, 21]. Of note, *MYCN* amplification is usually assessed by fluorescent in situ hybridization (FISH) and is considered as such with a *MYCN*/centromere ratio above 4. However, it can also be detected by whole exome sequencing (WES), a technique that enables the detection of CNVs, mutations and amplifications. FISH is a fast, simple, specific and highly sensitive technique to detect chromosomal rearrangements, gene amplifications and to analyze intra-tumor heterogeneity[22, 23, 26]. This technique is based on the hybridization of desoxyribonucleic acid (DNA) probes directly or indirectly labeled to a specific sequence in the gene of interest, followed by fluorescence microscopy[24, 25].

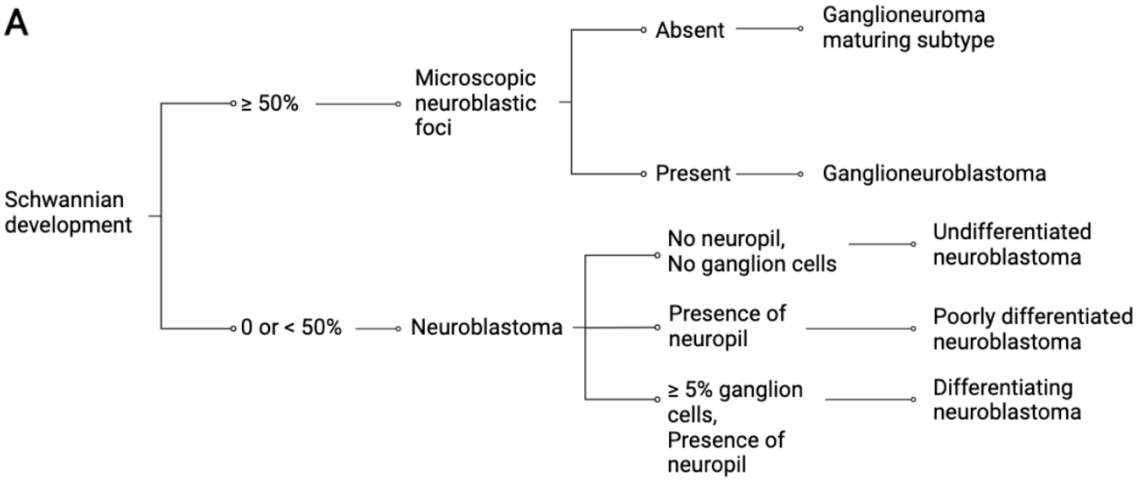
The presence of segmental chromosomal alterations (SCAs) was shown to be a strong predictor of poor prognosis and relapse in neuroblastic tumors[40, 41]. SCAs are defined as any partial chromosome loss or gain[41]. In pNTs, SCAs often take the form of deletions, the most common occurring in chromosomes containing tumor suppressor (TS) genes, such as 1p, 3p, 4p and 11q. Gains in chromosomes carrying putative oncogenes as 1q, 2p and 17q also occur on a recurrent basis and are commonly associated with a poor outcome in patients with neuroblastic tumors[18, 40]. SCAs in infants are correlated with higher risks of relapse in both localized unresectable and metastatic neuroblastic tumors without *MYCN* amplification[40]. Segmental chromosomal changes are detectable by genome-based approaches such as multiplex ligation-dependent probe amplification (MLPA) or single nucleotide polymorphism (SNP) arrays[42]. These techniques have demonstrated that patients with deletions such as the 11q deletion constitute a high-risk group in non-*MYCN*-amplified neuroblastomas[42]. Copy number variations (CNVs) may as well be assessed by MLPA or by single nucleotide polymorphism (SNP) array. MLPA enables the detection of copy number changes, provides a genetic profile of neuroblastic tumors and shows to be a reliable technique for risk stratification[23]. SNP arrays also enable the detection of CNVs that might predispose to tumorigenesis. Furthermore, SNP arrays also detect events such as aneuploidy and ploidy changes, as well as partial chromosomal imbalances. The reason why SNP arrays are now performed in many diagnostic laboratories and not MLPA is that they provide genome-wide rather than disease-specific data[35].

Additional markers of interest such as anaplastic lymphoma kinase (ALK) may be evaluated by immunohistochemistry (IHC) as it constitutes a potential therapeutic target. *ALK* is a receptor tyrosine kinase (RTK) and possible oncogenic driver being altered in about 14% of high-risk neuroblastomas by gain-of-function point mutations[27], but may also occur in non-high-risk NTs[30, 31]. Mutations in *ALK* were also assessed with targeted next generation sequencing (tg-NGS). tg-NGS is a DNA sequencing technique most commonly used for diagnostic purposes that focuses on highly specific areas of the genome based on a panel of genes. It also enables to limit the investigation to specific mutations: hotspot mutations, found within a gene of interest. It is particularly relevant when it comes to detecting small CNVs in genes known to be involved in neuroblastic tumors[44]. This technique provides detailed information and enables the visualization of DNA variations such as clinically relevant mutations, precise information for accurate diagnosis and tumor classification[33].

Finally, since 2018, the PMC has been the first hospital in The Netherlands to introduce whole transcriptome sequencing also known as RNA sequencing (RNA-seq) as a routine diagnostic technique followed by WES. RNA-seq enables the detection of fusion genes, one of the most common genomic alterations found in pediatric cancers. It also measures more accurately transcript levels and detects the overexpression of genes such as *MYCN*[36, 37, 43]. This technique relies on RNA extraction from tumor samples, complementary DNA (cDNA) library preparation, sequencing and data analysis (Figure S2). WES allows the detection of CNVs and DNA variants responsible for protein sequence alterations[43]. It targets genes encoding proteins, which represent about 3% of the whole genome[38]. However, whole transcriptome sequencing has additional benefits as it allows to generate reliable gene expression data, and gives the potential to classify tumors on the basis of transcriptome profiles. Together, these techniques help genetic and molecular diagnosis as well as tumor classification. Moreover, they enable the detection of potential therapeutic targets, improving treatment selection for both patients with cancer and other types of diseases[22-36].

Throughout our study we aimed to optimize and refine the RNA-seq-based classification of peripheral neuroblastic tumors. To do so, we explored the RNA-seq data of neuroblastic tumors to visualize the distribution of the samples in relation to each other, based on the diagnoses derived from this same RNA-seq data. The morphology of biopsy and resection samples from the Princess Máxima Center was reviewed and compared to the one determined by RNA-seq and a correlation between the RNA-seq data and the morphology of the tumor was found. We also sought to find differentially expressed genes and the enriched pathways they are involved in. This was carried out with the objective in the near future to put all these results in relation to outcome to facilitate the potential course of disease and outcome prediction and to make treatment stratification easier.

A



B

Pre-surgical staging system

INRGSS — International Neuroblastoma Risk Group Staging System

Stage	Description
L1	Localized tumor without any involvement of vital structures, confined to one body compartment (e.g neck, chest, abdomen, pelvis).
L2	Locoregional tumor with one or more image-defined risk factors (IDRFs).
M	Presence of distant metastasis.
MS	Metastatic disease in children under 18 months of age. Metastases confined to the skin, liver and/or bone marrow. The primary tumor can be in stage 1, 2 or 3 based on INSS staging.

Post-surgical staging system

INSS — International Neuroblastoma Staging System

Stage	Description	Prognosis
1	Localized tumor with complete gross excision, with or without microscopical residual disease. Homolateral lymph nodes microscopically negative for tumor. Lymph nodes adherent to the primary tumor may be positive.	Favorable
2A	Localized tumor with incomplete gross resection, homolateral non-adherent lymph nodes microscopically negative for tumor.	Favorable
2B	Localized tumor with or without incomplete gross resection, homolateral non-adherent lymph nodes microscopically positive for tumor. Enlarged contralateral lymph nodes, microscopically negative for tumor.	Favorable
3	Unresectable unilateral tumor infiltrating through the midline (vertebral column) with or without involved regional lymph nodes; or localized homolateral tumor with contralateral regional lymph node involvement or midline tumor with bilateral extension by infiltration (unresectable) or by lymph node involvement.	Unfavorable
4	Any primary tumor with dissemination to distant lymph nodes, liver, skin, bone, bone marrow and/or other organs, except as defined in stage 4S.	Unfavorable
4S	Localized primary tumor as described for stages 1, 2A or 2B with limited dissemination to skin, liver, bone and/or bone marrow in infants younger than 12 months. Bone marrow involvement should be minimal and more extensive bone marrow involvement should be considered stage 4 disease. Results of the metaiodobenzylguanidine (MIBG) scan should be negative for tumor in the bone marrow if performed.	Favorable

Figure 2. International classification and stagings of neuroblastic tumors

A. Simplified International Neuroblastoma Pathology classification. Modified from van Arendonk *et al*.[13] Histologically, pNTs are first classified based on the proportion of Schwannian stroma, then according to the presence or absence of microscopic neuroblastic foci, neuropil and ganglion cells[4, 11, 13]. Poorly differentiated tumors composed of 0 to 49% of Schwannian stroma (SS) and abundant neuroblasts are classified as neuroblastomas (Schwannian stroma-poor NBs)[4, 11]. NBs are then subclassified as follows: undifferentiated NB with absence of both neuropil and ganglion cells, poorly differentiated NB with presence of neuropil and, differentiating NB with at least 5% of ganglion cells and presence of neuropil. Intermediary to well-differentiated benign tumors composed of more than 50% of Schwannian stroma containing mature ganglion cells and scattered neuroblasts are classified as Schwannian stroma-rich GNB and Schwannian stroma-dominant GN[4, 13].
B. International Neuroblastoma Risk Group and International Neuroblastoma staging systems. Modified from Robbins & Cotran[2] and from the Physician Data Query (PDQ®).

MATERIALS AND METHODS

Collection, processing and storage of neuroblastic tumor samples

Core needle biopsy (CNB) and resection samples from the Princess Máxima Center were collected between October 2014 and December 2020 (n=150) for clinical purposes. Informed consent was given by patients or their parents. This study was performed under a waiver of the University Medical Center Utrecht medical ethical committee (non-WMO plichtig).
Patient material was processed by the Pathologist Assistant (PA) at the diagnostic laboratory to create frozen sections from biobank specimens, and paraffin blocks.

Processing and storage of samples

Frozen sections. Two pieces of the patient material were cryoembedded as follows : each sample was placed on a lens paper, transferred to the bottom of an aluminum mold in the cryoembedder PrestoCHILL (Milestone Medical) and covered in a cryo-embedding compound (Milestone Medical). A chunk was placed into the mold with a heat extractor and frozen at -42°C for one minute. The frozen blocks were transferred to a cryostat CryoStar™ (Fisher Scientific) set to create sections of 4µm. One of the blocks was trimmed in order to remove the lens paper and used to create 4µm-sections that were transferred to two glass microscope slides. Both frozen sections were immersed in a FineFIX (Milestone Medical) fixing solution and transferred to a fully automated sample processor/stainer PRESTO PRO (Milestone Medical) to be stained with hematoxylin and eosin (H&E). Slides were finally transferred into a ClearVue™ Coverslipper (Fisher Scientific) and stored at room temperature (RT).

Paraffin blocks. Two pieces of sample were fixed overnight before being paraffin-embedded, stained with H&E and stored at room temperature.

Containers. Samples in the containers were flash frozen at -80°C for two minutes to be stored for the biobank.

Once processed, the frozen sections were observed under a LED microscope by a pathologist in order to determine the tumor cell percentage and tumor morphology. Sections were then stored at RT.

Morphology review of H&E-stained neuroblastic tumor core needle biopsy and resection samples

Eighty H&E-stained frozen sections of biopsy (52) and resection (28) samples obtained from patients in the Princess Máxima Center were observed with a Leica DM3000 LED microscope (Figure 3). The proportions of vital tumor tissue (VTT), necrosis, neuroblasts and Schwannian stroma were assessed and compared to the simplified International Neuroblastoma Pathology Classification (INCP) (Figure 2A) to verify the diagnosis established with FFPE H&E slides. The relabelled cases were registered as frozen section-derived diagnosis or final diagnosis and summarized in Table S3.

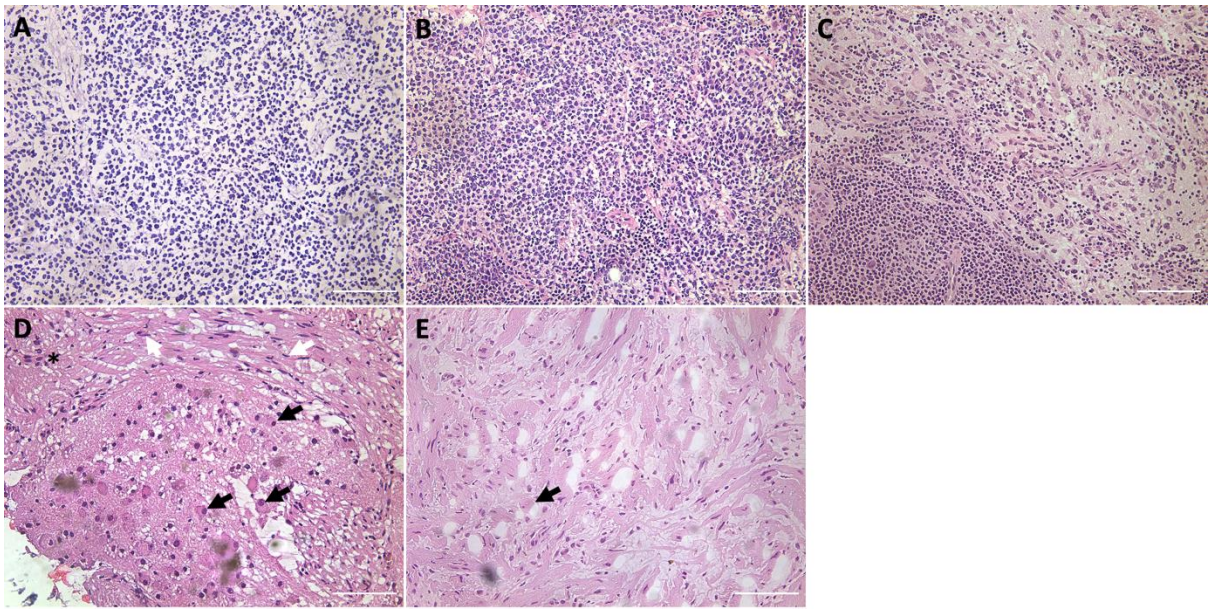


Figure description on the next page.

Figure 3. Pathology of peripheral neuroblastic tumors. The morphology of each sample was assessed based on the International Neuroblastoma Pathology Classification (INPC). All photomicrographs were taken on frozen sections used to perform RNA-seq at a 20X magnification. Scale bar: 100µm.
A. Resection sample of a lymph node located in the neck (left) with metastasis from a 6-year-old patient. The morphology is characteristic of an undifferentiated neuroblastoma, with a complete absence of neuropil and the exclusive presence of neuroblasts.
B. Resection sample of a lymph node located in the groin (left) with extensive metastasis from a 2-year-old patient with a MYCN-amplified tumor. The morphology is characteristic of a poorly differentiated neuroblastoma with small blue round cell-tumor and presence of neuropil.
C. Resection sample of a differentiating neuroblastoma in the retroperitoneal area of a 2-year-old patient. The morphology of this tumor is characteristic of a differentiating neuroblastoma with the presence of neuropil and ganglion cells (≥5%).
D. Bone marrow biopsy sample taken from a 11-year-old-patient. The morphology is representative of a ganglioneuroblastoma with more than 50% of Schwannian stroma and the presence of microscopic neuroblastic foci. Black and white arrows indicate mature ganglion cells and Schwannian stroma respectively, the asterisk indicate the presence of neuropil.
E. Tru-Cut and incisional biopsy sample of a left adrenal gland from a 15-year-old patient displaying a morphology characteristic of ganglioneuroma. This picture shows histology indicative of mature Schwannian stroma as well as mature ganglion cells. In addition, there is no presence of neuropil and neuroblasts.

Cohort establishment

The RNA-seq data of 107 peripheral neuroblastic and benign peripheral nerve sheath tumor (BPNST) samples was collected. Three samples of pretreated necrotic poorly differentiated neuroblastoma and one neuroblastoma sample collected from bone marrow were excluded from our study leaving 103 samples retained for the establishment of the **whole tumor spectrum cohort** and its visualization. The BPNST samples (six neurofibromas (NFs) and five Schwannomas (SCHs)) were then excluded to create the **peripheral neuroblastic tumor (pNT) cohort** gathering 1 NBUD, 62 PDNB, 14 DNB, 8 GNB and 7 GN for a total of 92 samples. After visualization of the pNT cohort, a **MYCN cohort** was established. This cohort exclusively gathered the data of 61 poorly differentiated neuroblastoma samples of which 16 carried a MYCN amplification and 45 did not. All three cohorts were created with the histology-based diagnoses. The composition of each cohort is summarized in Figure 4A and Table S4.

Data collection

Patient data were collected from the pathology archives of the PMC and clinical data retrieved from the hospital system. The data of interest were then entered into the Castor database.

RNA-sequencing

Neuroblastoma frozen samples were processed as shown in Figure S2. Part of the core needle biopsy and resection samples from the PMC collected between June 2018 and December 2021 for clinical purposes was frozen. Total RNA was extracted from these samples, after which ribosomal RNA (rRNA) and mitochondrial RNA (mtRNA) depletion step was performed to retain the RNA of interest: the mRNA. mRNAs were fragmented, random primers were added and cDNA synthesis was performed. Adaptors were added to the ends of the cDNA fragments, ligated and PCR amplification cycles were conducted. The samples were then sequenced with the NovaSeq™ 6000 Sequencing System from Illumina. This sequencing step was followed by data processing and analysis.

RNA-seq data analysis

Data visualization with UMAP and ggplot2

RNA-seq data was received as count and meta data and processed using both R and RStudio softwares. The data was normalized to counts per million (CPM) and then log transformed to provide a normal-like distribution. The variance and mean of each gene were calculated after which the 5000 most variable genes were used for further analysis. The data was then z-score normalized and a principal component analysis (PCA) was performed on the z-normalized data. A UMAP projection was then generated on the loadings of the PCA using 100.000 as seed. UMAP projections were generated either in one or two dimensions depending on the type of analysis. The UMAP projections were visualized using the ggplot2 R package.

Finding differentially expressed genes with DESeq2

Differentially expressed genes (DEGs) were assessed using the DESeq2 R package on our different data subsets (differentiated versus undifferentiated pNTs, differentiating versus undifferentiated pNTs, differentiating versus differentiated pNTs and MYCN-amplified PDNBs versus non-MYCN-amplified PDNBs). The DESeq function was used on each comparison. The generated results were ordered by adjusted p-values (padj).

Determining enriched signaling pathways with FGSEA

Gene set enrichment analysis was performed on the DESeq2 results using the FGSEA R package. We used the Hallmark gene set from MSigDB (h.all.v7.4.symbols.gmt.txt) and set a random seed of 100.000. The list of pathways was loaded and the fgsea algorithm was ran a thousand permutations. The results obtained were tidied, the normalized enrichment scores (NES) were plotted and each bar was colored to indicate whether or not the pathways were significant.

Key resource table

Softwares and algorithms

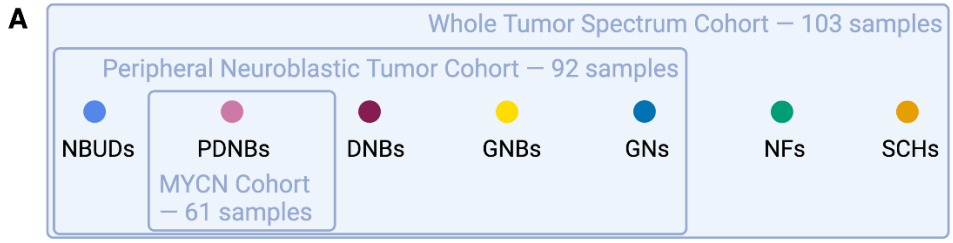
Resource	Identifier
BioRender	https://biorender.com/
Castor EDC	https://www.castoredc.com
Microsoft Excel	Microsoft
HiX Productie	ChipSoft
PACS IDS7 22.1 Pathologie	Sectra Medical
R	https://www.r-project.org
RStudio	https://www.rstudio.com
SymPathy	Tieto

RESULTS

Morphological characterization of peripheral neuroblastic tumors.

Morphological review of frozen sections resulted in sixteen (20%) cases being relabelled as compared to the original H&E-based diagnosis (Table S3). If no frozen section was available the H&E diagnosis was used for further analysis. We then assessed the degree of relabeling, and found that thirteen (81%) of the sixteen relabelled cases were first-degree relabelled. That is to say that there was only one step of differentiation between the first and the second labels such as a PDNB relabelled into a DNB and vice versa or a GNB relabelled into a GN and vice versa. Of the thirteen first-degree relabelled specimens, one NBUD was recategorized as a PDNB and one was inversely relabelled. Four PDNBs were reclassified as DNBs, and conversely, two DNBs were reclassified as PDNBs. Three GNB samples were relabelled as DNBs, one GNB as a GN and one central nervous system neuroblastoma (CNS-NB) was recategorized as a PDNB. The three (19%) remaining samples were second-degree relabelled. In other words, there were two steps of differentiation between the first and second labels. Of these three specimens one NBUD was relabelled as a DNB and two PDNBs were reclassified as GNBs. Having determined the frozen section-derived diagnoses, we sought to observe the distribution and location of our samples in relation to each other based on this second diagnosis.

The morphology review matches the whole tumor spectrum and the peripheral neuroblastic tumor cohort expression data and their visualization reveals that both follow a differentiation gradient.



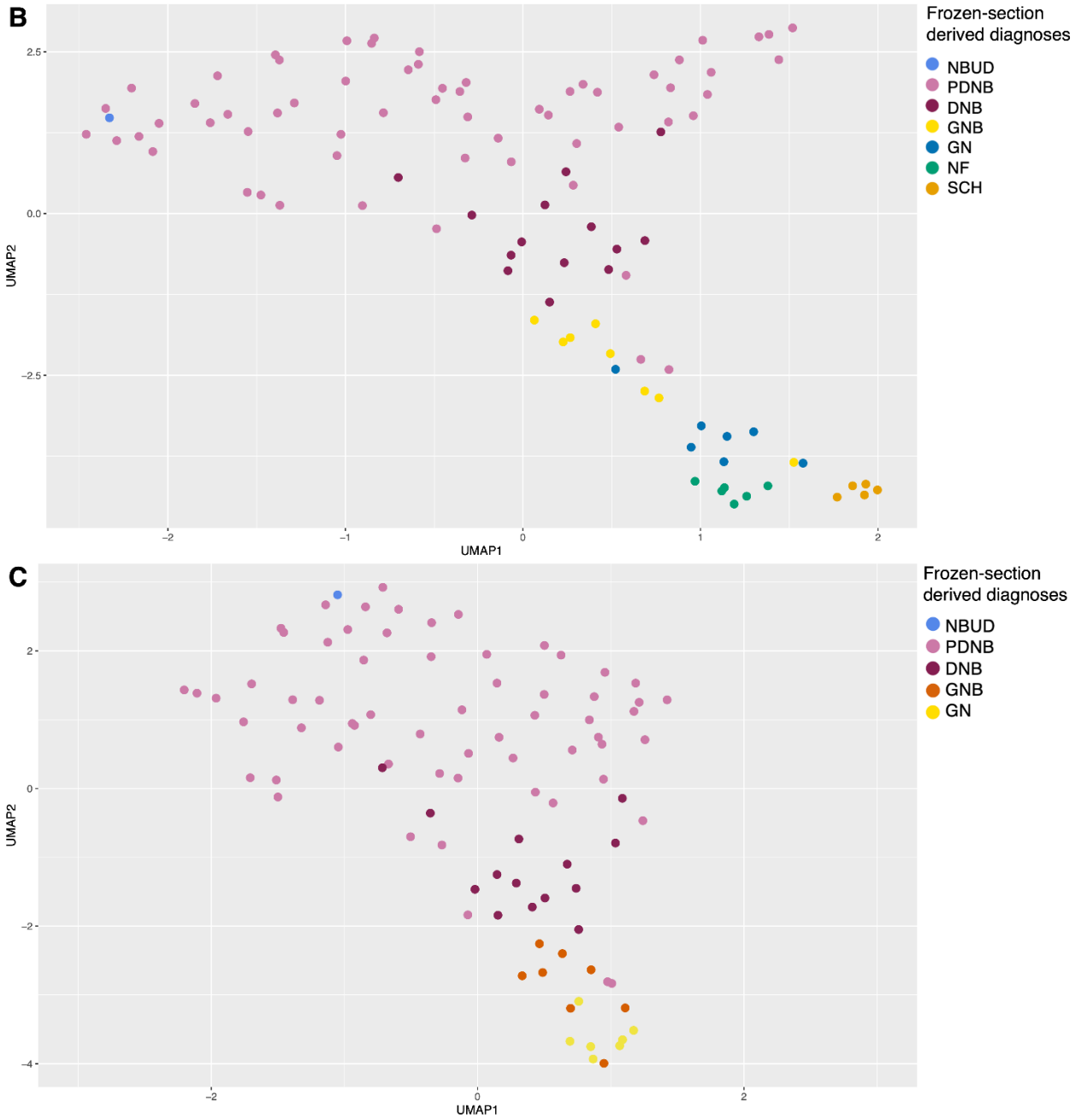


Figure 4. Visualizations of a whole tumor spectrum and a neuroblastic tumor dataset based on their morphology-based diagnoses.

Both cohorts were visualized in 2D plots after using the UMAP algorithm to reduce the dimension of the data. **A.** Overview of the three cohorts of which two were used for the two following 2D UMAP visualizations. The Whole tumor spectrum cohorts gathers 103 samples of which 92 are pNTs and 11 are BPNSTs, the Peripheral neuroblastic tumor cohort that gathers the 92 pNT samples and the MYCN cohort which gathers 61 PDNB samples of which 16 carry a MYCN amplification and 54 don't. **B.** 2D UMAP plot of the Whole tumor spectrum cohort based on the diagnoses obtained after the morphology of the available frozen samples was reviewed. **C.** 2D UMAP plot of the Peripheral neuroblastic tumor cohort based on the diagnoses obtained after morphological assessment.

Having determined the frozen section-derived diagnoses, we looked at the distribution and location of the specimens in relation to each other based on this final diagnosis. Visualizing the expression data of the samples gathered in the whole tumor spectrum cohort (Figure 4A) revealed that the tumors follow a differentiation gradient (Figure 4B, Figure S1B). Thus, the most malignant and undifferentiated tumors (e.g. NBUD and PDNBs) clustered on one end of the tumor spectrum, while the most benign and differentiated tumors (e.g. GNBs, GNs, NFs and SCHs) clustered to the lower right of the spectrum. We were able to visualize the same distribution of the samples included in the peripheral neuroblastic tumor cohort (Figure 4C), excluding the benign peripheral nerve sheath tumors.

Differentially expressed genes and enriched pathways in the pNT and MYCN cohorts.

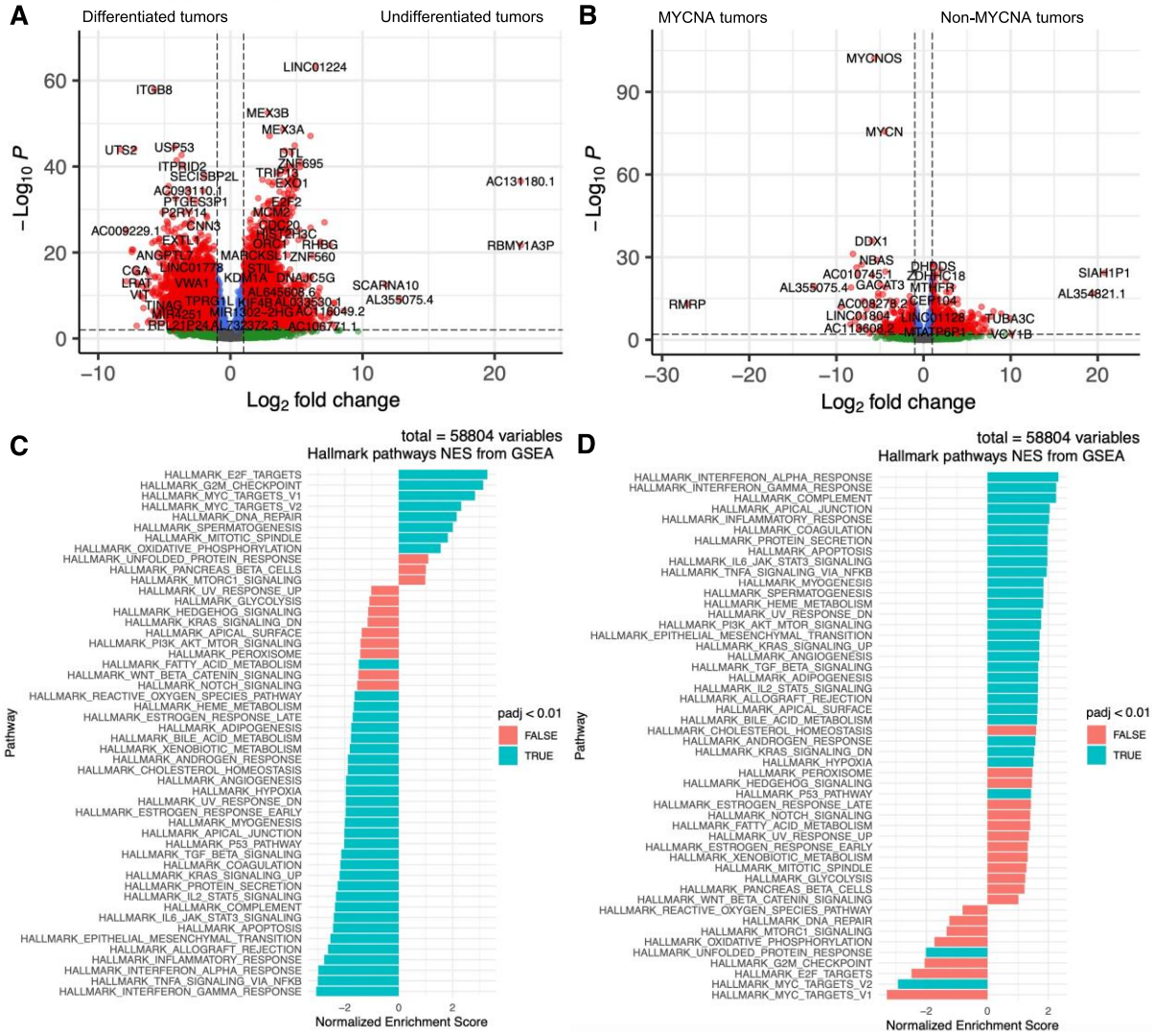


Figure 5. Differentially expressed genes and their corresponding enriched pathways in the pNT and MYCN cohorts.

A, **B**. Volcano plots showing Log2 fold change versus adjusted P-value obtained by using the DESeq2 package on the undifferentiated versus differentiated tumors (**A**) and on the MYCNA versus non-MYCNA-PDNBs (**B**), with an adjusted P-value threshold of 0.01. **C**, **D**. Visualization of normalized enriched scores versus pathways generated by using the fgsea package on the undifferentiated versus differentiated tumors (**C**) and on the MYCNA versus non-MYCNA-PDNBs (**D**). The pathways are colored based on the adjusted P-value (padj) threshold of 0.01.

For our further analyses the cases were merged into three groups (Table S5): the undifferentiated tumors (NBUD and PDNBs), the differentiating tumors (DNBs only) and the differentiated tumors (GNBs and GNs). The pairwise comparison of the RNA-seq expression data of these groups enabled us to find differentially expressed genes (DEGs) and identify the corresponding enriched pathways (Figure 5).

We first compared undifferentiated and differentiated tumors and showed high expression of stemness genes such as *MEX3A*, a key regulator of neuroblasts proliferation in neurogenesis and its homolog *MEX3B*, mini-chromosome maintenance genes such as *MCM10*, a protein involved in the initiation of genomic replication, and histone genes in undifferentiated tumors (Figure 5A). High levels of MEX3A and MCM10 proteins among others were associated with proliferative, metabolic, DNA damage and developmental hallmarks (Figure 5C). In the differentiated tumors, results showed high expression of *ITGB8* and *UTS2* which correlated with immune, proliferative and developmental hallmarks (Figure 5A, 5C).

Secondly, given the crucial impact of *MYCN* amplification in neuroblastic tumors, we compared MYCN-amplified (MYCNA) and non-MYCNA PDNBs. Results showed high expression of *MYCNOS*, *MYCN*, *DDX1* and *NBAS* genes in MYCNA PDNBs. (Figure 5B, 5D) These genes are all located on the same

locus and their expression level is a consequence of the MYCN amplification in the MYCNA PDNB samples. These results correlated with proliferative hallmarks. In the non-MYCNA PDNBs, differential expression of genes such as *S100PBP* correlated with enriched immune, developmental and metabolic pathways (Figure 5B, 5D).

Expression level of *MYCN* throughout the differentiation gradient of a neuroblastic tumor cohort.

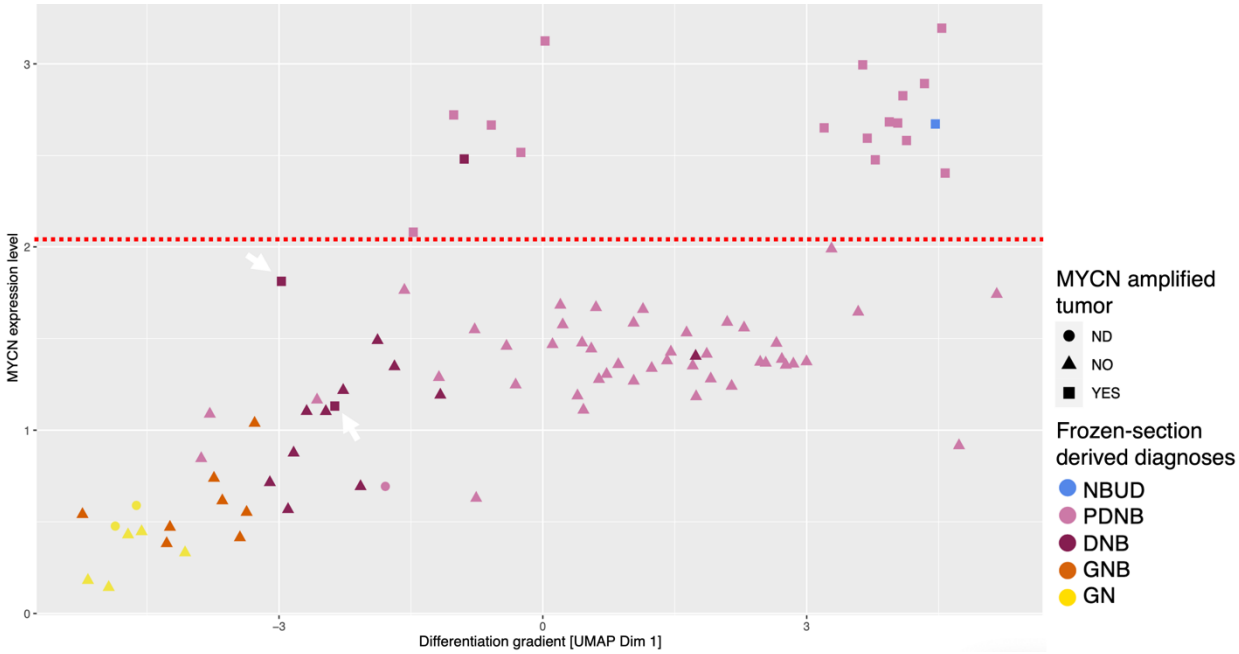


Figure 6. *MYCN* expression level throughout the differentiation gradient of neuroblastic tumors.
1D UMAP representation of the *MYCN* expression level throughout the differentiation gradient (Figure 4, Figure S1B) in the peripheral neuroblastic tumors cohort. The most differentiated samples are located on the left part of x axis and the most undifferentiated ones are located on the right side of the axis. The samples above the red dashed line are all *MYCN*-amplified neuroblastic tumors (represented by squares), all samples below the red dashed line do not carry this amplification (represented by triangles) except for the two samples indicated by a white arrow. The colors are based on the histology-based diagnoses. ND: no data available regarding the *MYCN* status.

MYCN amplification is a strong indicator for poor prognosis and therefore marks the most malign neuroblastic tumors. Therefore we wondered whether MYCN expression follows the differentiation gradient we observed in the transcriptome wide cluster analyses. To test this, we performed a new dimensionality reduction using UMAP dimension. This single dimension still reflects the differentiation gradient (Figure 6). Furthermore, expression level of MYCN increases along the differentiation gradient and is higher on the far right end of the plot where undifferentiated tumors cluster. Two samples indicated as *MYCN*-amplified and located under the red dashed line showed a lower *MYCN* expression level than the other *MYCN*-amplified samples. Most interestingly, two groups stood out within the *MYCN*-amplified tumors, with one group gathering six samples localized on a more differentiated part of the gradient and the other one including twelve samples localized on a more undifferentiated part of this same differentiation gradient.

CONCLUSION

In this study we used RNA expression data obtained from whole transcriptome sequencing to analyze pediatric neuroblastic tumors and investigate how the expression profile relates to the histology-based classification. First, we performed morphological review of frozen sections corresponding to the tissue used for RNA-seq analysis, to see if this diagnosis correlated with the original H&E-based diagnosis. Our results showed that pNT and BPNSTs predominantly cluster together according to the morphological diagnosis and follow a gradient related to the tumor differentiation (e.g. neuroblasts, Schwannian stroma). Furthermore, we have shown that these groups of neuroblastic tumors may be distinguished on the basis of their RNA-seq data. The abovementioned differentiation gradient appears to be based on DEGs as we have shown that genes such as *MEX3A* and *MCM10* are differentially

expressed in undifferentiated NTs and that *ITGB8* and *UTS2* are highly expressed in differentiated NTs. These expression profiles correlated with proliferative, metabolic and developmental hallmarks in undifferentiated tumors and with immune, proliferative and developmental ones in differentiated NTs. We also showed that MYCNA PDNBs had enriched proliferative pathways whereas non-MYCNA PDNBs displayed enriched immune, developmental and metabolic pathways. The DEGs that were found may provide new avenues for targeted treatment. Within the MYCN cohort, we found a correlation between the *MYCN* expression level and the differentiation gradient. We also highlighted two distinct groups of MYCNA PDNBs that require further investigation. This led us to conclude that peripheral neuroblastic tumors can be classified on the basis of their RNA-seq data. Our data should be linked to outcome data in order to potentially serve as a prediction model for the prognosis of these tumors.

DISCUSSION

In our study we reviewed the morphology of NTs taken in the PMC over a five-year period and investigated their expression profile which constitutes the signature of the tumors. We determined that frozen section-derived diagnosis correlated with H&E-based diagnosis (Figure 2A, Figure 3). We then showed that pNTs together with BPNSTs follow a differentiation gradient based on the histology-based diagnoses (Figure 4, Figure S1B). The most undifferentiated tumors clustered on one end of the plot while the most differentiated tumors clustered on the other end. Having established the distribution pattern of our samples, we searched for DEGs and the enriched pathways they were involved in (Figure 5). We found that undifferentiated NTs had enriched proliferative and developmental pathways whereas differentiated tumors displayed enriched immune pathways. We then turned to a different approach through which we visualized the expression level of *MYCN* (Figure 6). This showed that the expression level of *MYCN* throughout the specimens followed the differentiation gradient mentioned earlier and correlated with the MYCN status of each tumor with the exception of two resection samples.

Frozen sections of peripheral neuroblastic tumors obtained from patients of the Princess Máxima Center between 2017 and 2021 were classified according to the INPC (Figure 2A). In the pNT cohort, of ninety-two cases, frozen sections were available for eighty, of which the morphology review resulted in relabeling of sixteen cases. The twelve missing slides that were not revised might alter the results obtained as it might have prevented us from highlighting relevant cases of relabeling. However, it remained important to keep the frozen section-derived or final diagnoses for our analyses as most of the slides (87%) had been reviewed and were found to provide complementary information to that provided by expression data. In the future, we could choose to include only morphologically reviewed samples, either by microscopy or through images for larger cohorts in order to limit biases. Another important aspect of our cohort is that it contains both biopsy and resection samples. Some specimens were taken from the same patient at different disease stages, other samples were taken after treatment and this led to morphological therapy-related changes in the frozen sections (e.g. necrosis, calcification, maturation towards ganglion cells). This could distort our observations especially when, in the future, we want to put our data in relation to outcome data and should be taken into consideration.

Most importantly, frozen sections were used for both purposes, that is to say for morphology review and RNA-sequencing. In most cases, we were able to correlate the morphological images to the molecular data from exactly the same piece of tissue. For our study, the frozen section-derived diagnosis gave the best match as we were able to verify its accuracy by visualizing the expression data of the frozen sections used for the morphology review. However, the overall histopathology and the RNA-seq data did not necessarily match as neuroblastic tumors are very heterogeneous and frozen tissue doesn't necessarily represent the entire tumor. This should be considered in the event of an intention to replicate this study. In the further analyses of the data, we used the adjusted labels and looked at the pNT cohort in an unsupervised approach and at known genes. In the event that someone would be interested in replicating our study, it is important to note that the use of the UMAP algorithm generates random results and therefore different plots and that it requires the use of a seed (material and methods).

We searched for differentially expressed genes and their corresponding enriched pathways in undifferentiated versus differentiated neuroblastic tumors and secondly, in *MYCN*A and non-*MYCN*A PDNBs (Figure 5). These groups were established in order to get higher numbers per groups and reduce as much as possible the discrepancies in numbers between each diagnostic group for analytical reasons. It is necessary to bear in mind that the groups formed did not contain the same number of samples. Even though our cohort appears representative as compared to the literature, the small sample size might have affected our statistical analyses. Differential expression analysis (DEA)

between undifferentiated and differentiated tumors showed that undifferentiated NTs are characterized by enriched proliferative and developmental pathways whereas differentiated ones displayed enriched immune pathways. Enriched proliferative hallmarks could be related to the aggressivity of the undifferentiated NTs. DEA on MYCNA and non-MYCNA PDNBs showed high expression of *MYCNOS*, *MYCN*, *DDX1* and *NBAS* genes in MYCNA-PDNBs. As these genes are located on the same locus, on chromosome 2, their co-amplification is a consequence of the *MYCN* amplification found in the MYCNA group of the MYCN cohort. *MYCNOS* is a modulator of the *MYCN* locus and causes the amplification of the latter when it is itself amplified. In the MYCNA-PDNBs, differentially expressed genes correlated with proliferative hallmarks. Meanwhile, we identified high expression of the *S100PBP* gene in non-MYCNA PDNBs, which encodes a protein involved in the regulation of many cellular processes such as cell cycle progression and cell differentiation. These results correlated with enriched immune, developmental and metabolic pathways.

Most interestingly, two groups stood out within the MYCNA PDNBs (Figure 6), with one group gathering six samples localized on a more differentiated part of the gradient and the other one including twelve samples localized on a more undifferentiated part of this same differentiation gradient. This specific distribution of the *MYCN*-expression level throughout the differentiation gradient of NTs seems to correlate with the level of aggressiveness of the tumors and would require an investigation regarding the outcome of these patients. In fact, the more undifferentiated they are, the more malignant and the higher the *MYCN* expression level seems to be.

Since the data visualization of the pNT cohort showed two clusters among the MYCNA PDNB samples, we could look for DEGs and enriched pathways within these two groups. This exploration could help us unravel genes responsible for the formation of these clusters and determine with the help of outcome data whether one group is associated to a better prognosis than the other one.

Having examined the distribution and location of each sample throughout our cohorts regardless of the material type, it would be of great interest to retain only the biopsy and resection samples used to establish the primary diagnoses and to repeat our analyses on this new dataset. Thus, these new analyses would not be biased by treated tumors that might have differentiated into other tumor subtypes. Moreover, it would be particularly interesting to re-examine the morphology and perform single-cell RNA sequencing (scRNA-seq) on these same samples used for diagnostic purposes to look into possible immune infiltrates[49]. We could perform the necessary immunostainings and look for a correlation with morphology, single cell and RNA-seq data. After a cell sorting step, the scRNA-seq analysis could thus reveal the different cell populations (immune cells and fibroblasts[2]) present in the tumor micro-environment (TME) of patient specimens. By searching for differentially expressed genes and their corresponding pathways on this primary material cohort, we could look into other candidate genes and *MYCN*. Even though NTs were shown to be immunologically “cold”[49], the presence of tumor-infiltrating lymphocytes (TILs) constitutes a prognostic indicator in many tumors[50]. Mina *et al*. showed that low-risk neuroblastic tumors were characterized by higher levels of TILs than the ones associated with a poor prognosis[50] making this kind of investigation on our samples relevant to completement our study.

So far, we have investigated the RNA-seq data of a PMC cohort of neuroblastic tumors. In the future, we could bring some improvements discussed earlier and use our study workflow to predict prognosis. This requires more time to enlarge our cohort, collect sufficient data about patient outcome and put these two elements in relation after data analysis.

ACKNOWLEDGEMENTS

We thank the patients and their parents for giving their consent regarding the access to the specimens that were analyzed in this study, and we thank the Dutch biobank for providing these specimens. I am very grateful to Ronald de Krijger, MD, PhD, and thank him for giving me this amazing opportunity to learn about RNA-seq data analysis and for supporting me professionally, personally and financially, for helping me grow under his supervision together with the precious help of Lennart Kester, PhD. Many thanks to Jacob Rozowsky, BSc, for his help with R throughout my project, to Rixt Bruinsma, MSc for her patience and benevolence during the data collection process. I also thank Alissa Groenendijk, MSc for sharing educational materials with me. Special thanks to Marc van Tuil, BSc, for showing me the RNA-seq workflow in the lab (Figure S2), to Arie Maat for showing me how patient material is processed in the Pathology lab and to Rubina Moeniralam for her help in retrieving the patient materials of interest. Many thanks to all the other members of the PMC who helped me out during my project.

REFERENCES

1. Kattner P, Strobel H, Khoshnevis N, et al. Compare and contrast: pediatric cancer versus adult malignancies. *Cancer Metastasis Rev*. 2019;38(4):673-682. doi:10.1007/s10555-019-09836-y
2. Kumar, V., Abbas, A.K. and Aster, J.C. (2015) Robbins & Cotran Pathologic Basis of Disease. 9th Edition, Elsevier, Philadelphia
3. Newman EA, Abdessalam S, Aldrink JH, et al. Update on neuroblastoma. J Pediatr Surg. 2019;54(3):383-389. doi:10.1016/j.jpedsurg.2018.09.004
4. Ishola TA, Chung DH. Neuroblastoma. *Surg Oncol*. 2007;16(3):149-156. doi:10.1016/j.suronc.2007.09.005
5. M.A. Hayat (ed.), *Pediatric Cancer, Volume 4: Diagnosis, Therapy, and Prognosis*, Pediatric Cancer 4, 3 DOI 10.1007/978-94-007-6591-7_1, © Springer Science+Business Media Dordrecht 2013
6. Tsubota S, Kadomatsu K. Origin and initiation mechanisms of neuroblastoma. *Cell Tissue Res*. 2018;372(2):211-221. doi:10.1007/s00441-018-2796-z
7. Hayat MA. *Neuroblastoma*. Dordrecht: Springer; 2012. doi:10.1007/978-94-007-2418-1
8. Putnam A.R, Thompson K.S, Wallentine J.C. *Diagnostic Pathology Pediatric Neoplasms*. Second Edition. Elsevier, 2018. ISBN: 9780323548069
9. Luksch R, Castellani MR, Collini P, et al. Neuroblastoma (Peripheral neuroblastic tumours). *Crit Rev Oncol Hematol*. 2016;107:163-181. doi:10.1016/j.critrevonc.2016.10.001
10. Conces M.R. Peripheral neuroblastic tumors of the adrenal gland: clinicopathologic features and important molecular alterations. Diagnostic Histopathology, Volume 26, Issue 5, 2020. Pages 200-206. ISSN 1756-2317, https://doi.org/10.1016/j.mpdhp.2020.02.002.
11. Hiroyuki Shimada, Atsuko Nakagawa, Pathology of the Peripheral Neuroblastic Tumors, *Laboratory Medicine*, Volume 37, Issue 11, November 2006, Pages 684–689, https://doi-org.proxy.library.uu.nl/10.1309/0506C1BM8GBVV224
12. Parodi S, Haupt R. The Epidemiology of Neuroblastoma. New York: Nova Biomedical Books, 2008
13. Van Arendonk KJ, Chung DH. Neuroblastoma: Tumor Biology and Its Implications for Staging and Treatment. *Children (Basel)*. 2019;6(1):12. Published 2019 Jan 17. doi:10.3390/children6010012
14. Ahmed AA, Zhang L, Reddivalla N, Hetherington M. Neuroblastoma in children: Update on clinicopathologic and genetic prognostic factors. *Pediatr Hematol Oncol*. 2017;34(3):165-185. doi:10.1080/08880018.2017.1330375
15. Hiroyuki Shimada, Hideki Sano, Florette K. Hazard. Pathology of peripheral neuroblastic tumors. Clinical pediatric hematology-oncology. 27(2):73-86. doi:10.15264/cpho.2020.27.2.73
16. Wang LL, Suganuma R, Ikegaki N, et al. Neuroblastoma of undifferentiated subtype, prognostic significance of prominent nucleolar formation, and MYC/MYCN protein expression: a report from the Children's Oncology Group. Cancer. 2013;119(20):3718-3726. doi:10.1002/cncr.28251
17. Ikegaki N, Shimada H; International Neuroblastoma Pathology Committee. Subgrouping of Unfavorable Histology Neuroblastomas With Immunohistochemistry Toward Precision Prognosis and Therapy Stratification. JCO Precis Oncol. 2019;3:PO.18.00312. doi:10.1200/PO.18.00312
18. Janoueix-Lerosey I, Schleiermacher G, Delattre O. Molecular pathogenesis of peripheral neuroblastic tumors. Oncogene. 2010;29(11):1566-1579. doi:10.1038/onc.2009.518
19. Otte J, Dyberg C, Pepich A, Johnsen JI. MYCN Function in Neuroblastoma Development. Front Oncol. 2021;10:624079. Published 2021 Jan 27. doi:10.3389/fonc.2020.624079
20. Wang LL, Teshiba R, Ikegaki N, et al. Augmented expression of MYC and/or MYCN protein defines highly aggressive MYC-driven neuroblastoma: a Children's Oncology Group study. Br J Cancer. 2015;113(1):57-63. doi:10.1038/bjc.2015.188
21. Matthay, K., Maris, J., Schleiermacher, G. et al. Neuroblastoma. Nat Rev Dis Primers 2, 16078 (2016). https://doi-org.proxy.library.uu.nl/10.1038/nrdp.2016.78
22. Misra DN, Dickman PS, Yunis EJ. Fluorescence in situ hybridization (fish) detection of mycn oncogene amplification in neuroblastoma using paraffin-embedded tissues. Diagnostic molecular pathology : the american journal of surgical pathology, part b. 1995;4(2):128-135.
23. Costa RA, Seuánez HN. Investigation of major genetic alterations in neuroblastoma. Mol Biol Rep. 2018;45(3):287-295. doi:10.1007/s11033-018-4161-4
24. Ratan ZA, Zaman SB, Mehta V, Haidere MF, Runa NJ, Akter N. Application of Fluorescence In Situ Hybridization (FISH) Technique for the Detection of Genetic Aberration in Medical Science. Cureus. 2017;9(6):e1325. Published 2017 Jun 9. doi:10.7759/cureus.1325
25. Shakoori AR. Fluorescence In Situ Hybridization (FISH) and Its Applications. Chromosome Structure and Aberrations. 2017;343-367. Published 2017 Feb 10. doi:10.1007/978-81-322-3673-3_16
26. Huber D, Voith von Voithenberg L, Kaigala GV. Fluorescence in situ hybridization (fish): history, limitations and what to expect from micro-scale fish? Micro and nano engineering. 2018;1:15-24. doi:10.1016/j.mne.2018.10.006
27. Trigg RM, Turner SD. ALK in Neuroblastoma: Biological and Therapeutic Implications. Cancers (Basel). 2018;10(4):113. Published 2018 Apr 10. doi:10.3390/cancers10040113

28. Alexander N, Marrano P, Thorner P, et al. Prevalence and Clinical Correlations of Somatostatin Receptor-2 (SSTR2) Expression in Neuroblastoma. J Pediatr Hematol Oncol. 2019;41(3):222-227. doi:10.1097/MPH.0000000000001326
29. Cracolici V, Wang EW, Gardner PA, et al. SSTR2 Expression in Olfactory Neuroblastoma: Clinical and Therapeutic Implications [published online ahead of print, 2021 Apr 30]. Head Neck Pathol. 2021;10.1007/s12105-021-01329-1. doi:10.1007/s12105-021-01329-1
30. O'Donohue T, Gulati N, Mauguen A, et al. Differential Impact of ALK Mutations in Neuroblastoma. JCO Precis Oncol. 2021;5:PO.20.00181. Published 2021 Mar 19. doi:10.1200/PO.20.00181
31. Umapathy G, Mendoza-Garcia P, Hallberg B, Palmer RH. Targeting anaplastic lymphoma kinase in neuroblastoma. APMIS. 2019;127(5):288-302. doi:10.1111/apm.12940
32. Chami R, Marrano P, Teerapakpinyo C, et al. Immunohistochemistry for ATRX Can Miss ATRX Mutations: Lessons From Neuroblastoma. Am J Surg Pathol. 2019;43(9):1203-1211. doi:10.1097/PAS.0000000000001322
33. Behjati S, Tarpey PS. What is next generation sequencing?. Arch Dis Child Educ Pract Ed. 2013;98(6):236-238. doi:10.1136/archdischild-2013-304340
34. Capasso M., Diskin S.J. (2010) Genetics and Genomics of Neuroblastoma. In: Pasche B. (eds) Cancer Genetics. Cancer Treatment and Research, vol 155. Springer, Boston, MA. https://doi-org.proxy.library.uu.nl/10.1007/978-1-4419-6033-7_4
35. Bashton M, Hollis R, Ryan S, et al. Concordance of copy number abnormality detection using SNP arrays and Multiplex Ligation-dependent Probe Amplification (MLPA) in acute lymphoblastic leukaemia. Sci Rep. 2020;10(1):45. Published 2020 Jan 8. doi:10.1038/s41598-019-56972-0
36. Chang F, Lin F, Cao K, et al. Development and Clinical Validation of a Large Fusion Gene Panel for Pediatric Cancers. J Mol Diagn. 2019;21(5):873-883. doi:10.1016/j.jmoldx.2019.05.006
37. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57-63. doi:10.1038/nrg2484
38. Suwinski P, Ong C, Ling MHT, Poh YM, Khan AM, Ong HS. Advancing Personalized Medicine Through the Application of Whole Exome Sequencing and Big Data Analytics. Front Genet. 2019;10:49. Published 2019 Feb 12. doi:10.3389/fgene.2019.00049
39. Temple WC, Vo KT, Matthay KK, et al. Association of image-defined risk factors with clinical features, histopathology, and outcomes in neuroblastoma. Cancer Med. 2021;10(7):2232-2241. doi:10.1002/cam4.3663
40. Schleiermacher G, Mosseri V, London WB, et al. Segmental chromosomal alterations have prognostic impact in neuroblastoma: a report from the INRG project. Br J Cancer. 2012;107(8):1418-1422. doi:10.1038/bjc.2012.375
41. Jeison M, Ash S, Halevy-Berko G, et al. 2p24 Gain region harboring MYCN gene compared with MYCN amplified and nonamplified neuroblastoma: biological and clinical characteristics. Am J Pathol. 2010;176(6):2616-2625. doi:10.2353/ajpath.2010.090624
42. Carén H, Kryh H, Nethander M, et al. High-risk neuroblastoma tumors with 11q-deletion display a poor prognostic, chromosome instability phenotype with later onset. Proc Natl Acad Sci U S A. 2010;107(9):4323-4328. doi:10.1073/pnas.0910684107
43. O'Brien TD, Jia P, Xia J, et al. Inconsistency and features of single nucleotide variants detected in whole exome sequencing versus transcriptome sequencing: A case study in lung cancer. Methods. 2015;83:118-127. doi:10.1016/j.ymeth.2015.04.016
44. Roca I, González-Castro L, Fernández H, Couce ML, Fernández-Marmiesse A. Free-access copy-number variant detection tools for targeted next-generation sequencing data. Mutat Res Rev Mutat Res. 2019;779:114-125. doi:10.1016/j.mrrev.2019.02.005
45. Kattner P, Strobel H, Khoshnevis N, et al. Compare and contrast: pediatric cancer versus adult malignancies. *Cancer Metastasis Rev*. 2019;38(4):673-682. doi:10.1007/s10555-019-09836-y
46. Srinivasan A, Toh YC. Human Pluripotent Stem Cell-Derived Neural Crest Cells for Tissue Regeneration and Disease Modeling. Front Mol Neurosci. 2019;12:39. Published 2019 Feb 22. doi:10.3389/fnmol.2019.00039
47. Naef V, De Sarlo M, Testa G, et al. The Stemness Gene Mex3A Is a Key Regulator of Neuroblast Proliferation During Neurogenesis. *Front Cell Dev Biol*. 2020;8:549533. Published 2020 Sep 22. doi:10.3389/fcell.2020.549533
48. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425. doi:10.1016/j.cels.2015.12.004

REFERENCES R CODE

1. 1 and 2 dimensional UMAP codes: Lennart A. Kester, PhD, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands
2. DESeq2 analysis:

    a. https://www.youtube.com/watch?v=me22e26vDo4
    b. https://lashlock.github.io/compbio/R_presentation.html
    c. https://bioconductor.org/packages/release/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#differential-expression-analysis
3. FGSEA analysis:
    a. https://www.bioconductor.org/packages/devel/bioc/vignettes/fgsea/inst/doc/fgsea-tutorial.html
    b. https://stephenturner.github.io/deseq-to-fgsea/
    c. https://stephenturner.github.io/deseq-to-fgsea/#appendix
4. OncoPrint: https://jokergoo.github.io/ComplexHeatmap-reference/book/oncoprint.html
5. Remaining code such as mutate and oncoprint: Jacob Rozowsky, BSc, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands

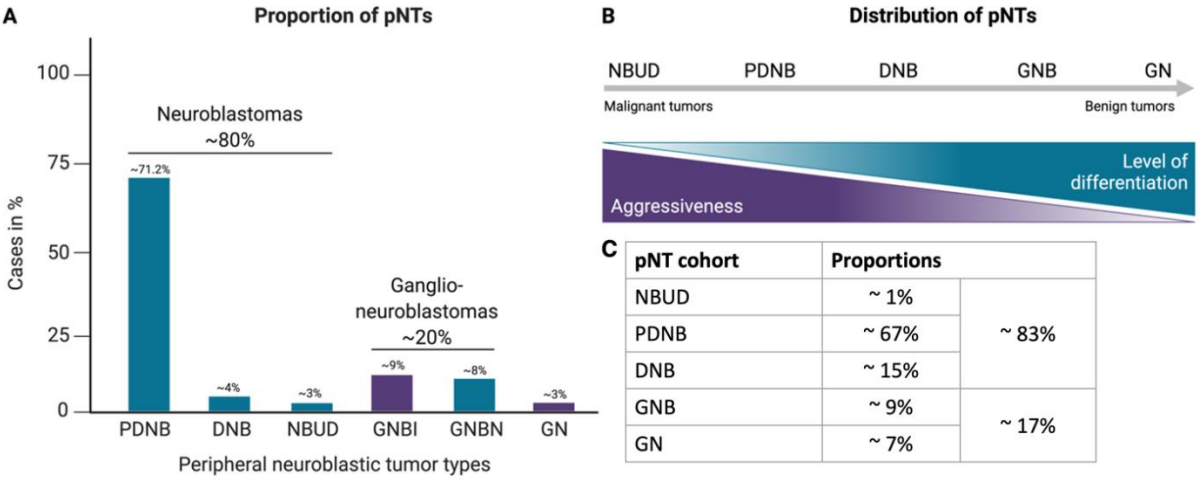
SUPPLEMENTAL DATA

List of abbreviations

ALK	Anaplastic lymphoma kinase
ALT	Alternate lengthening of telomere
ATRX	Alpha thalassemia/mental retardation syndrome X-linked
BPNST	Benign peripheral nerve sheath tumor
cDNA	Complementary desoxyribonucleic acid
CNB	Core needle biopsy
CNV	Copy number variation
CPM	Counts per million
DEA	Differential expression analysis
DEGs	Differentially expressed genes
DNA	Desoxyribonucleic acid
DNB	Differentiating neuroblastoma
DSS	Disease sub-specification
DSS2	Disease sub-specification 2
EFS	Event-free survival
FFPE	Formalin-fixed paraffin-embedded
FH	Favorable histology
FISH	Fluorescent in situ hybridization
GN	Ganglioneuroma
GNB	Ganglioneuroblastoma
GNBi	Ganglioneuroblastoma intermixed
GNBn	Ganglioneuroblastoma nodular
H&E	Hematoxylin and eosin
IDRF	Image-defined risk factor
IHC	Immunohistochemistry
INPC	International Neuroblastoma Pathology Classification
INRGSS	International Neuroblastoma Risk Group Staging System
INSS	International Neuroblastoma Staging System
OS	Overall survival
MKI	Mitosis-karyorrhexis index
MLPA	Multiplex ligation-dependent probe amplification
mtRNA	Mitochondrial RNA
NB	Neuroblastoma
NBUD	Undifferentiated neuroblastoma
NES	Normalized enrichment scores
NF	Neurofibroma
NGS	Next generation sequencing
NTs	Neuroblastic tumors
padj	Adjusted p-value
PDNB	Poorly differentiated neuroblastoma
PMC	Prinses Máxima Centrum
pNTs	Peripheral neuroblastic tumors
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
rRNA	Ribosomal ribonucleic acid
RT	Room temperature
RTK	Receptor tyrosine kinase
SCA	Segmental chromosomal alterations
SCH	Schwannoma

SNP array	Single nucleotide polymorphism array
SNS	Sympathetic nervous system
SS	Schwannian stroma
SSTRs	Somatostatin receptors
SSTR2A	Somatostatin receptor type 2A
tg-NGS	Targeted next generation sequencing
TS	Tumor suppressor
UH	Unfavorable histology
UMAP	Uniform Manifold Approximation and Projection
UT	Undifferentiated tumors
VTT	Vital tumor tissue
WES	Whole exome sequencing

Supplemental figures and tables



pNT cohort	Proportions	
NBUD	~ 1%	~ 83%
PDNB	~ 67%	
DNB	~ 15%	
GNB	~ 9%	~ 17%
GN	~ 7%	

Figure S1. Proportions and distribution of peripheral neuroblastic tumors and their subtypes

A. Bar plot representation of the proportions of the different peripheral neuroblastic tumors and their subtypes. PDNB: Poorly differentiated neuroblastoma, DNB: differentiating neuroblastoma, NBUD: undifferentiated neuroblastoma, GNBI: ganglioneuroblastoma intermixed, GNBN: ganglioneuroblastoma nodular and GN: ganglioneuroma. Neuroblastomas represent approximately 80% of pNTs with PDNB representing about 71%, DNB 4% and NBUD 3%. Ganglioneuroblastomas account for about 20% of pNTs with GNBI 9% and GNBN 8%. Finally, ganglioneuromas represent about 3% of all pNTs. In the bar plot, malignant tumors are represented by blue bars whereas benign tumors are represented by purple bars.
B. Schematic representation of the distribution of peripheral neuroblastic tumors. pNTs are ordered based on their aggressiveness, level of differentiation and malignancy. The purple gradient represents the aggressiveness of the pNTs and the blue gradient indicated the level of differentiation of the pNTs. NBUD is an undifferentiated and aggressive type of malignant neuroblastic tumor. PDNB constitutes a poorly differentiated and aggressive type of malignant neuroblastic tumor. DNB is a differentiating, aggressive and malignant tumor. Finally, both GNB and GN are well differentiated (GN being more differentiated than GNB) benign tumors.
C. Overview of the proportions of neuroblastic tumors in the peripheral neuroblastic tumor cohort.

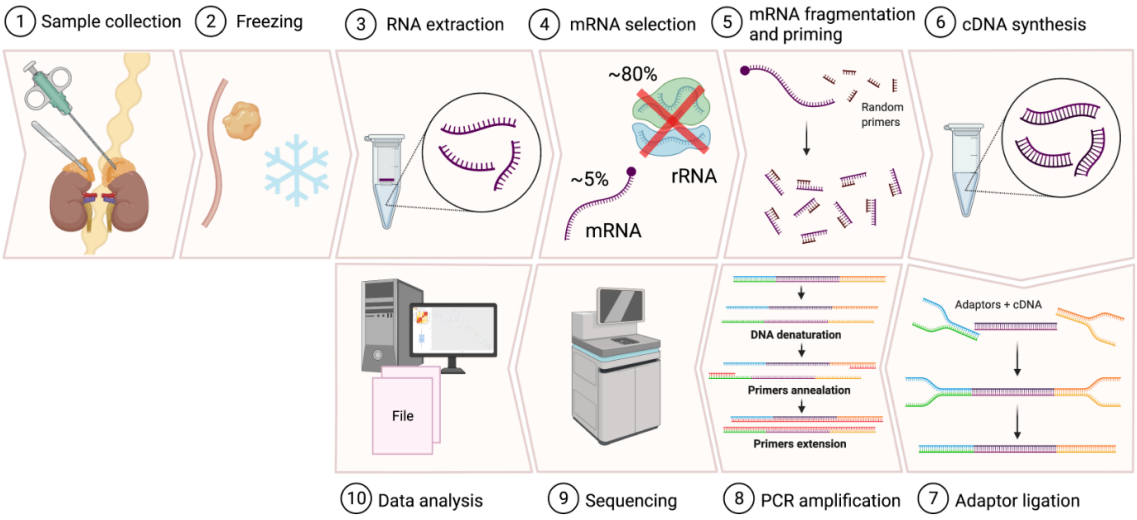


Figure description on the next page.

Figure S2. RNA-sequencing workflow

Schematic representation of the RNA-sequencing workflow. (**1**) Core needle biopsy and/or resection samples were collected. (**2**) Samples were frozen. (**3**) Total RNA was extracted from the samples and stored until library preparation steps. (**4**) rRNA as well as mitochondrial RNA (mtRNA) were depleted in order to select the mRNA of interest. (**5**, **6**) mRNA was fragmented and random primers were added to it, after what the cDNA synthesis took place. (**7**, **8**) The cDNA was mixed, annealed with adaptors and PCR amplification cycles were performed. (**9**) Samples were sequenced with the NovaSeq™ 6000 Sequencing System from Illumina. (**10**) Data analysis was performed in order to make sense of the data, to highlight fusion genes, gene amplification among other alterations.

	Diagnosis 1	Diagnosis 2	% VTT	% necrosis	% neuroblasts	%Schwannian stroma
1	GNB	GN	100	0	0	100
2	GNB	DNB	95	0	95	5
3	PDNB	DNB	70	10	90	0
4	PDNB	DNB	95	5	50	10
5	NBUD	PDNB	100	0	100	0
6	DNB	PDNB	80	20	100	0
7	CNS NB	PDNB	90	10	100	0
8	GNB	DNB	90	10	90	10
9	PDNB	GNB	25	0	90	10
10	PDNB	DNB	95	5	100	0
11	GNB	DNB	60	40	90	10
12	PDNB	GNB	85	15	50	50
13	PDNB	NBUD	25	75	100	0
14	NBUD	DNB	80	20	90	10
15	DNB	PDNB	80	15	95	5
16	PDNB	DNB	80	20	65	45

Table S3. Relabelled frozen sections stained with H&E.

This table summarizes the sixteen samples that were relabelled after scoring. The diagnosis 1 column refers to the diagnoses established on the basis of the H&E slides assessed by a pathologist. The diagnosis 2 column corresponds to the labels assigned to the same cases as in the diagnosis 1 based on the morphology review of the frozen sections that were used to perform the RNA-seq. The determination of the frozen section-based diagnosis (diagnosis 2) was determined with the help of the scores and the simplified International Neuroblastoma Pathology Classification (INPC) (Figure 2A).

Sample types	Whole tumor spectrum cohort	pNT cohort	MYCN cohort
NBUD	1	1	-
PDNB	62	62	-
DNB	14	14	-
GNB	8	8	-
GN	7	7	-
SCH	5	-	-
NF	6	-	-
MYCN-amplified PDNB	-	-	16
Non-MYCN-amplified PDNB	-	-	45

Table S4. Overview of the composition of the different cohorts.

Overview of the types and number of samples included in each cohort of our study. The whole tumor spectrum cohort gathered both peripheral neuroblastic tumors (NBUD, PDNBs, DNBs, GNBs and GNs) and benign peripheral nerve sheath tumors (SCHs and NFs). The pNT cohort included peripheral neuroblastic tumors and the MYCN cohort gathered only *MYCN*-amplified and non-*MYCN*-amplified poorly differentiated neuroblastomas.

Compared groups	Diagnosis 2 (number of samples)	Number of samples once merged
Undifferentiated tumors	Poorly differentiated neuroblastomas (62), Undifferentiated neuroblastomas (1)	63 samples
Differentiating tumors	Differentiating neuroblastomas (14)	14 samples
Differentiated tumors	Ganglioneuroblastomas (8), Ganglioneuromas (7)	15 samples
MYCNA — PDNBs	16 samples	16 samples
Non-MYCNA — PDNBs	45 samples	45 samples

Table S5. Overview of the composition of the different groups used to perform DEG and FGSEA analyses. This table recapitulates the data subsets used to perform DESeq2 and GSE analyses.

R packages	Uses of the package	Citation
ComplexHeatmap	Provides a flexible way to arrange multiple heatmaps.	Gu, Z. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics.
DESeq2	Differential gene expression analysis based on the negative binomial distribution	Love MI, Huber W, Anders S (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology, 15, 550. doi: 10.1186/s13059-014-0550-8.
dplyr	Function designed to manipulate dataframes.	Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.7. https://CRAN.R-project.org/package=dplyr
EnhancedVolcano	Publication-ready volcano plots with enhanced coloring and labeling	Kevin Blighe, Sharmila Rana and Myles Lewis (2021). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.12.0. https://github.com/kevinblighe/EnhancedVolcano
fgsea	R package for fast preranked gene set enrichment analysis	G. Korotkevich, V. Sukhov, A. Sergushichev. Fast gene set enrichment analysis. bioRxiv (2019), doi:10.1101/060012
ggplot2	Data visualization, creation of plots.	Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.
ggpubr	Package for elegant data visualization.	Alboukadel Kassambara (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr
ggrepel	Repulsive text labels.	Kamil Slowikowski (2021). ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.9.1. https://CRAN.R-project.org/package=ggrepel
org.Hs.eg.db	Genome wide annotation for Human	Marc Carlson (2021). org.Hs.eg.db: Genome wide annotation for Human. R package version 3.14.0.
RColorBrewer	Palette of colors.	Erich Neuwirth (2014). RColorBrewer: ColorBrewer Palettes. R package version 1.1-2. https://CRAN.R-project.org/package=RColorBrewer
readxl	Enables importation of Excel files in R.	Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl
tidyverse	Package gathering ggplot2, tidyr and dplyr among others.	Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
umap	Uniform manifold approximation and projection (UMAP). Algorithm for dimensional reduction.	Tomasz Konopka (2020). umap: Uniform Manifold Approximation and Projection. R package version 0.2.7.0. https://CRAN.R-project.org/package=umap
viridis	Colorblind-friendly color palette	Simon Garnier, Noam Ross, Robert Rudis, Antônio P. Camargo, Marco Sciaini, and Cédric Scherer (2021). Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.2.

Table S6. R packages used for the RNA-seq data analysis.