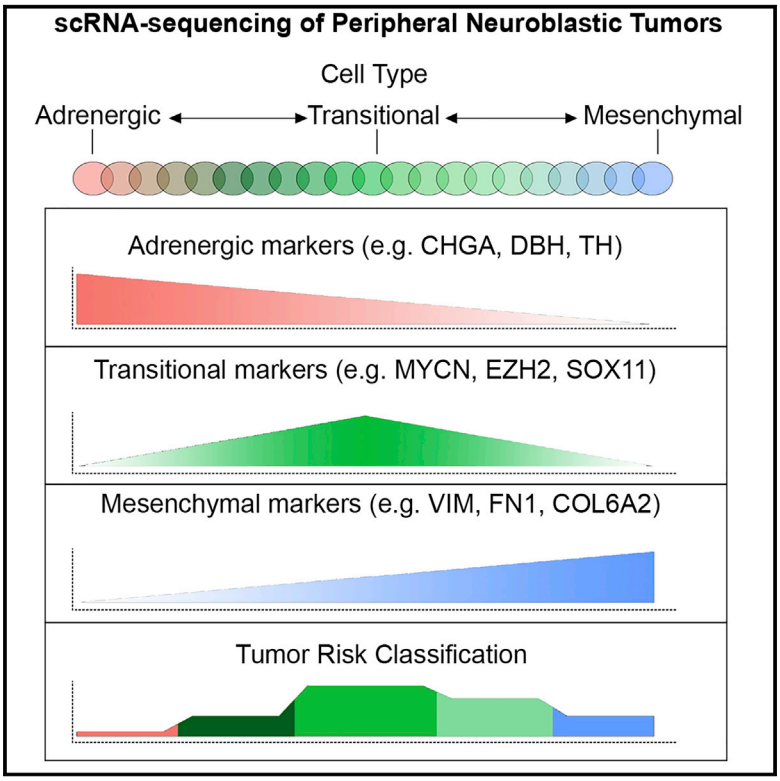
Cell Reports

Article

# Single-cell profiling of peripheral neuroblastic tumors identifies an aggressive transitional state that bridges an adrenergic-mesenchymal trajectory

## Graphical abstract




## Highlights

- Single-cell RNA-seq of 10 PNTs shows malignant and non-malignant cell heterogeneity
- Trend of fewer T cells and more macrophages in neuroblastoma compared with other PNTs
- A malignant transitional cell state bridges adrenergic and mesenchymal neuroblasts
- Transitional gene signatures associate with aggressive disease and poor prognosis

## Authors

Xiaojun Yuan, Janith A. Seneviratne, Shibei Du, ..., Weiyang Shi, Daniel R. Carter, Chao Zhang

## Correspondence

wshi@ouc.edu.cn (W.S.), daniel.carter@uts.edu.au (D.R.C.), zhangchao@tongji.edu.cn (C.Z.)

## In brief

Yuan et al. undertook single-cell RNA sequencing on 10 peripheral neuroblastic tumors (PNTs), including neuroblastoma, ganglioneuroblastoma, and ganglioneuroma. They used single-cell gene expression to model interconverting malignant cell transitions between the so-called adrenergic and mesenchymal cell types and identify an aggressive bridging cell type referred to as "transitional" cells.

Yuan et al., 2022, Cell Reports *41*, 111455
October 4, 2022 © 2022 The Authors.
https://doi.org/10.1016/j.celrep.2022.111455

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

# Single-cell profiling of peripheral neuroblastic tumors identifies an aggressive transitional state that bridges an adrenergic-mesenchymal trajectory

Xiaojun Yuan,[1,8] Janith A. Seneviratne,[2,8] Shibei Du,[1,8] Ying Xu,[3,8] Yijun Chen,[3] Qianya Jin,[1] Xuanxuan Jin,[3] Anushree Balachandran,[2] Shihao Huang,[1] Yanli Xu,[1] Yue Zhai,[3] Liumei Lu,[3] Mengjie Tang,[1] Yushuang Dong,[1] Belamy B. Cheung,[1,2,4] Glenn M. Marshall,[2,4,5] Weiyang Shi,[6,9,*] Daniel R. Carter,[2,3,4,7,9,10,*] and Chao Zhang[3,9,*]

[1]Department of Pediatric Hematology & Oncology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University, School of Medicine, Shanghai, China

[2]Children's Cancer Institute Australia for Medical Research, Lowy Cancer Research Centre, UNSW Sydney, Kensington, NSW, Australia

[3]Fundamental Research Center, Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Life Sciences and Technology, Tongji University, Shanghai 201619, China

[4]School of Women's and Children's Health, UNSW Sydney, Randwick, NSW 2031, Australia

[5]Kids Cancer Centre, Sydney Children's Hospital, Randwick, NSW 2031, Australia

[6]Ministry of Education Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao, China

[7]School of Biomedical Engineering, University of Technology Sydney, Sydney, NSW, Australia

[8]These authors contributed equally

[9]Senior author

[10]Lead contact

*Correspondence: wshi@ouc.edu.cn (W.S.), daniel.carter@uts.edu.au (D.R.C.), zhangchao@tongji.edu.cn (C.Z.)

https://doi.org/10.1016/j.celrep.2022.111455

## SUMMARY

Peripheral neuroblastic tumors (PNTs) represent a spectrum of neural-crest-derived tumors, including neuroblastoma, ganglioneuroblastoma, and ganglioneuroma. Malignant cells in PNTs are theorized to interconvert between adrenergic/noradrenergic and mesenchymal/neural crest cell states. Here, single-cell RNA-sequencing analysis of 10 PNTs demonstrates extensive transcriptomic heterogeneity. Trajectory modeling suggests that malignant neuroblasts move between adrenergic and mesenchymal cell states via an intermediate state that we term "transitional." Transitional cells express programs linked to a sympathoadrenal development and aggressive tumor phenotypes such as rapid proliferation and tumor dissemination. Among primary bulk tumor patient cohorts, high expression of the transitional gene signature is predictive of poor prognosis compared with adrenergic and mesenchymal expression patterns. High transitional gene expression in neuroblastoma cell lines identifies a similar transitional H3K27-acetylation super-enhancer landscape. Collectively, our study supports the concept that PNTs have phenotypic plasticity and uncovers potential biomarkers and therapeutic targets.

## INTRODUCTION

Peripheral neuroblastic tumors (PNTs) represent a spectrum of tumors derived from the neural crest and account up to 8%–10% of all pediatric malignancies. A salient feature of PNTs is a heterogeneous clinical course ranging from spontaneous regression to persistent disease progression (Matthay et al., 2016). Histologically, PNTs comprise four variants: neuroblastoma, ganglioneuroblastoma nodular (GNBn), ganglioneuroblastoma intermixed (GNBi), and ganglioneuroma (GN) (Matthay et al., 2016). GN and GNBi are low grade in nature and usually curable by surgical resection alone (De Bernardi et al., 2008). In contrast, the most common subtype, neuroblastoma, is often lethal. Despite intensive treatments, the long-term survival of high-risk neuroblastoma is less than 50% (Pinto et al., 2015). Around half of high-risk patients relapse after initial treatment response, and salvage therapies for relapsed patients are rarely effective (Matthay et al., 2016). Moreover, genomics studies comparing longitudinal samples from the same patient show that clonal evolution is a prominent feature of disease progression (Schramm et al., 2015; Eleveld et al., 2015; Padovan-Merhar et al., 2016). Therefore, a better understanding of tumor heterogeneity will be required to improve therapy for patients.

Recent single-cell studies in neuroblastoma have elegantly characterized malignant cells of neuroblastoma tumors in terms of normal cell counterparts of the sympathoadrenal development, in particular, aiming to provide insight to a neuroblastoma cell of origin (Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Hanemaaijer et al., 2021; Dong et al., 2020). Interestingly, these studies have not


This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

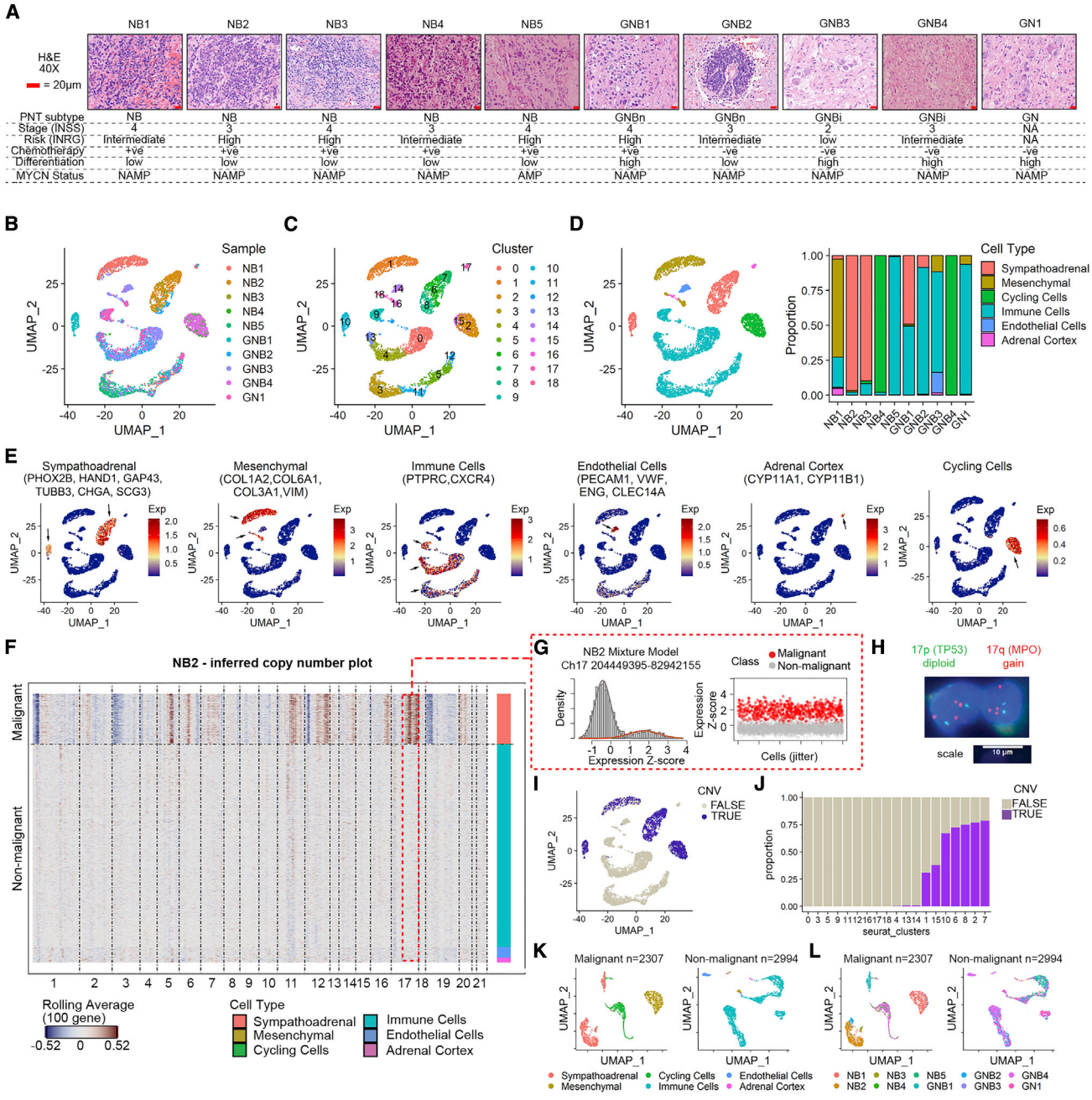



**Figure 1. scRNA-seq of peripheral neuroblastic tumors**

(A) Representative hematoxylin and eosin (H&E) staining in neuroblastoma (n = 5), GNB (n = 4), and ganglioneuroma (n = 1) tumor samples. Scale bar: 20 μm.

(B) Uniform manifold approximation and projection (UMAP) plots of 5,301 cells from 10 tumor samples colored by original sample.

(C) UMAP plots of 5,301 cells from 10 tumor samples colored by transcriptomic cluster.

(D) Left: UMAP plots of 5,301 cells classified by preliminary cell type based on marker gene expression from Figure 1E. Right: proportion of preliminary cell types within each sample.

(E) UMAP plots of 5,301 cells from all 10 tumor samples showing expression of marker gene signatures (blue, low; dark red, high). Arrows indicate regions of high expression.

(F) Heatmap of copy number variations (CNVs) in the NB2 sample for individual cells (rows), visualized by a rolling gene expression average centered on 100 gene windows at each chromosomal position (columns). Dark red, copy number gains; blue, copy number losses.

(G) Left: two-component Gaussian mixture models (GMMs) for NB2 *Ch17* region. GMM components are shown in different colors and overlaid on a density plot of 100-gene rolling average expression. Right: scatterplot visualization 100-gene rolling average expression for each cell. Malignant cells, red; non-malignant cells, gray.

*(legend continued on next page)*

entirely agreed on their description of cell types in neuroblastoma, which has most likely arisen due to disparate use of cell type markers and variable methods of cell type annotation or technologies/samples used (Ponzoni et al., 2022). Nevertheless, these studies generally have a consensus that the majority of neuroblastoma cells appear to have a developmental origin similar to committed sympathoadrenal cell types such as sympathoblasts or chromaffin cells (Ponzoni et al., 2022). There are observations, though, that link to less differentiated cell types such as undifferentiated chromaffin or sympathoblast subtypes, as well as more immature cells of neural crest development with mesenchymal-like features such as Schwann cell precursors, and so-called bridge cells (Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Hanemaaijer et al., 2021; Dong et al., 2020; Ponzoni et al., 2022). Moreover, a recent article suggested that high-risk neuroblastoma may arise within postnatal cholinergic progenitors, potentially explaining the long-recognized disparity between neuroblastoma in older (>18 months) and younger individuals (<18 months) in terms of genetic features and their long-term prognosis (Bedoya-Reina et al., 2021). Collectively, while creating valuable insights into the developmental origins of neuroblastoma, the variable interpretations from these studies also highlight the challenges associated with malignant cell type classification using normal cell counterparts as a reference.

Another feature of recent single-cell studies in neuroblastoma is that malignant cell types likely represent a continuum of differentiation intermediates rather than discrete classes and highlight potential for neuroblastoma cell plasticity. Variable cell phenotypes may arise due to a unique cell of origin in different tumors, processes of dedifferentiation driven by genetic alteration in tumor cells or using mechanisms akin to noradrenergic to mesenchymal transition or vice versa based on environmental exposure (Ponzoni et al., 2022; Gautier et al., 2021). Importantly, emerging evidence proposes that spontaneous tumor plasticity is readily observed in neuroblastoma cells and that cell identity can be reprogrammed with lineage-associated transcription factors PHOX2B, PRRX1, ASCL1, and NOTCH3 (van Groningen et al., 2017, 2019; Boeva et al., 2017; Wang et al., 2019). These studies conjointly suggest that neuroblastoma is composed of transdifferentiating malignant cells, for which unique core-regulatory transcription factor circuits control distinct epigenetic and transcriptomic landscapes. In these studies, tumors and neuroblastoma cell lines were subtyped according to a two-group classification conforming to an adrenergic/noradrenergic state or a mesenchymal/neural crest cell state. Importantly, cell state was seen to be relevant for therapeutic efficacy, with mesenchymal neuroblastic cells more resistant to conventional anticancer therapies and enriched in relapsed tumors (van Groningen et al., 2017; Gartlgruber et al., 2021; Westerhout et al., 2022).

Considering the variable interpretations from single-cell studies with respect to neuroblastoma cell of origin and cell identity (Ponzoni et al., 2022), here we present an alternative single-cell RNA sequencing (scRNA-seq)-based classification approach for primary neuroblastoma that does not utilize a normal adrenal cell reference to supervise annotation. Our focus is to describe malignant cells in terms of an observed transdifferentiation spectrum between adrenergic and mesenchymal cell types. For this, we conducted modified Smart-seq2-based scRNA-seq on more than 5,301 cells of fresh, surgically resected tissues from 10 PNT patients, which spanned all histological variants, including five neuroblastoma, two GNBn, two GNBi, and one GN lesion. Cell type abundance of PNT microenvironments differs between neuroblastoma and other PNT subtypes. Moreover, our results suggest that malignant neuroblasts can exist as an intermediate between adrenergic and mesenchymal transcriptional states via a previously undescribed transitional population. Transitional neuroblasts have an aggressive neurodevelopmental phenotype not ascribed as a particular feature of the sympathetic of chromaffin lineage, and are similar to highly proliferative, disseminated tumor cells. Finally, we found that transitional gene expression signatures predict poorer patient prognosis in large cohorts of neuroblastoma.

## RESULTS

### Single-cell transcriptomics analysis of PNTs

We conducted scRNA-seq using a modified 3′ unique molecular identifier (UMI) based version of the Smart-seq2 protocol (Picelli et al., 2014) on 10 PNTs (Figure S1A). We used viable single cells (DRAQ5+/Calcein Blue+) derived from five neuroblastomas, four ganglioneuroblastomas (GNBs), and one GN (Figure S1B). Samples were acquired from surgical resection and had a range of clinical and histological features (Figure 1A and Table S1). We used quality control measures, such as the number of detected genes/cell (500 > n > 750,000 UMI per cell, 500 < × < 9,500 unique genes per cell, where a given gene had to be expressed in at least three cells), proportion of mitochondrial gene counts per cell (<0.15), and proportion of ribosomal gene counts per cell (<0.3), to filter out poor-quality cells (Figures S1C–S1G). Following quality control, the median read depth was 27,228 UMIs per cell with 1,568 unique genes detected per cell (Figures S1D and S1E). Ultimately, we yielded 5,301 high-quality single-cell transcriptomes for downstream analysis across the 10 PNTs (range, 49–892 single cells per tumor). Using 4,000 highly variable genes, we implemented principal-component analysis (PCA) and selected the top 12 principal components for graph-based cell clustering (Figures S1H and S1I). This analysis identified 19 distinct

(H) Representative images for fluorescence *in situ* hybridization (FISH) of NB2 Ch17 copy number gain. Probes for MPO gene on *Ch17* q-arm (red) compared with *TP53* on *Ch17* p arm (green).
(I) UMAP plot of 5,301 cells colored by CNV classification. Confident CNV inferences were made for cells with a posterior probability >0.75 of CNV. Only FISH-validated CNVs were used.
(J) Proportion of cells within each cluster based on CNV classification. Clusters 1, 15, 10, 6, 8, 2, and 7 were considered to be malignant.
(K) UMAP plots showing preliminary cell type classification of left, malignant cells (n = 2,307), or right, non-malignant cells (n = 2,994).
(L) UMAP plots showing original sample identity of left, malignant cells (n = 2,307), or right, non-malignant cells (n = 2,994).

clusters (clusters 0–18), which were projected by uniform manifold approximation and projection (UMAP) (Figures 1B and 1C).

To assign each cluster to a cell phenotype, we undertook a preliminary sub-classification that considers differential expression and the expression of marker genes for major cell lineages (Figures 1D, 1E, S1J, and S1K). Established markers were used for these assignments, with the cycling cluster annotated based on high expression of a cell cycle gene set (Tirosh et al., 2016). Based on these cell type assignments, the proportion of cell types within individual tumors differed markedly (Figure 1D). Notably, most clusters conformed to broad lineage classifications that would be expected to occur in PNTs (Boeva et al., 2017; van Groningen et al., 2017; Pezzolo et al., 2007, 2011; Zeine et al., 2009; Asgharzadeh et al., 2012; Mina et al., 2015; Pelizzo et al., 2018), such as sympathoadrenal, cycling, immune, mesenchymal, endothelial, and adrenal cell types (Figures 1D and 1E).

## Identification of malignant neuroblasts by copy number variation inference

With the knowledge that malignant cells in neuroblastic tumors can exist in multiple differentiated forms (Boeva et al., 2017; van Groningen et al., 2017; Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Dong et al., 2020), classification of malignant cells using gene expression alone may be inaccurate. We therefore aimed to distinguish malignant cells from non-malignant cells using copy-number analysis in single-cell RNA-sequencing (CONICS) copy number estimation as an additional method of classification (Muller et al., 2018). This technique uses the expression of a sliding window of 100 genes across each chromosome to infer copy number variations (CNVs), so that malignant cells with greater copy number instability can be identified (Figures 1F and S1L). Two-component Gaussian mixture models were calculated to identify single-cell copy number estimations for various regions of the genome, using the cells of each tumor (for example, Figure 1G; Table S2). We verified predicted CNVs in samples by fluorescence *in situ* hybridization (FISH) (Figures 1H and S1M; Table S2). Individual cell CNV status was calculated on the posterior probability that a cell belonged to the predicted CNV for these FISH-validated regions. Projection of inferred CNVs on the transcriptomic UMAP plot showed that the CNV distribution was non-random and localized to specific clusters of cells (Figures 1I and 1J). Using these inferred CNVs, we therefore concluded that transcriptomic clusters (1, 2, 6, 7, 8, 10, and 15) corresponding to sympathoadrenal, cycling, and some mesenchymal clusters were malignant cells, while all other cells were considered non-malignant (Figures 1J and 1K). After splitting the cells into malignant (n = 2,307) and non-malignant (n = 2,994) groups, UMAP projection of malignant cells separated mostly into distinct subpopulations associated with individual patient tumors, suggesting pronounced intertumoral heterogeneity among cancer cells (Figure 1L, left panel). In contrast, non-malignant cells tended to cluster independent of tumor origin (Figure 1L, right panel), consistent with previous reports showing non-malignant cells cluster by cell type rather than the tumor that they derive from (Puram et al., 2017; Tirosh et al., 2016).

## Neuroblastomas and GNBs have distinct tumor microenvironmental cell types

Next, we examined gene expression characteristics of cells in the PNT microenvironment. In this cohort, the proportion of infiltrating immune cells showed a trend that indicated it was higher in the GNB and GN tumors compared with the five neuroblastoma samples, although significant variation existed even within each histological subtype (Figure 2A). When we excluded malignant cells and interrogated cell type abundance in each tumor as a proportion of all non-malignant cells, we found 12 distinct cell clusters (Figure 2B). Based on differential expression analysis and known markers of cell type, we then classified cells into subtypes, and found that main clusters conformed to recognized cell types from the immune, adrenal, fibroblast, smooth muscle, or endothelial lineages (Figures 2C, 2D, and S2A–S2C). When comparing the relative abundance of different cell types, most of the neuroblastomas showed a trend that indicated they had fewer total infiltrating T cells, including cytotoxic T cells (measured by CD8 and presence of cytolytic effector genes) than the non-neuroblastoma samples (Figures 2D, S2B, and S2C, upper panels). Consistent with these findings, immunohistochemistry of tissue sections from each of the 10 tumors showed that pan-T (CD3)-positive cells were more abundant in GNB2, GNB3, and GN samples than most neuroblastomas (Figure S2D). Evaluation of other cell types revealed that the majority of neuroblastoma samples had more macrophages, including non-inflammatory macrophages, with increased expression of the M2-polarization marker CD163 (Figures 2D, S2B, and S2C, lower panels). Our analysis of the cellular makeup of the tumor microenvironment (TME) showed a trend of increased T cell infiltration in GNBs and increased macrophage infiltration in neuroblastoma. This suggests a potential role of immune evasion and a pro-tumorigenic microenvironment in the more aggressive neuroblastomas, compared with non-neuroblastomas, although more samples and more specific technologies would be better suited to confirm this finding.

## Modeling of adrenergic/mesenchymal transdifferentiation identifies a transitional cell phenotype

For the malignant cell analysis, we then evaluated specific expression patterns. Given that cells detected as CNV positive included those cells annotated as cycling, we accounted for this by regressing cell cycle scores during normalization (Tirosh et al., 2016). To evaluate cell phenotype among malignant cells, we conducted differential expression analysis and identified divergent cell transcriptomes highlighted by expression of adrenergic (CHGA, CHGB, PHOX2B, DBH) and mesenchymal marker genes (COL1A1, COL1A2) (Figures S3A and S3B). We therefore investigated the possibility that malignant cells exist in either an adrenergic/noradrenergic or mesenchymal/neural crest cell state, as proposed by Boeva et al. (2017) and van Groningen et al. (2017). We first examined expression patterns of signatures for these cell states. UMAP projection of these two gene signatures showed expression that was enriched in distinct clusters

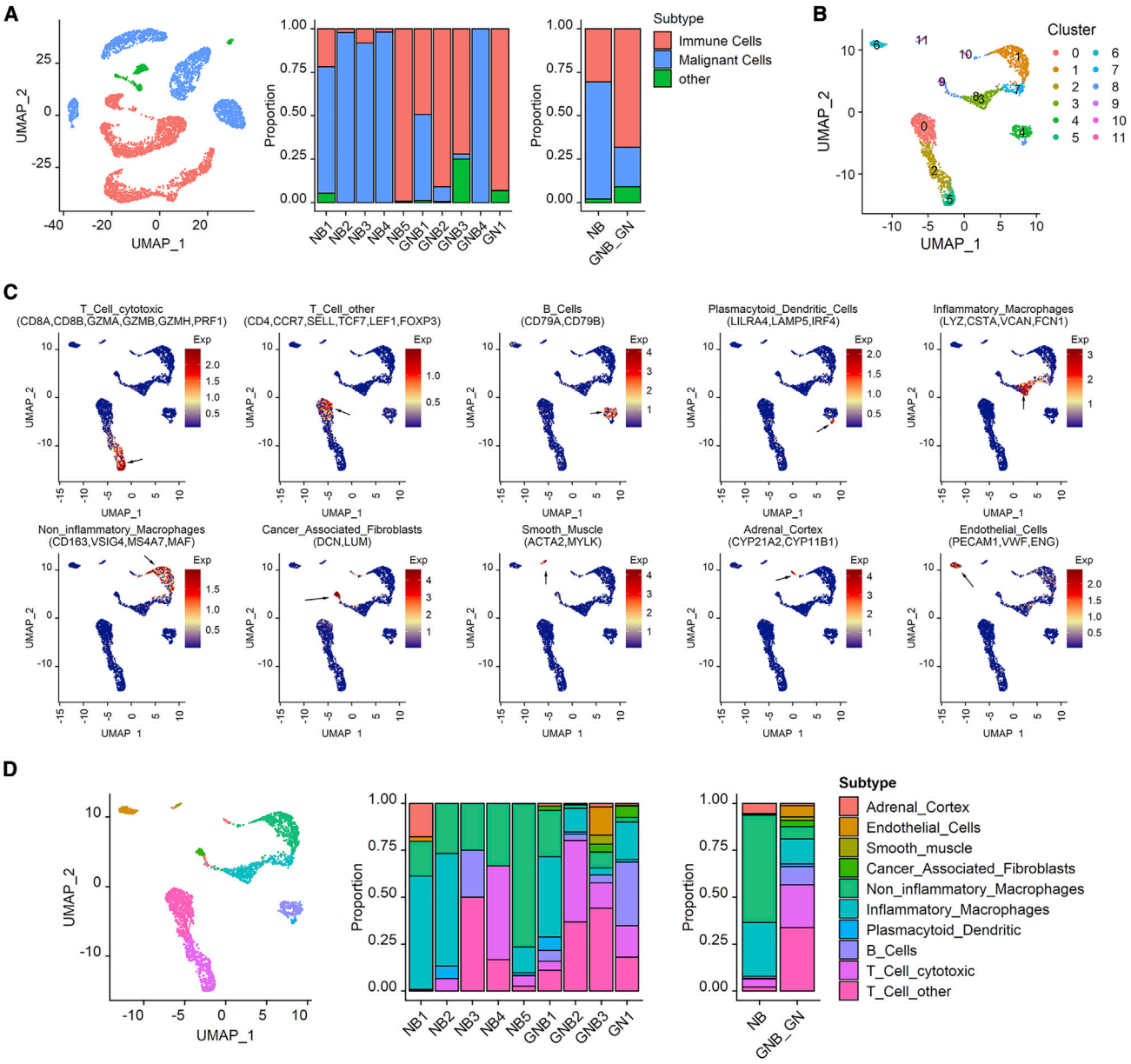



**Figure 2. Neuroblastomas and GNBs have distinct tumor microenvironmental cell types**

(A) UMAP plot showing malignant (n = 2,307), immune (n = 2,684), and other (n = 310) cell types from 10 tumors. The proportion of cell types within each tumor and major tumor types are also shown. NB, neuroblastoma; GNB_GN, GNB plus ganglioneuroma.

(B) UMAP plot of 2,994 non-malignant cells from 10 tumors colored based on transcriptomic cluster.

(C) UMAP plot of 2,994 non-malignant cells from 10 tumors showing expression of marker gene signatures (blue, low; dark red, high). Arrows indicate regions of high expression.

(D) Left: UMAP plot showing non-malignant cell subtypes of non-malignant cells (n = 2,994) from 10 tumors. Right: the proportion of non-malignant cell types within each tumor and major tumor types.

whether using the van Groningen (adrenergic/mesenchymal) or Boeva signatures (noradrenergic/neural crest cell) (Figures 3A–3D). When comparing directly, the expression patterns of adrenergic and mesenchymal signatures showed a strong inverse association, with cells existing in a continuum between high-expressing adrenergic and mesenchymal cells (Figures S3C and S3D). This suggests that cells may have an identity that reflects various points of transdifferentiation between adrenergic and mesenchymal cell states, similar to the model proposed by van Groningen et al. (2017) and Boeva et al. (2017). To explore this hypothesis further, we conducted pseudotemporal ordering of cells (Trapnell et al., 2014) to model cell transitions that would occur during adrenergic-mesenchymal transdifferentiation (Figure 3E). Interestingly, the predicted trajectory was more complex than a

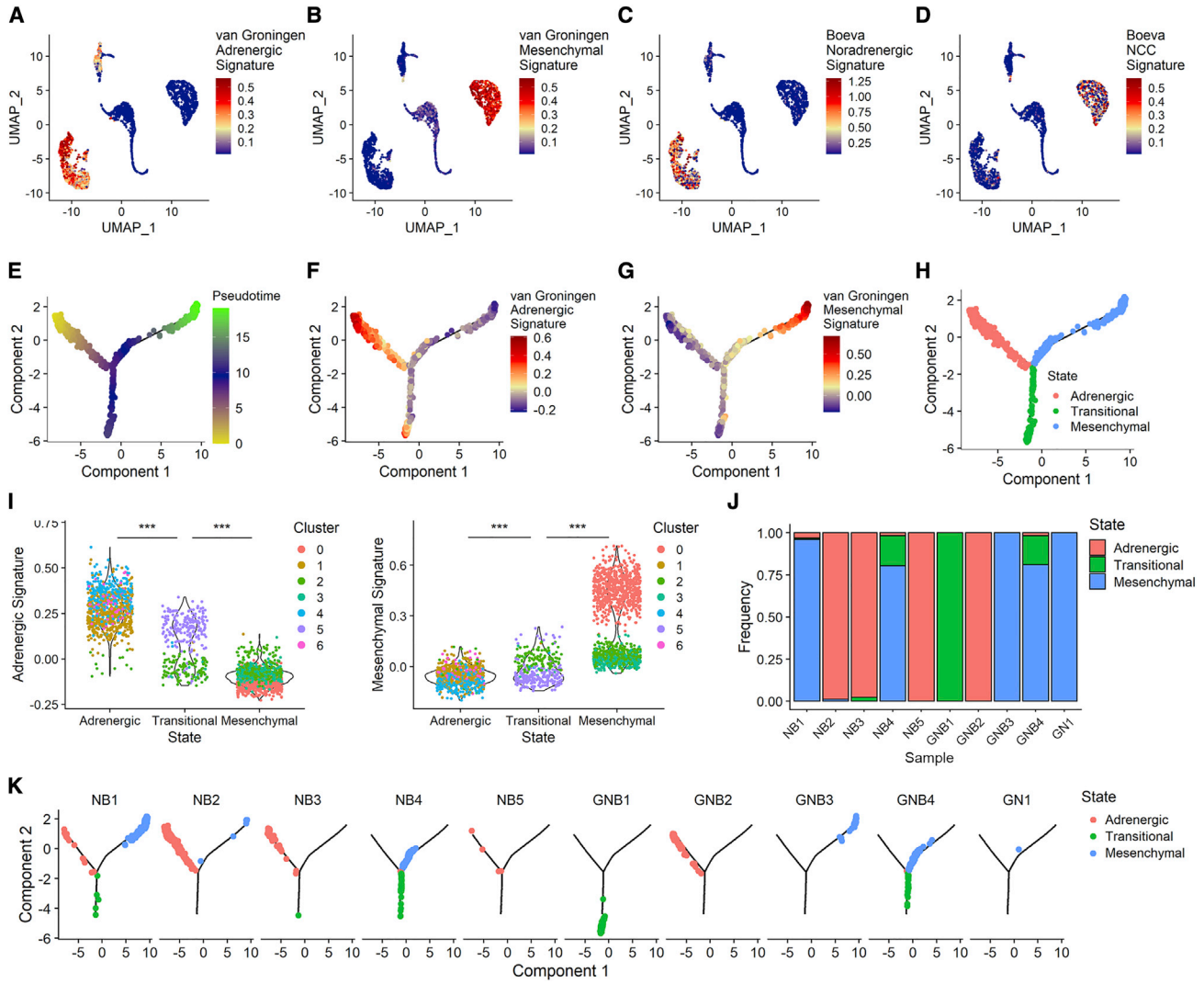



**Figure 3. Trajectory modeling of malignant neuroblasts identifies a transitional phenotype as an intermediate state between adrenergic and mesenchymal states**

(A–D) UMAP of 2,307 malignant cells showing expression of previously published marker gene signatures (van Groningen et al., 2017; Boeva et al., 2017) (blue, low; dark red, high) (van Groningen et al., 2017) (Boeva et al., 2017) (Boeva et al., 2017).

(E–G) Trajectory-based inference of malignant cells (n = 2,307) using Monocle2 (DDRTree) (Trapnell et al., 2014). (E) Pseudotime was modeled based on expressed genes from previously defined signatures (van Groningen et al., 2017) (yellow: low, green: high), (F) Adrenergic marker gene signature expression (van Groningen et al., 2017) (blue, low; dark red, high), (G) Mesenchymal marker gene signature expression (van Groningen et al., 2017) (blue, low; dark red, high).

(H) Cells colored by inferred cell state.

(I) Violin plots of adrenergic (left) and mesenchymal (right) gene signatures for all malignant cells split by inferred cell state in Figure 3H. Cells are colored according to transcriptomic cluster in Figure S3A, and p values determined using t test (***$p < 0.001$).

(J) Proportion of malignant cells that are either in an adrenergic, transitional, or mesenchymal cell state, within each tumor.

(K) Trajectory plots split by tumor sample, showing malignant cell state (in H).

simple linear path running between adrenergic and mesenchymal cell states. Examination of the adrenergic and mesenchymal signatures showed a unique state as a separate arm between the high adrenergic and high mesenchymal expression states, which we refer to as transitional cells (Figures 3F–3H). Gene expression analysis of this transitional cell state demonstrated an intermediate expression level between adrenergic and mesenchymal signatures that mostly corresponded to cluster 2, 3, and 5 cells (Figure 3I). Moreover, since transitional cells predominantly formed as a separate cluster, they showed similar proportion of CNV-positive cells as other malignant clusters (Figure 1J, see cluster 10). We also considered whether cells derived from each tumor contained cells from more than one neuroblastic state, as would be expected if cells had the capacity to transdifferentiate. Indeed, most tumors had evidence of cells in more than one neuroblastic state, with a tendency that one subtype would be dominant with a secondary minor population (Figures 3J, 3K, and S3E). These data suggest that extensive cell phenotype heterogeneity exists

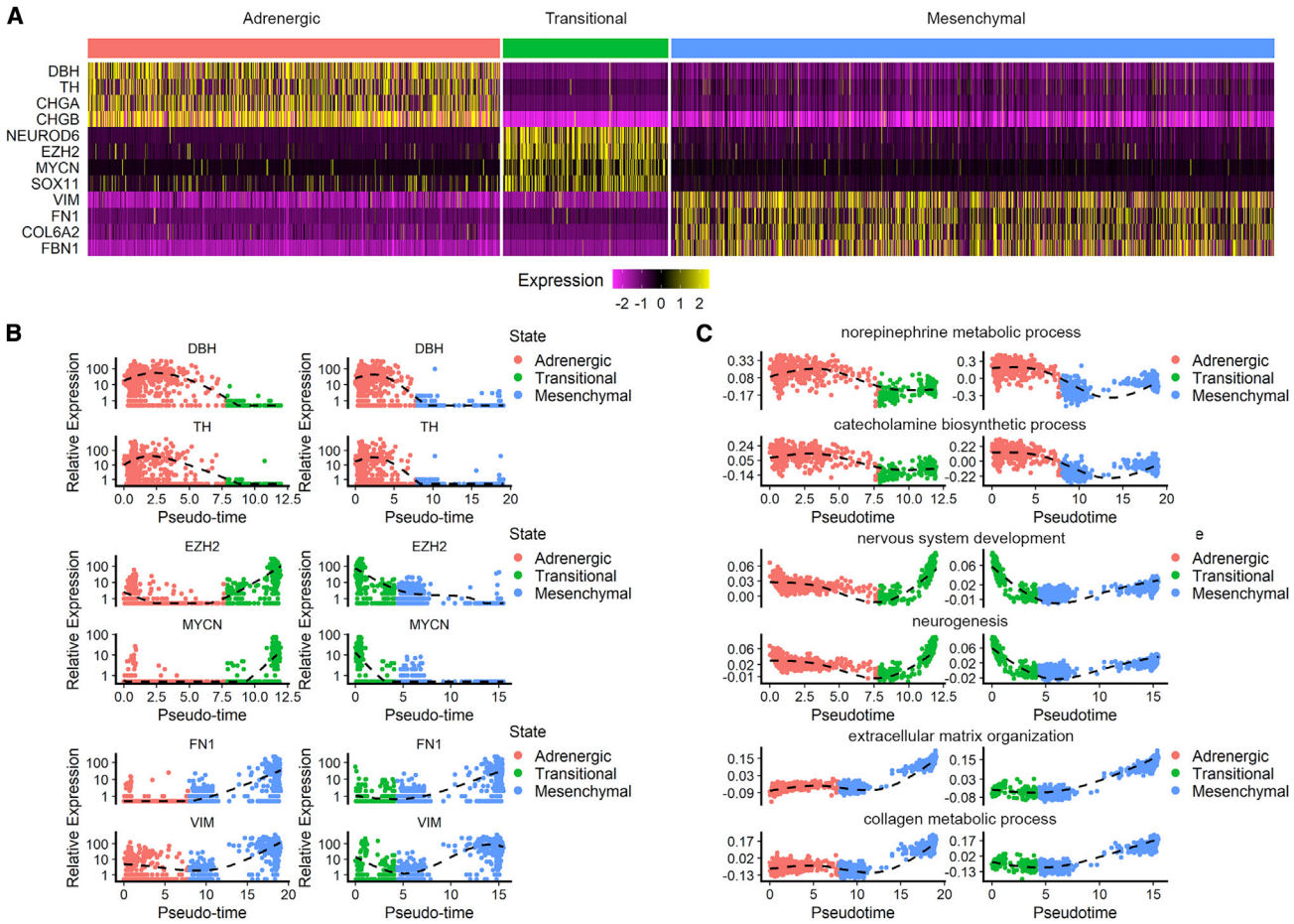



**Figure 4. Malignant neuroblasts in PNTs have gene expression characteristics representative of various differentiated stages in the developing neural crest**

(A) Gene expression heatmap of the representative differentially expressed genes (rows) of cells in either an adrenergic, transitional, or mesenchymal cell state (annotated columns). Gene expression *Z* scores (purple, low; yellow, high).

(B and C) Gene (B) or gene ontology (GO) (C) signature expression of individual cells ordered by pseudotime along arms of the trajectory in Figure 3H. Dashed line represents natural cubic splines fit. Pseudotime was calculated considering different arms of the trajectory.

in PNTs, with cell state corresponding to various points along a modeled transdifferentiation trajectory between adrenergic and mesenchymal cell types.

## Transitional neuroblasts are defined by an aggressive neurodevelopmental phenotype

In order to clarify transcriptomic change during cell state transition, we identified marker genes in each of the three cell states, adrenergic, transitional, and mesenchymal (Figures 4A, S4A, and S4B). Evaluation of changes in gene expression using pseudotime to visualize cell trajectories showed that the expression of these genes progressively changed when comparing different states (Figures 4B and S4C). Key noradrenergic and catecholaminergic enzymes expressed in the adrenergic state, DBH and tyrosine hydroxylase (TH), were downregulated along trajectories to either the transitional or mesenchymal arm (Figure 4B, top panel). The mesenchymal state showed upregulation of mesenchymal neuroblast markers FN1 and VIM, consistent with a role in extracellular matrix production expected in this cell type (Figure 4B, bottom panel). In transitional neuroblasts, we found an overexpression of well-known neuroblastoma genes EZH2 and MYCN (Figure 4B, middle panel). While *MYCN* is an established driver gene in the context of neuroblastoma when gene amplified (Schwab et al., 1983), this was an interesting finding since *MYCN* is non-amplified in all but one of the tumors of this cohort, and, in this case (NB5), cells were restricted to the adrenergic arm (Figures 3J and 3K). This upregulation of MYCN seems mostly to have arisen from overexpression in GNB1, the only sample dominated by a transitional cell phenotype (Figures 3J and 3K). EZH2, on the other hand, is a core component of polycomb repressive complex 2 and is responsible for the catalysis of H3K27 tri-methylation. EZH2 plays an essential role in tumorigenesis in neuroblastoma, and aberrant expression of EZH2 is strongly associated with poor prognosis (Chen et al., 2018; Li et al., 2018; Tsubota et al., 2017). Next, we employed Gene Ontology (GO) analysis on gene signatures created for each of the three cell states. GO-term enrichment identified gene sets as expected related to

neurotransmitter production and extracellular matrix as part of the adrenergic and mesenchymal states, respectively (Figures 4C, S4D, and S4E). Transitional neuroblasts, however, had gene sets relating to neurogenesis (Figures 4C, S4D, and S4E), suggesting that differentiation via transitional neuroblasts may be a developmentally co-opted process. Indeed examination of scRNA-seq data in the E12.5 murine neural crest compartment (Furlan et al., 2017) shows that the transitional signature has relatively high expression in the bridge cells that act as differentiation intermediates between the Schwann cell precursors of the neural crest (mesenchymal-like) and the terminal chromaffin cells and neuroblasts of the adrenal medulla (adrenergic-like) (Figures S4F and S4G). However, exploration of the developmental origins of neuroblastoma cells have been covered more extensively in other recent single-cell studies, which show varying interpretations (Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Hanemaaijer et al., 2021; Dong et al., 2020; Ponzoni et al., 2022), so our conclusions around developmental origins of the cells in our cohort are tentative.

We next investigated whether the three cell states related to biological processes are important in tumorigenesis. To gain more insight into the cell cycle status of the malignant cells across the different phenotypes, we determined their cell cycling states based on the average expression levels of genes within the S and G2/M (dividing) and G1 (non-dividing) gene sets previously published (Tirosh et al., 2016). This analysis revealed a markedly higher proportion of dividing cells among the transitional neuroblasts compared with the other two cell states (Figure S4H). Next, we evaluated whether any of the three cell states related to a metastatic phenotype. For this, we calculated a disseminated tumor cell (DTC) score for gene expression in each cell state by determining the difference between the expression of significantly upregulated and downregulated genes of a previously published DTC gene set (Rifatbegovic et al., 2018). Interestingly, the transitional malignant cells shared the highest similarity with DTCs from the bone marrow, which is the most common distant metastatic site of neuroblastoma (Figure S4I). Consistent with this, the tumor with the highest abundance of transitional neuroblasts, GNB1, was graded as stage 4 with distant metastasis by the International Neuroblastoma Staging System (INSS), further suggesting a possible link between the transitional phenotype and metastasis. Collectively, these data suggest the presence of distinct malignant cell subpopulations within individual tumors, which demonstrate divergent differentiation status, varying transcriptional signatures, and potential for malignant clinical behavior.

## Tumors expressing transitional neuroblast signatures are associated with poor prognosis in neuroblastoma patients

We next studied expression data from large primary neuroblastoma tumor cohorts where gene expression was derived from total tumor or bulk gene expression data (Kocak cohort, n = 649. GEO: GSE45480) (Kocak et al., 2013). We classified tumors based on adrenergic, transitional, and mesenchymal gene signatures (Figure 5A). The signatures separated five tumor groups, ranging from adrenergic, transitional, and mesenchymal, as well as two mixed classes of adrenergic-transitional, and transitional-mesenchymal states (Figure 5A). This is consistent with the notion that, in some cases, bulk tumors can have mixed proportions of the three cell states identified in our scRNA-seq data. To explore whether the three cell states had prognostic relevance, we conducted Kaplan-Meier analysis for the five-tumor subgroups defined by these signatures. Transitional tumor classes were shown to predict poorer outcome than adrenergic and mesenchymal subgroups, which was driven mostly by poorer prognosis in the transitional-only and transitional-mesenchymal subgroup (Figure 5B). Additionally, higher expression of transitional neuroblast signatures was observed in patient subgroups defined by poor prognostic factors: *MYCN* amplification and INSS 4 disease (Figure 5C). This suggests that transitional cell state signatures are associated with tumors with high risk of resistance to conventional therapy and relapse. Consistent with these findings, Kaplan-Meier analyses for gene signatures subdividing the cohort at the median showed a similar association between transitional neuroblasts and poor prognosis, as did similar analyses in a separate cohort, albeit statistical significance was dependent on the analysis conducted (Versteeg cohort, n = 88. GEO: GSE16476) (Molenaar et al., 2012) (Figures S5A–S5E).

## Neuroblastoma cell lines with high transitional gene signature expression have an intermediate super-enhancer landscape between adrenergic and mesenchymal states

Previous studies used H3K27-acetylation (H3K27ac) landscapes to define two distinct super-enhancer states, either adrenergic/noradrenergic or mesenchymal/neural crest (Boeva et al., 2017; van Groningen et al., 2017). Since our trajectory analysis using gene expression in single cells identified a transitional population between these two identities, we investigated whether transitional gene expression signatures in neuroblastoma cell lines similarly identify intermediate super-enhancer profiles. As before, we used adrenergic, transitional, and mesenchymal signatures to classify 33 neuroblastoma cell lines with paired bulk RNA sequencing and H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) data (Boeva et al., 2017) into five groups: either adrenergic, adrenergic-transitional, transitional, transitional-mesenchymal, or mesenchymal (Figure S6A). We then conducted comparisons to identify super-enhancers that were enriched in either the adrenergic class or the mesenchymal class. Of the 5,975 previously annotated super-enhancer loci across these 33 cell lines (Boeva et al., 2017), we found 531 were enriched in the adrenergic state (versus mesenchymal cell lines), and 2,004 were enriched in the mesenchymal state (versus adrenergic cell lines) (Figures 6A and S6B). To evaluate whether a progressive shift in H3K27ac occurs through cell lines according to our cell state classification, we investigated whether adrenergic or mesenchymal-associated super-enhancers showed any trend across different cell classes. H3K27ac levels at adrenergic-associated super-enhancers progressively decreased through the spectrum of adrenergic, adrenergic-transitional, transitional, transitional-mesenchymal, and mesenchymal cell lines (Figure 6A, left panel; Figures 6B and S6B, upper panel; Figure S6C). In contrast, H3K27ac levels

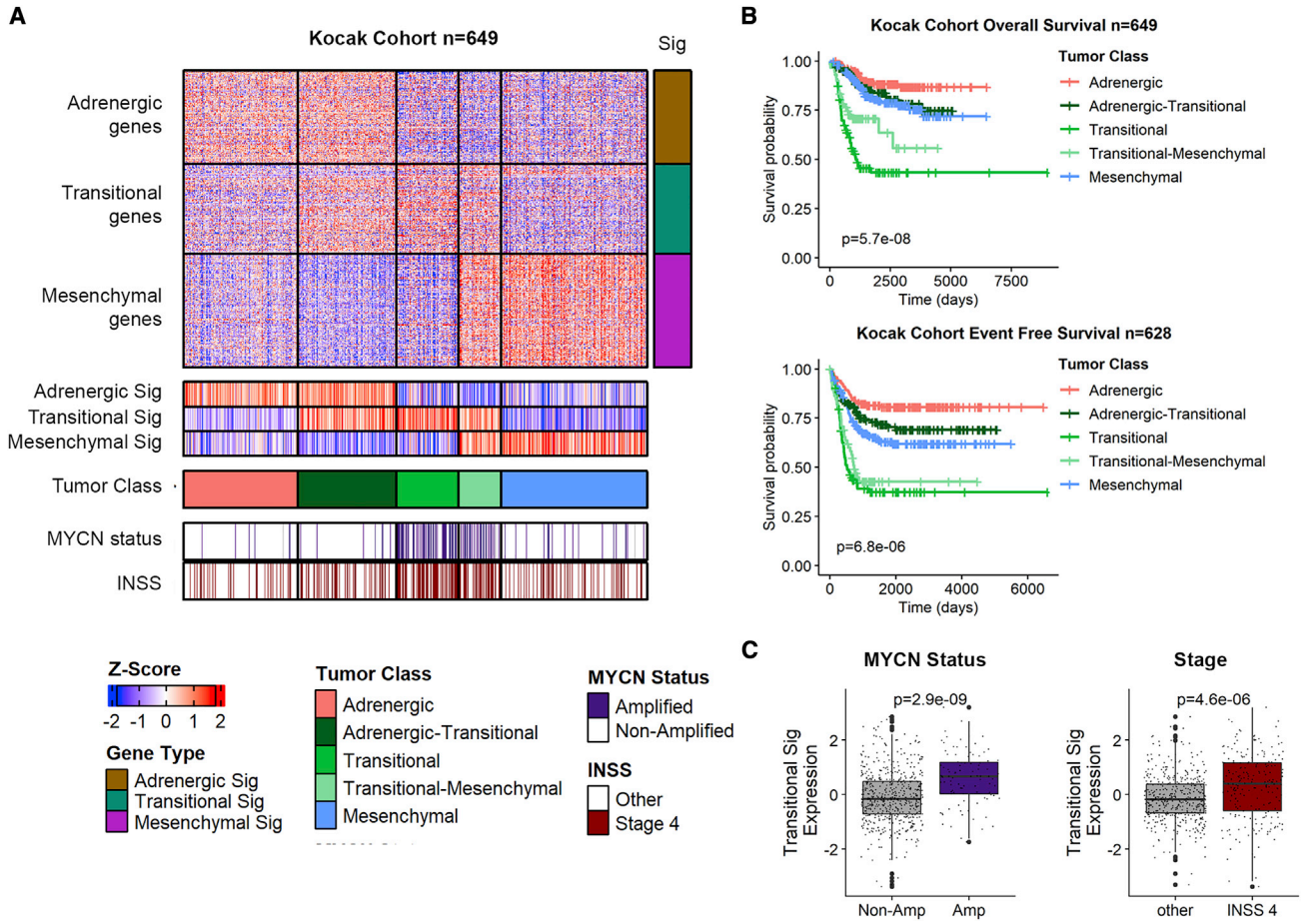



**Figure 5. Tumors expressing transitional neuroblast signatures are associated with poor prognosis in neuroblastoma**

(A) Heatmap showing the classification of neuroblastoma tumors from bulk mRNA expression (Kocak, n = 649 patients; Kocak et al., 2013) based on adrenergic, transitional, and mesenchymal gene signatures. Column averages were used for each signature (shown immediately below the heatmap) to classify tumors to tumor class subgroups. *MYCN* status (amplified) and neuroblastoma staging (INSS 4) for each tumor is annotated. Scale bar indicates gene expression *Z* scores.

(B) Kaplan-Meier plots of overall survival (top) and event-free survival (bottom) in the Kocak cohort colored based on tumor class classifications made in Figure 5A. The p value was calculated using log rank tests, comparing all transitional subgroups combined with adrenergic and mesenchymal combined.

(C) Boxplots of transitional gene signature expression in patients of the Kocak cohort, split by *MYCN* status (left, non-amplified versus amplified) or INSS staging (right, stage 1, 2, 3, and 4S versus stage 4). The p values are reported from two-sided t tests.

at mesenchymal-enriched super-enhancers progressively increased through the same spectrum (Figure 6A, right panel; Figures 6B and S6B, lower panel; Figure S6C). Moreover, H3K27ac tracks show a similar trend in super-enhancers linked to core-regulatory circuit transcription that were previously found to define adrenergic and mesenchymal identities (Boeva et al., 2017) (Figure 6C). Interestingly, transitional-enriched super-enhancers did exist, albeit only in the minority, with 395 unique super-enhancers significantly enriched in transitional cells compared with adrenergic and mesenchymal cells (Figure S6D). There were, however, some examples of super-enhancer loci significantly upregulated exclusively in transitional cell lines (Figure S6E). These results suggest that expression of transitional cell state signatures in neuroblastoma cell lines was indeed associated with intermediate super-enhancer profiles that lie between adrenergic/nor-adrenergic and mesenchymal/neural crest cell differentiation states. Notably however, these results show that, while transitional cells lie between adrenergic and mesenchymal cell states, they tended to have an H3K27ac profile toward the adrenergic end of the transdifferentiation spectrum (Figures 6A and 6B), similar to what was seen in our single-cell trajectory analysis (Figures 3F and 3G).

## Validation of core findings in an external single-cell cohort

To confirm that our findings are generalizable to larger cohorts, we next undertook an analysis from three recent scRNA-seq datasets for neuroblastoma (Slyper et al., 2020; Kildisiute et al., 2021; Dong et al., 2020). After quality control and data harmonization (Korsunsky et al., 2019), in total we analyzed 193,555 cells from 40 distinct tumors (Figures S7A and S7B). Using CONICS copy number estimation, we showed that predicted CNVs

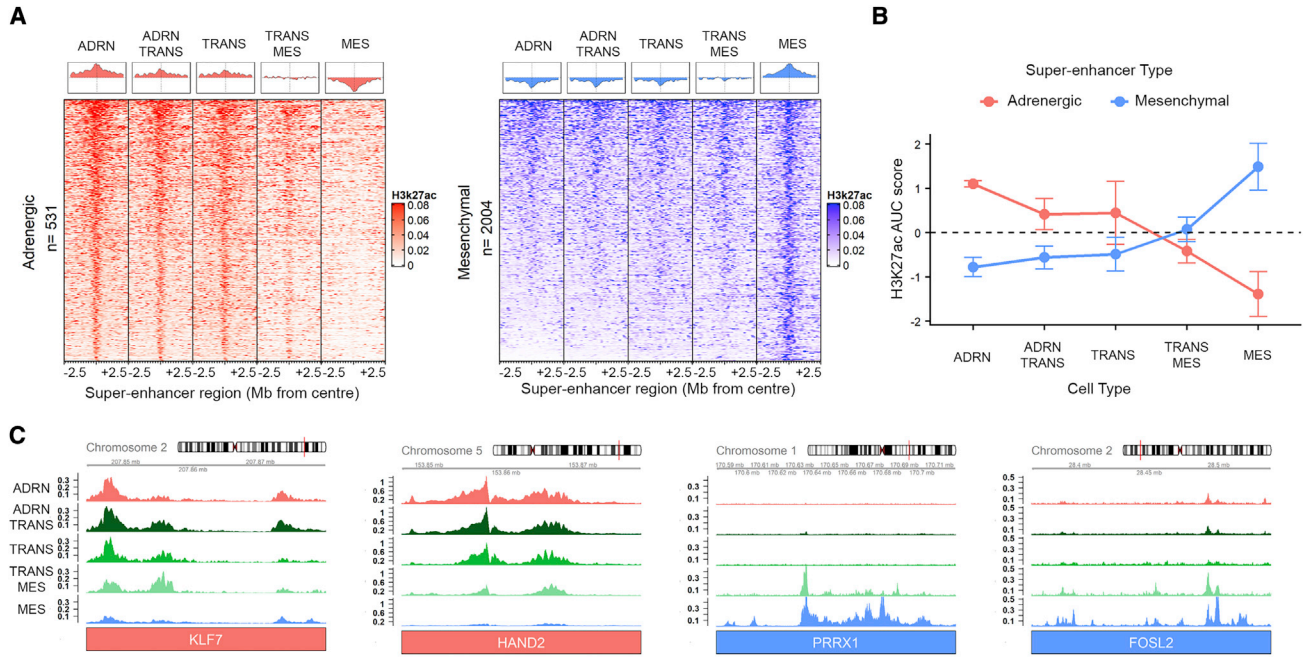



**Figure 6. Neuroblastoma cell lines expressing transitional neuroblast signatures are at the junction between adrenergic and mesenchymal super-enhancer states**

(A) Global normalized H3K27-acetylation (H3K27ac) values averaged for cell lines classified to five subgroups (see Figure S6A for classification). Super-enhancer loci enriched in adrenergic cell lines (left, red), or mesenchymal cell lines (right, blue). H3K27ac values represent 0.5 Mb centered on the super-enhancer locus in that region. For individual cell line traces, see Figure S6B. Above each set of traces is the average H3K27ac *Z* score value for that binned region. Bin size for each trace is 1 kb.

(B) *Z* score value of the total area under the H3K27ac curve for adrenergic super-enhancers (red) and mesenchymal super-enhancers (blue) for five subgroups of cell lines. Error bars represent the standard deviation of individual cell lines in each subgroup.

(C) Representative H3K27ac for each group of classified cell lines for core-regulatory circuit transcription factor (CRC TF) super-enhancer loci. Red are examples of adrenergic CRC TFs and blue are examples of mesenchymal CRC TFs.

were enriched in specific clusters enabling separation of malignant and non-malignant cells (Figures S7C and S7D). Exploring malignant cells, immune cells, and all other cells (using CNV and gene-expression-based classification), this cohort showed a similar trend, with enrichment of immune cells in those tumors with favorable histology and lower risk, and, interestingly, a marked relative depletion of malignant cells in tumors that had already been treated with induction chemotherapy (Figure S7E). Notably, this UMAP projection was consistent across datasets, particularly in cell types classified as immune or other, supporting that there was no dataset bias created during integration (Figure S7F).

We next explored the non-malignant cells for this 40-sample cohort (Figures S7G and S7H). Gene signatures for non-malignant cells identified exclusive expression patterns and was used for cell type classification (Figure S7I). Relative abundance of non-malignant cell types was heterogeneous across the cohort and demonstrated differences in the tumors with favorable compared with unfavorable histology using the Shimada grading (Figure S7J). In particular, the trend observed was a higher abundance of T cells and lower abundance of macrophages being noted in tumors with favorable histology, similar to the findings in our cohort, comparing GNB with neuroblastoma. However, since only two GNB samples were available from the 40 samples of this cohort, this finding could not be confirmed, and further research is required to establish non-malignant cell type abundance comparing GNB with neuroblastoma.

We next undertook a malignant cell analysis in this 40-sample cohort (Figures S7K and S7L). While the focus of these datasets and other recent publications was to relate malignant cell types to the normal cell types of the sympathoadrenal system they derived from (Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Hanemaaijer et al., 2021; Dong et al., 2020), we instead explored the concept we have introduced around potential for cell type transdifferentiation. We first explored expression of the signatures for our three-state model (adrenergic, transitional, and mesenchymal). Expression of all of these gene signatures was noted in a relatively exclusive manner; however, it was not entirely specific to particular clusters, suggesting that there is inter-tumoral differences in control of expression and potentially differences in applying gene signatures across different sequencing technologies (Figure S7M). So, we instead used a label transfer classification algorithm (Stuart et al., 2019) using the 10-sample cohort in this study as a reference. This showed that tumors of this cohort are indeed heterogeneous, and cells of this large dataset can exist as either adrenergic, transitional, or mesenchymal subtype

using our classification model (Figure S7N). Interestingly, similar to our findings, this showed a trend that transitional cells were more abundant in high-risk tumors and *MYCN*-amplified tumors (Figure S7N). Moreover, mesenchymal cells in this 40-sample cohort had a trend that they mostly arose in tumors that had undergone prior chemotherapy, again suggesting potential for cell type adaptation and plasticity in response to environmental exposure (Figure S7N), and consistent with previous reports that mesenchymal expression is associated with prior chemotherapy treatment (van Groningen et al., 2017; Gartlgruber et al., 2021). Collectively, this analysis supports the concept that transdifferentiation between adrenergic and mesenchymal cells creates an intermediate cell profile akin to the transitional cells we see in our dataset, but, notably, extensive heterogeneity suggests that differentiation programs are likely tumor context specific.

## DISCUSSION

Patients with PNTs are characterized by extensive inter- and intratumoral heterogeneity (Eleveld et al., 2015; Schulte et al., 2018; Uemura et al., 2019). Recent studies have reported that PNTs were composed of different types of transdifferentiating malignant cells with dissimilative epigenetic landscapes and gene expression profiles (Boeva et al., 2017; van Groningen et al., 2017). Moreover, the TME plays an important role in tumor biology and adds more complexity to intratumoral heterogeneity (Borriello et al., 2016; Wang et al., 2018). Here we provide comprehensive analysis of the cellular heterogeneity of PNTs through single-cell transcriptomic analysis of 5,301 cells obtained from 10 surgically resected tumors (Figures S8A and S8B).

Among non-malignant cells, we found that most non-neuroblastoma samples have a higher proportion of infiltrating immune cells compared with neuroblastoma samples, especially cytotoxic T cells. While our validation cohort did not contain sufficient GNB samples to confirm this finding, a similar trend was observed when comparing Shimada favorable histology (FH) versus unfavorable histology (UH). This is of interest since cytotoxic T cells are an important target of immunotherapy and considered as an important prognostic indicator in many cancers, including neuroblastoma (Wei et al., 2018; Jiang et al., 2018). The opposite trend was observed with regard to relative macrophage abundance, in which neuroblastoma showed a higher relative proportion compared with GNB/GN in our cohort, or Shimada FH versus UH in the validation cohort, particularly when considering non-inflammatory macrophages expressing M2-polarization marker CD163. This is potentially important due to the role of M2 macrophages in the promotion of tumor cell proliferation and invasion (Komohara and Takeya, 2017; Wei et al., 2018). Together, these findings highlight potential for unique TMEs in PNTs and support efforts to exploit this for therapeutic advantage (Liu and Joshi, 2020; Joshi, 2020).

Recent single-cell studies of neuroblastoma have been key in understanding developmental origins (Bedoya-Reina et al., 2021; Kameneva et al., 2021; Jansky et al., 2021; Kildisiute et al., 2021; Hanemaaijer et al., 2021; Dong et al., 2020). These studies however, have some different interpretations, particularly relating to malignant cell identity being that of chromaffin or sympathoblast. This disparity possibly arose due to different reference datasets of adrenal development, using mouse or human, for instance (Ponzoni et al., 2022). Moreover, a recent article showed that a postnatal adrenal reference dataset could lead to varying interpretations around a postnatal cholinergic progenitor population being a potential cell of origin for high-risk neuroblastoma (Bedoya-Reina et al., 2021), while some articles have identified the presence of mesenchymal features in tumors that could be indicative of earlier progenitors as a cell of origin or evidence of dedifferentiation in cancer cells (Gartlgruber et al., 2021; Kameneva et al., 2021; Jansky et al., 2021). While we strongly support these continued efforts to resolve the developmental origins of neuroblastoma using single-cell technology, this led us to take a somewhat simpler approach and describe malignant cells according to their inter-relationship without using a normal adrenal cell reference dataset. Since cancer-specific alterations could be driving phenotypical plasticity, particularly drawing upon developmental reprogramming between mesenchymal progenitors and downstream derivatives (Kameneva et al., 2021), this simple model may offer an alternative that accounts for potential cancer-specific cell plasticity, dedifferentiation, and interconverting cell types.

Among malignant populations, we detected the expression of genes related to mesenchymal/neural crest cell and adrenergic/noradrenergic neuroblastoma cells, consistent with previous reports (van Groningen et al., 2017; Boeva et al., 2017; Wang et al., 2019; Lecca et al., 2018; Kameneva et al., 2021; Jansky et al., 2021). Notably, however, we found that some tumors display an additional transitional state with gene expression characteristics associated with aggressive disease, rapid proliferation, and metastasis. Consistent with these findings, transitional signatures were also associated with more aggressive clinical behavior in bulk transcriptome cohorts, such as enrichment in INSS 4/*MYCN*-amplified patients and a strong association with poor prognosis. Enrichment with *MYCN* amplification was particularly intriguing since the tumors containing transitional cells in our cohort, while having MYCN overexpression, were *MYCN* non-amplified. This is more evidence that high-risk neuroblastoma is often an MYC-driven disease regardless of whether the tumor is *MYCN* amplified or *MYCN* non-amplified (Zimmerman et al., 2018; Westermann et al., 2008; Carter et al., 2015). Since our cohort lacked sufficient samples to assess a direct relationship between *MYCN* amplification and transitional cells, we assessed relative cell type proportions in a 40-sample validation cohort, and, interestingly, *MYCN*-amplified samples had a higher proportion of transitional cells. However, while an association can be observed between *MYCN* amplification and the transitional subgroup, this does not seem to be a prerequisite for cell identity, since transitional cells are enriched in some *MYCN* non-amplified samples across both cohorts.

Another interesting finding from our malignant cell analysis was that a minority of tumors were strongly mesenchymal, while all other mesenchymal cells and perhaps some transitional cells showed subtle expression patterns more similar to the description proposed by Jansky et al. (2021) as having "mesenchymal features." Since mesenchymal dominance is relatively rare in neuroblastoma tumors, certainly compared with cell lines

(Gautier et al., 2021), our data at least suggest that, in rarer cases, mesenchymal neuroblastoma tumors can exist and may depend on unique genetic alterations that drive phenotype. Indeed, our predicted chromosome 5 gain validated by FISH in mesenchymal tumor NB1, while not common in neuroblastoma, could lead to this apparent variation in cell phenotype and explain this less common observation. Interestingly, though, while most tumors usually demonstrate a dominant population, there is cell state variability through all three cell states using our trajectory modeling approach, suggesting that many neuroblastomas demonstrate a degree of plasticity regardless of dominant cell type.

Prior studies have speculated that adrenergic/noradrenergic cells transdifferentiate to mesenchymal/neural crest cells and vice versa (van Groningen et al., 2017, 2019; Boeva et al., 2017; Wang et al., 2019; Lecca et al., 2018). This theory was primarily based on spontaneous state polarization of isogenic pairs of cells and forced overexpression experiments using strong lineage-associated transcription factors to reprogram cell phenotype such as PHOX2B, PRRX1, ASCL1, and NOTCH3 (van Groningen et al., 2017, 2019; Boeva et al., 2017; Wang et al., 2019). Our trajectory modeling using scRNA-seq data, and our reclassification approach for H3K27ac landscapes in neuroblastoma cell lines, is in support of this theory, albeit by a slightly more complex mechanism involving transitional cells as a transdifferentiation intermediate cell state. This model is akin to the long-held observation that neuroblastoma cell lines demonstrate three morphological variants with capacity for interconversion: neuroblastic (N), substrate-adherent (S), or intermediate (I) phenotypes (Ross et al., 1995; Ciccarone et al., 1989; Veschi et al., 2019). N-type cells are said to resemble embryonic sympathoblasts (similar to the adrenergic phenotype); S-type cells resemble Schwannian, glial, or melanocytic progenitors (similar to the mesenchymal phenotype); and I-type cells have an intermediate phenotype with the potential to differentiate toward either cell type (similar to our proposed transitional phenotype) (Veschi et al., 2019; Ciccarone et al., 1989; Ross et al., 1995). Moreover, examination of scRNA-seq data derived from the embryonic murine sympathoadrenal system shows expression patterns that suggest that differentiation through a transitional state may be developmentally co-opted. This phenotype has distinct gene expression markers of significant note, such as MYCN and EZH2, as well as upregulation of neurodevelopment and neurogenesis-related pathways and activation of some distinct super-enhancer loci. We suggest that the transitional state is likely transient in the respect that the cells have greater cell plasticity and are more likely to adapt to changing environments. We hypothesize that this state could act as a ''pit-stop'' during transdifferentiation depending on subtle environmental exposures, which may also give them a fitness advantage to survive anti-cancer barriers (i.e., anti-cancer therapies, and/or homeostatic anti-cancer mechanisms). Extending gene sets for our three-state model to a larger external cohort revealed a similar trend, but patterns of expression were not entirely specific, suggesting that each tumor likely has unique differentiation trajectories depending on cancer-specific genetic alterations, environmental exposures, and cell of origin context. Future research will benefit from functional genomics studies on key transitional genes and lineage tracing studies that can follow phenotypical adaptation in response to environmental stimuli. This will be required to elucidate the precise interconnections between all cell states and how these states relate to other established models of neuroblastoma cell phenotype, such as morphological and epigenetic classification methods.

Another key finding from recent studies is that tumor cells convert to a mesenchymal/neural crest cell phenotype in relapsed neuroblastoma tumors and upon drug exposure *in vitro*, suggesting that mesenchymal differentiation is a drug-resistance mechanism in neuroblastoma (van Groningen et al., 2017; Boeva et al., 2017; Naiditch et al., 2015). Interestingly, though, our analysis of bulk gene expression from diagnostic tumors paradoxically shows that the mesenchymal-only class is not a predictor of poor outcome, and indeed is generally predictive of favorable outcome. It is possible that this is a consequence of bulk profiling in which mesenchymal expression may stem from normal Schwannian stromal cells in FH tumors (Ponzoni et al., 2022). However, since our data showed that transitional signatures are more predictive of poor outcome, we suggest that a mesenchymal phenotype is not a high-risk phenotype per se, but rather that transitional cells may have the capacity to escape cytotoxic therapy by transdifferentiation to mesenchymal phenotype under drug selection pressure.

In conclusion, our analysis has uncovered extensive intra- and inter-individual, functional, and transcriptomic heterogeneity in neuroblastic tumor cells and associated microenvironmental cells. We identify interconnected mechanisms of phenotypical heterogeneity, highlighting the importance of considering cell plasticity and transdifferentiation potential in new therapeutic strategies for PNTs.

## Limitations of the study

One limitation of our study is that, in some instances, primary tumors were obtained from patients who had received prior chemotherapy. This may be important with respect to comparing neuroblastoma and GNB, since induction chemotherapy is more common in the more aggressive forms of PNT. Therefore, it is unknown to what extent these treatments may have influenced intratumoral heterogeneity. This is intriguing, since our validation cohort showed an enrichment of mesenchymal cells in samples treated with prior chemotherapy, consistent with models of mesenchymal selection or adaptation as a survival mechanism (van Groningen et al., 2017; Boeva et al., 2017; Naiditch et al., 2015). Future studies that pair pre-treatment biopsy specimens with post-chemotherapy surgical resections and relapsed tumors will be valuable to unveil tumor phenotype adaptation in the context of drug-induced tumor evolution. Another limitation within our study is the small number of samples profiled, although we aimed to address this by extending to a larger validation cohort. The addition of more samples with a diverse molecular background will be invaluable to the single-cell field, such as ALKmut, ATRXmut, PHOX2Bmut, TERT rearrangements, RASmut, p53mut, and other segmental chromosomal alterations recurrent to neuroblastoma (Ackermann et al., 2018; Pugh et al., 2013). This will facilitate the generation of molecular-associated single-cell gene expression profiles that will be important to distinguish genetic and non-genetic facets of cell phenotype.

Moreover, validation of our findings related to differences in immune cell proportions between different histotypes will be necessary to draw conclusions. Finally, while we have aimed to assess tumor cellular compositions in a quantitative manner, it will be important to validate these findings with potential that there was a bias in cell collection that arises when enriching viable cells. In particular, the changes in proportion of certain cell types when comparing neuroblastoma versus GNB were only trends and not statistically validated, suggesting larger cohorts and better methods other than scRNA-seq are required to confirm this quantitative difference. Moreover, a notable cell type that was not identified was the Schwann cells when there was histological evidence of Schwannian stroma.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Immunohistochemical staining
  - Fluorescence *in situ* hybridization
  - Single cell preparation and flow cytometry
  - Single cell library preparation and RNA sequencing
  - Single cell RNA-seq data analysis
  - Bulk tumor analyses
  - H3k27ac ChIP-seq analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111455.

## ACKNOWLEDGMENTS

The work was supported by grants from the National Key Research and Development Program of China (grant no. 2017YFA0103902); the National Natural Science Foundation of China (grant nos. 81570760, 32271165 and 31771283); Project of Shanghai Committee of Science and Technology, China (grant no. 16411962500); Translational Medicine Innovation Fund of Shanghai Jiao Tong University School of Medicine, China (grant no. 15ZH2005); and the Fundamental Research Funds for the Central Universities of Tongji University. W.S. was supported grants from National Key Research and Development Program of China (2018YFD0900600), the National Natural Science Foundation of China (41676119), and the Fundamental Research Funds for the Central Universities from the Ocean University of China. D.C. received Cancer Institute NSW Early Career Fellowship #DC000594.

## AUTHOR CONTRIBUTIONS

X.Y. and C.Z. designed this study with input from B.B.C., G.M.M., and D.R.C. X.Y., S.D., Q.J., S.H., Yanli Xu, M.T. and Y.D. performed the surgery and collected the specimen. Ying Xu, Y.C., X.J., Y.Z., and L.L. performed the scRNA-seq experiment. J.A.S., A.B., and D.R.C. analyzed the data. D.R.C. wrote the draft. X.Y., J.A.S., S.D., Ying Xu, D.R.C., and C.Z. contributed to the manuscript preparation and revision. W.S. contributed the reagent and technical support. All authors have read and approved the final version of the manuscript and agree with the order of presentation of the authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 17, 2020
Revised: April 16, 2022
Accepted: September 14, 2022
Published: October 4, 2022

## REFERENCES

Ackermann, S., Cartolano, M., Hero, B., Welte, A., Kahlert, Y., Roderwieser, A., Bartenhagen, C., Walter, E., Gecht, J., Kerschke, L., et al. (2018). A mechanistic classification of clinical phenotypes in neuroblastoma. Science *362*, 1165–1170.

Asgharzadeh, S., Salo, J.A., Ji, L., Oberthuer, A., Fischer, M., Berthold, F., Hadjidaniel, M., Liu, C.W.Y., Metelitsa, L.S., Pique-Regi, R., et al. (2012). Clinical significance of tumor-associated inflammatory cells in metastatic neuroblastoma. J. Clin. Oncol. *30*, 3525–3532.

Bedoya-Reina, O.C., Li, W., Arceo, M., Plescher, M., Bullova, P., Pui, H., Kaucka, M., Kharchenko, P., Martinsson, T., Holmberg, J., et al. (2021). Single-nuclei transcriptomes from human adrenal gland reveal distinct cellular identities of low and high-risk neuroblastoma tumors. Nat. Commun. *12*, 5309.

Boeva, V., Louis-Brennetot, C., Peltier, A., Durand, S., Pierre-Eugène, C., Raynal, V., Etchevers, H.C., Thomas, S., Lermine, A., Daudigeos-Dubus, E., et al. (2017). Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. Nat. Genet. *49*, 1408–1413.

Borriello, L., Seeger, R.C., Asgharzadeh, S., and Declerck, Y.A. (2016). More than the genes, the tumor microenvironment in neuroblastoma. Cancer Lett. *380*, 304–314.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Carter, D.R., Murray, J., Cheung, B.B., Gamble, L., Koach, J., Tsang, J., Sutton, S., Kalla, H., Syed, S., Gifford, A.J., et al. (2015). Therapeutic targeting of the MYC signal by inhibition of histone chaperone FACT in neuroblastoma. Sci. Transl. Med. *7*, 312ra176.

Chen, L., Alexe, G., Dharia, N.V., Ross, L., Iniguez, A.B., Conway, A.S., Wang, E.J., Veschi, V., Lam, N., Qi, J., et al. (2018). CRISPR-Cas9 screen reveals a MYCN-amplified neuroblastoma dependency on EZH2. J. Clin. Invest. *128*, 446–462.

Ciccarone, V., Spengler, B.A., Meyers, M.B., Biedler, J.L., and Ross, R.A. (1989). Phenotypic diversification in human neuroblastoma cells: expression of distinct neural crest lineages. Cancer Res. *49*, 219–225.

De Bernardi, B., Gambini, C., Haupt, R., Granata, C., Rizzo, A., Conte, M., Tonini, G.P., Bianchi, M., Giuliano, M., Luksch, R., et al. (2008). Retrospective study of childhood ganglioneuroma. J. Clin. Oncol. *26*, 1710–1716.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dong, R., Yang, R., Zhan, Y., Lai, H.D., Ye, C.J., Yao, X.Y., Luo, W.Q., Cheng, X.M., Miao, J.J., Wang, J.F., et al. (2020). Single-cell characterization of malignant phenotypes and developmental trajectories of adrenal neuroblastoma. Cancer Cell *38*, 716–733.e6.

Eleveld, T.F., Oldridge, D.A., Bernard, V., Koster, J., Colmet Daage, L., Diskin, S.J., Schild, L., Bentahar, N.B., Bellini, A., Chicard, M., et al. (2015). Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. Nat. Genet. *47*, 864–871.

Furlan, A., Dyachuk, V., Kastriti, M.E., Calvo-Enrique, L., Abdo, H., Hadjab, S., Chontorotzea, T., Akkuratova, N., Usoskin, D., Kamenev, D., et al. (2017). Multipotent peripheral glial cells generate neuroendocrine cells of the adrenal medulla. Science *357*, eaal3753.

Gartlgruber, M., Sharma, A.K., Quintero, A., Dreidax, D., Jansky, S., Park, Y.G., Kreth, S., Meder, J., Doncevic, D., Saary, P., et al. (2021). Super enhancers define regulatory subtypes and cell identity in neuroblastoma. Nat. Cancer *2*, 114–128.

Gautier, M., Thirant, C., Delattre, O., and Janoueix-Lerosey, I. (2021). Plasticity in neuroblastoma cell identity defines a noradrenergic-to-mesenchymal transition (NMT). Cancers *13*, 2904.

Gu, Z., Eils, R., Schlesner, M., and Ishaque, N. (2018). EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. BMC Genom. *19*, 234.

Hahne, F., and Ivanek, R. (2016). Statistical Genomics: Methods and Protocols - Visualizing Genomic Data Using Gviz and Bioconductor (Springer).

Hanemaaijer, E.S., Margaritis, T., Sanders, K., Bos, F.L., Candelli, T., Al-Saati, H., Van Noesel, M.M., Meyer-Wentrup, F.A.G., Van De Wetering, M., Holstege, F.C.P., and Clevers, H. (2021). Single-cell atlas of developing murine adrenal gland reveals relation of Schwann cell precursor signature to neuroblastoma phenotype. Proc. Natl. Acad. Sci. USA *118*. e2022350118.

Jansky, S., Sharma, A.K., Körber, V., Quintero, A., Toprak, U.H., Wecht, E.M., Gartlgruber, M., Greco, A., Chomsky, E., Grünewald, T.G.P., et al. (2021). Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. Nat. Genet. *53*, 683–693.

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., Li, Z., Traugh, N., Bu, X., Li, B., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat. Med. *24*, 1550–1558.

Joshi, S. (2020). Targeting the Tumor Microenvironment in Neuroblastoma: Recent Advances and Future Directions. Cancers *12*, 2057.

Kameneva, P., Artemov, A.V., Kastriti, M.E., Faure, L., Olsen, T.K., Otte, J., Erickson, A., Semsch, B., Andersson, E.R., Ratz, M., et al. (2021). Single-cell transcriptomics of human embryos identifies multiple sympathoblast lineages with potential implications for neuroblastoma origin. Nat. Genet. *53*, 694–706.

Kildisiute, G., Kholosy, W.M., Young, M.D., Roberts, K., Elmentaite, R., Van Hooff, S.R., Pacyna, C.N., Khabirova, E., Piapi, A., Thevanesan, C., et al. (2021). Tumor to normal single-cell mRNA comparisons reveal a pan-neuroblastoma cancer cell. Sci. Adv. *7*, eabd3311.

Kocak, H., Ackermann, S., Hero, B., Kahlert, Y., Oberthuer, A., Juraeva, D., Roels, F., Theissen, J., Westermann, F., Deubzer, H., et al. (2013). Hox-C9 activates the intrinsic pathway of apoptosis and is associated with spontaneous regression in neuroblastoma. Cell Death Dis. *4*, e586.

Komohara, Y., and Takeya, M. (2017). CAFs and TAMs: maestros of the tumour microenvironment. J. Pathol. *241*, 313–315.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296.

Lecca, M.C., Jonker, M.A., Kulsoom Abdul, U., Küçükosmanoglu, A., Van Wieringen, W., and Westerman, B.A. (2018). Adrenergic to mesenchymal fate switching of neuroblastoma occurs spontaneously in vivo resulting in differential tumorigenic potential. J. Mol. Clin. Med. *1*, 219–226.

Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. Cell Stem Cell *20*, 891–892.e4.

Li, Z., Takenobu, H., Setyawati, A.N., Akita, N., Haruta, M., Satoh, S., Shinno, Y., Chikaraishi, K., Mukae, K., Akter, J., et al. (2018). EZH2 regulates neuroblastoma cell differentiation via NTRK1 promoter epigenetic modifications. Oncogene *37*, 2714–2727.

Liu, K.X., and Joshi, S. (2020). "Re-Educating" tumor associated macrophages as a novel immunotherapy strategy for neuroblastoma. Front. Immunol. *11*, 1947.

Matthay, K.K., Maris, J.M., Schleiermacher, G., Nakagawara, A., Mackall, C.L., Diller, L., and Weiss, W.A. (2016). Neuroblastoma. Nat. Rev. Dis. Primers *2*, 16078.

Mina, M., Boldrini, R., Citti, A., Romania, P., D'alicandro, V., De Ioris, M., Castellano, A., Furlanello, C., Locatelli, F., and Fruci, D. (2015). Tumor-infiltrating T lymphocytes improve clinical outcome of therapy-resistant neuroblastoma. OncoImmunology *4*, e1019981.

Molenaar, J.J., Koster, J., Zwijnenburg, D.A., Van Sluis, P., Valentijn, L.J., Van Der Ploeg, I., Hamdi, M., Van Nes, J., Westerman, B.A., Van Arkel, J., et al. (2012). Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. Nature *483*, 589–593.

Müller, S., Cho, A., Liu, S.J., Lim, D.A., and Diaz, A. (2018). CONICS integrates scRNA-seq with DNA sequencing to map gene expression to tumor subclones. Bioinformatics *34*, 3217–3219.

Naiditch, J.A., Jie, C., Lautz, T.B., Yu, S., Clark, S., Voronov, D., Chu, F., and Madonna, M.B. (2015). Mesenchymal change and drug resistance in neuroblastoma. J. Surg. Res. *193*, 279–288.

Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J., Thomas, S., et al. (2012). BEDOPS: high-performance genomic feature operations. Bioinformatics *28*, 1919–1920.

Padovan-Merhar, O.M., Raman, P., Ostrovnaya, I., Kalletla, K., Rubnitz, K.R., Sanford, E.M., Ali, S.M., Miller, V.A., Mossé, Y.P., Granger, M.P., et al. (2016). Enrichment of targetable mutations in the relapsed neuroblastoma genome. PLoS Genet. *12*, e1006501.

Pelizzo, G., Veschi, V., Mantelli, M., Croce, S., Di Benedetto, V., D'angelo, P., Maltese, A., Catenacci, L., Apuzzo, T., Scavo, E., et al. (2018). Microenvironment in neuroblastoma: isolation and characterization of tumor-derived mesenchymal stromal cells. BMC Cancer *18*, 1176.

Pezzolo, A., Parodi, F., Corrias, M.V., Cinti, R., Gambini, C., and Pistoia, V. (2007). Tumor origin of endothelial cells in human neuroblastoma. J. Clin. Oncol. *25*, 376–383.

Pezzolo, A., Parodi, F., Marimpietri, D., Raffaghello, L., Cocco, C., Pistorio, A., Mosconi, M., Gambini, C., Cilli, M., Deaglio, S., et al. (2011). Oct-4+/Tenascin C+ neuroblastoma cells serve as progenitors of tumor-derived endothelial cells. Cell Res. *21*, 1470–1486.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

Pinto, N.R., Applebaum, M.A., Volchenboum, S.L., Matthay, K.K., London, W.B., Ambros, P.F., Nakagawara, A., Berthold, F., Schleiermacher, G., Park, J.R., et al. (2015). Advances in risk classification and treatment strategies for neuroblastoma. J. Clin. Oncol. *33*, 3008–3017.

Ponzoni, M., Bachetti, T., Corrias, M.V., Brignole, C., Pastorino, F., Calarco, E., Bensa, V., Giusto, E., Ceccherini, I., and Perri, P. (2022). Recent advances in the developmental origin of neuroblastoma: an overview. J. Exp. Clin. Cancer Res. *41*, 92.

Pugh, T.J., Morozova, O., Attiyeh, E.F., Asgharzadeh, S., Wei, J.S., Auclair, D., Carter, S.L., Cibulskis, K., Hanna, M., Kiezun, A., et al. (2013). The genetic landscape of high-risk neuroblastoma. Nat. Genet. *45*, 279–284.

Puram, S.V., Tirosh, I., Parikh, A.S., Patel, A.P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C.L., Mroz, E.A., Emerick, K.S., et al. (2017). Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell *171*, 1611–1624.e24.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rifatbegovic, F., Frech, C., Abbasi, M.R., Taschner-Mandl, S., Weiss, T., Schmidt, W.M., Schmidt, I., Ladenstein, R., Ambros, I.M., and Ambros, P.F. (2018). Neuroblastoma cells undergo transcriptomic alterations upon dissemination into the bone marrow and subsequent tumor progression. Int. J. Cancer *142*, 297–307.

Ross, R.A., Spengler, B.A., Domènech, C., Porubcin, M., Rettig, W.J., and Biedler, J.L. (1995). Human neuroblastoma I-type cells are malignant neural crest stem cells. Cell Growth Differ. *6*, 449–456.

Schramm, A., Köster, J., Assenov, Y., Althoff, K., Peifer, M., Mahlow, E., Odersky, A., Beisser, D., Ernst, C., Henssen, A.G., et al. (2015). Mutational dynamics between primary and relapse neuroblastomas. Nat. Genet. *47*, 872–877.

Schulte, M., Köster, J., Rahmann, S., and Schramm, A. (2018). Cancer evolution, mutations, and clonal selection in relapse neuroblastoma. Cell Tissue Res. *372*, 263–268.

Schwab, M., Alitalo, K., Klempnauer, K.H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., and Trent, J. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature *305*, 245–248.

Slyper, M., Porter, C.B.M., Ashenberg, O., Waldman, J., Drokhlyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nat. Med. *26*, 792–802.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Tsubota, S., Kishida, S., Shimamura, T., Ohira, M., Yamashita, S., Cao, D., Kiyonari, S., Ushijima, T., and Kadomatsu, K. (2017). PRC2-Mediated transcriptomic alterations at the embryonic stage govern tumorigenesis and clinical outcome in MYCN-driven neuroblastoma. Cancer Res. *77*, 5259–5271.

Uemura, S., Ishida, T., Thwin, K.K.M., Yamamoto, N., Tamura, A., Kishimoto, K., Hasegawa, D., Kosaka, Y., Nino, N., Lin, K.S., et al. (2019). Dynamics of minimal residual disease in neuroblastoma patients. Front. Oncol. *9*, 455.

van Groningen, T., Akogul, N., Westerhout, E.M., Chan, A., Hasselt, N.E., Zwijnenburg, D.A., Broekmans, M., Stroeken, P., Haneveld, F., Hooijer, G.K.J., et al. (2019). A NOTCH feed-forward loop drives reprogramming from adrenergic to mesenchymal state in neuroblastoma. Nat. Commun. *10*, 1530.

van Groningen, T., Koster, J., Valentijn, L.J., Zwijnenburg, D.A., Akogul, N., Hasselt, N.E., Broekmans, M., Haneveld, F., Nowakowska, N.E., Bras, J., et al. (2017). Neuroblastoma is composed of two super-enhancer-associated differentiation states. Nat. Genet. *49*, 1261–1266.

Veschi, V., Verona, F., and Thiele, C.J. (2019). Cancer stem cells and neuroblastoma: characteristics and therapeutic targeting options. Front. Endocrinol. *10*, 782.

Wang, L., Livak, K.J., and Wu, C.J. (2018). High-dimension single-cell analysis applied to cancer. Mol. Aspects Med. *59*, 70–84.

Wang, L., Tan, T.K., Durbin, A.D., Zimmerman, M.W., Abraham, B.J., Tan, S.H., Ngoc, P.C.T., Weichert-Leahey, N., Akahane, K., Lawton, L.N., et al. (2019). ASCL1 is a MYCN- and LMO1-dependent member of the adrenergic neuroblastoma core regulatory circuitry. Nat. Commun. *10*, 5622.

Wei, J.S., Kuznetsov, I.B., Zhang, S., Song, Y.K., Asgharzadeh, S., Sindiri, S., Wen, X., Patidar, R., Najaraj, S., Walton, A., et al. (2018). Clinically relevant cytotoxic immune cell signatures and clonal expansion of T-cell receptors in high-risk MYCN-not-amplified human neuroblastoma. Clin. Cancer Res. *24*, 5673–5684.

Westerhout, E.M., Hamdi, M., Stroeken, P., Nowakowska, N.E., Lakeman, A., Van Arkel, J., Hasselt, N.E., Bleijlevens, B., Akogul, N., Haneveld, F., et al. (2022). Mesenchymal-type neuroblastoma cells escape ALK inhibitors. Cancer Res. *82*, 484–496.

Westermann, F., Muth, D., Benner, A., Bauer, T., Henrich, K.O., Oberthuer, A., Brors, B., Beissbarth, T., Vandesompele, J., Pattyn, F., et al. (2008). Distinct transcriptional MYCN/c-MYC activities are associated with spontaneous regression or malignant progression in neuroblastomas. Genome Biol. *9*, R150.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).

Zappia, L., and Oshlack, A. (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience *7*, giy083.

Zeine, R., Salwen, H.R., Peddinti, R., Tian, Y., Guerrero, L., Yang, Q., Chlenski, A., and Cohn, S.L. (2009). Presence of cancer-associated fibroblasts inversely correlates with Schwannian stroma in neuroblastoma tumors. Mod. Pathol. *22*, 950–958.

Zhao, C., Hu, S., Huo, X., and Zhang, Y. (2017). Dr.seq2: a quality control and analysis pipeline for parallel single cell transcriptome and epigenome data. PLoS One *12*, e0180583.

Zimmerman, M.W., Liu, Y., He, S., Durbin, A.D., Abraham, B.J., Easton, J., Shao, Y., Xu, B., Zhu, S., Zhang, X., et al. (2018). MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. Cancer Discov. *8*, 320–335.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| Anti -CD3 Rabbit mAb | Servicebio | Cat# GB13014; RRID: AB_2920578 |
| HRP conjugated Goat Anti-Rabbit IgG (H + L) | Servicebio | Cat# GB23303; RRID: AB_2811189 |
| **Biological samples** | | |
| Fresh peripheral neuroblastic tumor tissues | Xinhua Hospital affiliated to Shanghai Jiao Tong University School of Medicine | N/A |
| **Chemicals, peptides, and recombinant proteins** | | |
| Sodium citrate antigen retrieval solution (pH 6.0) | Servicebio | Cat# G1202 |
| Dako REAL$^{TM}$ EnVision$^{TM}$ Detection System, Peroxidase/DAB+, Rabbit/Mouse | Dako | Cat# K5007 |
| Water Soluble Hematoxylin Stain Solution | Servicebio | Cat# G1004 |
| Proteinase K stock solution | LBP Medicine | Cat# F01239 |
| EGR1/TERT Fish Probe | LBP Medicine | Cat# F.01171-01 |
| TP53/17q22 Dual Color Probe | LBP Medicine | Cat# F.01030-01 |
| CCND1/CSP 11 Dual Color Probe | LBP Medicine | Cat# F.01023-01 |
| 20q11 Fish Probe | LBP Medicine | Cat# F.01344 |
| CEP20 Fish Probe | VividFISH | Cat# FP097 |
| MYCN/2q11 Dual Color Probe | LBP Medicine | Cat# F.01013-01 |
| CBFB/MYH11 t(16; 16); inv (16) Fusion Probe | LBP Medicine | Cat# F.01094-01 |
| DAPI | LBP Medicine | Cat# F01238 |
| Liberase$^{TM}$ TM Research Grade | Roche | Cat# 5401127001 |
| CellTrace$^{TM}$ Calcein Blue, AM - Special Packaging | Invitrogen | Cat# C34853 |
| DRAQ5 | Cell Signaling Technology | Cat# 4084L |
| Recombinant RNase Inhibitor | Takara Bio | Cat# 2313B |
| Triton X-100 | Sigma- Aldrich | Cat# X100-100ML |
| dNTP mix | Takara Bio | Cat# 4019 |
| KAPA HiFi HotStart PCR Kit | Roche | Cat# KK2502 |
| AMPure XP bead | Beckman | Cat# A63882 |
| Qubit dsDNA HS Assay Kit | Invitrogen | Cat# Q32854 |
| **Critical commercial assays** | | |
| Nextera XT DNA Sample Preparation Kit | Illumina | Cat # FC-131-1024 |
| HiSeq X Ten Reagent Kit v2.5 | Illumina | Cat# FC-501-2501 |
| **Deposited data** | | |
| Data files for single-cell RNA sequencing | This paper | GSE192906 |
| Human neuroblastoma scRNA-seq data (Dong et al., 2020) | GEO | GSE137804 |
| Human neuroblastoma scRNA-seq data (Kildisiute et al., 2021) | https://www.neuroblastomacellatlas.org/ | N/A |
| Human neuroblastoma scRNA-seq data (Slyper et al., 2020) | GEO | GSE140819 |
| H3k27ac ChIP-seq and RNA-seq data for 33 neuroblastoma samples (Boeva et al., 2017) | GEO | GSE90683 |
| Human neuroblastoma microarray data (Kocak et al., 2013) | GEO | GSE45480 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Human neuroblastoma microarray data (Molenaar et al., 2012) | GEO | GSE16476 |
| **Software and algorithms** | | |
| Dr.Seq2 | (Zhao et al., 2017) | V1 |
| STAR | (Dobin et al., 2013) | V3 |
| Seurat | (Butler et al., 2018) | V3.2.3 |
| Clustree | (Zappia and Oshlack, 2018) | V0.4.3 |
| Monocle2 | (Trapnell et al., 2014) | V2.16 |
| CONICSmat | (Muller et al., 2018) | V0.0.0.1 |
| Bedtools | (Quinlan and Hall, 2010) | V2.30.0 |
| Bedops | (Neph et al., 2012) | V2.4.40 |
| EnrichedHeatmap | (Gu et al., 2018) | V1.18.2 |
| Gviz | (Hahne and Ivanek, 2016) | V1.32.2 |
| FlowJo | BD | V10 |
| HiSeq Control Software | Illumina | V1.6 |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Daniel Carter (Daniel.Carter@uts.edu.au).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Single cell RNA-seq data have been deposited at the Gene expression omnibus: GSE192906 and are publicly available as of the date of publication. Accession numbers are listed in the Key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

Fresh peripheral neuroblastic tumor (PNT) tissue samples were obtained from 10 patients enrolled at the Xinhua Hospital affiliated to Shanghai Jiao Tong University School of Medicine (Shanghai, China). This study was approved by the ethics committee of Xinhua Hospital and was conducted in accordance with the Declaration of Helsinki. Patients were enrolled for this study based on pathological diagnosis. All patients were enrolled under informed consent. Staging and risk assessment were performed according to the International Neuroblastoma Staging System Committee (INSS) system and the International Neuroblastoma Risk Group (INRG) respectively, based on tumor histology and *MYCN*-fluorescent *in situ* hybridization (FISH) analysis. Tumors were resected from the primary site either prior to the patient undergoing chemotherapy or post-induction chemotherapy. Treatment regimen included "treatment protocol CCCG-NB-2014". Detailed clinical information on individual patients in this study are outlined in Table S1.

## METHOD DETAILS

### Immunohistochemical staining

All PNTs specimens were fixed, paraffin-embedded, sectioned, and stained with hematoxylin and eosin (H&E) following routine method of Xinhua Hospital's Pathology. Immunohistochemical (IHC) studies employed 5 μm paraffin-embedded slides. Sections were incubated in 3 changes of xylene for each 15 min, 2 changes of pure ethanol for each 5 min, 85% ethanol for 5 min and 75% ethanol for 5 min. Then sections were washed in distilled water. Antigen was retrieved by sodium citrate antigen retrieval solution (pH 6.0) (Servicebio, Cat# G1202) in the microwave oven. After natural cooling, the sections were washed thrice with PBS (pH 7.4) for 5 min each. Endogenous peroxidase was inactivated by incubation in 3% $H_2O_2$ for 25 min. After using 3% BSA to block

nonspecific sites for 30 min, slides were incubated with Anti -CD3 Rabbit mAb (1:100, Servicebio, Cat# GB13014, RRID: AB_2920578) and HRP conjugated Goat Anti-Rabbit IgG (H + L) (1:200, Servicebio, Cat# GB23303, RRID: AB_2811189) to assess the presence of infiltrating T lymphocytes. Slides were washed three times with PBS (pH 7.4) for 5 min each. Dako REAL™ EnVision™ Detection System, Peroxidase/DAB+, Rabbit/Mouse (Dako, Cat# K5007) was used for immunohistochemistry visualization. Nuclei are stained by Water Soluble Hematoxylin Stain Solution (Servicebio, Cat# G1004). After dehydration and sealing, images were captured by XSP-C204 Student Binocular Microscope (COIC, China).

### Fluorescence *in situ* hybridization

Slides were baked in the oven at 65°C for more than 2 h. The slides were dewaxed twice in preheated xylene for 10 min each. Slides were immersed sequentially in 100% ethanol for 5 min, 100% ethanol for 3 min, 85% ethanol for 3 min, 70% ethanol for 3 min and deionized water (pH = 7.0-7.2) for 5 min. Next, slides were placed in deionized water at 97°C for 25 min 0.2 mL of Proteinase K stock solution (LBP Medicine, Cat# F01239) is dissolved in 40 mL of 2×SSC (pH 7.0-7.2) to obtain proteinase K working solution (100 μg/mL). The slides are digested by Proteinase K working solution at 37°C for 10 min. After digestion, the slides were rinsed in 2× SSC solution for 10 min. Tissues were incubated in 70% ethanol and 100% ethanol three times in sequence and then dried naturally. The tissue was covered with 10 μL probe, placed a coverslip, and sealed the edge. The probes include *EGR1/TERT* Fish Probe (LBP Medicine, Cat# F.01171-01), *TP53/17q22* Dual Color Probe (LBP Medicine, Cat# F.01030-01), CCND1/CSP 11 Dual Color Probe (LBP Medicine, Cat# F.01023-01), *20q11* Fish Probe (LBP Medicine, Cat# F.01344), *CEP20* Fish Probe (VividFISH, Cat# FP097), *MYCN/2q11* Dual Color Probe (LBP Medicine, Cat# F.01013-01) and CBFB/*MYH11 t(16;16);inv (16)* Fusion Probe (LBP Medicine, Cat# F.01094-01). The slides are placed on the hybridizer, denatured at 82°C for 8 min, and held at 42°C overnight. 2× SSC was preheated at 46°C for 10 min, rinse slides in 2× SSC for 3 min, remove the coverslip gently and rinse again in 2× SSC for 3 min. The slides are placed in new 2× SSC for 3 min, 70% ethanol for 3 min and 100% ethanol for 3 min in turn. After natural drying, 10 μL DAPI (LBP Medicine, Cat# F01238) was added to the tissue and place a coverslip. After 5 min, the slides were observed by the BX53 microscope (Olympus, Japan). All FISH images were validated by a pathologist to determine whether a CNV was present.

### Single cell preparation and flow cytometry

Following surgical resection, fresh tissue samples were immediately transferred to DMEM/F12 medium (MULTICELL, Cat# 319-075-CL), supplemented with 2% Fetal Bovine Serum (Gemini Bio-Products, Cat# 900-108) and were delivered on ice within 60 min to the laboratory to be processed. Tissue processing was completed within 90 min of collection. Samples were first rinsed in ice-cold PBS (BBI life science, Cat# E607008-0500) with 2% penicillin-streptomycin (BBI life science, Cat# E607011-0100) and then minced into 1mm$^3$ pieces using curved scissors under sterile conditions. The fragments were further enzymatically dissociated into single cells using Liberase TM Research Grade (Roche, Cat# 5,401,127,001) with a final concentration of 0.26 U/mL and a final volume of 25mL Liberase solution/0.5g tissue on a shaker with a speed of 85 rpm for 45 min at 37°C. The Liberase was diluted in Hyclone L-15 Leibovitz media. DMEM/F12 media supplemented with 10% FBS was used to stop dissociation. The resulting single-cell suspension was filtered through a 70μm nylon cell strainer (BD Falcon, Cat# 352,350) and then centrifuged for 5 min at 300g at room temperature. The cell pellet was resuspended in 500μL of PBS supplemented with 0.1% BSA (BBI life science, Cat# A600332-0100) and passed through a 40μm nylon cell strainer (BIOFIL, Cat# CSS013040). This cell suspension was stained with Calcein-Blue (Invitrogen, Cat# C34853) and DRAQ5 (Cell Signaling Technology, Cat# 4084L) prior to fluorescence-activated cell sorting (FACS) in order to isolate live and nucleated cells (Calcein-Blue + DRAQ+). Single cells were sorted using a BD Becton Dickinson FACSAriaII into 96-well PCR plates. Each well of the 96-well plate contained 3μL lysis buffer (10U Recombinant RNase Inhibitor (Takara Bio, Cat# 2313B), 0.2% Triton X-100 (Sigma- Aldrich, Cat# X100-100ML), 3mM dNTP mix (Takara Bio, Cat# 4019)), and reverse transcription reagents (see below). Immediately following FACS, plates were briefly centrifuged and stored at −80°C.

### Single cell library preparation and RNA sequencing

scRNA-seq was conducted using the Smart-seq2 protocol (Picelli et al., 2014) with some modifications being made to incorporate unique molecular identifiers (UMIs) into the 3′ end of transcripts (Li et al., 2017). Reverse transcription was performed directly in the 96-well plates by incubation at 72°C for 5 min, after which the plate was replaced on ice to allow the oligo-dT primer to hybridize to the poly(A) tail of the mRNA molecules. The oligo(dT) primer used in reverse transcription included an additional 8bp cell barcode, 9bp unique molecular identifiers (UMIs) and template-switching oligo sequence. PCR amplification was performed by adding 15μL PCR mix containing 0.5U KAPA HiFi HotStart (Roche, Cat# KK2502), 1x KAPA Buffer (Roche, Cat# KK2502), 12.5mM MgCl2 (Roche, Cat# KK2502), 5μM ISPCR Primer and 7.5mM dNTP mix (Takara Bio, Cat# 4019). PCR amplification was performed in a thermal cycler (BIORAD C1000 Touch Thermal Cycler) at 98°C for 3 min, 24 cycles of 98°C for 20s, 67°C for 15s, and 72°C for 6 min, and a final incubation at 72°C for 5 min. PCR products were purified using 1X AMPure XP bead (Beckman, Cat# A63882) and Qubit dsDNA HS Assay Kit (Invitrogen, Cat# Q32854). TN5 tagmentation and library amplification realized 3′end fragments by Nextera XT DNA Sample Preparation Kit (Illumina, Cat# FC-131-1024) according to the manufacturer's instructions while P5_TSO and Nextera_N7xx took the place of the custom primers. Pooled single cell libraries were sequenced at an average depth of 0.5 million reads per cell on an Illumina HiSeq X Ten instrument (Illumina, San Diego, CA, USA) using a 2 × 150bp paired end HiSeq X Ten Reagent Kit v2.5 (Illumina, Cat# FC-501-2501). Image analysis and base calling were conducted by the HiSeq Control Software (HCS) + OLB + GAPipeline-1.6 (Illumina) on the HiSeq instrument.

### Single cell RNA-seq data analysis

Raw sequencing data was processed using the Dr.Seq2 pipeline (Zhao et al., 2017). Reads were aligned to the human genome (hg38) using STAR (Dobin et al., 2013). In each read, cell barcodes were between 9:16bp and UMIs were between 17:26bp. Following the Dr.Seq2 pipeline we converted the resultant aligned sam file to a bed file using a custom script. We then generated a gene annotation file and annotated the aligned bed file. We then reproduced the aligned sam file to contain gene annotations, cell barcode and UMI information for each read. UMI counts were calculated by removing duplicate reads which had an identical genomic location, cell barcode and UMI sequence, resulting in the final UMI count matrices used for downstream analyses. Cells with fewer than 500 genes or more than 9500 genes were removed (genes were only considered if they were expressed (UMI >1) in at least 3 cells). Furthermore, cells with more than 30% ribosomal gene content and more than 15% mitochondrial gene content were filtered. Downstream normalization and scaling was performed using the R package Seurat with the SCTransform function using 4000 variable genes (Seurat version 3) (Butler et al., 2018) Datasets with all cells (n = 5301) were normalized using SCTransform without regressing any variables. Datasets with malignant cells (n = 2307) and non-malignant cells (2994) were normalized using SCTransform regressing G2M-scores and S-scores created using the CellCycleScoring function in Seurat using published cell cycle gene sets (Tirosh et al., 2016). The SCTransform data output was used to perform a principal component analysis (PCA), and the variance in each principal component (standard deviation) was visualized on an elbow plot which further used to determine the PC-cutoff for downstream clustering and dimensionality reduction analysis. Additionally, clustering trees were generated using the R Clustree Package (Zappia and Oshlack, 2018) to identify a stable clustering resolution parameter. After identifying the number of PCs and resolution to use for downstream analysis, Uniform Manifold Approximation and Projection (UMAP)-based dimensional reduction was performed using Seurat (Butler et al., 2018). All single-cell expression signatures were created using AddModuleScore function in Seurat and visualization of data was undertaken using ggplot2, Seurat, Monocle2, and CONICSmat plotting functions (Butler et al., 2018; Muller et al., 2018; Trapnell et al., 2014; Wickham, 2016).

Copy number variations were inferred from single cell RNA-seq data using the CONICSmat R-package (Muller et al., 2018) with some minor modifications. Rather than using chromosomal arms for regions of mixture model assessment, a custom script was used to define a regions of interest. This script defined regions based on the difference between the 100 gene rolling average of inferred malignant and non-malignant cells (for this, ''sympathetic'' and ''mesenchymal'' clusters were assumed to be malignant in preliminary supervision of mixture models, while the non-malignant clusters were assumed to be all other clusters). Statistically significant mixture-models were then assessed using these new differentially defined regions. All reported mixture models were chosen based on: a 2-component model being more likely than a 1 component model using Bayesian Information Criterion (BIC) ratio of >1.01 (see Table S2, BIC ratio = BIC.1component/BIC.2.component). After regions for CNVs were chosen, FISH was used to validate candidate CNVs (see above for STAR methods). If a CNV region was validated by FISH, a posterior probability was calculated whether an individual cell in each tumor possesses the candidate CNV or not (threshold >0.75 to considered to be CNV positive). All CNV heatmaps were generated using the identified malignant and non-malignant clusters from the above analyses and were plotted using an expression threshold to keep only genes with average expression greater than the 10% quantile of all genes. For visualization, CNVs were color mapped using a lower bound of 1 standard deviation and upper bound of 2.5 standard deviations from the mean of the global rolling average. Projection of CNVs to UMAP plots was done using the Seurat (Butler et al., 2018).

Trajectory analyses of single cells was performed using the Monocle2 R package (version 2.16) (Trapnell et al., 2014) with the CellDataSet created using UMI count values. Pseudotime and trajectory calculation was undertaken using expressed genes from the Adrenergic and Mesenchymal signatures previously published (van Groningen et al., 2017), where genes part of cell cycle were removed to account the influence of cell cycle genes on the trajectory (Tirosh et al., 2016). Gene expression pseudotime plots were plotted using plot_genes_in_pseudotime function and gene expression pseudotime heatmaps were plotted using plot_genes_branched_heatmap function in the Monocle package. Cell state signatures were created using the top 150 genes ranked by adjusted p value using FindAllMarkers for overexpressed genes (average log fold change>0.25) that are expressed in at least 25% of cells for that state designation (Table S3).

Gene ontology analysis was performed using the topGO package. Significantly enriched gene ontology terms were identified using the ''classic'' algorithm and the ''fisher'' test. Representative GO Terms were chosen that had a p value < 0.001.

Analysis of the 40 sample external dataset was from three external sources Dong et al., GEO: GSE137804 (Dong et al., 2020), Kildisiute et al. from https://www.neuroblastomacellatlas.org/ (Kildisiute et al., 2021) and Slyper et al., GEO: GSE140819 (Slyper et al., 2020). Analysis was similar as described above. One exception was that data integration was required using R package Harmony, using the study each sample derived from to account for batch effects. Labels for adrenergic, transitional and mesenchymal classification were achieved using FindTransferAnchors in Seurat (Stuart et al., 2019) using our original 10 sample dataset as the reference.

### Bulk tumor analyses

Bulk tumor analyses were conducted on previously described tumor microarray cohorts: Kocak (Kocak et al., 2013) and Versteeg (Molenaar et al., 2012). Average gene expression signatures (i.e. Adrenergic, Transitional and Mesenchymal) were created from the average of all z-scores for genes in each signature. Kaplan-Meier analyses were conducted based on a median expression cutoff to stratify high and low expression groups. p-values for survival analysis were calculated using log rank tests. To classify tumors into 5

"Tumor Class" subgroups, average z-scores for each of the Adrenergic, Transitional and Mesenchymal signatures were again *Z* score scaled and subgroups were created based on the following criteria:

| Adrenergic signature expression | Transitional signature expression | Mesenchymal signature expression | Assignment |
|---|---|---|---|
| >0 | <0 | <0 | Adrenergic |
| >0 | >0 | <0 | Adrenergic-transitional |
| <0 | >0 | <0 | Transitional |
| <0 | >0 | >0 | Transitional-Mesenchymal |
| <0 | <0 | >0 | Mesenchymal |
| <0 | <0 | <0 | Either Adrenergic, Transitional or Mesenchymal groups, based on which signature had the highest value |
| >0 | <0 | >0 | Either Adrenergic or Mesenchymal groups, based on which signature had the highest value |
| >0 | >0 | >0 | Either Adrenergic, Transitional or Mesenchymal groups, based on which signature had the highest value |

### H3k27ac ChIP-seq analysis

bigWig files containing H3k27ac ChIP-seq read densities for 33 neuroblastoma samples were retrieved directly from the Gene Expression Omnibus: GSE90683 (Boeva et al., 2017). bigWig files were converted to the bedGraph format using the UCSC bigWigToBedGraph tool, and then combined into one bedGraph using bedtools (unionbed) (Quinlan and Hall, 2010). Each sample was then normalized for total read density and multiplied by a 10^6 scaling factor. The bedops tool was then used to find overlapping read bins with the 5975 previously annotated superenhancer regions (Boeva et al., 2017; Neph et al., 2012). Matched RNA-seq profiles for these 33 samples were directly retrieved from GEO: GSE90683 (Boeva et al., 2017). Average gene expression signatures (i.e. Adrenergic, Transitional and Mesenchymal) were created from the average of all z-scores for genes in each signature. To classify samples into the same 5 subgroups as described above for bulk tumor analyses, average z-scores for each of the Adrenergic, Transitional and Mesenchymal signatures were again *Z* score scaled and subgroups were created based on the criteria described previously. Normalized ChIP-seq read densities in the bedGraph file were then averaged across matching genomic bins for samples belonging to each of the five subgroups determined by RNA-seq. Five bigWigs representing averaged H3k27ac ChIP-seq profiles across the five subgroups were generated from the bedGraph file using bedtools (bdg2bw) (Quinlan and Hall, 2010). Differentially enriched super-enhancers were then identified using 2-sided t-tests for the normalized ChIP-seq read densities for previously annotated super-enhancer regions (Boeva et al., 2017) between cells belonging to each Tumor Class. All p values were then adjusted using Benjamini & Hochberg correction and enriched super-enhancers were identified if $p < 0.05$. Enriched super-enhancers for the adrenergic cell lines were identified by comparing to the mesenchymal and vice versa. Transitional enriched super-enhancers were identified by comparing cell lines of any combination of Adrenergic-Transitional, Transitional or Transitional-mesenchymal cell lines compared with Adrenergic & Mesenchymal cell lines (see Tables S4A and S4B for comparisons and statistics). Downstream visualization was undertaken using the EnrichedHeatmap package on all unique enriched super-enhancers identified by differential testing based on 1kb bins through the 0.5Mb region around the center of each annotated super-enhancer locus (Gu et al., 2018). All H3K27ac *Z* score plots were created by taking the mean of all z-scores of all enriched superenhancers designated as either Adrenergic, Transitional or Mesenchymal for each binned region for each cell line and plotted using ggplot2 (Wickham, 2016). H3K27ac *Z* score plots for Adrenergic and Mesenchymal enriched super-enhancers were then quantified for each tumor class by calculating the total area under curve using the auc function in R package MESS using default parameters considering each cell line belonging to the 5 Tumor Classes (error bars among cell lines from each class are represented as standard deviation). H3K27ac traces on core-regulatory transcription factors and markers of transitional state were plotted using the Gviz package (Hahne and Ivanek, 2016) using regions corresponding previously annotated super-enhancers (Boeva et al., 2017).

## QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were undertake using R software. If not detailed above, specific statistical analyses undertaken in this manuscript are detailed in the respective figure legends.