


RESEARCH

Open Access



Transcriptional immunogenomic analysis reveals distinct immunological clusters in paediatric nervous system tumours

Arash Nabbi^{1†}, Pengbo Beck^{2,3†}, Alberto Delaidelli^{4,5}, Derek A. Oldridge^{6,7}, Sumedha Sudhaman⁸, Kelsey Zhu¹, S. Y. Cindy Yang¹, David T. Mulder¹, Jeffrey P. Bruce¹, Joseph N. Paulson⁹, Pichai Raman¹⁰, Yuankun Zhu¹⁰, Adam C. Resnick¹⁰, Poul H. Sorensen^{4,5}, Martin Sill^{2,3}, Sebastian Brabetz^{2,3}, Sander Lambo^{2,3}, David Malkin⁸, Pascal D. Johann^{2,3,11}, Marcel Kool^{2,3,12}, David T. W. Jones^{2,13}, Stefan M. Pfister^{2,3,11}, Natalie Jäger^{2,3*} and Trevor J. Pugh^{1,14,15*} 

Abstract

Background Cancer immunotherapies including immune checkpoint inhibitors and Chimeric Antigen Receptor (CAR) T-cell therapy have shown variable response rates in paediatric patients highlighting the need to establish robust biomarkers for patient selection. While the tumour microenvironment in adults has been widely studied to delineate determinants of immune response, the immune composition of paediatric solid tumours remains relatively uncharacterized calling for investigations to identify potential immune biomarkers.

Methods To inform immunotherapy approaches in paediatric cancers with embryonal origin, we performed an immunogenomic analysis of RNA-seq data from 925 treatment-naïve paediatric nervous system tumours (pedNST) spanning 12 cancer types from three publicly available data sets.

Results Within pedNST, we uncovered four broad immune clusters: Paediatric Inflamed (10%), Myeloid Predominant (30%), Immune Neutral (43%) and Immune Desert (17%). We validated these clusters using immunohistochemistry, methylation immune inference and segmentation analysis of tissue images. We report shared biology of these immune clusters within and across cancer types, and characterization of specific immune cell frequencies as well as T- and B-cell repertoires. We found no associations between immune infiltration levels and tumour mutational burden, although molecular cancer entities were enriched within specific immune clusters.

Conclusions Given the heterogeneity of immune infiltration within pedNST, our findings suggest personalized immunogenomic profiling is needed to guide selection of immunotherapeutic strategies.

Keywords Paediatric neuro-oncology, Immunogenomics, CNS tumours, Tumour microenvironment, Neuroblastoma

[†]Arash Nabbi and Pengbo Beck contributed equally to this work.

*Correspondence:

Natalie Jäger
n.jaeger@kitz-heidelberg.de
Trevor J. Pugh
trevor.pugh@utoronto.ca

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Cancer immunotherapies have been clinically and experimentally investigated in paediatric oncology with a wide range of response rates. Objective responses to immune checkpoint inhibitors (ICI) as a monotherapy have been limited to 5–11% of paediatric cancers [1–3]. Addition of a monoclonal antibody targeting disialoganglioside GD2, granulocyte–macrophage colony-stimulating factor (GM-CSF) and interleukin-2 to standard therapy (isotretinoin) has been shown to improve overall survival of children with high-risk neuroblastoma treated with intensive multimodal therapy [4, 5]. More recently, anti-GD2 Chimeric Antigen Receptor (CAR) T-cell therapy has been investigated in diffuse midline gliomas in four patients, three of whom showed clinical improvement [6]. Experimentally, novel immunotherapies have been proposed and tested in preclinical models including CAR-T targeting immune checkpoint protein, B7-H3 [7], and targeting the myeloid compartment with anti-CSF1R [8] or anti-CD47 [9]. Considering the wide range of response rates for existing and emerging paediatric cancer immunotherapies, biomarkers for patient stratification are needed to identify potential candidates for clinical trials.

Immunogenomic analysis of tumours has been a major focus of biomarker discovery for immunotherapy. A prominent outcome of such studies is the FDA approval of tumour mutation burden (TMB) and microsatellite instability as the first immunotherapy-related biomarkers in adult cancers [10]. Several other biomarkers have been derived from large-scale genomic and transcriptomic datasets with a major focus on adult extracranial tumours [11–15]. Despite recent single-cell RNA sequencing (RNA-seq) studies in neuroblastoma [16], high-grade glioma [17], ependymoma [18, 19] and medulloblastoma [20], a comparison of the immune microenvironment across paediatric nervous system cancers and implications for informing immunotherapeutic interventions or patient selection have not been systematically analysed. In this study, we sought to characterize the immune microenvironment in 925 treatment-naïve paediatric central and peripheral nervous system cancers, as they share embryonal origins and similar levels of TMB [21]. We showed discrepancies across existing immune deconvolution tools when applied to paediatric nervous system tumours. To address this, we performed a consensus geneset analysis using publicly available datasets to identify specific immune genes with no tumour cell expression, followed by further characterization of pathway and gene expression, TMB and genomic alterations, T- and B-cell repertoire and cellular interactions.

Methods

Human subjects

We curated a total of 925 tumours from the Children's Brain Tumour Network [22] (CBTN, $n=581$), the Therapeutically Applicable Research To Generate Effective Treatments [23] (NCI TARGET, $n=149$) and the International Cancer Genome Consortium [24] (ICGC, $n=195$). We selected tumours from paediatric patients (median age 7 years) who had RNA-seq data generated from their primary tumours. We focused this study to major types of paediatric CNS tumours and neuroblastoma; ETMR ($n=9$), neurofibroma ($n=11$), choroid plexus tumours ($n=16$), meningioma ($n=13$), schwannoma ($n=14$), craniopharyngioma ($n=27$), ATRT ($n=31$), ependymoma ($n=65$), pedHGG ($n=83$), neuroblastoma ($n=151$), medulloblastoma ($n=208$) and pedLGG ($n=298$). As adult cancer comparator, we included TCGA participants with common types of adult cancers; glioblastoma multiforme (GBM, $n=153$), low-grade glioma (LGG, $n=507$), skin cutaneous melanoma (SKCM, $n=102$), colorectal adenocarcinoma (COAD, $n=298$), ovarian serous adenocarcinoma (OV, $n=373$), prostate adenocarcinoma (PRAD, $n=497$) and lung adenocarcinoma (LUAD, $n=522$).

Tumour datasets

We used RNA-seq datasets from the TARGET [23] (<https://ocg.cancer.gov/programs/target>), CBTN [1] (<https://cbtn.org/research/specimendata/>), ICGC [24] (<https://icgc.org/icgc/cgp/62/345/822>) and TCGA [25] (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>) in this study. For the ICGC dataset, we chose 195 primary CNS tumour samples with matched RNA-seq and WGS, and complete clinical data [21]. For the TARGET dataset, we chose 149 primary neuroblastoma samples with matched RNA-seq and WES data. For the CBTN dataset [22], we refined the original dataset ($n=996$) and excluded data from 22 cell lines, 153 progression samples, 70 recurrences and 16 secondary malignancies. We further removed 98 samples from rare (defined as $\leq 10\%$ of the cohort) or unannotated tumours. To enable comparison at the transcriptional level, we performed unsupervised clustering on primary samples and removed 51 samples that did not match their corresponding pathological annotation (Additional file 1: Fig. S1B-C). This sample curation resulted in 581 samples from the CBTN consortium (Additional file 1: Fig. S1A). These resulted in a cohort of 925 paediatric CNS tumours and neuroblastomas.

We accessed CBTN, TARGET and TCGA data through the Kids First Data Resource Centre [22] (<https://portal.kidsfirstdrc.org/>) and used the CAVATICA [22] platform

for data processing and analysis (<https://cavatica.sbggenomics.com/>). In total, we analysed 925 cases of paediatric nervous system tumours (CNS and neural crest tumours) (CBTN, TARGET and ICGC). We included RNA-seq data from 79 PDX models (2 ATRT, 10 ependymoma, 10 HGG, 14 neuroblastoma, 43 medulloblastoma) obtained from ITCC-P4 as control for immune infiltration and immune cell-specific geneset analysis (<https://www.itcc-consortium.org/>). We used a neuroblastoma dataset as a validation set for immune checkpoint profiling that was available on CAVATICA pounder accession number phs001436.c1 [26]. This dataset was not included in the pedNST cohort, as tumour types (primary, relapsed, etc.) were not known.

Transcriptome data processing

RNA-seq reads from TCGA and TARGET were aligned to human genome 38 (hg38) as described on the GDC website (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline). We used RSEM output for the TCGA and TARGET from Toil [27]. The CBTN raw RNA-seq reads were aligned to hg38 using the STAR v2.5.2b [28] and quantified using RSEM v1.2.28 [29] (detailed workflow can be accessed under <https://github.com/kids-first/kf-RNA-seq-workflow>). The ICGC data were aligned and quantified using STAR [28] v2.7.6a and RSEM [29] v1.2.28, respectively. With the exception of the ICGC, all data processing was performed on the CAVATICA data analysis platform [22]. For immune inference, we used transcripts per million (TPM), as per instructions [30]. For gene set enrichment scores and consensus clustering, we used \log_2 transformed TPM values corrected for batch effects (data source) using ComBat function from the *sva* [31] R package. To adjust for cancer type differences, we normalized \log_2 transformed batch-corrected TPM values using median expression in each cancer type.

Variant and fusion calling

For the TCGA and TARGET, variant calls were obtained from whole exome sequencing (WES) while the ICGC and CBTN variant calls were from whole genome sequencing (WGS). For the TCGA and TARGET, we restricted our analyses to those with MuTect2 [32] calls available on the GDC (details available at https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/DNA_Seq_Variant_Calling_Pipeline). For the ICGC, we used SNV and Indels calls as previously described [21]. CBTN WGS data were processed using MuTect2 [32] and Strelka2 [33] (detailed workflows at <https://github.com/kids-first/kf-somatic-workflow>.) and only overlapping calls were used for downstream analysis. For all datasets, we excluded somatic variants with less than

3% variant allele frequency. The Arriba [34] workflow was used for gene fusion calling on CBTN and ICGC datasets (<https://github.com/suhrig/arriba>). Detailed workflows are available at <https://github.com/DKFZ-ODCF/RNAseqWorkflow> [35] and <https://github.com/kids-first/kf-rnaseq-workflow> [36].

TMB and oncogenic pathways

To ensure all variants in our study were covered across WES and WGS datasets, we generated a common region list by intersecting bed files used in exome experiments with 50 base pairs padding. The exome kits included Agilent Custom V2 Exome Bait (TCGA and TARGET), Agilent SureSelect All Exon 38 Mb V2 (TCGA), Agilent SureSelect All Exon 50 Mb (TCGA), SeqCap EZ Exome Probes v3.0 (TCGA), SeqCap EZ Human Exome Library v2.0 (TARGET) and SeqCap EZ HGSC VCRome 2.1 (TCGA). The final common bed file consisted of 30,028,393 base pairs. We then used this region list to subset the WGS variant calls. We included non-synonymous coding SNV and Indels in the TMB calculations.

For molecular pathways and their associations with immune clusters, we identified samples with at least one alteration in genes involved in ten TCGA oncogenic pathways [37]. The Receptor Tyrosine Kinase (RTK) pathway contained 71 affected genes with a total of 419 alterations in 274 samples that encompassed 71% of pedLGG. Twenty six genes in the PI3-kinase pathway were altered in 74 pedNST samples, 32% of which were pedHGG. The Wnt pathway was disrupted in 70 samples, of which 36 and 20% were medulloblastoma and craniopharyngioma. We found 155 alterations in 51 genes involved in the Notch signalling pathway across 66 samples, primarily medulloblastoma (45%) and pedHGG (23%). Cancer entities with alterations in 32 genes involved in the HIPPO pathway included neuroblastoma (25%), pedHGG (22%) and medulloblastoma (20%). Five core genes involved in the TP53 pathway accounted for 56 mutations across 51 pedNST samples. At least one of 10 cell cycle genes was mutated in 16 pedNST samples, primarily in pedHGG (56%). Acknowledging that *MYCN* amplifications were not included in this analysis, we found 26 alterations in 10 genes related to MYC signalling in 15 samples. Nine samples showed alterations in the TGF- β pathway that consisted of 7 altered genes. NRF2 pathway contained three genes (*KEAP1*, *CUL3* and *NFE2L2*) with 7 mutations across 6 samples. We studied the effects of somatic alterations in the Mismatch Repair (MMR) pathway on immune microenvironment by comparing samples with at least one somatic alteration in *MLH1*, *MSH2*, *MSH6*, *PMS2*, *POLE* or *POLD1* to those without any alterations.

Immune infiltration and in silico simulations

We used ESTIMATE with default parameters to measure overall immune infiltration (<https://bioinformatics.mdanderson.org/public-software/estimate/>) [38]. We performed in silico simulations using bulk RNA-seq data from PDX models derived from medulloblastoma and RNA-seq data of four immune cell types (CD8+T-, CD4+T-, B-cell and NK cells) downloaded from ENCODE portal under identifiers ENCSR861QKF [39] (CD8+T-cells), ENCSR463JBR [40] (CD4+T-cells), ENCSR449GLL [41] (B-cells), ENCSR357XTU [42] (NK cells)). For each simulation, an equal percentage of reads from four immune cell types was randomly sampled using SAMtools [43] v1.9, converted to fastq files with BEDTools [44] v2.27.1 and then concatenated together with randomly sampled reads from PDX RNA-seq data to a total of 100 million reads. The generated pseudo-samples were then processed with Kallisto and Sleuth [45, 46], and TPM values were used as input for the ESTIMATE.

Immune microenvironment analysis

We used TIMER2 [30] web interface (<http://timer.cistrome.org/>) for comprehensive analysis of immune cell composition using six computational tools, CIBERSORT [47], EPIC [48], QUANTISEQ [49], MCPOUNTER [50], TIMER [51] and XCELL [52]. We used the CRI-Atlas Shiny app (<https://isb-cgc.shinyapps.io/iatlas/>) to cluster the pedNST samples with the immune subtype classifier previously published for adult cancers [14]. To estimate immune cell composition using ICGC methylation array data, we used EpiDISH (Epigenetic Dissection of Intra-Sample Heterogeneity) R package that uses cell-type-specific DNase Hypersensitive Site (DHS) data and robust partial correlations to infer immune cell compositions [53]. We used a reference signature matrix previously constructed and used for immune deconvolution analysis of paediatric CNS tumours [54].

Development of immune cell-type gene sets and consensus clustering

To identify genes specific to the immune system, we leveraged four data sources: (1) We obtained median gene expression for 18 purified cell types, as previously reported [13]. We found 9897 genes were expressed >75th percentile of normalized expression (median centred and Median Absolute Deviation (MAD) scaled) in at least one cell type within this dataset. We selected 1958 genes with ≥ 2 MAD difference between immune and non-immune cell populations. (2) We downloaded the human protein atlas data v20.1 (<https://www.proteinatlas.org/about/download>). The Human Protein Atlas consisted of single-cell data from 51 cell populations [55]. 15,302

genes were expressed >75th percentile in at least one cell population. (3) In the Human Protein Atlas and across 37 tissues, 11,069 genes were annotated as specific to at least one tissue, determined as normalized expression ≥ 1 (median normalized expression 33.9). We found 7206 genes specific to only one tissue, 1257 of which were specific to blood, lymphoid tissue or bone marrow. (4) We included the ESTIMATE immune signature consisting of 141 genes [38]. In total, we found 3041 unique genes across four data sources with evidence of specificity to immune cell populations.

Next, we sought to identify genes that may be expressed in paediatric cancer and other non-immune cells. We compiled a paediatric cancer geneset using three data sources. (1) We used protein-coding gene expression data from 79 PDX models. We found 13,788 genes expressed >75th percentile in at least one PDX. (2) We used data from 22 cell lines collected by CBTN. These consisted of one ependymoma, 18 pedHGG and 3 medulloblastoma. Across these cell lines, 11,195 protein-coding genes were expressed >75th percentile in at least one cell line. (3) We leveraged published single-cell datasets from paediatric cancers, neuroblastoma [16] and ependymoma [18, 19] and used original cell annotations as reported by authors. As single-cell datasets are generally sparse, we focused our analysis to genes with maximum expression >75th percentile in at least one cell population in each dataset. We then determined non-immune genes as those with ≥ 2 MAD difference between non-immune and immune cell populations in each dataset. This analysis resulted in 3036, 12,109 and 5040 non-immune genes from Jansky et al. [16], Gojo et al. [18] and Gillen et al. [19] datasets, respectively. Intersecting the immune and non-immune genes resulted in 791 immune-specific genes.

To accurately assign the identified genes to immune cell types, we compiled genesets from seven sources and conducted a consensus approach to assign genes to specific immune cell types. For this analysis, we aggregated immune cell subtypes (e.g. Tregs) into the following major immune cell types: T, B, NK cells, dendritic cells, monocytes, macrophages, granulocytes and myeloid cells. These sources included gene annotations from five immune deconvolution tools, MCPOUNTER [50], QUANTISEQ [49], CIBERSORT [47], EPIC [48] and TIMER [51]. We used all 264 genesets that were used to develop XCELL [52]. We used two additional sources: (1) scaled average counts of 7172 genes specific to 28 purified immune cell types derived from 416 healthy donors and patients with immune diseases (Immunexut) [56]. We chose 3896 genes with scaled count >75th percentile in at least one cell type, of which 2555 genes had ≥ 2 MAD difference between one cell type compared to all others. (2)

The human protein atlas annotated 5934 genes as blood cell specific [57], determined as genes with normalized expression ≥ 1 in at least one cell type. We chose 4902 genes that were specific to one immune cell type. Our consensus analysis revealed 216 genes that were assigned to one specific immune cell type in at least two independent sources. We excluded myeloid cell geneset due to the low number of genes ($n=6$) that may lead to inaccurate enrichment scores. To calculate enrichment scores, we applied single-sample geneset enrichment analysis (ssGSEA) from GSVA R package [58] to batch-corrected \log_2 -transformed TPM values across pedNST. We then applied consensus clustering to the scaled (median centred and MAD scaled) immune cell-type enrichment scores using the ConsensusClusterPlus R package [59]. We performed k -means clustering based on Euclidean distance with 200 subsamples using 80% of samples and 100% of features for 2 to 8 clusters. This analysis revealed four major immune clusters across pedNST.

For further delineate types of infiltrating T-cells, we derived enrichment scores of 40 T-cell subtypes using top 50 signature genes derived from recent pan-cancer analysis of tumour-infiltrating lymphocytes [60]. We focused this analysis on Paediatric Inflamed to avoid overestimation as a result of low overall immune infiltration in other clusters. We scaled enrichment scores within each T-cell subtype across Paediatric Inflamed and performed consensus clustering with parameters as previously described. For myeloid cell subtypes, we used 13 myeloid signatures derived from single-cell analysis of adult cancer patients [61]. We chose 9 gene signatures consisting of genes with ≥ 1.5 fold change in expression between any given myeloid cell-type and other clusters. We scaled enrichment scores from myeloid gene signatures within each myeloid cell subtype across Myeloid Predominant and performed consensus clustering with parameters as previously described. For microglia, we used signatures from two previous reports and derived enrichment scores in the pedCNS subset of Myeloid Predominant using ssGSEA [62, 63].

Immunohistochemistry

Formalin-fixed, paraffin-embedded TMA sections were analysed for CD4, CD8 and CD19 expression. The IHC staining was performed using the Ventana Discovery platform. IHC was optimized and performed with CD4 (Abcam Ab183685), CD8 (Leica NCL-L-CD8-4B11) and CD19 (e-Bioscience 14-0194) with dilutions of 1:500, 1:100 and 1:500 respectively. In brief, tissue sections were incubated in Tris EDTA buffer (cell conditioning 1; CC1 standard) at 95 °C for 1 h to retrieve antigenicity, followed by incubation with the

respective primary antibody for 1 h. Bound primary antibodies were incubated with the respective secondary antibodies (Jackson Laboratories) with 1:500 dilution, followed by Ultramap HRP and Chromomaps DAB detection. For staining optimization and to control for staining specificity, normal tonsil was used as control. Intensity scoring was done on a common 4-point scale. Descriptively, 0 represents no staining, 1 represents low but detectable degree of staining, 2 represents clearly positive staining and 3 represents strong expression. Expression was quantified as H-score, the product of staining intensity and percentage of stained cells. The TMAs used in the study were obtained from the Children's Oncology Group (COG) and contains 9 ATRT, 13 pedHGG, 20 ependymoma, 64 medulloblastoma and 33 neuroblastoma cases, all of which are represented in duplicate cores.

To study protein levels of TIM3 and LAG3 in neuroblastoma, we purchased two serial TMA slides consisting of 26 neuroblastoma cases all in duplicate cores with 10 cores of normal peripheral nerve tissue (Biomax, NB642a). Briefly, TMA sections were dewaxed and rehydrated through an ethanol series to water and endogenous peroxidases were blocked using 3% H_2O_2 in PBS for 15 min at room temperature. Antigen retrieval was performed using 10 mM sodium citrate pH 6.0. Primary antibodies were incubated for 45 min at room temperature (TIM3 CST45208 1:150, LAG3 ab40466 1:200) followed by several washes with PBS. Secondary antibodies were incubated for 30 min at room temperature (BA-1000 and BA-9200 at 1:500 dilutions Vectors Labs). ABC kit (PK-6100, Vectors Labs) was applied for 25 min with DAB for 4 min (SK-4100) followed by PBS washes. Stained slides were scanned using a Nanozoomer 2.0HT (Hamamatsu Photonics) at 20 \times or 40 \times .

Computational tumour-infiltrating lymphocyte (TIL) analysis was performed on digitally scanned haematoxylin and eosin (H&E)-stained whole-slide images (WSI) from the Children's Brain Tumour Network (CBTN) as previously described [64]. Briefly, a pre-trained Inception-V4 deep learning model was applied to each WSI, returning a model probability of TIL presence within 50 \times 50 micron patches tiling across the entire WSI. Image patches that did not contain tissue were filtered out using a color standard deviation threshold of 18 (computed on an 8-bit [0,255] color scale), and a mean TIL score was computed as the average TIL probability on the remaining patches for each WSI. Upon manual review, WSI with poor scanning quality and/or with a high false positive rate of TIL detection were removed from the analysis, including medulloblastoma and other tumours arising in the cerebellum/posterior fossa. One outlier sample was removed (Grubbs test).

HLA-typing and neoantigen prediction

We used the Optitype [65] tool available on the CAVATICA platform to determine HLA-A, B and C types from the TARGET and CBTN RNA-seq datasets. To predict neoantigens from gene mutations, we used Mutant Peptide eXtractor and Informer (MuPeXI) tool using MuTect2 variant caller as input [66]. We predicted 8-11mer peptides for all HLA class I types. We determined strong or weak binding peptide-MHC complexes as previously described (percentile rank ≤ 0.5 for strong binding, > 0.5 and ≤ 2 for weak binding peptides) [67]. We included peptides predicted from mutations that overlapped between MuTect2 and Strelka2.

T- and B-cell repertoire analysis

To recover T- and B-cell clonotypes, we used MiXCR v2.1.12 [68] with default parameters for RNA-seq data processing. We applied the framework iNterpolation/EXTrapolation (iNEXT) to study immune repertoire diversity [69]. This method was primarily developed for diversity estimates in ecology and aimed to bring together asymptotic estimates and rarefaction/extrapolation methods of estimation for samples with different sizes. Three most popular diversity indices have already been translated to immune repertoire studies; richness, the total number of species, Shannon index (entropy), which puts moderate weight on abundance and Simpson index, which measures diversity of abundant species. The robustness of these indices is dependent on sample size and could be biased in under-sampled experiments, which is a known problem in biodiversity studies and attempts have been reported to reduce such empirical biases [70]. Chao et al. provided a more accurate estimation of diversity by integrating slopes obtained from accumulation curves into entropy formula followed by bootstrapping method for variance [70]. We applied similar principles on RNA-seq datasets to infer immune diversity. We found that the estimates were severely affected in samples with less than three clonotypes recovered from bulk RNA-seq or in samples with multiple clonotypes of the same clonal fractions and excluded them from our analysis. Note that these samples may be (A) samples with low number of immune cells and therefore low number of clonotypes that could not be reflected on shallow RNA-seq data, or (B) samples with sufficient immune infiltration, but highly uneven clonal distribution, therefore only highly abundant clonotypes were seen at shallow coverage, yet the true diversity could not be robustly estimated.

For B-cell repertoire, we used constant regions of immunoglobulin heavy chain (IGH C segments) to study the distribution of immunoglobulin isotypes. Due to somatic hypermutation in the B-cell repertoire,

we removed sequences with ≤ 2 reads. As B-cell clones originating from a naïve B-cell may have mismatches in CDR3 sequences via somatic hypermutation events, we clustered IGH CDR3 octamers allowing for one mismatch, as described previously [71]. We considered sequences shorter than 8 amino acids as individual CDR3 sequences. Only samples with > 3 CDR3 sequences were included from this analysis and we used the gini index of inequality as a measure for uneven distribution of B-cell clusters.

TCR sequencing

To validate the T-cell diversity estimates from RNA-seq, we applied CapTCR-seq [72] hybrid capture protocol to RNA-seq libraries of adult nasopharyngeal carcinoma ($n=33$), paediatric samples from the PRecision Oncology For Young peopLE (PROFYLE) programme (www.profy.le.ca) ($n=10$) and ten paediatric GBM samples (ICGC). We compared estimated values of TCR β richness, Shannon and Simpson diversities from bulk RNA-seq data with observed richness, Shannon and Simpson diversities from CapTCR-seq experiment.

Statistics and visualization

Statistical tests, p -values and other details are noted in text and figure legends. For tumour subtypes, we used one-sided Fisher's exact test (alternative = 'greater' in R) comparing each subtype to all others within each immune cluster. We used analysis of covariance (ANCOVA) to compare continuous variables such as TMB or age at diagnosis across immune clusters adjusting for cancer entities. For survival analysis, we defined overall survival (OS) as follows. For NCI TARGET, we used 'Overall Survival Time in Days' and 'vital status' reported in the metadata. For CBTN, we used 'Last Known Status' with 'Deceased-due to disease' recoded as 'Event'. We calculated OS time using 'Age at Collection' of the first collected specimen ('Initial CNS Tumor') and 'Age At Last Known Status' of the last collected specimen, as reported in the metadata. For ICGC, we defined OS as the time between the date of first diagnosis and date of death or last follow-up. We defined progression-free survival (PFS) as follows. For NCI TARGET, we used 'Event Free Survival Time in Days' and 'First Event' as reported in the metadata with 'Event', 'Death', 'Progression', 'Relapse' and 'Second Malignant Neoplasm' recoded as 'Event'. For CBTN cases with one surgical resection, we used 'Last Known Status' with 'Deceased-due to disease' recoded as 'Event'. For these cases, we calculated PFS time using 'Age at Collection' and 'Age At Last Known Status' of the collected specimen, as reported in the metadata. For cases with multiple surgical resections, we calculated PFS time using 'Age at Collection' of the first collected specimen

(‘Initial CNS Tumor’) and ‘Age at Collection’ of the second collected specimen (‘Progressive’ or ‘Recurrence’), as reported in the metadata. For ICGC, we calculated PFS as the time between date of first diagnosis and date of first relapse/progression or death (if death was the first event), or date of last follow-up (if no event).

We used log-rank tests for Kaplan–Meier analyses. We used Cox proportional hazards models to adjust for cancer entities and gender in multivariable analyses. We used Student’s *t* test to compare scaled values across groups and rank sum test to compare values with non-normal distribution. We applied the Cochran–Mantel–Haenszel (CMH) test to compare samples with mutations in oncogenic pathways across immune clusters controlling for the effects of cancer entities. In all boxplots, boxes show median and IQR and whiskers represent 1.5 times IQR. Boxplots are shown for groups with more than three datapoints. We annotated statistical significance levels as follows: **p* < 0.05, ***p* < 0.01 and ****p* < 0.001. We performed all analyses and visualizations in R v4.0. We used Adobe Illustrator v24.0.1 for aesthetic edits and figure alignments.

Results

Establishment of paediatric nervous system tumour cohort for immunogenomic analysis

To study immune attributes across paediatric nervous system tumours (pedNST), we compiled a cohort of bulk RNA-seq data from three consortia: the Children’s Brain Tumour Network (CBTN), the National Cancer Institute Therapeutically Applicable Research To Generate Effective Treatments initiative (NCI TARGET) [23] and the International Cancer Genome Consortium (ICGC) [24]. To better represent primary pedNST and to enable comparison at the transcriptional level, we performed unsupervised clustering on the CBTN RNA-seq dataset (*n* = 581, Additional file 1: Fig. S1A–C). We further annotated CBTN medulloblastoma and ATRT tumour

subtypes using unsupervised clustering for downstream analysis (Additional file 1: Fig. S1D–E, ‘Methods’). We included 195 primary samples from the ICGC [21] with matched RNA-seq, whole genome sequencing (WGS) and DNA methylation data and 149 primary neuroblastomas from the NCI TARGET [23] with matched RNA-seq and whole exome sequencing (WES) data [23].

The final, aggregated non-overlapping paediatric dataset for immunogenomic analysis consisted of 925 samples with primary locations in the central nervous system (CNS) or peripheral nervous system: embryonal tumours with multilayered rosettes (ETMR, *n* = 9), neurofibroma (NFB, *n* = 11), choroid plexus tumours (CP, *n* = 16), meningioma (MNG, *n* = 13), schwannoma (SCHW, *n* = 14), craniopharyngioma (CPH, *n* = 27), atypical teratoid/rhabdoid tumours (ATRT, *n* = 30), ependymoma (EPN, *n* = 65), paediatric high-grade glioma (pedHGG, *n* = 83), neuroblastoma (NBL, *n* = 151), medulloblastoma (MB, *n* = 208), and paediatric low-grade glioma (pedLGG, *n* = 298) (Fig. 1A) (Additional file 2: Table S1). We included data from 79 patient-derived xenograft (PDX) models as part of the Innovative Therapies for Children with Cancer Pediatric Preclinical Proof-of-Concept Platform (ITCC-P4) project as negative controls lacking immune or stromal infiltration [73, 74] (Fig. 1B).

Immune infiltration analysis reveals high variability across and within cancer types

To determine the overall levels of immune infiltration across pedNST, we first assessed the performance of the deconvolution tool ESTIMATE [38] using in silico simulations and converted immune scores to ‘immune read percentage’ using a non-linear regression model (Additional file 1: Fig. S2B). To enable comparison with adult brain tumours and cancers in which ICI agents have been clinically studied [75–79], we performed a parallel analysis of 2452 primary adult tumour samples from The Cancer Genome Atlas (TCGA) [25, 80]:

(See figure on next page.)

Fig. 1 Transcriptional analysis of 925 paediatric nervous system tumours (pedNST) reveals four distinct immune clusters. **A** Overview of cohorts and sample size in the present study (PDX: patient-derived xenografts, ETMR: embryonal tumour with multilayered rosettes, NFB: neurofibroma, MNG: meningioma, SCHW: schwannoma, CP: choroid plexus tumours, CPH: craniopharyngioma, ATRT: atypical teratoid/rhabdoid tumour, EPN: ependymoma, pedHGG: paediatric high-grade glioma, NBL: neuroblastoma, MB: medulloblastoma, pedLGG: paediatric low-grade glioma). **B** Distribution of percentage immune reads based on ESTIMATE immune score across paediatric and adult cancers. PDX samples serve as a negative control for immune infiltration. PRAD: prostate adenocarcinoma, LGG: low-grade glioma, OV: ovarian serous cystadenocarcinoma, SKCM: skin cutaneous melanoma, COAD: colorectal adenocarcinoma, GBM: glioblastoma multiforme, LUAD: lung adenocarcinoma. **C** Heatmap representing consensus clustering using enrichment scores derived from six immune-cell-specific genesets across the pedNST cohort. Cell-type scores correspond to normalized gene set enrichment scores. **D** Heatmap illustrating proportion of samples in each immune cluster across cancer entities. Barplots show the total number of samples for each row (cancer type) and column (immune cluster). **E** Heatmap showing fraction of samples in each CRI-iAtlas cluster across immune clusters in pedNST. Barplot shows the total number of samples in each CRI-iAtlas cluster. **F** Boxplot showing average tumour-infiltrating lymphocyte (TIL) scores as determined by segmentation analysis of pathological images across the paediatric CNS tumour samples (CBTN). Three sample images are shown representing 1% (bottom), 10% (middle) and 15% (top) TIL scores corresponding to lower quartile, mean and higher quartile, respectively. Boxes show median and interquartile range (IQR) and whiskers represent 1.5 times IQR. Two-sided rank sum test, **p* < 0.05. **G** Barplots showing frequency (top barplot) and fraction (stacked barplot) of tumour subtypes across immune clusters (SHH: Sonic Hedgehog, TYR: Tyrosine, MYCN-NA: MYCN non-amplified, MYCN-A: MYCN amplified, WT: wildtype)

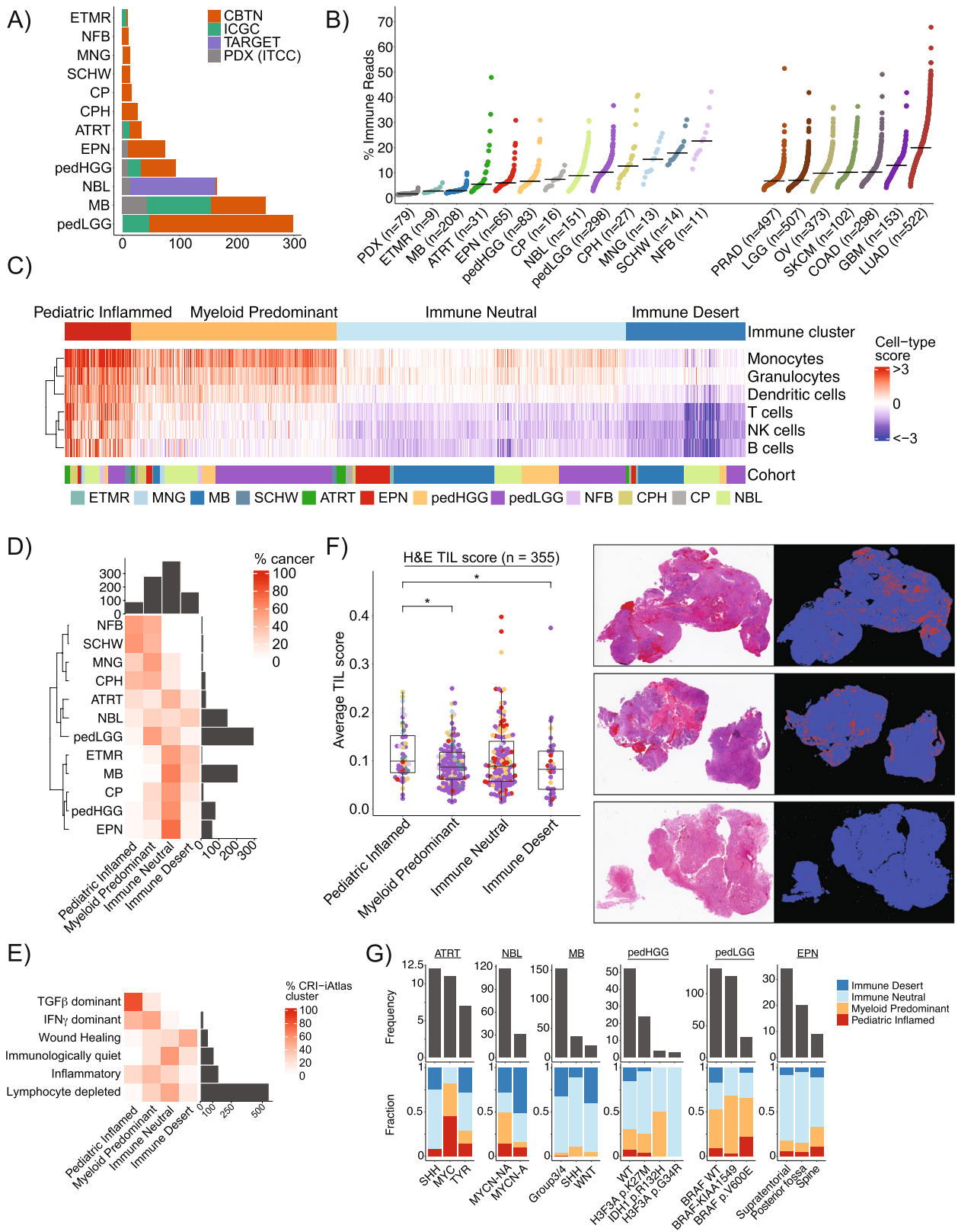


Fig. 1 (See legend on previous page.)

glioblastoma multiforme (GBM, $n=153$), low-grade glioma (LGG, $n=507$), skin cutaneous melanoma (SKCM, $n=102$), colorectal adenocarcinoma (COAD, $n=298$), ovarian serous adenocarcinoma (OV, $n=373$), prostate adenocarcinoma (PRAD, $n=497$), lung adenocarcinoma (LUAD, $n=522$).

Immune read percentages varied considerably within and across pedNST (median 7.2%, range 1.9–47.9%) and in similar range compared to adult CNS tumours (median 7.8%, range 2.9–41.8%). Common non-CNS adult cancers showed higher overall levels of immune infiltration (median 10.3%, range 2.5–67.8%). In pedNST, the highest values were in extracranial entities including neurofibroma, schwannoma and meningioma (medians 22.6, 17.8 and 15.3%, respectively). Paediatric brain tumour entities ETMR and medulloblastoma had the lowest median immune infiltration of all cancers analysed (medians 2.7 and 2.8%) [81–83] (Fig. 1B). Choroid plexus tumours had consistently low variation of immune read percentage (range 3.1–13%, median 7.3%). In contrast, ATRT exhibited the widest distribution of immune infiltration (range 2.3–47.9%, median 5.5%), consistent with known associations between infiltration and ATRT subgroups [84, 85]. Similarly, a wide range of immune read percentages were observed in craniopharyngioma and neuroblastoma (range 3.9–40.9%, median 12.6% and 2.6–30.7%, median 8.8%). These data revealed notable examples of highly infiltrated and immune-excluded samples in paediatric nervous system tumours and identified general trends of immune infiltration.

To investigate whether the wide variability in immune infiltration could be recapitulated at the protein level, we applied immunohistochemistry (IHC) to a tissue microarray (TMA) from an independent cohort of 139 paediatric cancers, obtained from Children's Oncology Group (COG). The median CD8 staining was greatest in ATRT followed by neuroblastoma and pedHGG, while medulloblastoma and ependymoma had median H-score (defined as product of staining intensity and percentage of stained cells) of zero (Additional file 1: Fig. S3A-B). We found CD4 staining in eight samples (H-score > 0, 3 ATRTs, 3 ependymoma, 2 medulloblastoma), while other samples showed no CD4 staining. CD19 staining was variable among cancer types ranging from 3% (medulloblastoma) to 51% (neuroblastoma) (Additional file 1: Fig. S3A-B). Summing H-scores across all three markers, we found highest staining in ATRT followed by neuroblastoma (medians 17 and 15) and medulloblastoma showing the lowest staining. These results confirm the high variability of immune cell-type infiltration within and across pedNST, as well as generalized trends inferred from gene expression.

Consensus gene set clustering reveals four distinct immune clusters in pedNST

In light of the variability of immune infiltrates within each tumour type, we sought to categorize immune microenvironments across pedNST. We detected unexpected immune signals inferred by various immune deconvolution algorithms when applied to PDX RNA-seq data (Additional file 1: Fig. S4A-G, Additional file 2: Table S2), which should not yield a (human) immune cell signature. This indicated wide discordance and non-specific signals across existing immune deconvolution tools. To address the non-specific immune signal that may originate from tumour cells, we sought to identify immune cell-specific genes that lack expression in paediatric nervous system cancer cells. We compiled 3041 immune-related genes by incorporating data from four sources: ESTIMATE signature [38], an immune cell compendium [13], the Human Protein Atlas (HPA) single cell [55] and tissue specificity datasets [86] (Methods). We excluded genes with evidence of expression in paediatric cancer cells using data from the PDX models (ITCC-P4), single-cell RNA-seq [16, 18, 19] and established cell lines (CBTN). This analysis identified that 791 of the 3041 immune genes are not expressed in paediatric nervous system tumour cells. We performed consensus analysis using gene sets from aforementioned immune deconvolution tools along with gene expression profiling of 28 immune cell types [56], and the HPA blood-cell-specific dataset [57]. Using this approach, 216 genes were assigned to a single cell-type in at least two data sources. These included genes specific for T-cells ($n=41$), B-cells ($n=32$), NK cells ($n=20$), monocytes ($n=18$), dendritic cells ($n=28$) and granulocytes ($n=71$) (Additional file 2: Table S3). Consensus clustering of normalized enrichment scores for these cell-specific genesets identified four distinct immune clusters (Fig. 1C,D) that we characterized further and designated Paediatric Inflamed, Myeloid Predominant, Immune Neutral, and Immune Desert based on differential expression of aforementioned genesets.

To place these clusters in context of prior immunogenomics work in adult tumours, we applied the CRI-Atlas, a comprehensive classification method performed on TCGA dataset consisting of over 10,000 tumour samples spanning 33 adult tumour types [14], to pedNST. We found that a lower proportion of paediatric brain tumours were grouped in 'Immunologically quiet' or 'Lymphocyte depleted' clusters compared to adult counterparts (85%, 635/749 vs 98%, 654/668). In contrast, 16.5% of paediatric extracranial tumours (neuroblastoma, neurofibroma and schwannoma, 29/176) harboured cold immune microenvironment, compared to 10.5% of adult extracranial tumours (888/8458). Across pedNST, we found 19% of pedNST ($n=174$) belonged to the

'Inflammatory' or 'IFN- γ dominant' clusters, while 7% ($n=64$) were grouped in 'Wound healing'. Twenty per cent of pedLGG (60/298) were grouped in the 'Inflammatory' cluster, in contrast to adult LGG, 98% of which showed a cold immune microenvironment. Almost all pedHGG samples (82/83) were grouped in the 'Immunologically quiet' or 'Lymphocyte depleted' clusters, similar to adult GBM (151/154). Overall, we found 72% of pedNST ($n=664$) were grouped in the 'Immunologically quiet' or 'Lymphocyte depleted' clusters indicating a generally cold immune microenvironment in pedNST (Fig. 1E).

The Paediatric Inflamed cluster ($n=90$, 9.7%) had the highest immune read percentage across pedNST. 57.1% of schwannoma ($n=8$) and 54.5% of neurofibroma ($n=6$) were clustered in this group followed by 40.7% of craniopharyngioma ($n=11$) and 30.8% of meningioma ($n=4$) (Fig. 1D). Five of six cases grouped in the 'TGF- β dominant' cluster belonged in this cluster, followed by 42.3% of 'IFN- γ dominant' samples ($n=11$) (Fig. 1E). This cluster was devoid of samples in the 'Immunologically quiet' or 'Lymphocyte depleted' cluster indicating that ~10% of pedNST samples harbour T-, B- and NK cells, monocytes, granulocytes and dendritic cells.

The Myeloid Predominant cluster ($n=279$, 30.1% of pedNST) scored highly for monocyte, dendritic cell and granulocyte gene sets while harbouring lower levels of lymphoid cell types. 53.8% of meningioma ($n=7$) and 53% of pedLGG ($n=158$) cases clustered in Myeloid Predominant followed by 44.5% of craniopharyngioma ($n=12$) (Fig. 1D). 50% of the samples within the CRI-iAtlas 'IFN- γ dominant' ($n=13$) and 40% of the 'Inflammatory' samples ($n=59$) were clustered in this group (Fig. 1E), suggesting the inflammatory component of this cluster is driven primarily by myeloid cells.

The Immune Neutral cluster ($n=393$, 42.5%) had myeloid and lymphoid cell infiltration scores near the median of the entire pedNST cohort. This cluster contained a large fraction of the cancer types with low immune read percentage including ETMR ($n=5$, 55.6%), medulloblastoma ($n=136$, 65.4%) and ependymoma ($n=47$, 72.3%) (Fig. 1D). This cluster included 56.4% of the 'Immunologically quiet' ($n=62$) and 48.7% of the 'Lymphocyte depleted' ($n=270$) tumours (Fig. 1E), indicating that intermediate immune infiltration across pedNST corresponds to cold immune microenvironment compared to adult cancers, which may contribute to intrinsic resistance to immune checkpoint inhibitors.

The Immune Desert cluster ($n=163$, 17.6%) received low immune inference scores for all immune cell types (Fig. 1C). This cluster included 32.5% of neuroblastoma ($n=49$) and 29.8% of medulloblastoma ($n=62$) and none of the immune infiltrated cancer entities (Fig. 1D). Purity

scores inferred from copy number alterations [87] in 156 ICGC samples revealed that samples in this cluster were of highest cancer cell content compared to Immune Neutral or Myeloid Predominant (Additional file 1: Fig. S5A-B).

To validate lymphocyte infiltration levels across the immune clusters, we leveraged a deep learning method of tissue image analysis for 355 pedCNS samples from CBTN with haematoxylin and eosin (H&E)-stained tissue slides and matched RNA-seq [64], as well as 195 ICGC samples with matched DNA methylation arrays and RNA-seq. Consistent with the RNA-seq immune inference analysis, choroid plexus tumours received the lowest tumour-infiltrating lymphocyte (TIL) scores determined by H&E tissue analysis (median 0.08), while meningioma and neurofibroma scored highest across cancer entities (medians 0.14 and 0.11, Additional file 1: Fig. S6A). TIL scores were variable across CNS tumour sites ranging from a median of 0.07 in cerebral hemispheres to a median of 0.18 in case of meninges (Additional file 1: Fig. S6B). Paediatric Inflamed showed significantly higher TIL score compared to Myeloid Predominant and Immune Desert (medians, 0.1, 0.09 and 0.08, respectively, two-sided rank sum test, $p=0.02$ and 0.03, Fig. 1E, Additional file 2: Table S4). This difference was more pronounced in samples with immune read percentage >5% (medians 0.1 and 0.03 in Paediatric Inflamed and Immune Desert) or in those collected from cerebral hemispheres (medians 0.09 and 0.05), suggesting that high immune infiltration and composition of tumour microenvironment may influence the H&E-based inference of TILs. These findings were recapitulated by immune inference analysis utilizing the ICGC DNA methylation array data [53, 54]. Paediatric Inflamed tumours displayed higher levels of T-cell infiltration (pairwise two-sided Student's t test with Bonferroni correction, $p<0.01$, Additional file 1: Fig. S6D). Although B-cell estimates were not significantly different between Paediatric Inflamed and Myeloid Predominant, lower levels of B-cells were estimated in Immune Neutral and Immune Desert compared to Myeloid Predominant ($p<0.001$, Additional file 1: Fig. S6D). Similarly, Immune Neutral and Immune Desert showed depletion of DNA methylation-based estimates for NK cells and monocytes compared to Myeloid Predominant, in agreement with our RNA-seq clusters ($p<0.001$, Additional file 1: Fig. S6D). We compared levels of CD8+ T-, CD4+ T- effector and regulatory T-cell (Treg) subtypes estimated from DNA methylation deconvolution across immune clusters. Paediatric Inflamed showed elevated levels of CD8+ T-cell and Treg estimates, while depleted in CD4+ T effector estimates. Conversely, Myeloid Predominant showed high levels of Treg and CD4+ T effector estimates with minimal CD8+ T-cells (Additional

file 1: Fig. S6E). These results suggest presence of a lymphocyte population in Myeloid Predominant that is undetectable in bulk RNA-seq.

As pedNSTs consist of distinct molecular entities and subgroups [88, 89], we next investigated associations between cancer subtypes and immune clusters (Fig. 1G). MYC-like ATRTs appeared to be enriched in Paediatric Inflamed or Myeloid Predominant compared to non-MYC-like ATRT, and 8 out of 12 SHH-like ATRT tumours clustered into Immune Neutral, although it did not reach statistical significance (one-sided Fisher’s exact test with Bonferroni correction, $p > 0.05$) [84]. In neuroblastoma ($n = 148$), *MYCN* non-amplified neuroblastoma cases were enriched in Myeloid Predominant ($n = 41$, $p = 0.007$), while neuroblastoma *MYCN* amplified showed a trend toward Immune Desert ($n = 16$, $p = 0.1$) [90, 91]. SHH medulloblastoma showed a trend toward Myeloid Predominant and Immune Neutral ($n = 4$ and 28 , $p > 0.05$), while 95% of WNT medulloblastoma were grouped into Immune Neutral and Immune Desert (19 out of 20). In pedLGG ($n = 298$), *BRAF* wildtype samples were enriched in Immune Desert ($n = 24$, $p = 5 \times 10^{-6}$), *BRAF-KIAA1549* fusion-positive tumours were grouped into Myeloid Predominant ($n = 82$, $p = 0.005$) and *BRAF*

p.V600E samples showed a trend toward Paediatric Inflamed ($n = 7$, $p = 0.07$). In summary, our findings reveal associations between tumour-intrinsic characteristics and the immune microenvironment, although it remains to be established whether the microenvironment is sculpted by the cancer cells or whether specific cancer subtypes thrive in specific microenvironments.

We next investigated whether there were associations between patient characteristics or outcome with the immune clusters. We found an association between male sex and pedNST, more prominently seen in medulloblastoma and ependymoma (Additional file 1: Fig. S7A). A logistic regression model adjusting for cancer type indicated a trend toward an association between male genetic sex and Immune Desert (Wald test, $p = 0.05$) (Fig. 2A). There were no significant associations between race or age and immune clusters independent of cancer type (Fig. 2B, C, Additional file 1: Fig. S7B-C). Kaplan–Meier analysis indicated significant differences in overall survival (OS) and progression-free survival (PFS) among the immune clusters (log-rank test, $p = 0.001$ and 0.02) (Fig. 2D, E). We observed poorer OS for Immune Neutral and Immune Desert compared to Paediatric Inflamed in a Cox proportional hazards model adjusting for cancer

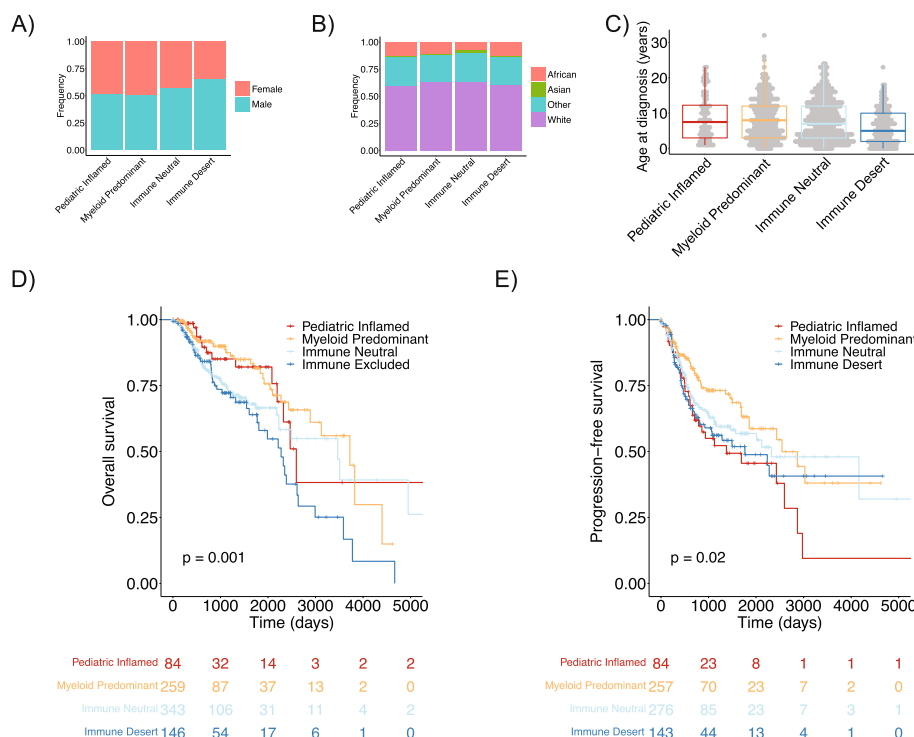


Fig. 2 Associations of immune clusters with patients’ clinical parameters and survival. **A, B** barplots showing distributions of gender (**A**) and race (**B**) across immune clusters in pedNST. ANCOVA, not significant. **C** boxplot depicting age distribution across immune clusters. ANCOVA, not significant. Boxes show median and interquartile range (IQR) and whiskers represent 1.5 times IQR. **D, E** Kaplan–Meier curves and risk tables for overall survival (**D**) and progression-free survival (**E**) among immune clusters. Log-rank test p -values are denoted

type and sex (hazard ratio (HR)=1.56 and 1.97, confidence interval (CI) 1.13–2.14 and 1.39–2.79, $p=0.006$ and 0.0001) (Additional file 2: Table S5). Myeloid Predominant cases showed a trend toward improved PFS compared to Paediatric Inflamed (HR=0.68, CI 0.44–1.06, $p=0.08$) (Additional file 2: Table S6). This trend may be a result of low but differential presence of T-cells in Myeloid Predominant compared to Paediatric Inflamed (Additional file 1: Fig. S6E). Our results suggest that independent of cancer type, infiltration of specific immune cell populations (Paediatric Inflamed and Myeloid Predominant) provides a survival advantage over low infiltration (Immune Neutral and Desert) that may indicate poor immune recognition.

Gene expression analysis reveals differential molecular pathways and immunoregulatory genes in immune clusters

To understand the cellular pathways underlying each immune cluster, we conducted quantitative set analysis for gene expression (QuSAGE) [92] comparing samples in one cluster against all others while adjusting for data source and cohort. Leveraging the Molecular Signatures Database (MSigDB) Hallmark genesets [93], we found that immune-related pathways such as IFN- α , IFN- γ and allograft rejection were significantly enriched in Paediatric Inflamed and Myeloid Predominant and depleted in Immune Neutral and Immune Desert (Fig. 3A, Additional file 2: Table S7). Mitotic spindle and G2M checkpoint pathways were significantly enriched in Immune Desert indicating a high proliferation rate with low amounts of non-malignant cells. To investigate whether these associations could be recapitulated at the protein level, we leveraged data from 147 samples with matched RNA-seq and proteomic profiles [62]. Compared to Paediatric Inflamed, Immune Neutral and Immune Desert showed a significantly lower average protein z -score in eight immune pathways identified using QuSAGE (pairwise two-sided Student's t test with Bonferroni correction, $p<0.05$, Fig. 3B). These findings reveal active immune pathways in Paediatric Inflamed and Myeloid Predominant consistent with the inflammatory and IFN- γ dominant adult clusters [14].

We further sought to determine key differential genes underlying each immune microenvironment. Using differential gene expression analysis [94], we found cell-type-specific genes upregulated in Paediatric Inflamed including B-cell (*JCHAIN*, *MZB1* and *CD79A*), macrophages (*MARCO*) and granulocyte-specific genes (*ADGRG3* and *FPR2*) (Fig. 3C). Consistently, differentially expressed genes comparing Paediatric Inflamed to other samples included genes specific to B-cells, plasma cells and macrophages (Additional file 1: Fig. S8A). In

Myeloid Predominant, we found only one significantly upregulated gene, *CHIT1*, which encodes for chitotriosidase secreted by active macrophages [95] (Fig. 3C). The lack of other significantly upregulated myeloid-specific genes may suggest heterogeneity among myeloid cell types in different cancer entities. *JCHAIN* and *CXCL10* were downregulated along with *MMP9* in Immune Neutral (Fig. 3C). Immune-related genes such as *JCHAIN*, *CD3* chains, *IL6* and cytotoxicity genes *GZMK* and *GZMA* were significantly downregulated in the Immune Desert (Fig. 3C). Altogether, these results reveal core immunological pathways and genes driving immune clusters.

We sought to determine possible immunoregulatory players in pedNST. Comparison of expression levels of 59 genes with known regulatory functions [96] revealed differential expression across immune clusters after adjusting for cancer type and the total immune infiltrate (Analysis of Covariance (ANCOVA)). Members of CD28 superfamily receptor (*PDCDI*, *CTLA4*, *BTLA* and *ICOS*) were upregulated in Paediatric Inflamed ($p=6.61\times 10^{-9}$, 9.98×10^{-12} , 3.04×10^{-13} and 1.67×10^{-17} , respectively, Fig. 3D). Genes encoding PD-L1 and PD-L2 (*CD274* and *PDCD1LG2*) were upregulated in Paediatric Inflamed ($p=0.03$ and 6.34×10^{-4}). Genes involved in regulatory T-cell (Tregs) function and activation (*IDO1*, *STAT3* and *FOXP3*) were upregulated in Paediatric Inflamed ($p=1.63\times 10^{-8}$, 0.002 and 2.31×10^{-11} , respectively). These findings reveal high expression levels of immune checkpoint genes as well as genes involved in immunosuppression in Paediatric Inflamed.

In Myeloid Predominant, *TGFB3* encoding for TGF- β 3 was significantly upregulated (ANCOVA, $p=0.007$) (Fig. 3D). *LGALS3* encoding for Galectin-3 with negative regulatory functions in macrophages [97] was upregulated in Myeloid Predominant suggesting the prominent macrophage influence in this cluster ($p=0.004$). Other myeloid-specific genes with immunoregulatory functions were upregulated in both Paediatric Inflamed and Myeloid Predominant and included *IL4I1* [98], *CCL2* [99] and *FGF2* [100] (Fig. 3D). These findings show preferential expression of myeloid-related regulatory genes in Myeloid Predominant.

In Immune Desert, we found that *CD276* encoding immune checkpoint protein B7-H3 to be highly expressed ($p=9.61\times 10^{-4}$, Fig. 3D). Unexpectedly, among immune checkpoint genes, *LAG3* was expressed at higher levels in Immune Desert ($p=0.004$). Further inspection of immune checkpoint genes in PDX data confirmed high expression of *LAG3* (Additional file 1: Fig. S8B). Across medulloblastoma samples in Immune Desert, WNT subgroup expressed *LAG3* at significantly higher levels compared to SHH and there was

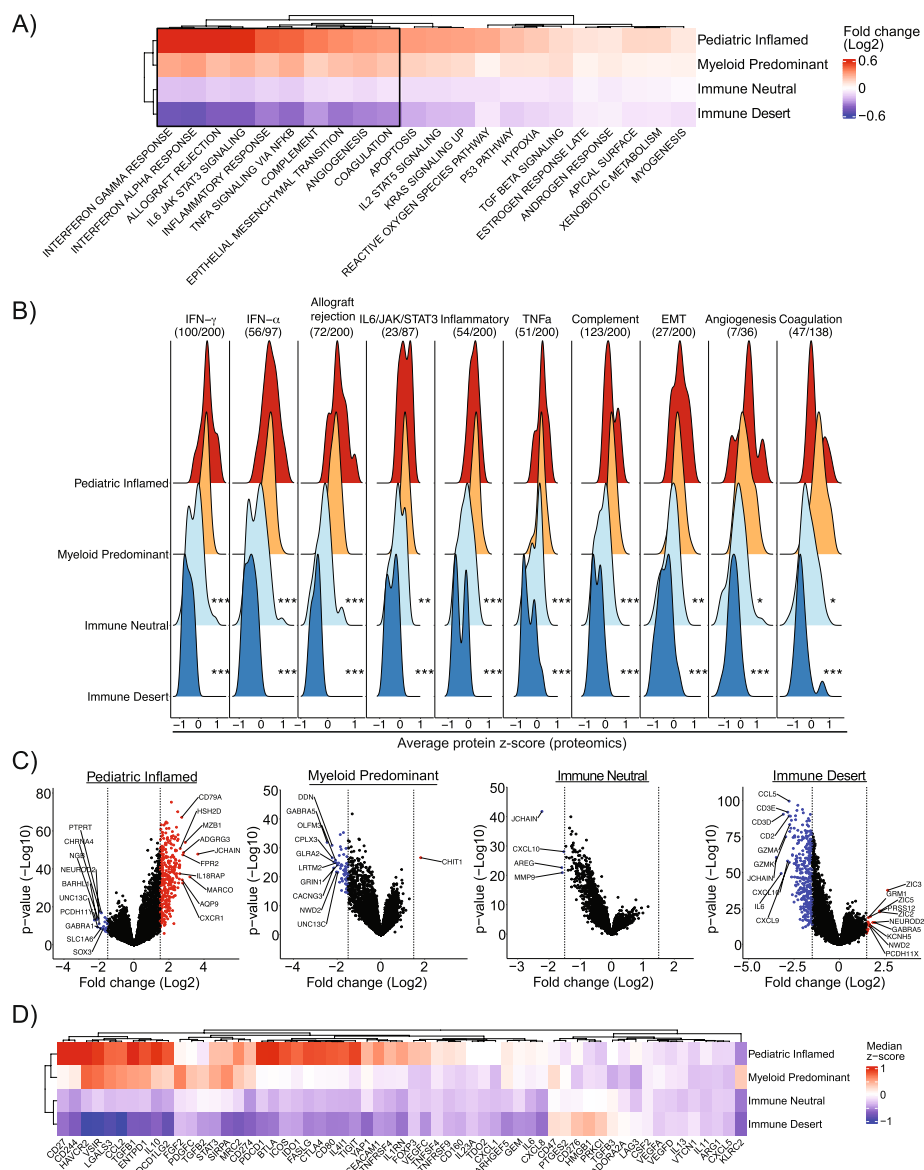


Fig. 3 Pathway and differential gene expression analysis confirm immune features across immune clusters. **A** Heatmap showing log-fold change in gene enrichment scores (derived from QuSAGE) in each immune cluster compared to all others in pedNST. Columns show MSigDB pathways with false discovery rate (FDR) < 0.1 in at least three immune clusters. Black box outlines pathways validated by protein in (B). **B** Ridge plots illustrating average z-scores of proteins [62] involved in immune pathways from A across immune clusters in 141 pedCNS samples (CBTN). Fraction indicates the number of protein readouts from the number of genes in each pathway. Pairwise two-sided Student’s *t* test with Bonferroni correction, * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. Significance levels are shown for comparison to Paediatric Inflamed. **C** Volcano plots for differentially expressed genes (derived from DESeq2) in each immune cluster compared to other samples. Up- or downregulated genes with absolute log-fold change > 1.5 and FDR < 0.1 are shown in red or blue, respectively. Dashed line shows the log-fold change threshold. Top ten differentially expressed genes are annotated. **D** Heatmap showing median z-score expression of 59 genes with known immunoregulatory functions across immune clusters in pedNST

a non-significant trend between WNT and Group3/4 (two-sided Student’s *t* test, $p = 0.02$ and 0.08 , Additional file 1: Fig. S8C). These results suggest *LAG3* expressed by tumour cells in pedNST calling for further validation of its function in the brain and identify B7-H3 as

a possible mechanism of immune exclusion in subsets of pedNST.

In neuroblastoma, profiling expression of five immune checkpoint genes revealed ~10% of samples had exclusively elevated expression of *HAVCR2* gene encoding

TIM3 (Additional file 1: Fig. S8D). We confirmed this finding using a non-overlapping neuroblastoma dataset of 209 immune infiltrated tumours. In this cohort, we found 13 cases (6.2%) that showed elevated expression of *HAVCR2* and low expression of *LAG3* (Additional file 1: Fig. S8D). Immunohistochemical staining for TIM3 and LAG3 proteins using an independent TMA showed that TIM3 was detectable in 14 samples without LAG3 staining (Additional file 1: Fig. S8E). These findings identify an *HAVCR2* expression in a subset of neuroblastomas suggesting a distinct mode of immunosuppression.

Immune microenvironment associations with tumour-intrinsic genomic alterations and tumour mutational burden

We next asked whether immune clusters are associated with tumour mutational burden (TMB). We did not find a statistically significant difference in the total number of non-synonymous single-nucleotide variants (SNV) per coding megabase (SNV/Mb) or SNV + Insertion/deletion (Indel) /Mb (SNV + Indel /Mb) across immune clusters (ANCOVA comparing to Paediatric Inflamed, Fig. 4A,B, Additional file 2: Tables S9, S10). However, when we looked at 63 pedHGG samples, we found that Myeloid Predominant and Immune Neutral harboured significantly higher TMB compared to Immune Desert (two-sided rank sum test, $p=0.02$, Fig. 4C). Four of 6 cases with >5 SNV + Indel/Mb and germline variations in *MLH1*, *MSH2*, *MSH6*, *PMS2*, *POLE* or *POLD1* belonged to Myeloid Predominant, suggesting higher levels of immune infiltration in patients with biallelic Mismatch Repair Deficiency (bMMRD) syndrome. Our findings otherwise revealed no associations between the immune microenvironment and TMB in non-hypermutant pedNST, suggesting TMB, with the exception of MMR-associated tumours, may not be an appropriate biomarker for immune checkpoint inhibitors in this population.

To elucidate whether antigen presentation may be associated with immune clusters, we investigated the MHC presentation potential of somatic mutations in pedNST, we determined patients' HLA class I types [65] (Additional file 1: Fig. S9) and performed the mutant peptide extractor and informer (MuPeXI) [66] analysis. We identified 7591 strong binding and 21,680 weak binding peptides, as defined previously [67]. The number of predicted strong binding peptides was highest in neuroblastoma and pedHGG (medians 19 and 17) and lowest in craniopharyngioma (median 2.5), consistent with its low TMB (median 0.07 SNV + Indel/Mb, Additional file 1: Fig. S10A-B). Differences in the number of strong or weak binding peptides did not reach statistical significance after adjusting for cancer type and data

source (ANCOVA, Fig. 4D, Additional file 1: Fig. S10C, Additional file 2: Table S11). Consistent with our TMB observation, pedHGG showed a significantly higher number of strong or weak binding peptides in Myeloid Predominant compared to Immune Neutral or Immune Desert (Fig. 4E, Additional file 1: Fig. S10D). While these data indicate neopeptides are not independently associated with immune infiltration across cancer entities in the pedNST, higher numbers of neopeptides in pedHGG cases clustered in the Myeloid Predominant suggest associations with immune infiltration and potentially immune recognition in this cancer entity.

We hypothesized that disruptions in oncogenic pathways may be associated with distinct immune clusters. We identified samples with alterations (SNVs, Indels and fusions) in at least one gene in the ten TCGA oncogenic pathways [37] (Fig. 4F). Tumours with somatic alterations in members of the receptor tyrosine kinase (RTK) pathway were most frequently found in Myeloid Predominant (Cochran-Mantel-Haenszel (CMH) test, $p=6.1 \times 10^{-5}$, Fig. 4F). Paediatric Inflamed showed a higher percentage of samples with mutations in the Mismatch Repair (MMR) pathway, although this difference did not reach statistical significance across cancer entities ($p=0.41$, Fig. 4F). Among common driver mutations (Fig. 4G) [101], we found that *BRAF*-altered samples were differentially clustered among immune clusters, with *BRAF-KIAA1549* fusion-positive samples grouped primarily in Myeloid Predominant ($p=3 \times 10^{-6}$, Fig. 4G). While the causal relationship between tumour-intrinsic genomic alterations and immune microenvironment remains unclear, our results reveal associations of altered molecular pathways and microenvironmental features.

T- and B-cell repertoire analysis suggest antigen presentation and clonal outgrowth in immune infiltrated pedNST

Adaptive immune system consisting of T- and B-cells is a critical determinant of anti-tumour immunity with clonal expansion suggesting antigen exposure. To infer potential immune reactivity across pedNST, we sought to determine the extent of clonal diversity for T- and B-cells. Using an immune repertoire processing framework [68], we recovered a total of 23,842 complementarity-determining region (CDR3 β) sequences from 582 pedNST samples. To validate the diversity estimates, we used the CapTCR-seq method [72] to enrich T-cell receptor (TCR) sequences in RNA-seq libraries from adult and paediatric cancer samples ($n=26$). TCR β estimated Shannon diversity inferred from bulk RNA-seq showed the highest correlation with the observed Shannon diversity obtained from CapTCR-seq (Fig. 5A, Additional file 1: Fig. S11A-B). Comparing the diversity

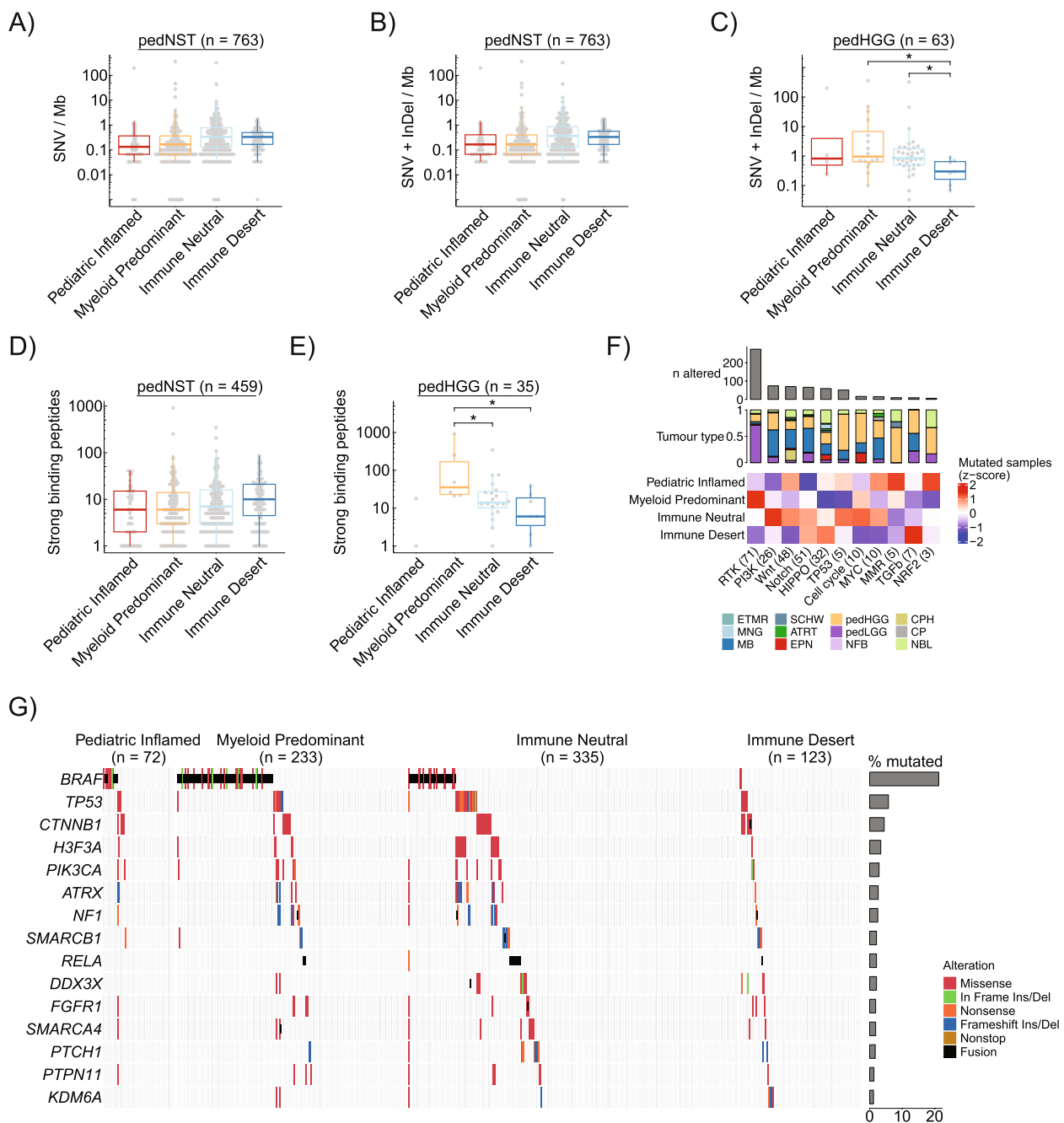


Fig. 4 No associations between TMB or predicted neopeptides across immune clusters. **A, B** Boxplots showing TMB defined as SNV per megabase (Mb) (**A**) or SNV+Indel per Mb (**B**) across immune clusters in pedNST. ANCOVA, not significant. **C** Boxplot showing TMB for pedHGG samples across immune clusters. Two-sided rank sum test, $*p < 0.05$. **D** Boxplot showing number of predicted strong binding peptides (defined as binding affinity ≤ 0.5) for pedNST samples across immune clusters. ANCOVA, not significant. **E** Boxplot showing number of predicted strong binding peptides for pedHGG samples across immune clusters. Two-sided rank sum test, $*p < 0.05$. **F** Heatmap illustrating scaled number of samples (z-score) with at least one non-synonymous SNV/Indels in genes involved in oncogenic pathways, as defined by TCGA. Barplot shows the total number of samples with alterations in each pathway. Stacked barplot shows proportion of tumour types present in samples with alterations in each pathway. Numbers in brackets indicate the number of altered genes in each pathway. **G** Oncoprint illustrating the distribution of somatic mutations in the top 15 most commonly altered genes in pedNST across the four immune clusters. In all boxplots, boxes show median and IQR and whiskers represent 1.5 times IQR

estimates and total number of TCR β reads showed a linear correlation (adjusted $r^2=0.72$, Fig. 5B). However, 12 samples displayed lower diversity relative to their number of TCR β reads, suggesting a T-cell clonal expansion in these samples. Conversely, eight samples harboured outlier high diversity indicating several individual clones infiltrating these tumours (Fig. 5B). These findings suggest a method of T-cell diversity inference and identification of potential polyclonal and clonal repertoires using bulk RNA-seq data.

Using estimated Shannon diversity, we analysed 361 pedNST samples for TCR β diversity (Additional file 1: Fig. S11C, 'Methods'). Neuroblastoma samples in Paediatric Inflamed showed significantly higher diversity compared to neuroblastoma samples in Immune Neutral and Desert (two-sided Student's t test with Bonferroni correction, $p=0.03$ and 2.6×10^{-6} , Fig. 5C). Within pedCNS, Paediatric Inflamed had higher diversity compared to Myeloid Predominant and Immune Neutral ($p=3.3\times 10^{-7}$ and $<2\times 10^{-16}$, Fig. 5C). To determine microenvironmental changes that co-vary with T-cell diversity independent of T-cell infiltration, we calculated residuals from a linear regression model for TCR β diversity and reads (Fig. 5B). In Paediatric Inflamed, samples with residuals in lower quartile, corresponding to uneven T-cell repertoire, had significantly higher dendritic cell scores ($p=0.01$, Fig. 5D). Conversely, within Myeloid Predominant, samples with residuals in lower quartile harboured significantly higher levels of monocytes ($p=0.03$, Fig. 5D). These results reveal associations of myeloid compartment with clonal T-cell repertoire, suggesting the possibility of clonal outgrowth as a consequence of interactions with antigen-presenting cells.

We used a similar immune repertoire inference tool as for the T-cell repertoire analysis [68] and recovered 197,769 unique immunoglobulin heavy chain (IGH) CDR3 sequences. Across pedNST, we found the highest number of IGH isotypes in craniopharyngioma and

neuroblastoma samples (Additional file 1: Fig. S11D). Consistent with findings in adult cancers [71], IGHG1 constituted the largest proportion of B-cell repertoire relative to total number of isotypes (Additional file 1: Fig. S11D). We found IGHG1 and IGHG3 were significantly enriched in Myeloid Predominant compared to Immune Neutral and Immune Desert (two-sided Student's t test, $p=0.03$ and 0.008 , Fig. 5E), suggesting that potential antibody-producing cells infiltrated these tumours.

To determine the extent of B-cell clonal expansion in pedNST that may suggest antigen recognition [102], we used the gini index as a measure of uneven B-cell cluster distribution in each sample [103] (Additional file 1: Fig. S11E, 'Methods'). Across neuroblastoma samples, Immune Desert had significantly lower gini index compared to Paediatric Inflamed (medians 0.29, and 0.53, two-sided Student's t test with Bonferroni correction, $p=0.05$, Fig. 5F). Within pedCNS, Paediatric Inflamed had the higher levels of gini index compared to Myeloid Predominant and Immune Neutral (medians 0.36, 0.26 and 0.23, respectively, $p=0.03$ and 0.002 , Fig. 5F). Within Paediatric Inflamed, samples with clonal T-cell repertoire (residuals ≤ 25 th) also had clonal B-cell repertoire and expressed higher levels of *IGHG1* and *IGHG3* compared to samples with polyclonal T-cell repertoire (residuals ≥ 75 th) ($p=0.02$, 0.01 and 0.01 , Fig. 5G). We did not find this association in Myeloid Predominant that harboured lower levels of T-cells (Fig. 5G), suggesting that both the extent and clonality of B-cell infiltration may contribute to T-cell clonal expansion.

Lymphoid and Myeloid subtypes provide insights into mechanisms of immune evasion

To further characterize T-cell subtypes in the Paediatric Inflamed cluster with high levels of lymphocyte infiltration, we obtained 40 gene signatures of CD4+ and CD8+ T-cell subpopulations from the single-cell profiling of tumour-infiltrating lymphocytes [60]. Consensus

(See figure on next page.)

Fig. 5 T-cell repertoire analysis reveals associations of T-cell clonal expansion with B-cell repertoire and immune microenvironment in Paediatric Inflamed. **A** Scatterplot depicting a linear correlation between TCR β Shannon diversity estimated from RNA-seq and measured by capturing all TCR sequences from the same RNA-seq libraries in adult and paediatric cancer samples using CapTCR-seq. **B** Scatterplot showing correlation between TCR β estimated Shannon diversity and total number of TCR β reads. Blue line shows fitted linear regression. Red and blue dots represent polyclonal and clonal T-cell repertoires defined as residuals greater than two absolute standard deviations. Circle plots to the right illustrate two examples of polyclonal (top) and clonal (bottom) T-cell repertoires. Each circle is one T-cell clone and circle diameters are proportional to TCR β reads. **C** Boxplots showing differences in log-transformed estimated TCR β Shannon diversity across immune clusters for neuroblastoma (left) and pedCNS (right). Two-sided Student's t test with Bonferroni correction, $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$. **D** Boxplots comparing levels of T-cells, dendritic cells or monocytes, as determined in Fig. 1C, in samples with TCR β residuals (obtained from the linear regression in **B**) ≤ 25 th percentile and ≥ 75 th percentile of Paediatric Inflamed (left) or Myeloid Predominant (right). Two-sided Student's t test, $*p < 0.05$, NS: not significant. **E** Boxplot showing proportion of specific immunoglobulin isotypes in B-cell repertoires across immune clusters in pedNST. Two-sided Student's t test, $*p < 0.05$ and $***p < 0.01$. **F** Boxplots comparing immunoglobulin clonality (gini index) across immune clusters for neuroblastoma (left) and pedCNS (right). Two-sided Student's t test with Bonferroni correction, $*p < 0.05$ and $**p < 0.01$. **G** Boxplots comparing levels of gini index or tumour-type normalized expression of *IGHG1* or *IGHG3* in samples with TCR β residuals (obtained from the linear regression in **B**) ≤ 25 th percentile and ≥ 75 th percentile of Paediatric Inflamed (left) or Myeloid Predominant (right). Two-sided Student's t test, $*p < 0.05$, NS: not significant. In all boxplots, boxes show median and IQR and whiskers represent 1.5 times IQR

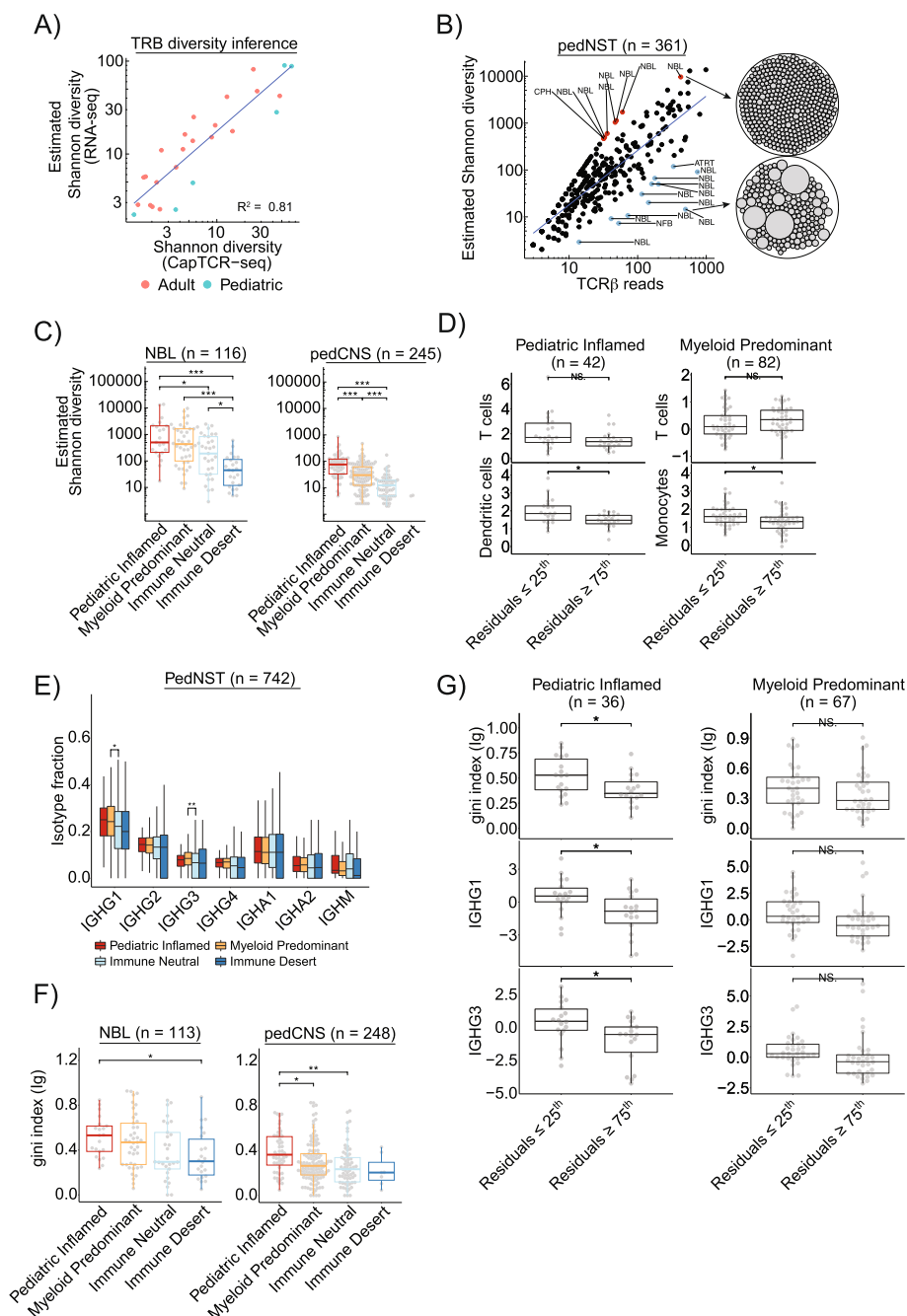


Fig. 5 (See legend on previous page.)

clustering revealed five distinct T-cell groups (TGs) within Paediatric Inflamed with notable patterns (Fig. 6A). Nine neuroblastoma samples were characterized by elevated enrichment of naïve T cells (TG2). TG3 consisted of 6 samples with enrichment of Naive and NME1+ T cell signatures while depletion of other T cell subtypes. TG4 ($n=30$) showed intermediate enrichment for T-cell signatures, while TG5 had highest enrichment

of T-cell signatures related to memory (Tm), effector memory (Tem), T follicular helper/helper (Tfh/h), Tregs and exhausted (Tex) T cells. These data identify subset of samples within Paediatric Inflamed that are infiltrated with naïve (TG2/TG3) or differentiated (TG4/TG5) T cells.

In Paediatric Inflamed, Cox proportional hazards model revealed that TG5, highly infiltrated with

differentiated T cells, had worse PFS compared to TG1 (HR=9, 95% CI 1.5–54, $p=0.02$, Additional file 2: Table S12). TG5 also showed lower T-cell diversity compared to TG2 and TG3, which were enriched with T cell naïve and proliferating NME1+T-cell signatures (Fig. 6A) (Student's t test with Bonferroni correction, $p=0.002$, Fig. 6B, left). While TG2 harboured significantly higher TMB compared to TG1 (Student's t test, $p=0.02$), TG5 did not show any significant difference in TMB compared to other groups (Fig. 6B, right). TG5 harboured higher infiltration of T, NK and myeloid cell infiltration suggesting antigen recognition by myeloid cells followed by clonal expansion and T-cell differentiation (Fig. 6C). TG5 showed higher expression of HLA class I genes (*HLA-A*, *HLA-B*, *HLA-C*), inhibitory cytokines and chemokines (*CXCL8*, *CXCL1*, *IL6*), genes encoding transcription factor AP-1 subunits (*FOS*, *FOSB*, *JUN*, *JUNB*, *ATF3*), markers of T-cell activation (*CD86*, *CD69*) and Toll-like receptors (*TLR2*, *TLR4*) (Fig. 6D). Conversely, TG2 showed higher levels of B-cell marker genes (*CD19*, *CD22*), immune checkpoint gene (*BTLA*) and Fc receptors (*FCER2*, *FCRL2*) (Fig. 6D) suggesting a B-cell-mediated mechanism of T-cell suppression. These findings suggest two immunosuppression mechanisms in Paediatric Inflamed: suppression of T-cell effector functions (TG5) and inhibition of T-cell activation and differentiation (TG2).

To delineate myeloid cell subtypes in Myeloid Predominant, we performed consensus clustering using single-cell signatures for tumour-infiltrating myeloid cells [61, 104] ('Methods'). MG1 consisted of 22 samples with high enrichment of mast cells, conventional dendritic cells type 2 (cDC2), monocytes and macrophages (Fig. 6E). MG2 ($n=89$) showed low enrichment of dendritic cells but high levels of *PLTP*-expressing macrophages. MG3 ($n=70$) was characterized by elevated enrichment of plasmacytoid and conventional dendritic cells (pDC and cDC1), and tumour-associated macrophages (TAM). MG4 ($n=75$) showed depletion of mast and macrophage signatures with moderate enrichment of monocytes and cDC1 signatures, while MG5 ($n=23$) had the

lowest enrichment of all myeloid signatures (Fig. 6E). Scoring microglia-specific gene signatures [62, 63] in CNS tumours showed pedLGG samples had the highest levels of microglia geneset enrichment, primarily in MG3 (Fig. 6F). These findings reveal two subgroups within Myeloid Predominant, characterized by high levels of mast cells, macrophages and monocytes (MG1) and high levels of pDC and TAMs (MG3).

Cox proportional hazards model adjusting for gender and cohort revealed that low Mast/Macro groups (MG3, MG4 and MG5) had a trend toward improved PFS compared to MG1 (HR=0.4, 0.38 and 0.37, 95% CI 0.14–1.14, 0.14–1.02 and 0.12–1.12, $p=0.09$, 0.05 and 0.08, respectively, Additional file 2: Table S13). In MG1, differential gene expression analysis showed genes associated with cytotoxic T cells (*FGFBP2*) and leukocyte trafficking (*TSPAN8*) [105] were expressed at higher levels compared to MG3 and MG5 (Fig. 6G). *PENK*, encoding an opioid receptor proenkephalin and a precursor for neuropeptides involved in immune-neural cross-talk [106] was expressed at higher levels (Fig. 6G). MG1 also showed high expression of *ANGPTL7*, a regulator of angiogenesis [107], and *F2RL2*, encoding proteinase-activated receptor 3 (PAR3), that may trigger myeloid infiltration [108]. *IL34*, a brain-specific ligand for CSF1R, was expressed at higher levels in MG1 suggesting a role in monocyte-macrophage differentiation in MG1 [109, 110]. In MG3, genes involved in endothelial regulation (*ESM1*, *APLN*, *ALCAM*) and ligands for $\alpha V\beta 3$ integrin (*EDIL3* and *POSTN*) were expressed at higher levels compared to MG1 and MG5 (Fig. 6G), the latter of which recruit tumour-associated macrophages (TAM) in adult glioblastoma [111, 112]. Notably, MG3 exhibited higher expression of Histamine receptor H1 (*HRH1*) and *HNMT* involved in histamine metabolism (Fig. 6G) suggesting immune evasion and T-cell dysfunction [113]. While further validation is warranted, these findings suggest cellular interactions influencing the myeloid compartment (periostin-integrins, IL34-CSFR1) and immune exclusion and evasion (histamine and collagens) in Myeloid Predominant.

(See figure on next page.)

Fig. 6 Analysis of immune-cell subtypes in Paediatric Inflamed and Myeloid Predominant provides insights into mechanisms of immunosuppression. **A** Heatmap showing scaled gene set enrichment scores for 40 T-cell subtypes as annotated in [60] in the Paediatric Inflamed (Tn: Naïve T-cells, Th: T helper, ISG: Interferon-stimulated genes, Tm: memory T-cells, Trm: Tissue-resident memory T-cells, Tem: effector memory T-cells, Temra: effector memory re-expressing CD45RA T-cells, Treg: regulatory T-cells, Tfh: follicular helper T-cells, Th17: T helper 17, Tex: exhausted T-cells, KIR: Killer immunoglobulin-like receptor). **B** Boxplots showing log-transformed TCR β estimated Shannon diversity (left) and TMB (right) across T-cell groups (TG). Pairwise two-sided Student's t test with Bonferroni correction, * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$ (left). Two-sided rank sum test, * $p < 0.05$ (right). **C** Boxplots showing differences in immune-cell infiltration, as determined in Fig. 1C, across T-cell groups (TG). Pairwise two-sided Student's t test with Bonferroni correction, * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. **D** Heatmap illustrating median tumour-type normalized expression of selected genes in TG2 and TG5. **E** Heatmap depicting scaled gene set enrichment scores for 9 myeloid cell subtypes as annotated in [61, 104] (DC: dendritic cell, Mono: monocyte, Mac/Macro: macrophage, TAM: tumour-associated macrophage). **F** Heatmap showing scaled gene set enrichment scores for microglia [62, 63] in the pedCNS subset of Myeloid Predominant. **G** Heatmap illustrating median tumour-type normalized expression of selected genes in MG1, MG3 and MG5. In all boxplots, boxes show median and IQR and whiskers represent 1.5 times IQR

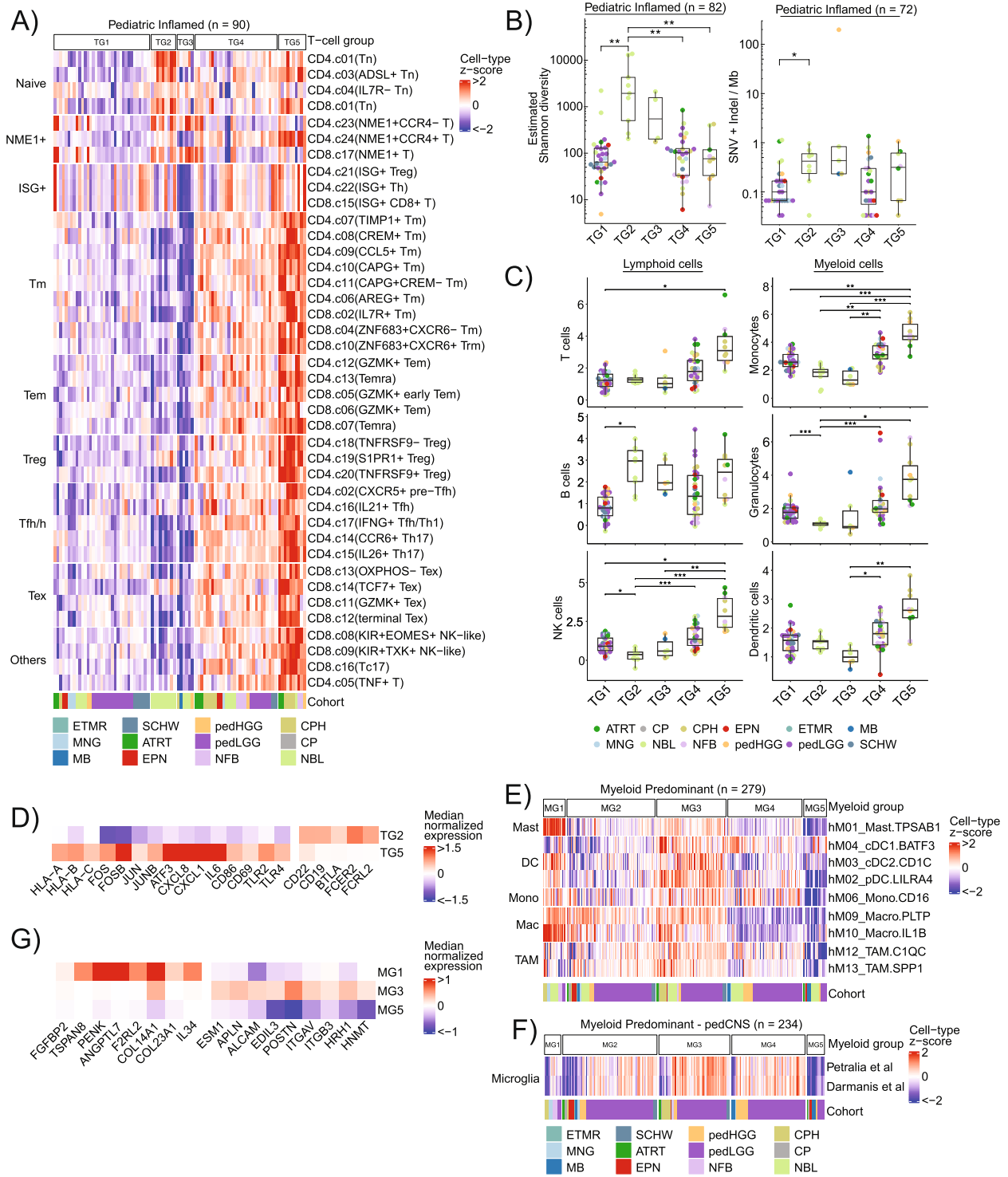


Fig. 6 (See legend on previous page.)

Discussion

We report universal immune microenvironment groups across 12 cancer entities highlighting tumour-agnostic immunological similarities in paediatric primary nervous system tumours. Across our compendium, we found that ~72% of samples harboured generally cold immune microenvironments in higher frequency compared to adult tumour microenvironments [14]. Within our cohort, we found that 10% of samples showed an inflamed microenvironment harbouring high levels of lymphoid and myeloid cell infiltration and diverse T- and B-cell repertoire. In this cohort of non-hypermutant cancers, TMB was not independently associated with an inflamed microenvironment. This finding is in contrast to bMMRD cancers where samples with high CD8+ T-cell infiltration had higher SNV/Mb [114]. Although we could not directly study the associations between TMB and ICI response, our findings do not support TMB as a biomarker for use in non-hypermutant paediatric cancers. The dynamic range of cellular interactions in immune-infiltrated tumours, some of which we suggested in the presented association studies, reflects the existing challenge in defining a robust biomarker of anti-tumour immune response in paediatric tumours.

Analysis of immunoregulatory genes and immune cell subtypes provided insights into the ligand-receptor interactions with translational implications. In Paediatric Inflamed, a number of upregulated genes overlapped with those differentially expressed in post-treatment samples collected from melanoma patients responding to nivolumab [115]. These include immune checkpoint genes (*PDCD1*, *TIGIT*, *CTLA4* and *BTLA*) as well as genes involved in T-cell cytotoxic functions (*IL4I1*, *FASLG*, *TNFRSF9*, *TNFRSF4*, *CD244*, *CD27*, *CD80* and *ICOS*) and suggest samples in this cluster may be good candidates for ICI therapy. The Myeloid Predominant cluster showed elevated expression of two immunoregulatory genes with myeloid-specific functions: *TGFB1* [116, 117] and *SIRPA*, encoding for SIRP α that negatively regulates phagocytosis [118]. Blocking this interaction with anti-CD47 promotes cell killing in preclinical models [9] and a phase I clinical study, the efficacy testing is ongoing (NCT02216409). Our results suggest disrupting cellular interactions involving the lymphoid and myeloid compartments may be beneficial among a subset of pedNST to reshape the tumour microenvironment and elicit anti-tumour immunity.

Although this study provides insights into immunological aspects of paediatric nervous system tumours, there are important limitations to note. We may have missed specific regions with significant immune infiltration due to biased sampling. The IHC analysis on tissue microarrays may not capture tumour heterogeneity accurately.

Immune deconvolution analysis, based on bulk RNA-seq data, can be prone to overestimation due to overlapping gene expression in different immune cell types. Due to this limitation, we could not confirm the cell origin of LAG3 and TIM3 gene expression in the Immune Desert, which may have implications for immune modulation. Higher-resolution techniques like single-cell and spatial RNA and protein analysis can offer a more comprehensive understanding of the tumour microenvironment.

Conclusions

We identified distinct immune microenvironment clusters across paediatric nervous system tumours and proposed microenvironmental mechanisms of immune dysfunction and suppression. With immunotherapy becoming more widely available in the paediatric oncology armamentarium, our findings highlight the value of immunogenomic approaches to guide patient stratification and inform precision oncology programmes.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-023-01219-x>.

Additional file 1. Supplementary figures for the manuscript.

Additional file 2. Supplementary tables for the manuscript.

Acknowledgements

We thank Ben Wang and Andrew Elia of Tumour Immunotherapy Program at Princess Margaret Cancer Centre for performing immunohistochemistry experiments (pm-tumourimmunotherapy.ca). We thank Wei Xu, Jingyue Huang and Osvaldo Espin-Garcia (Dana Lana School of Public Health, University of Toronto) for statistical assistance. The data and materials used for the immune repertoire analysis were made available by Nada Jabado, and the PRrecision Oncology For Young peoPLE (PROFYLE) programme and through funds from the Terry Fox Research Institute and all other funders supporting PROFYLE. We thank Pierre Antoine (Princess Margaret Cancer Centre), Angela Bik-Yu Hui (Stanford Cancer Institute) and Fei-Fei Liu (Princess Margaret Cancer Centre) for providing nasopharyngeal carcinoma samples for immune diversity measurements. We also thank the staff of the Princess Margaret Genomics Centre (www.pmggenomics.ca) and Bioinformatics and High-Performance Computing Core for their expertise in generating the immune repertoire sequencing data used in this study.

Authors' contributions

AN and PB performed data analysis and interpretation. AD and PHS performed TMA staining, scoring and analysis. DOA performed computational TIL analysis. SS, KZ, SYCY, PR and YZ provided bioinformatics support. DTM and JPB performed TCR sequencing and analysis of data from nasopharyngeal carcinoma. JNP provided support for diversity inference. MS provided methylation data and analysis. DM led and managed the PROFYLE programme. ACR led and managed CBTN and CAVATICA. PDJ, SB, SL, DTWJ, MK and SMP provided patient data and contributed to data curation. DTWJ, MK, SMP, NJ and TJP contributed to manuscript design and interpretation of data. AN, PB, NJ and TJP wrote the manuscript with comments and contributions from all authors. All authors read and approved the final manuscript.

Funding

This research was supported by a Terry Fox Research Institute New Investigator Award (TJP), the Gabriella Miller's Kids First Data Resource Center (NIH common funds), the Dragon Masters Foundation and the ITCC-P4 consortium

(Innovative Medicines Initiative2 Joint Undertaking, under grant agreement No. 116064). Additional funding support came from the PedBrain Tumour Project contributing to the International Cancer Genome Consortium (ICGC) funded by German Cancer Aid (109252) and the German Federal Ministry of Education and Research (BMBF, NGFN^{plus} #01GS0883). AN was supported by a Princess Margaret Post-doctoral Fellowship. PB is funded by the DKFZ International PhD Program (Annemarie Poustka fellowship). TJP is supported by the Canada Research Chair in Translational Genomics, a Senior Investigator Award from the Ontario Institute for Cancer Research, and the Princess Margaret Cancer Foundation Gattuso-Slaight Personalized Cancer Medicine Fund. DAO is supported by NHLBI R38 HL143613, NCI T32 CA009140 and the Parker Institute for Cancer Immunotherapy.

Availability of data and materials

CBTN RNA and WGS data, variant and fusion calls are available at the Kids First data repository (<https://www.kidsfirstdrc.org>). Controlled access data can be obtained upon application to Data Access Committee (<https://cbtn.org/pediatric-brain-tumor-atlas>). TCGA and TARGET can be accessed through Genomic Data Commons (GDC) data portal (<https://gdc.cancer.gov/>). Controlled access data can be obtained from the database of Genotypes and Phenotypes (dbGaP) upon application to Data Access Committee (accession numbers: phs000467 and phs000178). Gene expression matrices, clinical metadata, T- and B-clonotype files, HLA types, predicted neopeptides, variant and fusion datasets and deconvolution outputs are available upon free registration at <https://cavatica.sbggenomics.com/u/pughlab/immpedcan> [119]. Custom codes and input data are available at https://github.com/pughlab/immunogenomics_pedNST [120]. Main figures can be reproduced via CodeOcean capsule: <https://doi.org/10.24433/CO.0339991.v1> [121].

Declarations

Ethics approval and consent to participate

The Research Ethics Board at University Health Network (UHN REB) approved this study (CAPCR/REB number 18–5248). All participants provided written informed consent to participate as well as to use and share data for the study. The UHN REB operates in compliance with the Tri-Council Policy Statement, ICH Guideline for Good Clinical Practice E6 (R1), Ontario Personal Health Information Protection Act (2004), Part C Division 5 of the Food and Drug Regulations, and Part 4 of the Natural Health Products Regulations and the Medical Devices Regulations of Health Canada. This research conformed to the Declaration of Helsinki.

Consent for publication

All participant information was anonymized and informed consent for publication is obtained for all datasets.

Competing interests

J.N.P. is an employee of Genentech, Inc. (Stock and Other Ownership Interests: F. Hoffmann-La Roche AG). The remaining authors declare that they have no competing interests.

Author details

¹Princess Margaret Cancer Centre, University Health Network, Princess Margaret Cancer Research Tower, Room 9-305, MaRS Centre, 101 College Street, Toronto M5G 1L7, Canada. ²Hopp Children's Cancer Center Heidelberg (KiTZ), Heidelberg, Germany. ³Division of Pediatric Neurooncology and German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), B062, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. ⁴Department of Molecular Oncology, British Columbia Cancer Agency, Vancouver, Canada. ⁵Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada. ⁶Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁷Center for Computational and Genomic Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁸Division of Hematology/Oncology, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Canada. ⁹Department of Biostatistics, Genentech Inc, San Francisco, CA, USA. ¹⁰Division of Neurosurgery, Center for Childhood Cancer Research, Department of Biomedical and Health Informatics and Center for Data-Driven Discovery in Biomedicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ¹¹Department of Pediatric Hematology and Oncology, Heidelberg University

Hospital, Heidelberg, Germany. ¹²Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands. ¹³Division of Pediatric Glioma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. ¹⁴Department of Medical Biophysics, University of Toronto, Toronto, Canada. ¹⁵Ontario Institute for Cancer Research, Toronto, Canada.

Received: 24 October 2022 Accepted: 7 August 2023

Published online: 07 September 2023

References

- Geoerger B, Zwaan CM, Marshall LV, Michon J, Bourdeaut F, Casanova M, et al. Atezolizumab for children and young adults with previously treated solid tumours, non-Hodgkin lymphoma, and Hodgkin lymphoma (iMATRIX): a multicentre phase 1–2 study. *Lancet Oncol.* 2020;21:134–44.
- Geoerger B, Kang HJ, Yalon-Oren M, Marshall LV, Vezina C, Pappo A, et al. Pembrolizumab in paediatric patients with advanced melanoma or a PD-L1-positive, advanced, relapsed, or refractory solid tumour or lymphoma (KEYNOTE-051): interim analysis of an open-label, single-arm, phase 1–2 trial. *Lancet Oncol.* 2020;21:121–33.
- Davis KL, Fox E, Merchant MS, Reid JM, Kudgus RA, Liu X, et al. Nivolumab in children and young adults with relapsed or refractory solid tumours or lymphoma (ADVL1412): a multicentre, open-label, single-arm, phase 1–2 trial. *Lancet Oncol.* 2020;21:541–50.
- Yu AL, Gilman AL, Ozkaynak MF, London WB, Kreissman SG, Chen HX, et al. Anti-GD2 antibody with GM-CSF, Interleukin-2, and isotretinoin for neuroblastoma. *N Engl J Med.* 2010;363:1324–34.
- Kushner BH, Morgenstern DA, Nysom K, Bear MK, Tornøe K, Losic N, et al. Efficacy of naxitamab in patients with refractory/relapse (R/R) high-risk neuroblastoma (HR-NB) by bone/bone marrow (BM) evaluation, potential sites of residual disease. *JCO.* 2021;39:10022–10022.
- Majzner RG, Ramakrishna S, Yeom KW, Patel S, Chinnasamy H, Schultz LM, et al. GD2-CAR T cell therapy for H3K27M-mutated diffuse midline gliomas. *Nature.* 2022;603(7903):934–41.
- Majzner RG, Theruvath JL, Nellan A, Heitzeneder S, Cui Y, Mount CW, et al. CAR T cells targeting B7-H3, a pan-cancer antigen, demonstrate potent preclinical activity against pediatric solid tumors and brain tumors. *Clin Cancer Res.* 2019;25(8):2560–74.
- Pyonteck SM, Akkari L, Schuhmacher AJ, Bowman RL, Sevenich L, Quail DF, et al. CSF-1R inhibition alters macrophage polarization and blocks glioma progression. *Nat Med.* 2013;19:1264–72.
- Gholamin S, Mitra SS, Feroze AH, Liu J, Kahn SA, Zhang M, et al. Disrupting the CD47-SIRPα anti-phagocytic axis by a humanized anti-CD47 antibody is an efficacious treatment for malignant pediatric brain tumors. *Sci Transl Med.* 2017;9:eaa2968.
- Marcus L, Fashoyin-Aje LA, Donoghue M, Yuan M, Rodriguez L, Gallagher PS, et al. FDA approval summary: pembrolizumab for the treatment of tumor mutational burden-high solid tumors. *Clin Cancer Res.* 2021;27:4685–9.
- Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest.* 2017;127:2930–40.
- Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 2015;160:48–61.
- Bagaev A, Kotlov N, Nomie K, Svekolkina V, Gafurov A, Isaeva O, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell.* 2021;39:845–865.e7.
- Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Yang T-HO, et al. The immune landscape of cancer. *Immunity.* 2018;48:812–830.e14.
- Litchfield K, Reading JL, Puttick C, Thakkar K, Abbosh C, Bentham R, et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell.* 2021;184:596–614.e14.
- Jansky S, Sharma AK, Körber V, Quintero A, Toprak UH, Wecht EM, et al. Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nat Genet.* 2021;53:683–93.
- Filbin M, Monje M. Developmental origins and emerging therapeutic opportunities for childhood cancer. *Nat Med.* 2019;25:367–76.

18. Gojo J, Englinger B, Jiang L, Hübner JM, Shaw ML, Hack OA, et al. Single-Cell RNA-Seq reveals cellular hierarchies and impaired developmental trajectories in pediatric ependymoma. *Cancer Cell*. 2020;38:44-59.e9.
19. Gillen AE, Riemondy KA, Amani V, Griesinger AM, Gilani A, Venkataraman S, et al. Single-Cell RNA sequencing of childhood ependymoma reveals neoplastic cell subpopulations that impact molecular classification and etiology. *Cell Rep*. 2020;32: 108023.
20. Hovestadt V, Smith KS, Bihannic L, Filbin MG, Shaw ML, Baumgartner A, et al. Resolving medulloblastoma cellular architecture by single-cell genomics. *Nature*. 2019;572:74-9.
21. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. *Nature*. 2018;555:321-7.
22. Lilly JV, Rokita JL, Mason JL, Patton T, Stefankiewicz S, Higgins D, et al. The children's brain tumor network (CBTN) - Accelerating research in pediatric central nervous system tumors through collaboration and open science. *Neoplasia*. 2023;35: 100846.
23. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet*. 2013;45:279-84.
24. International network of cancer genome projects. *Nature*. 2010;464:993-8.
25. Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell*. 2018;173:283-5.
26. dbGaP/database of Genotypes and Phenotypes/ National Center for Biotechnology Information, National Library of Medicine (NCBI/NLM). accession phs001436.c1. Available from: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001436.v1.p1. Accessed 2021.
27. Vivian J, Rao A, Nothhaft FA, Ketchum C, Armstrong J, Novak A, et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol*. 2017;35:314-6.
28. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15-21.
29. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
30. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48:W509-14.
31. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28:882-3.
32. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43:491-8.
33. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15:591-4.
34. Uhrig S, Ellermann J, Walthert T, Burkhardt P, Fröhlich M, Hutter B, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res*. 2021;31(3):448-60.
35. German Cancer Research Center (DKFZ). RNAseq processing workflow . DKFZ - ODCF; [cited 2023 Aug 2]. Available from: <https://github.com/DKFZ-ODCF/RNAseqWorkflow>.
36. Kids-First DRC. Kids First RNA-Seq Workflow . Kids First Data Resource Center; [cited 2023 Aug 2]. Available from: <https://github.com/kids-first/kf-rnaseq-workflow>.
37. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell*. 2018;173:321-337.e10.
38. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
39. Stamatoyannopoulos J. ENCSR861QKF . 2013 [cited 2023 Aug 2]. Available from: <https://www.encodeproject.org/experiments/ENCSR861QKF/>.
40. Stamatoyannopoulos J. ENCSR463JBR . 2013 [cited 2023 Aug 2]. Available from: <https://www.encodeproject.org/experiments/ENCSR463JBR/>.
41. Stamatoyannopoulos J. ENCSR449GLL . 2013 [cited 2023 Aug 2]. Available from: <https://www.encodeproject.org/experiments/ENCSR449GLL/>.
42. Stamatoyannopoulos J. ENCSR357XTU . 2013 [cited 2023 Aug 2]. Available from: <https://www.encodeproject.org/experiments/ENCSR357XTU/>.
43. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;10:giab008.
44. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841-2.
45. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525-7.
46. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods*. 2017;14:687-90.
47. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12:453-7.
48. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Valencia A, editor. *eLife*. 2017;6:e26476.
49. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Medicine*. 2019;11:34.
50. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17:218.
51. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res*. 2017;77:e108-10.
52. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18:220.
53. Teschendorff AE, Breeze CE, Zheng SC, Beck S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics*. 2017;18:105.
54. Grabovska Y, Mackay A, O'Hare P, Crosier S, Finetti M, Schwalbe EC, et al. Pediatric pan-central nervous system tumor analysis of immune-cell infiltration identifies correlates of antitumor immunity. *Nat Commun*. 2020;11:4324.
55. Karlsson M, Zhang C, Méar L, Zhong W, Digre A, Katona B, et al. A single-cell type transcriptomics map of human tissues. *Sci Adv*. 2021;7:eabh2169.
56. Ota M, Nagafuchi Y, Hatano H, Ishigaki K, Terao C, Takeshima Y, et al. Dynamic landscape of immune cell-specific gene regulation in immune-mediated diseases. *Cell*. 2021;184:3006-3021.e17.
57. Uhlen M, Karlsson MJ, Zhong W, Tebani A, Pou C, Mikes J, et al. A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science*. 2019;366:eaax9198.
58. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013;14:7.
59. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26:1572-3.
60. Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science*. 374:abe6474.
61. Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell*. 2021;184:792-809.e23.
62. Petralia F, Tignor N, Reva B, Koptyra M, Chowdhury S, Rykunov D, et al. Integrated proteogenomic characterization across major histological types of pediatric brain cancer. *Cell*. 2020;183:1962-1985.e31.
63. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-cell RNA-Seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep*. 2017;21:1399-410.
64. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep*. 2018;23:181-193.e7.

65. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*. 2014;30:3310–6.
66. Bjerregaard A-M, Nielsen M, Hadrup SR, Szallasi Z, Eklund AC. MuPeXI: prediction of neo-epitopes from tumor sequencing data. *Cancer Immunol Immunother*. 2017;66:1123–30.
67. Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med*. 2016;8:33.
68. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12:380–1.
69. Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, et al. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol Monogr*. 2014;84:45–67.
70. Chao A, Wang YT, Jost L. Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods Ecol Evol*. 2013;4:1091–100.
71. Hu X, Zhang J, Wang J, Fu J, Li T, Zheng X, et al. Landscape of B cell immunity and related immune evasion in human cancers. *Nat Genet*. 2019;51:560–7.
72. Mulder DT, Mahé ER, Dowar M, Hanna Y, Li T, Nguyen LT, et al. CapTCR-seq: hybrid capture for T-cell receptor repertoire profiling. *Blood Adv*. 2018;2:3506–14.
73. Brabetz S, Leary SES, Gröbner SN, Nakamoto MW, Şeker-Cin H, Girard EJ, et al. A biobank of patient-derived pediatric brain tumor models. *Nat Med*. 2018;24:1752–61.
74. Hidalgo M, Amant F, Biankin AV, Budinská E, Byrne AT, Caldas C, et al. Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov*. 2014;4:998–1013.
75. Matulonis UA, Shapira-Frommer R, Santin AD, Lisyanskaya AS, Pignata S, Vergote I, et al. Antitumor activity and safety of pembrolizumab in patients with advanced recurrent ovarian cancer: results from the phase II KEYNOTE-100 study. *Ann Oncol*. 2019;30:1080–7.
76. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010;363:711–23.
77. Hellmann MD, Ciuleanu T-E, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med*. 2018;378:2093–104.
78. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;372:2509–20.
79. Sharma P, Pachynski RK, Narayan V, Fléchon A, Gravis G, Galsky MD, et al. Nivolumab plus ipilimumab for metastatic castration-resistant prostate cancer: preliminary analysis of patients in the CheckMate 650 Trial. *Cancer Cell*. 2020;38:489–499.e3.
80. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell*. 2018;173:371–385.e18.
81. Griesinger AM, Birks DK, Donson AM, Amani V, Hoffman LM, Waziri A, et al. Characterization of distinct immunophenotypes across pediatric brain tumor types. *J Immunol*. 2013;191:4880–8.
82. Bockmayr M, Mohme M, Klauschen F, Winkler B, Budczies J, Rutkowski S, et al. Subgroup-specific immune and stromal microenvironment in medulloblastoma. *Oncoimmunology*. 2018;7: e1462430.
83. Northcott PA, Robinson GW, Kratz CP, Mabbott DJ, Pomeroy SL, Clifford SC, et al. Medulloblastoma. *Nat Rev Dis Primers*. 2019;5:11.
84. Chun H-JE, Johann PD, Milne K, Zapatka M, Buellesbach A, Ishaque N, et al. Identification and analyses of extra-cranial and cranial rhabdoid tumor molecular subgroups reveal tumors with cytotoxic T cell infiltration. *Cell Rep*. 2019;29:2338–2354.e7.
85. Leruste A, Tosello J, Ramos RN, Tauziède-Espariat A, Brohard S, Han Z-Y, et al. Clonally expanded T cells reveal immunogenicity of rhabdoid tumors. *Cancer Cell*. 2019;36:597–612.e8.
86. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347:1260419.
87. Kleinheinz K, Bludau I, Hübschmann D, Heinold M, Kensche P, Gu Z, et al. ACESeq – allele specific copy number estimation from whole genome sequencing. *bioRxiv*; 2017 [cited 2022 Mar 4]. p. 210807. Available from: <https://www.biorxiv.org/content/10.1101/210807v1>.
88. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018;555:469–74.
89. Gajjar A, Pfister SM, Taylor MD, Gilbertson RJ. Molecular insights into pediatric brain tumors have the potential to transform therapy. *Clin Cancer Res*. 2014;20:5630–40.
90. Wei JS, Kuznetsov IB, Zhang S, Song YK, Asgharzadeh S, Sindiri S, et al. Clinically relevant cytotoxic immune cell signatures and clonal expansion of T-cell receptors in high-risk MYCN-not-amplified human neuroblastoma. *Clin Cancer Res*. 2018;24:5673–84.
91. Layer JP, Kronmüller MT, Quast T, van den Boorn-Konijnenberg D, Efferm M, Hinze D, et al. Amplification of N-Myc is associated with a T-cell-poor microenvironment in metastatic neuroblastoma restraining interferon pathway activity and chemokine expression. *Oncoimmunology*. 2017;6: e1320626.
92. Yaari G, Bolen CR, Thakar J, Kleinstein SH. Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. *Nucleic Acids Res*. 2013;41: e170.
93. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1:417–25.
94. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
95. Boot RG, Renkema GH, Strijland A, van Zonneveld AJ, Aerts JM. Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. *J Biol Chem*. 1995;270:26252–6.
96. Samstein RM, Krishna C, Ma X, Pei X, Lee K-W, Makarov V, et al. Mutations in BRCA1 and BRCA2 differentially affect the tumor microenvironment and response to checkpoint blockade immunotherapy. *Nat Cancer*. 2020;1:1188–203.
97. Farhad M, Rolig AS, Redmond WL. The role of Galectin-3 in modulating tumor growth and immunosuppression within the tumor microenvironment. *Oncoimmunology*. 2018;7: e1434467.
98. Marquet J, Lasoudris F, Cousin C, Puisse M-L, Martin-Garcia N, Baud V, et al. Dichotomy between factors inducing the immunosuppressive enzyme IL-4-induced gene 1 (IL4I1) in B lymphocytes and mononuclear phagocytes. *Eur J Immunol*. 2010;40:2557–68.
99. Yoshimura T, Robinson EA, Tanaka S, Appella E, Kuratsu J, Leonard EJ. Purification and amino acid analysis of two human glioma-derived monocyte chemoattractants. *J Exp Med*. 1989;169:1449–59.
100. Im JH, Buzzelli JN, Jones K, Franchini F, Gordon-Weeks A, Markelc B, et al. FGF2 alters macrophage polarization, tumour immunity and growth and can be targeted during radiotherapy. *Nat Commun*. 2020;11:4064.
101. Jones DTW, Banito A, Grünewald TGP, Haber M, Jäger N, Kool M, et al. Molecular characteristics and therapeutic vulnerabilities across paediatric solid tumours. *Nat Rev Cancer*. 2019;19:420–38.
102. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. *Blood*. 2008;112:1570–80.
103. Bashford-Rogers RJM, Palsler AL, Huntly BJ, Rance R, Vassiliou GS, Follows GA, et al. Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. *Genome Res*. 2013;23:1874–84.
104. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell*. 2020;181:442–459.e29.
105. Heo K, Lee S. TSPAN8 as a novel emerging therapeutic target in cancer for monoclonal antibody therapy. *Biomolecules*. 2020;10:388.
106. Ovidia H, Magenheimer Y, Behar O, Rosen H. Molecular characterization of immune derived proenkephalin mRNA and the involvement of the adrenergic system in its expression in rat lymphoid cells. *J Neuroimmunol*. 1996;68:77–83.
107. Carbone C, Piro G, Merz V, Simionato F, Santoro R, Zecchetto C, et al. Angiopoietin-like proteins in angiogenesis, inflammation and cancer. *Int J Mol Sci*. 2018;19:431.
108. Galmiche A, Rak J, Roumenina LT, Saidak Z. Coagulome and the tumor microenvironment: an actionable interplay. *Trends in Cancer*. 2022;8:369–83.
109. Muñoz-García J, Cochoneau D, Télétchéa S, Moranton E, Lanoe D, Brion R, et al. The twin cytokines interleukin-34 and CSF-1:

- masterful conductors of macrophage homeostasis. *Theranostics*. 2021;11:1568–93.
110. Boulakirba S, Pfeifer A, Mhaidly R, Obba S, Goulard M, Schmitt T, et al. IL-34 and CSF-1 display an equivalent macrophage differentiation ability but a different polarization potential. *Sci Rep*. 2018;8:256.
 111. Park SY, Piao Y, Jeong KJ, Dong J, de Groot JF. Periostin (POSTN) regulates tumor resistance to antiangiogenic therapy in glioma models. *Mol Cancer Ther*. 2016;15:2187–97.
 112. Zhou W, Ke SQ, Huang Z, Flavahan W, Fang X, Paul J, et al. Periostin secreted by glioblastoma stem cells recruits M2 tumor-associated macrophages and promotes malignant growth. *Nat Cell Biol*. 2015;17:170–82.
 113. Li H, Xiao Y, Li Q, Yao J, Yuan X, Zhang Y, et al. The allergy mediator histamine confers resistance to immunotherapy in cancer patients via activation of the macrophage histamine receptor H1. *Cancer Cell*. 2022;40:36–52.e9.
 114. Das A, Sudhaman S, Morgenstern D, Coblenz A, Chung J, Stone SC, et al. Genomic predictors of response to PD-1 inhibition in children with germline DNA replication repair deficiency. *Nat Med*. 2022;28:125–35.
 115. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell*. 2017;171:934–949.e15.
 116. Ye X, Xu S, Xin Y, Yu S, Ping Y, Chen L, et al. Tumor-associated microglia/macrophages enhance the invasion of glioma stem-like cells via TGF- β 1 signaling pathway. *J Immunol*. 2012;189:444–53.
 117. Battle E, Massagué J. Transforming growth factor- β signaling in immunity and cancer. *Immunity*. 2019;50:924–40.
 118. Barclay AN, Van den Berg TK. The interaction between signal regulatory protein alpha (SIRP α) and CD47: structure, function, and therapeutic target. *Annu Rev Immunol*. 2014;32:25–50.
 119. Nabbi A, et al. Immunogenomics of Pediatric Cancers. CAVATICA; Available from: <https://cavatica.sbggenomics.com/u/pughlab/impedcan>. Accessed 2021.
 120. Nabbi A, et al. Immunogenomic analysis of pediatric nervous system tumours. Github; Available from: https://github.com/pughlab/immunogenomics_pedNST. Accessed 2021.
 121. Nabbi A, et al. Immunogenomic analysis of paediatric nervous system tumors. CodeOcean; 2022. Available from: <https://doi.org/10.24433/CO.0339991.v1>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

