



HAL
open science

Integrative bioinformatics analysis of intra-tumoral heterogeneity in neuroblastoma at single cell level

Jaydutt Bhalshankar

► **To cite this version:**

Jaydutt Bhalshankar. Integrative bioinformatics analysis of intra-tumoral heterogeneity in neuroblastoma at single cell level. Genomics [q-bio.GN]. Université Paris sciences et lettres, 2024. English. NNT : 2024UPSL002 . tel-04924811

HAL Id: tel-04924811

<https://theses.hal.science/tel-04924811v1>

Submitted on 1 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à U830 INSERM, Institut Curie

**INTEGRATIVE BIOINFORMATICS ANALYSIS OF INTRA-TUMORAL
HETEROGENEITY IN NEUROBLASTOMA AT SINGLE CELL LEVEL**

**ANALYSE BIOINFORMATIQUE INTÉGRATIVE DE L'HÉTÉROGÉNÉITÉ
INTRA-TUMORALE DANS LE NEUROBLASTOME AU NIVEAU DE LA
CELLULE UNIQUE**

Soutenue par

Jaydutt D. Bhalshankar

Le 22 01 2024

Ecole doctorale n° 582

**Cancérologie : biologie-
médecine-santé (CBMS)**

Spécialité

Sciences du Cancer

Préparée au

**Equipe Recherche
Translationnelle en Oncologie
Pédiatrique**

Composition du jury :

Jessica Zucman-Rossi

Professeur, Université Paris Cité
INSERM UMR 1162, Paris (France)

Président du jury

François Radvanyi

Professeur, Institut Curie
CNRS UMR 144, Paris (France)

Examineur

Jan Koster

Professeur, Amsterdam UMC
Amsterdam (Pays-Bas)

Examineur

Benjamin Ory

Professeur, Nantes Université
CRCI²NA, Team 9, Nantes (France)

Rapporteur

Frank Westermann

Professeur, German Cancer Research
Center, DKFZ, Heidelberg (Allemagne)

Rapporteur

Gudrun Schleiermacher

Docteure, Institut Curie, Paris (France)

Directrice de thèse

Andrei Zinovyev

Principal Scientist
Evotek R&D, Toulouse (France)

Co-directeur de thèse

Table of Contents

Acknowledgements.....	VI
List of Abbreviations.....	VIII
List of Tables.....	X
List of Figures	XI
1 Introduction.....	1
1.1 Neuroblastoma	1
1.1.1 Clinical Features and Risk Stratification	2
1.1.2 Pathogenesis and molecular alterations	3
1.1.2.1 Familial neuroblastomas	3
1.1.2.2 Recurrent focal amplification	4
1.1.2.3 Chromosomal copy number alterations	5
1.1.2.4 <i>ALK</i> alterations	5
1.1.2.5 RAS-MAPK and p53 signaling pathway alterations	6
1.1.2.6 Activation of telomere maintenance mechanisms	7
1.2 Intra-tumoral heterogeneity (ITH)	8
1.2.1 Spatial Intra-Tumor Heterogeneity	8
1.2.1.1 Variation in <i>MYCN</i> Amplification	8
1.2.1.2 Presence of Genetic Mutations and Chromosomal Alterations	8
1.2.2 Temporal Intra Tumoral Heterogeneity.....	9
1.2.2.1 Evolution of Genetic Mutations Post-Diagnosis	9
1.2.2.2 Development of Chemotherapy Resistance:.....	10
1.2.2.3 Shifts in Tumor Cell Populations Over Time:	10
1.3 Models of genetic tumor evolution	11
1.3.1 Linear Evolution Model.....	12
1.3.2 Branching Evolution Model.....	13

1.3.3	Macroevolution or Punctuated Evolution Model.....	13
1.3.4	Neutral Evolution Model	14
1.4	Molecular mechanisms involved in genetic evolution: DNA replication	15
1.4.1	Cell cycle and DNA replication timing (RT)	15
1.4.2	Cancer cell and DNA replication	17
1.4.3	Computational tools to study replication in single cell.....	18
1.5	Quantifying genetic changes in cancer evolution	19
1.5.1	Leveraging cohort-level information to study cancer evolution.....	19
1.6	Methods for evaluating genetic tumor evolution	21
1.6.1	The role of Single cell sequencing.....	24
1.6.1.1	First-generation single cell DNA sequencing	24
1.6.1.2	Single Cell Copy-Number Profiling of tumors.....	25
1.6.1.3	High-throughput single-cell DNA Sequencing methods	26
1.6.1.4	Copy Number Inference from high-throughput scDNA-seq	26
1.6.2	Interpretation complexities of tumor evolutionary models.....	27
1.7	Integrative scDNA-seq and scRNA-seq data analysis deciphering complex relationships between genotype and phenotype.....	27
2	Thesis objectives	30
3	Results.....	33
3.1	Part A. Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing suggest clone specific therapeutic vulnerabilities in neuroblastoma.....	33
3.1.1	Manuscript: Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing suggest clone specific therapeutic vulnerabilities in neuroblastoma.	35
3.1.1.1	Abstract.....	36
3.1.1.2	Introduction.....	37
3.1.1.3	Material and Methods	39
3.1.1.3.1	Preparation of the single-cell suspension for DNA sequencing.....	39
3.1.1.3.2	Single-cell DNA library generation	39

3.1.1.3.3	Single cell RNA sequencing library generation.....	40
3.1.1.3.4	In-vivo experiments for the study of clonal evolution.....	40
3.1.1.3.5	Tumor DNA sampling, processing, and sequencing.....	40
3.1.1.3.6	Single-cell DNA sequence data possessing	41
3.1.1.3.7	CNV calling at single cell level and inference of genomic architecture.....	42
3.1.1.3.8	Integer copy number and consensus profiles	42
3.1.1.3.9	Phylogenetic reconstruction of clonal lineages.....	43
3.1.1.3.10	Inference of most recent common ancestor (MRCA) profile	43
3.1.1.3.11	Bulk whole exome and pseudobulk WGS data analysis.....	43
3.1.1.3.12	Mutation timing or Mutation multiplicity analysis.....	43
3.1.1.3.13	Replication timing analysis using S-phase cell population identified from scWGS	44
3.1.1.3.14	Single-Cell RNA Sequencing: Genotyping and Copy Number Variation Analysis	44
3.1.1.3.15	Bulk whole exome sequencing (WES) data analysis to study subclonal dynamics under targeted treatment	44
3.1.1.3.16	Inference of pre and post treatment subclonal dynamics.....	45
3.1.1.3.17	Gene set overlapping analysis.....	45
3.1.1.3.18	Mutational Signature analysis.....	45
3.1.1.4	Results.....	46
3.1.1.4.1	Study overview	46
3.1.1.4.2	Genetic alterations defined at bulk level and summery of single cell DNA sequencing data analysis	49
3.1.1.4.3	Genomic Copy Number Architecture in Neuroblastoma Tumors.....	52
3.1.1.4.4	Longitudinal single cell sequencing shows distinct evolutionary patterns of SCNA from diagnosis to progression.....	57
3.1.1.4.5	Replication timing study using single cell DNA S-phase cell population in neuroblastoma.	60

3.1.1.4.5.1	Dissecting replication timing variations between <i>MYCN</i> -amplified and non-amplified neuroblastomas	60
3.1.1.4.5.2	Clone specific replication timing (RT) and its consequences towards ITH	61
3.1.1.4.6	Parallel clonal evolution and subclone specific expression	65
3.1.1.4.7	In Vivo Investigations of Clonal Evolution During Treatment.....	68
3.1.1.4.7.1	Analysis of treatment efficacy	68
3.1.1.4.7.2	Subclonal evolutionary dynamic under treatment	69
3.1.1.5	Discussion.....	73
3.1.1.6	Supplementary data	79
3.1.1.6.1	In vivo experiments for the study of clonal evolution under treatment	145
3.1.1.6.1.1	Ethical considerations of in vivo experimentation.....	145
3.1.1.6.1.2	Treatment schedules	145
3.1.1.6.1.3	In vivo treatment efficacy	146
3.1.1.6.1.4	Tumor Growth Inhibition (TGI) determination	146
3.1.1.6.1.5	Overall response rate (ORR) determination.....	147
3.1.1.6.1.6	Analysis of treatment efficacy	147
3.2	Part B: Further analysis of genetic heterogeneity in neuroblastoma	149
3.2.1	Paper I: Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma	149
3.2.1.1	Introduction.....	149
3.2.1.2	Materials and Methods	149
3.2.1.3	Results.....	149
3.2.1.4	Conclusion	150
3.2.2	Paper II: Impact of <i>ALK</i> Amplifications and Mutations in the European Neuroblastoma Study Group (SIOPEN) High-Risk Neuroblastoma Trial (HR-NBL1).....	153
3.2.2.1	Introduction.....	153
3.2.2.2	Materials and Methods	153
3.2.2.3	Results.....	153

3.2.2.4	Conclusion	154
4	Discussion and perspectives.....	155
5	Conclusion	158
	References.....	160
	Appendices.....	179

ACKNOWLEDGEMENTS

I express my gratitude to my co-supervisor, Andrei Zinovyev, whose valuable discussions and insights have significantly contributed to my research and academic development. I wish to express my sincere gratitude to the jury members, Prof. Jessica Zucman Rossi, Prof. François Radvanyi, Prof. Jan Koster, Prof. Benjamin Ory and Prof. Frank Westermann for their valuable time and thoroughly reviewing my work.

I want to extend my deepest appreciation to my team, particularly Angela Bellini for her crucial role in our single cell DNA study design and sequencing, and Irene Jiménez, whose dedication to the PDX treatment study and sequencing was invaluable. Their partnership was essential to our project's success.

Thanks to Yasmine Iddir for her consistent bioinformatics support and assistance with various administrative tasks. I'm also grateful for the insightful bioinformatics discussions with Mamy Andrianteranagna, especially as I navigated the complexities of scRNA sequencing data analysis. A big thank you to all RTOP team members for their support, friendship, and shared joys during my PhD journey.

I am sincerely thankful to Isabelle Janoueix-Lerosey and her team for sharing their scRNA-seq data and for our enriching scientific interactions. Their contributions significantly enhanced my research.

My gratitude also extends to the Curie and Doctoral school administrative staff, whose help was indispensable throughout my PhD.

I am particularly grateful to Olivier Delattre for his continual encouragement and support, not just in my scientific pursuits but also in navigating administrative challenges like VISA issues.

A special note of thanks to Dr. Gudrun Schleiermacher for accepting me into her team, offering the chance to pursue a PhD, and providing incredible mentorship and guidance. Her support, especially during the difficult times of 2023, was invaluable.

Finally, to my family, especially my parents: your constant support and unconditional love, despite the distance, have been the cornerstone of my journey. Your belief in me and my goals has been my constant motivation, making the challenges worthwhile.

LIST OF ABBREVIATIONS

ADRN	Adrenergic cells
ALT	Alternative mechanism of telomerase
ASCN	Allele-specific copy number alteration
BAF	b-allele frequency
CCF	Cancer cell fraction
CN	Copy number
CNA	Copy number alteration
CN-LoH	Copy-neutral loss of heterogeneity
CNV	Copy number variation
DMR	DNA mis-match repair
EMG	Epigenetic modifier genes
gDNA	Genomic DNA
HR-NB	High-risk neuroblastoma
HSCN	Haplotype-specific copy number alteration
ITH	Intra-tumoral heterogeneity
LoF	Loss of function
LoH	Loss of heterogeneity
MAF	Mutation allele fraction
MES	Mesenchymal cells
MRCA	Most recent common ancestor
NB	Neuroblastoma
PDX	Patient derived xenograft.
RT	Replication timing
RTV	Relative tumor volume
SCA	Segmental chromosomal alterations
scDNA-seq	Single cell DNA sequencing
SCNA	Somatic copy number alteration
scRNA-seq	Single cell RNA sequencing
scRT	Single cell replication timing
scWGS	Single cell whole genome sequencing

SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
TGI	Tumor growth inhibition
TMB	Tumor mutation burden
UMAP	Uniform Manifold Approximation and Projection
VAF	Variant allele fraction
WES	Whole exome sequencing
WGD	Whole genome duplication
WGS	Whole genome sequencing

LIST OF TABLES

Table 3.1.1.4.1 Clinical information of the 18 patients.....	47
Table 3.1.1.6.1 Single cell DNA sequencing run summary details.....	102
Table 3.1.1.6.2 Single-Cell Phasing and Sample Selection for Replication Timing Analysis in Neuroblastoma Study.....	103
Table 3.1.1.6.3. In Vivo Investigation of Clonal Evolution in High-Risk Neuroblastoma Using Patient-Derived Xenograft (PDX) Models and Targeted Therapies.	103
Table 3.1.1.6.4. Tumor Growth Inhibition Efficacy Across Various PDX Models Under Distinct Treatment Regimens.	148

LIST OF FIGURES

Figure 1.1.2.5.1: Recurrent somatic alterations by age group in neuroblastoma.....	7
Figure 1.2.2.3.1: Models of tumor evolution	12
Figure 1.2.2.3.1: The basic principles of genome-wide single-cell DNA replication profiling.....	18
Figure 3.1.1.4.1 Comprehensive analysis of neuroblastoma (NB) genetic heterogeneity using single cell whole genome sequencing (scWGS).....	49
Figure 3.1.1.4.2: Summary of genetic alterations	51
Figure 3.1.1.4.3 Clonal substructure and evolutionary analysis of clonal lineages in NB patient tumors and PDX models	57
Figure 3.1.1.4.4 Investigating temporal heterogeneity by comparing genetic alterations at diagnosis and progression.	60
Figure 3.1.1.4.5 Single cell DNA replication timing (RT) profiling.....	64
Figure 3.1.1.4.6 GR-NB4 PDX model showing parallel clonal evolution.	67
Figure 3.1.1.4.7 Clonal Evolution During Treatment	72
Figure 3.1.1.6.1 Monoclonal genomes with minimal subclonal diversity.	79
Figure 3.1.1.6.2 Clonal substructure of neuroblastoma Patient tumor (P2) and 7 PDXs samples.	80
Figure 3.1.1.6.3: Clonal substructure and evolutionary analysis of clonal lineages in NB patient tumors P3.	81
Figure 3.1.1.6.4: Allele- and haplotype-specific copy numbers for 865 cells in GR-NB4 PDX model.....	82
Figure 3.1.1.6.5: Allele- and haplotype-specific copy numbers for 359 cells in GR-NB4 PDX model.....	83
Figure 3.1.1.6.6: Allele- and haplotype-specific copy numbers for 84 cells in GR-NB5 PDX model.....	84
Figure 3.1.1.6.7: Allele- and haplotype-specific copy numbers for 257 cells in GR-NB10 PDX model.....	85
Figure 3.1.1.6.8: Allele- and haplotype-specific copy numbers for 326 cells in HSJD-NB-003 PDX model.....	86
Figure 3.1.1.6.9: Allele- and haplotype-specific copy numbers for 323 cells in HSJD-NB-009 PDX model.....	87
Figure 3.1.1.6.10: Allele- and haplotype-specific copy numbers for 170 cells in HSJD-NB-011 PDX model.....	88

Figure 3.1.1.6.11: Allele- and haplotype-specific copy numbers for 523 cells in IC-pPDX-17 PDX model.....	89
Figure 3.1.1.6.12: Allele- and haplotype-specific copy numbers for 255 cells in IC-pPDX-75 PDX model.....	90
Figure 3.1.1.6.13: Allele- and haplotype-specific copy numbers for 962 cells in IC-pPDX-109 PDX model.....	91
Figure 3.1.1.6.14: Allele- and haplotype-specific copy numbers for 884 cells in IC-pPDX-112 PDX model.....	92
Figure 3.1.1.6.15: Allele- and haplotype-specific copy numbers for 359 cells in NB1572 patient tumor.	93
Figure 3.1.1.6.16: Allele- and haplotype-specific copy numbers for 458 cells in NB1576 patient tumor.	94
Figure 3.1.1.6.17: Allele- and haplotype-specific copy numbers for 663 cells in NB1583 patient tumor.	95
Figure 3.1.1.6.18 Exome mutation counts according to clonality.....	96
Figure 3.1.1.6.19: Analysis of Loss of Heterozygosity (LoH) regions.....	96
Figure 3.1.1.6.20 GR-NB4, phylogeny tree with seven clones.	97
Figure 3.1.1.6.21 Single cell ASCN profiles from GR-NB4 PDX model representing clone 6 and 7 with and without chromosomes 2 and 12 gain.....	98
Figure 3.1.1.6.22 Bioinformatics data analysis workflow.	99
Figure 3.1.1.6.23 Comparative Analysis of bin-to-bin CN ratio variance Between G1/G2 and S-Phase Cell Populations in Neuroblastoma.	100
Figure 3.1.1.6.24 Comparative Analysis of Copy Number Breakpoints Between G1/G2 and S-Phase Cell Populations in Neuroblastoma.	101
Figure 3.1.1.6.25. GR-NB4, a <i>ALK</i> -amplified NB PDX model Relative tumor volume (mean).	104
Figure 3.1.1.6.26: GR-NB4, a <i>ALK</i> amplified NB PDX model The impact of targeted treatment (lorlatinib) alone on different clonal populations within a tumor.....	104
Figure 3.1.1.6.27: GR-NB4, a <i>ALK</i> amplified NB PDX model The impact of targeted treatment (lorlatinib) in combination with chemotherapy-A.	105
Figure 3.1.1.6.28. GR-NB4, a <i>ALK</i> amplified NB PDX model The impact of chemotherapy-A treatment.....	106
Figure 3.1.1.6.29: GR-NB4, a <i>ALK</i> amplified NB PDX model The impact of chemotherapy-B treatment.....	107

Figure 3.1.1.6.30: GR-NB4 model: SNVs from s1 and s2 subclone mapped to WES derived clones.	108
Figure 3.1.1.6.31. IC-pPDX-75, a <i>ALK</i> -mutated NB PDX model, growth curves.....	111
Figure 3.1.1.6.32. IC-pPDX-75, <i>ALK</i> mutated NB PDX model The impact of targeted treatment (lorlatinib) alone.....	112
Figure 3.1.1.6.33. IC-pPDX-75, <i>ALK</i> mutated NB PDX model The impact of targeted treatment (lorlatinib) in combination with chemotherapy-A	113
Figure 3.1.1.6.34. IC-pPDX-75, <i>ALK</i> mutated NB PDX model The impact of targeted treatment (lorlatinib) in combination with chemotherapy-B	114
Figure 3.1.1.6.35: IC-pPDX-75, <i>ALK</i> mutated NB PDX model The impact of chemotherapy-A treatment alone.....	115
Figure 3.1.1.6.36: IC-pPDX-75, <i>ALK</i> mutated NB PDX model The impact of chemotherapy-B treatment alone.....	116
Figure 3.1.1.6.37. IC-pPDX-109, a <i>HRAS</i> -mutated NB PDX model, growth curves	118
Figure 3.1.1.6.38: IC-pPDX-109, <i>HRAS</i> mutant NB PDX model The impact of Targeted treatment (trametinib) alone.....	119
Figure 3.1.1.6.39: IC-pPDX-109, <i>HRAS</i> mutant NB PDX model The impact of Targeted treatment (trametinib) in combination with chemotherapy-A	120
Figure 3.1.1.6.40. IC-pPDX-109, <i>HRAS</i> mutant NB PDX model The impact of Targeted treatment (trametinib) in combination with chemotherapy-B treatment.	121
Figure 3.1.1.6.41. IC-pPDX-109, <i>HRAS</i> mutant NB PDX model The impact of chemotherapy-A treatment alone.....	122
Figure 3.1.1.6.42. IC-pPDX-109, <i>HRAS</i> mutant NB PDX model The impact of chemotherapy-B treatment alone.....	123
Figure 3.1.1.6.43. GR-NB10, a <i>NF1</i> -mutated NB PDX model, growth curves	125
Figure 3.1.1.6.44. GR-NB10, a <i>NF1</i> -mutated NB PDX model The effect of targeted treatment (trametinib).....	126
Figure 3.1.1.6.45. GR-NB10, a <i>NF1</i> -mutated NB PDX model The impact of targeted treatment (trametinib) in combination with chemotherapy-A	127
Figure 3.1.1.6.46. GR-NB10, a <i>NF1</i> -mutated NB PDX model The impact of targeted treatment (trametinib) in combination with chemotherapy-B.....	128
Figure 3.1.1.6.47. GR-NB10, a <i>NF1</i> -mutated NB PDX model The impact of chemotherapy-A treatment alone.....	129

Figure 3.1.1.6.48. GR-NB10, a NF1-mutated NB PDX model The impact of chemotherapy-B treatment alone.....	130
Figure 3.1.1.6.49. IC-pPDX-17, a CDK4-amplified model, growth curves.....	132
Figure 3.1.1.6.50. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) alone.....	132
Figure 3.1.1.6.51. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) in combination with chemotherapy-A	133
Figure 3.1.1.6.52. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) in combination with chemotherapy-B	134
Figure 3.1.1.6.53. IC-pPDX-17, CDK4 amplified model. Impact of chemotherapy-A treatment	135
Figure 3.1.1.6.54. IC-pPDX-17, CDK4 amplified model. Impact of chemotherapy-B treatment	136
Figure 3.1.1.6.55. HSJD-NB-05, a ARID1A-mutated model, growth curves.	138
Figure 3.1.1.6.56. HSJD-NB-005, ARID1A mutant model Impact of targeted treatment (Tazemetostat) alone	139
Figure 3.1.1.6.57. HSJD-NB-005, ARID1A mutant model Impact of targeted treatment (Tazemetostat) in combination with chemotherapy-A.....	140
Figure 3.1.1.6.58. HSJD-NB-005, ARID1A mutant model Impact of targeted treatment (Tazemetostat) in combination with chemotherapy-B	141
Figure 3.1.1.6.59. HSJD-NB-005, ARID1A mutant model Impact of chemotherapy-A treatment	142
Figure 3.1.1.6.60. HSJD-NB-005, ARID1A mutant model Impact of chemotherapy-B treatment	143
Figure 3.2.1.4.1: Independent Origin and Convergent Parallel Evolution of Tumors in Patient 1 with SMARCA4 mutation.	151
Figure 3.2.2.3.1	154
Figure 3.2.2.3.2: OS according to the type of ALK alteration in the cohort of 571 patients with known ALK amplification and ALK mutation status	154

1 INTRODUCTION

As cancer progresses, it undergoes continuous evolution, leading to increasing genetic heterogeneity (Hanahan 2022). This heterogeneity manifests in the bulk tumor as a diverse set of cells, each with unique molecular characteristics and varying sensitivities to treatment. This variability manifests as spatial heterogeneity, where genetically distinct tumor-cell subpopulations are unevenly distributed across different disease sites, and temporal heterogeneity, characterized by changes in the molecular composition of cancer cells over time (McGranahan and Swanton 2017). Heterogeneity is a key driver of resistance to treatment, making its accurate assessment crucial for developing effective therapies. To dissect the complex clonal architecture of cancers, recent approaches aim at integrating more numerous or sequential tumor samples such as multiregional sequencing (Gerlinger et al. 2012), autopsy sample analysis, and longitudinal liquid biopsy analysis (Peneder et al. 2021); furthermore, emerging technologies including single cell sequencing approaches offer significant promise (Navin et al. 2011). These advanced methods provide deeper insights into the dynamic and multifaceted nature of cancer cells in a given tumor, paving the way for more targeted and effective treatment strategies. This thesis focusses on the study of genetic heterogeneity in neuroblastoma.

1.1 Neuroblastoma

Neuroblastoma, the most prevalent extracranial solid tumor in children, manifests in 90% of patients before age 10 and is typically diagnosed at a median age of 18 months (Matthay et al. 2016). Neuroblastomas primarily develop from embryonal (developing cells) neural crest cells in the sympathetic nervous system, arising in the adrenal medulla or within the sympathetic ganglia, with most tumors occurring within the abdomen (> 75%). Metastasis occurs by lymphatic or hematological spread to either locoregional or distant lymph nodes, or bone, bone marrow or less frequently involved sites such the liver, lung or the CNS.

Accordingly, symptoms at diagnosis may differ according to the site of the primary tumor, and metastatic involvement. Upon diagnosis, clinical and radiological investigations (TDM, MRI and MIBG scan) are required to define local and metastatic extent of disease, and a tumor biopsy is realised whenever possible for pathological and biological analyses (Maris et al. 2007).

1.1.1 Clinical Features and Risk Stratification

Neuroblastoma is characterized by a heterogeneous clinical course that ranges from spontaneous regression to aggressive progression. This heterogeneity presents a diagnostic and therapeutic challenge due to its broad spectrum of clinical behavior. Its management relies heavily on risk categorization, taking into account the main clinical and biological prognostic factors, the latter of which can be impacted by the observation of tumor heterogeneity.

The contemporary staging system, the International Neuroblastoma Risk Group Staging System (INRGSS), stratifies patients based on image-defined risk factors (IDRFs) into stages L1, L2, M, and MS. L1 indicates Localized tumor not involving vital structures as defined by the list of image-defined risk factors and confined to one body compartment, L2 indicates Locoregional tumor with presence of one or more image defined risk factors, M indicates Distant metastatic disease (except stage MS), and MS, in which metastases are confined to the skin, liver, and/or bone marrow in children younger than 18 months of age (Monclair et al. 2009). Depending on age, stage, and the copy number status of the *MYCN* oncogene (encoding the transcription factor N-MYC), neuroblastoma is categorized into three risk groups - low, intermediate, and high, each associated with distinct survival outcomes. Low-risk patients typically exhibit L1 staging and have over a 95% 5-year survival rate, often requiring minimal treatment. Intermediate-risk patients display either L1 or L2 with certain unfavorable biologic features or are infants with stage M without *MYCN* amplification; their 5-year survival rate is approximately 90-95%, with treatment balanced between efficacy and long-term effects. In High-risk patients, usually over 12 or 18 months with stage M or with *MYCN* amplification, 5-year survival have increased from <20% to >50% over the past few decades owing to progressive improvements in the efficacy of multimodal therapies evaluated in large scale clinical trials. In Europe, these trials are developed within the SIOPEN network (<https://www.siopen.net/>)

Pathologically neuroblastoma characterized as a small round blue cell tumor, a classification that includes a group of aggressive cancers with similar histological appearance. These neoplasms are primarily composed of small, densely packed cells with hyperchromatic nuclei and scant cytoplasm, giving them a "blue" appearance on hematoxylin and eosin (H&E) staining. The morphology of neuroblastoma cells is indicative of their origin from primitive sympathetic nervous system cells. A critical aspect of neuroblastoma pathology is the expression of *PHOX2B*, a transcription factor crucial for the development of the autonomic nervous system. Immunohistochemical staining for *PHOX2B* serves as a valuable diagnostic marker for neuroblastoma, as it specifically highlights tumor cells derived from neuroblastic lineage.

PHOX2B-positive cells in neuroblastoma indicate the tumor's origin from neural crest cells, contributing to its distinct classification and aiding in differential diagnosis from other small round blue cell tumors.

1.1.2 Pathogenesis and molecular alterations

Neuroblastomas typically occur sporadically, yet heritable germline mutations can alter the risk of disease development. Compared to many other solid tumors, neuroblastomas have fewer somatic mutations, with a frequency of less than one mutation per mega base. However, they often exhibit extensive copy number alterations, either affecting whole chromosomes (numeric alterations) or spanning large chromosome segments (segmental alterations).

1.1.2.1 Familial neuroblastomas

Most familial neuroblastomas, which represent 2–3% of all cases, are linked to germline mutations with high penetrance. Germline activating mutations in *ALK* (a receptor tyrosine kinase) account for 80% of these, while inactivating mutations in *PHOX2B* (which encodes paired mesoderm homeobox 2B transcription factor and is a master regulator of neural crest development) are also implicated. Both genes are integral to development of the sympathoadrenal lineage system, and neuroblastomas stemming from *PHOX2B* mutations frequently present with pathological features consistent with neural crest abnormalities (Mossé et al. 2008) (Bourdeaut et al. 2005) (Amiel et al. 2003). Both the *ALK* and *PHOX2B* gene play integral roles in the development of the sympathoadrenal system. Neuroblastomas that develop due to mutations in *PHOX2B* often present with clinical manifestations linked to neural crest pathologies, including conditions such as Hirschsprung disease and congenital central hypoventilation syndrome.

Beyond these genes, genome-wide association studies (GWAS) have uncovered a multitude of lower penetrance polymorphisms that can collectively influence the initiation and progression of neuroblastoma. These genetic variations can subtly modify the risk or behavior of the disease. Notable among these are polymorphisms found in several genes, including *LINC00340*, *BARD1*, *LMO1*, *DUSP12*, *DDX4/IL31RA*, *HSD17B12*, *LIN28B*, and *HACE1* (Maris et al. 2007). Each of these genetic alterations contributes to a more nuanced understanding of the complex genetic landscape of neuroblastoma, highlighting the importance of both major genetic mutations and more subtle genetic variations at a germline level in the disease's development and progression.

1.1.2.2 Recurrent focal amplification

Oncogene *MYCN* amplification on chromosome 2p24 is a hallmark of high-risk neuroblastoma and serves as a critical prognostic indicator. The frequency of *MYCN* amplification has been found to be at 20-25% of all neuroblastomas and in ~40% of high-risk neuroblastomas (Berbegall et al. 2016). As a pivotal transcriptional regulator, N-myc orchestrates cell growth, metabolism, and differentiation. Early research identified that neuroblastomas with a marked increase in *MYCN* copies, often exceeding a tenfold amplification, were indicative of a more aggressive form of the disease (Seeger et al. 1985) (Brodeur et al. 1984). At a cytogenetic level, *MYCN* amplification corresponds to the integration of numerous copies of the *MYCN* genes in the chromosomes, as homogeneously staining regions, or outside chromosomes as double minutes. Such amplifications contribute to both the intrinsic properties of cancer cells promoting proliferation and an undifferentiated state and the tumor microenvironment, such as promoting immunosuppression and new blood vessel formation. A mutation in *MYCN*, specifically P44L, is present in about 1% of high-risk cases, but its prognostic significance remains unclear, though it may enhance N-myc stability (Pugh et al. 2013) (Liu et al. 2017).

N-myc is a key member of a protein family, which includes the closely associated c-Myc protein, encoded by the *MYC* gene and commonly referred to as Myc. While N-myc expression is usually tightly controlled and specific to certain developmental stages and cell types, the ubiquitous Myc is upregulated in various cancers. About 10% of high-risk neuroblastomas without *MYCN* amplification exhibit Myc overexpression. This is not typically due to *MYC* gene amplification but rather the result of transcriptional activation through amplified distal enhancer regions or the rearrangement of chromosomes that brings Myc under the control of active enhancer sequences.

Beyond *MYCN*, focal amplification of other chromosomal regions also correlates with very poor prognoses, particularly through the overexpression of oncogenic proteins like *CDK4* and *MDM2*. In neuroblastoma, particularly among high-risk patients, Myc overexpression is a critical factor, identified in about 10% of cases lacking *MYCN* amplification (Zimmerman et al. 2018). Notably, direct *MYC* gene amplifications are infrequent in neuroblastoma. Instead, the overexpression of Myc often results from strategic transcriptional upregulation. This upregulation is typically orchestrated either through the amplification of distal enhancer elements or by rerouting highly active enhancer elements via chromosomal translocation (Zimmerman et al. 2018). These mechanisms effectively boost Myc expression, driving the cancer process in neuroblastoma through pathways that do not involve direct *MYC* gene amplification.

1.1.2.3 Chromosomal copy number alterations

In neuroblastoma, chromosomal changes are remarkably prevalent, affecting about 90% of patients, with the nature of these alterations carrying significant prognostic implications (Janoueix-Lerosey et al. 2009). In young patients with neuroblastoma lacking *MYCN* amplification, the occurrence of hyper diploidy due to whole-chromosome gains or losses is generally indicative of a more favorable prognosis compared to a near-diploid state (George et al. 2005). Conversely, segmental chromosomal alterations (SCAs), such as the loss of 1p or 11q and the gain of 17q, correlate with poorer outcomes, regardless of any co-existing whole-chromosome anomalies. Rare germline SCAs including deletions of 11q or 1p are observed in neuroblastoma cases (Biegel et al. 1993) (Egolf et al. 2019). Other less common chromosomal changes associated with adverse prognosis include the loss of chromosomal regions 3p, 4p, 5p, 6q, and 9p. Recurrent losses of 1p or 6q are more commonly seen at relapse, supporting the theory that loss of tumor suppressor genes on these segments contributes to tumor progression. For example, *CHD5*, *CAMTA1*, *KIF1B*, *CASZ1* and *ARID1A* on 1p chromosomal arm. *ARID1A* is known for its tumor suppressor activity, and its functional disruption is linked to enhanced tumorigenesis in N-myc-driven preclinical models (Shi et al. 2020) (Sausen et al. 2013). Ongoing preclinical research aims to elucidate the oncogenic roles of SCAs and identify actionable targets for therapy.

1.1.2.4 *ALK* alterations

Only a few genes recurrently altered directly by coding sequencing mutations have been described in NB. The most frequent somatic mutations concern the *ALK* gene, a receptor tyrosine kinase with a neural lineage expression, triggers mitogenic signaling through the RAS–MAPK and PI3K-AKT pathways upon activation. Sporadic neuroblastomas of all risk categories present somatic *ALK* mutations in 8–12% of cases, while ~2% of high-risk cases exhibit *ALK* amplifications (Bellini et al. 2021). These activating mutations are predominantly situated within the kinase domain, with F1174, R1275, and F1245 as common mutation sites, and can occur both at a clonal or subclonal level. However, the interpretation of mutant allele frequency could be influenced by tumor cell content. Co-occurrences of *ALK* mutations and amplifications with *MYCN* amplifications are significant, with 44% of *ALK* mutations and 97% of *ALK* amplifications found in the HR-NBL1 study's *MYCN*-amplified tumors. Moreover, high-risk neuroblastoma patients with *ALK* alterations represent a subset with an especially poor prognosis. A 2021 clinical trial revealed that patients with *ALK* amplifications or high MAF mutations (>20%) had significantly lower 5-year event-free and overall survival rates (28% vs 51%; $P < 0.01$) compared to those without these alterations (Bellini et al. 2021). Although the prognostic significance of isolated *ALK* mutations or

amplifications is complex, its role as a neuroblastoma driver is evident. Synergistic effects of activated *ALK* and N-myc on tumorigenesis have been demonstrated in zebrafish and mouse neuroblastoma models.

1.1.2.5 RAS-MAPK and p53 signaling pathway alterations

Drivers of aggressive neuroblastoma are frequently linked to aberrations in the RAS–MAPK and p53 signaling pathways (Ackermann et al. 2018), which are associated with treatment resistance and relapse. Relapsed neuroblastoma samples show higher incidences of mutations in these pathways compared to initial disease presentations. While *MYCN* amplifications tend to be clonal, indicative of their presence in all or most tumor cells, alterations in RAS–MAPK and p53 pathways are usually subclonal, with emergence as the disease progresses (Gundem et al. 2023). Such mutations are observed in neuroblastoma at every stage and within all risk categories. Mutations in RAS and p53 pathway genes occurred in both high- and non–high-risk tumors, although at lower frequencies in the latter group (21.3% versus 13.3%) (Ackermann et al. 2018) while other research group suggested 18 of the 23 relapse tumors (78%) showed mutations predicted to activate the RAS-MAPK signaling pathway (Eleveld et al. 2015). The presence of these mutations, particularly when they co-occur with telomere maintenance defects, often signals a dire prognosis.

ALK mutations, leading to MAPK pathway activation, are especially prevalent in neuroblastoma. Other notable changes include *PTPN11* inactivating mutations, which deactivate a phosphatase that typically curtails mitogenic signaling. Germline mutations in *PTPN11*, which are known to cause Noonan syndrome, a disorder affecting multiple organs are linked to an elevated neuroblastoma risk. Less commonly, activating mutations in genes associated with the RAS–MAPK pathway or upstream receptor tyrosine kinases, like *HRAS*, *NRAS*, *FGFR1*, and *BRAF*, are also detected in neuroblastoma cases (Ackermann et al. 2018) (Brady et al. 2020).

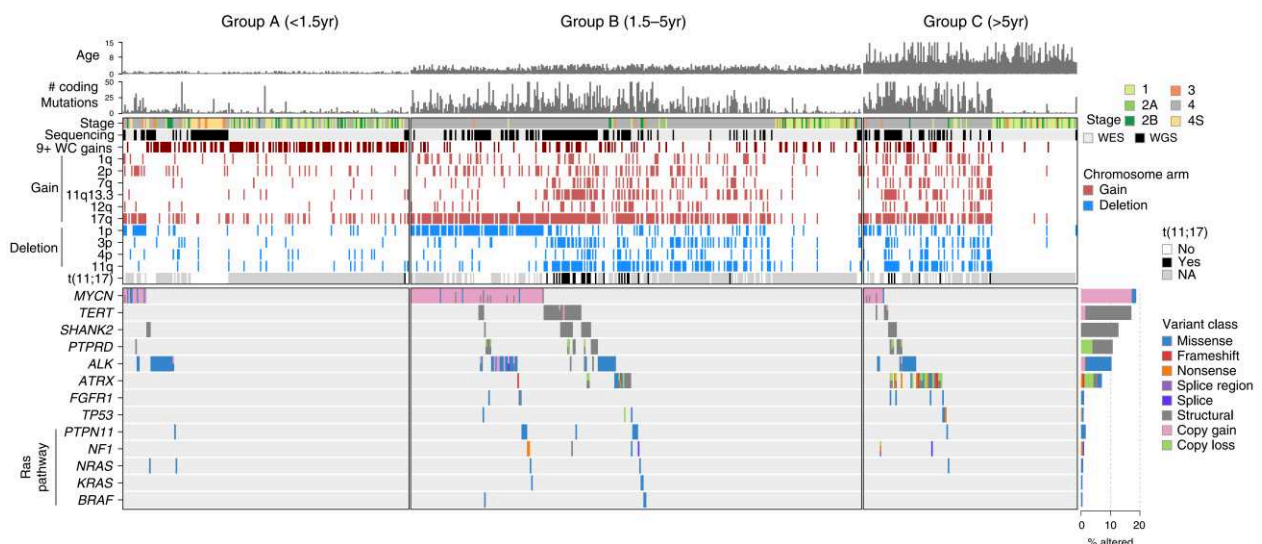


Figure 1.1.2.5.1: Recurrent somatic alterations by age group in neuroblastoma

Top, age at diagnosis and number of coding mutations in each of 685 neuroblastoma samples (662 diagnosis, 23 relapse) sequenced by Whole Genome Sequencing (WGS) or Whole Exome Sequencing (WES), with samples categorized into <18 months of age at diagnosis (group A, n = 206), 18 months to 5 years (group B, n = 325), and 5 or more years (group C, n = 154). Middle, segmental chromosome copy changes which were statistically significant per GISTIC analysis, and structural variants; blue indicates segmental copy loss, red indicates segmental copy gain, white indicates no change. 9+ WC gains samples (gain of nine or more whole chromosomes) are shown in dark red. Bottom, somatic variants in driver genes (Brady et al. 2020).

1.1.2.6 Activation of telomere maintenance mechanisms

High-risk neuroblastoma patients often display genomic alterations that impact telomere maintenance, typically associated with poor clinical outcomes. Telomere length in neuroblastoma cells is preserved via distinct mechanisms: telomerase activation due to *TERT* overexpression or through alternative lengthening of telomeres (ALT) pathway. Approximately a quarter of high-risk cases exhibit structural rearrangements that drive *TERT* overexpression by bringing super-enhancers in proximity to the *TERT* locus. Unlike in other cancers, *TERT* promoter mutations are absent in neuroblastoma. Although *TERT* amplification is rare, *MYCN* amplification is linked to higher *TERT* expression levels, yet some *MYCN*-amplified tumors exhibit low *TERT* expression, suggesting that *MYCN* amplification is not the sole driver of telomere maintenance activation. ALT mechanisms are identified in a similar proportion of high-risk cases, with mutations in *ATRX*, which governs epigenetic regulation, being the most common alteration linked to ALT, occurring in roughly 10% of these patients. These *ATRX* mutations, which may lead to loss of function or generate novel fusion proteins, are more common in patients over 10 years old, less so in younger children, and non-existent in infants under 18 months. In a study of 134 high-risk neuroblastoma, 20% had structural variants leading to *TERT* overexpression and 23.4% showed ALT activity, with

60% of the ALT-positive tumors carrying *ATRX* mutations. This leaves 40% of ALT tumors without known gene alterations related to ALT. This underlines the prognostic importance of telomere maintenance activation in neuroblastoma, which is now being prospectively examined in clinical trials. Additionally, targeted therapies to inhibit this aberrant activation are being tested preclinically.

1.2 Intra-tumoral heterogeneity (ITH)

Neuroblastomas are characterized by a high degree of heterogeneity which is reflected in clinical presentations spanning from spontaneous regression or differentiation to treatment-refractory progression regardless of intensive multimodal therapies. Although spontaneous regression occurs in some patients, the majority of neuroblastomas are diagnosed as high risk and most of these patients have metastatic disease already at diagnosis suggesting that induction of metastasis is an early event.

1.2.1 Spatial Intra-Tumor Heterogeneity

Spatial Intra-Tumor Heterogeneity (ITH) refers to the variation in the genetic, epigenetic, and phenotypic characteristics of cancer cells within a given tumor. This heterogeneity is a result of the tumor's evolutionary process, where different regions of the tumor accumulate distinct mutations and adaptations over time. Spatial ITH is a crucial aspect of cancer biology as it can influence the tumor's response to treatment, its progression, and the risk of metastasis. Spatial Intra-Tumor Heterogeneity (ITH) in neuroblastoma manifests in various ways, reflecting the diverse genetic and phenotypic profiles within different regions of the same tumor.

1.2.1.1 Variation in *MYCN* Amplification

In neuroblastoma, a common feature is the amplification of the *MYCN* gene. However, spatial ITH is observed in the distribution of *MYCN* amplification within the tumor. Some regions of the tumor may exhibit high levels of *MYCN* amplification, contributing to aggressive tumor behavior, while other regions may have normal *MYCN* levels or no amplification at all. This heterogeneity can significantly impact the overall aggressiveness and treatment response of the tumor (Theissen et al. 2009).

1.2.1.2 Presence of Genetic Mutations and Chromosomal Alterations

Within a single neuroblastoma tumor, different regions can exhibit various genetic mutations and chromosomal alterations. One area might show mutations in genes like *ALK* or amplifications in

TERT, while another region might present different alterations such as loss of heterozygosity at certain chromosome arms. These genetic discrepancies within the tumor can lead to varied behavior and treatment responses from different parts of the same tumor (Schmelz et al. 2021).

1.2.2 Temporal Intra Tumoral Heterogeneity

Temporal Intra-Tumor Heterogeneity (ITH) in neuroblastoma can be exemplified in several ways, as the disease progresses and evolves over time, especially in response to treatment or due to natural progression. Following three examples that illustrate temporal ITH in neuroblastoma:

1.2.2.1 Evolution of Genetic Mutations Post-Diagnosis

In neuroblastoma, subclonal *ALK* mutations can be present at diagnosis with subsequent clonal expansion at relapse (Schleiermacher et al. 2014). (Gundem et al. 2023) studied clonal transitions in high-risk neuroblastoma (HR-NB) patients experiencing consecutive relapses. Analyzing 114 clonal transitions in 72 patients, they identified three distinct evolutionary patterns:

- I. Accumulation of Genetic Changes (72% of cases): Most cases showed additional copy number alterations (CNAs) or mutations/structural variations (SVs), reflecting linear and branched evolutionary progressions in HR-NB.
- II. Relapse by Identical Clones (24% of cases): A notable portion of relapses originated from the same clone as the primary tumor, indicating the expansion of pre-existing clones without new genetic changes.
- III. Relapse by Earlier Clones (4% of cases): A smaller group showed relapses driven by earlier-stage clones, suggesting that dormant clones can later become dominant and cause relapse.

This implies that some clones, which may have been minor or dormant in earlier stages, gained prominence and led to relapse at a later stage. Notably, these patterns were consistent from diagnosis to first relapse, across consecutive relapses, and irrespective of disease subtypes, suggesting that while new driver subclones emerge and supersede existing ones, certain biological themes remain preserved. In the same study, in high-risk neuroblastoma (HR-NB), particularly among patients with a history of three or more consecutive relapses, a pattern of multiple waves of clonal successions over extended periods has been observed. This complex clonal evolution is exemplified by two distinct patient cases:

1. Early and Frequent Relapses: In a young patient with stage-I *MYCN* non-amplified neuroblastoma, a specific clone with a *SPRY2* deletion and *IGF2BP3* amplification

consistently reemerged in multiple recurrences over a short span, later transitioning to different mutant subclones, highlighting a stable yet evolving clonal landscape.

2. **Delayed and Diverse Relapses:** Another case, an older child with the same neuroblastoma type, showed early relapses driven by a PIK3CA-mutant subclone, with later relapses years post-diagnosis driven by a distinct MTOR-mutant clone. Both clones were present initially, suggesting a pattern of long-term clonal dormancy and later emergence.

1.2.2.2 Development of Chemotherapy Resistance:

NB cells can demonstrate inherent primary resistance to drugs or develop multi-drug resistance (MDR) after prolonged chemotherapy.

Neuroblastoma comprises two epigenetically distinct cell types: undifferentiated mesenchymal (MES) cells and committed adrenergic (ADRN) cells (Boeva et al. 2017). Notably, MES cells, characterized by the expression of the stem cell marker CD133, exhibit high migratory capabilities and heightened resistance to chemotherapy. These cells are prevalently found in tumors that reemerge post-chemotherapy. The resistance mechanism in NB encompasses a comprehensive array of factors. Beyond genetic (mutations, amplifications) and epigenetic alterations (DNA hypermethylation, histone modifications), resistance is thought to be fueled by a multitude of factors: elevated expression of drug efflux transporters, aberrant miRNA expression, cancer cell stemness, autophagy, influences from the tumor microenvironment, extracellular vesicles, hyperactivation of MEK/ERK signaling, and altered internalization of anti-disialoganglioside antibodies.

1.2.2.3 Shifts in Tumor Cell Populations Over Time:

Neuroblastomas can exhibit shifts in the prevalence of different cell types over time. For example, a neuroblastoma might initially consist predominantly of noradrenergic (ADRN) tumor cells (Thirant et al. 2023), but as the disease progresses, there might be an increase in the proportion of mesenchymal (MES) cells, which are often associated with more aggressive disease and resistance to certain therapies. Most adrenal NB tumors analyzed in the study by (Jansky et al. 2021) were classified as ADRN type, yet the malignant cells of three high-risk tumors showed increased MES signature expression and reduced ADRN signature expression and were termed high-risk NBs with MES features, as previously defined Groningen et al. This shift can be due to differential growth rates, treatment responses, or microenvironmental influences. However, these results were primarily derived from studying neuroblastoma cell lines and less conclusive in patient tumor at diagnosis and relapse.

1.3 Models of genetic tumor evolution

A fundamental question in cancer research involves deciphering the evolutionary process of tumors. The inherent genomic instability (Hanahan 2022) characteristic of tumors creates an ideal backdrop, marking a trail of changes across successive generations. This instability is a defining feature of tumor cells, leaving discernible genomic footprints that can be traced and studied over time. (Körber et al. 2023). The foundational theory proposed by Nowell posits that a normal cell transforms into a neoplastic one through the acquisition of a genetic alteration. As tumors evolve, they sequentially accumulate genetic alterations often in a stepwise fashion, under selective pressures. This process results in the survival and dominance of the most adaptive cell population (Nowell 1976). These alterations impact a cell's propensity to proliferate, thereby influencing its evolutionary fitness. Some alterations are detrimental, reducing a cell's fitness, while others, known as driver alterations/mutations, confer a selective advantage. Additionally, there are passenger mutations, which neither enhance nor diminish fitness but may expand within the tumor population (Beerenwinkel et al. 2015).

Tumor cells typically proliferate in a clonal manner. Within the tumor's ecological niche, various clones engage in competition for resources, collaboration for invasive capacity, and strategies to evade immune system detection (Black and McGranahan 2021). Understanding tumor evolution requires comprehensive models that encapsulate the intricate interactions and adaptations of cancer cells within the tumor microenvironment. These models are crucial for devising effective cancer treatments. The following sections provide concise summaries of several key models in tumor evolution (Vendramin, Litchfield, and Swanton 2021).

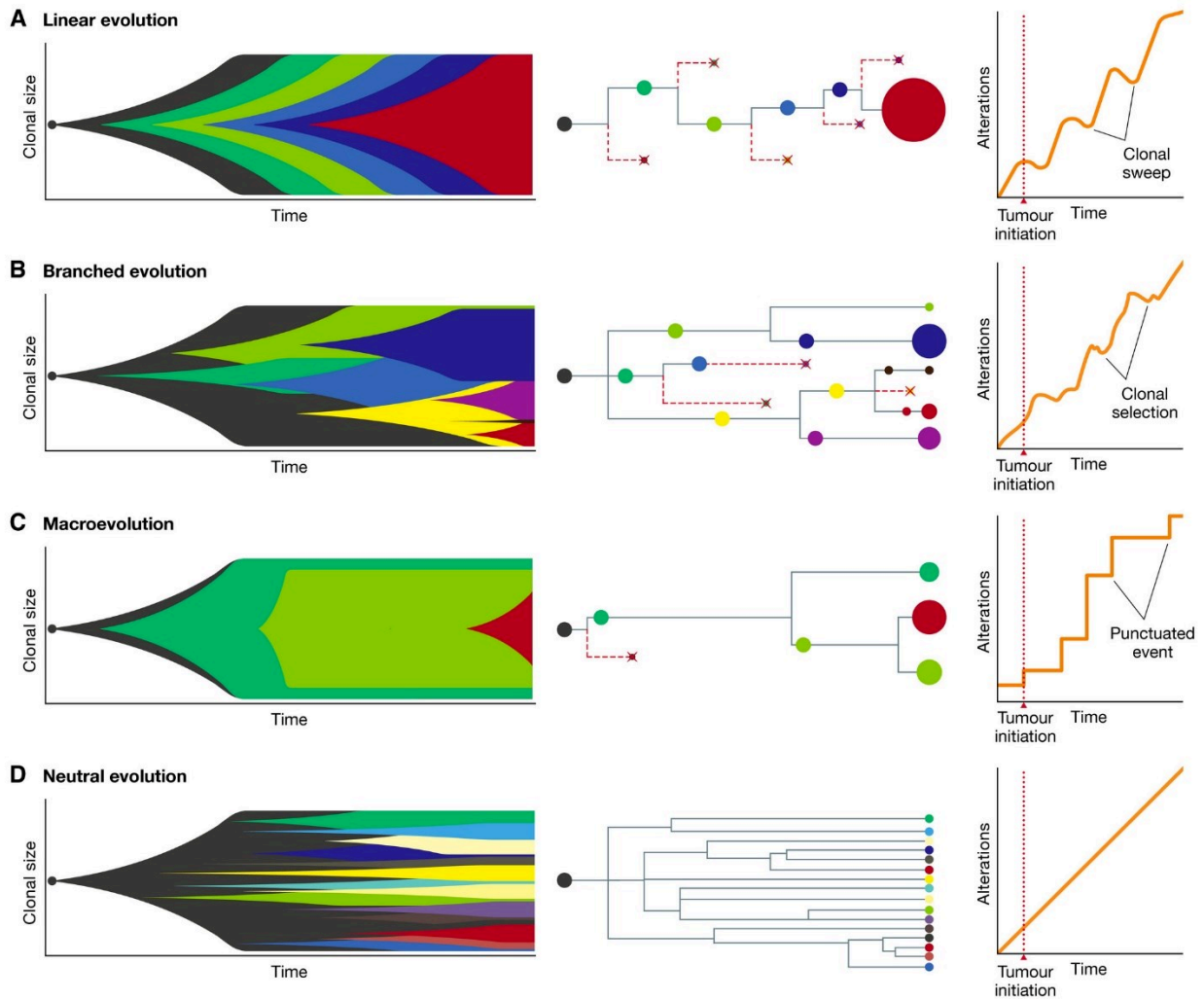


Figure 1.2.2.3.1: Models of tumor evolution

Models of linear evolution (A) branched evolution (B), macroevolution or punctuated evolution (C) and neutral evolution (D) described by Muller plots representing dynamic changes in clonal size over time (left), clonal lineages and phylogenetic trees (centre) and changes in the number of alterations over time (right). Colors indicate different clones (Vendramin, Litchfield, and Swanton 2021).

1.3.1 Linear Evolution Model

The linear evolution model theorizes that tumors develop through the acquisition of mutations in a stepwise manner (Nowell 1976). This model posits that mutations in key driver genes occur one after the other, each mutation providing a fitness advantage that leads to its prevalence within the tumor population through a process known as a clonal sweep. A clonal sweep takes place if a subclone outcompetes its neighboring cells, thus ultimately leading to a reduction of diversity and a homogeneous cancer cell population (Black and McGranahan 2021). This model's validity is supported by evidence from breast cancer research (Ullah et al. 2018) and more recently in lung cancer (Martínez-Ruiz et al. 2023). Their studies illustrated a sequential progression of mutations in oncogenes and tumor suppressor genes, tracing linear evolution to successive metastasis. While

this model offers valuable insights into the evolutionary trajectory of cancer, it is important to note that its support primarily stems from studies focused on single genes. Consequently, there is a relative paucity of comprehensive, genome-wide experimental evidence underpinning the linear evolution model (Gatenby and Vincent 2003).

1.3.2 Branching Evolution Model

The branching evolution model proposes that tumor evolution occurs through the parallel development of multiple clones. This model differs from the linear model primarily in its characterization of fewer clonal sweeps (Davis, Gao, and Navin 2017). In this framework, phylogenetic trees exhibit significant heterogeneity, including the presence of intermediate ancestor. The branching model is extensively supported by a wide range of studies, both in bulk sequencing and single-cell analyses, across various types of cancer. Notable examples include lung cancer (Jamal-Hanjani et al. 2017), breast cancer (Darlan C. Minussi et al. 2021), liver cancer (Su et al. 2021), ovarian cancer (Schwarz et al. 2015), renal cancer (Turajlic et al. 2018), pediatric cancers (Andersson et al. 2020) and non-small cell lung cancer metastases (Al Bakir et al. 2023)

The evidence supporting this model is robust and well-documented in diverse tumor types. For instance, a comparison of multiregional sampling studies in pediatric cancers like neuroblastoma, rhabdomyosarcoma and Wilms tumor have all provided evidence for the branching pattern of tumor evolution (Andersson et al. 2020; Schmelz et al. 2021). A particularly compelling illustration of this model's prevalence is its identification in a pan-cancer cohort study employing whole-genome sequencing (Dentro et al. 2021). This study revealed that branching evolution is a common evolutionary pattern in tumors, even in those exhibiting low mutation rates.

1.3.3 Macroeolution or Punctuated Evolution Model

The punctuated evolution model is characterized by a brief phase of heightened genomic instability (R. Gao et al. 2016), leading to a multitude of genomic changes and distinct, punctuated evolutionary steps (Fittall and Van Loo 2019). This instability often results in extensive structural genomic alterations, including gains or losses of chromosomal arms or whole chromosome, TP53 mutations, and instances of whole genome duplication (WGD). Unlike the linear, neutral, and branching models, which are primarily supported by single nucleotide variations, the punctuated evolution model is predominantly evidenced by copy number alterations and structural variations (Davis, Gao, and Navin 2017).

A key element in punctuated evolution appears to be the mutation of the TP53 gene (R. Gao et al. 2016). TP53 mutations are often observed in the early clonal stages of various tumor types,

seemingly playing a critical role in this evolutionary pattern. Tumors with TP53 mutations are found to be more likely to experience whole-genome doubling – a hallmark of punctuated evolution – compared to those with wild-type TP53 (Gerstung et al. 2020). This phenomenon is notable, with synchronous copy number events indicative of punctuated evolution observed in a significant proportion of tumors that have undergone whole-genome doubling. Even in tumors that remain near diploid, there is evidence of synchronous gains in a majority of cases (Gerstung et al. 2020). The distinct burst of evolutionary activity characteristic of the punctuated model leads to the formation of unique phylogenetic trees. These trees typically show large truncal distances and a notable absence of intermediate taxa (R. Gao et al. 2016).

However, the interpretation of these trees requires careful consideration, particularly in the context of strong selective pressures and clonal expansion. Evidence of punctuated evolution has been observed in a variety of cancer types, including breast cancer (R. Gao et al. 2016; Navin et al. 2011), prostate cancer (Baca et al. 2013), uveal melanoma (Field et al. 2018), and colorectal cancer (Cross et al. 2018). This suggests that punctuated evolution may be a more widespread phenomenon across different cancers than previously thought.

1.3.4 Neutral Evolution Model

The neutral evolution model in cancer biology posits that tumor development is characterized by extensive branching and a high degree of intra-tumor heterogeneity. This heterogeneity, according to the model, results not from selective evolutionary pressures but from random genetic drift (Davis, Gao, and Navin 2017). Consequently, in a neutral evolutionary scenario, there is no significant expansion of any particular subclonal population despite the observed diversity.

Support for the neutral evolution model comes from detailed studies in various cancers. For instance, an analysis involving microdissection of 286 samples from a single hepatocellular carcinoma case revealed about 20 subclones with high genetic diversity across different tumor regions (Ling et al. 2015). This finding aligns with the predictions of the neutral model. Additional support for this model has been provided by analyzing allele frequencies across 14 different tumor types. Out of 904 samples examined, 323 were found to better fit the neutral evolution model (Williams et al. 2016). However, these findings have faced criticism due to the use of simplifying assumptions in the analysis, which may have led to the misclassification of some non-neutral tumors as neutral (Tarabichi et al. 2018).

A related concept is the big bang evolution model, which describes an early, single expansion phase in the tumor's development. This expansion, not governed by strict selective pressures, leads to the

formation of a diverse population within the tumor. Predominantly observed in colorectal cancer studies, the big bang model (Sottoriva et al. 2015) suggests that the initial clonal expansion sets the stage for the subsequent diversity observed in the tumor, without ongoing selective pressures shaping the tumor's evolution.

In neuroblastoma both linear and branched evolutionary model are frequently reported previously (Karlsson et al. 2018) (Andersson et al. 2020) (Schmelz et al. 2021) (Gundem et al. 2023).

However, as of today's date there is no research in neuroblastoma highlighted variable replication timing and its potential role in genetic evolution.

1.4 Molecular mechanisms involved in genetic evolution: DNA replication

Genetic alterations in cancer cells can arise due to extrinsic factors that cause DNA damage. These alterations are then addressed by the cell's DNA repair mechanisms, which can either successfully repair the damage or fail to do so. The outcome of these repair processes contributes significantly to the genetic diversity observed within tumors. Additionally, DNA replication itself is a crucial factor in the development of intratumor heterogeneity. As cells replicate their DNA, errors can occur, leading to genetic variations. Distinct replication profiles are another key element that characterizes different cellular subclones within a tumor. These replication profiles can vary from one subclone to another, contributing to the complexity and diversity of the tumor's genetic landscape. Early replicating regions of the genome in certain subclones might be associated with rapid cell division and aggressive tumor behavior, while late replicating regions in other subclones might correspond to slower growth or different cellular characteristics. Understanding these nuances of DNA replication and repair, and how they contribute to intratumor heterogeneity, is crucial for a comprehensive understanding of cancer biology.

1.4.1 Cell cycle and DNA replication timing (RT)

DNA replication, essential for cell division and maintaining genomic integrity, is a fundamental biological process. The cell cycle includes several distinct phases: G1 (Gap 1), S (Synthesis), G2 (Gap 2), and M (Mitosis). During cell division, a human cell is programmed to replicate its entire DNA in a highly coordinated manner within the S-phase's constrained timeframe. This replication process initiates from multiple regions scattered across the genome, referred to as DNA replication origins (Akerman et al. 2020). Specific genomic loci exhibit distinct changes in their replication

timing during the process of differentiation, often aligning with the expression status of the genes they harbor.

Early-replicating domains of the genome are typically associated with active transcription. These regions are often found in euchromatin and characterized by markers indicative of open and active chromatin, which is less condensed and more transcriptionally active. Genes located in early-replicating domains are usually actively transcribed, and this early replication is thought to facilitate efficient transcriptional processes. The open chromatin structure in these regions allows easy access for transcription machinery and regulatory elements.

In contrast, domains that replicate later in the S-phase are generally enriched with markers of closed, inactive heterochromatin, which is more condensed and transcriptionally inactive. Genes in these regions are often silenced or expressed at low levels. The late replication of these areas may contribute to the maintenance of a repressive chromatin state, limiting the accessibility of transcription factors and other regulatory proteins.

A notable example of this phenomenon is observed in the X chromosome of mammalian females. The inactive X chromosome in these cases becomes predominantly late replicating as its transcriptional activity is significantly reduced or shut down. This observation underscores the intricate relationship between chromatin status, gene expression, and replication timing (Weiner et al. 2023).

The replication timing (RT) can vary across different cell types and developmental stages, reflecting the complexity and adaptability of the cell replication process. Moreover, replication timing is closely linked to genome stability. There is a correlation between the timing of replication and the rate of mutations, suggesting that when DNA replicates can influence its susceptibility to changes.

Neuroblastoma, a common pediatric solid tumor, frequently involves unbalanced chromosome translocations, particularly affecting chromosomes 1 and 17. A study using high-resolution arrays revealed that over 50% of breakpoints in neuroblastoma cell lines occur in early replicating regions of the genome. This suggests a significant link between these translocations and early genome replication, likely due to break-induced replication mechanisms (Janoueix-Lerosey et al. 2005).

1.4.2 Cancer cell and DNA replication

In cancer, where cell cycle regulation is frequently disrupted, a thorough understanding of DNA replication is crucial. Dysregulation in DNA replication and cell cycle control contributes to increased cellular proliferation and a heightened mutation burden (Sanders et al. 2021) (Sima and Gilbert 2014). These changes can lead to genomic instability, a common characteristic in cancer cells, underscoring the significance of replication timing (RT) in maintaining genomic integrity.

This disruption results in replication stress and genomic instability, leading to somatic copy number alterations (CNAs) and intratumoral heterogeneity, driving the evolution of cancer. The timing of DNA replication during the S-phase of the cell cycle is closely associated with various epigenomic features, including 3D nuclear organization, chromatin state, and transcription, all of which impact the cellular phenotype. Furthermore, structural variations and CNAs can affect these epigenomic states and potentially influence RT, highlighting the complex interplay between DNA replication, genomic structure, and cancer progression.

Specific genomic alterations can provide fitness advantages to certain cancer subclones, leading to distinct proliferation rates and a faster progression through the cell cycle. Single-cell whole genome sequencing (scWGS) has emerged as a potent tool for analyzing clonal heterogeneity and CNAs. It holds promise for enhancing our understanding of DNA replication dynamics, particularly in aneuploid cell populations. However, the computational challenge of identifying S-phase cells and distinguishing replicating from non-replicating loci is significant, primarily due to the difficulty in differentiating inherited somatic CNAs from transient changes in DNA replication. Deciphering these two signals is crucial for a more accurate study of replication timing and proliferation rates of individual genetic subclones. Such insights would significantly advance our understanding of the interplay between DNA replication, genomic instability, and cancer progression. This knowledge is key to developing targeted interventions that can effectively manage or inhibit the uncontrolled proliferation characteristic of cancerous cells. Moreover, some chemotherapies, like cisplatin, specifically target S-phase cells, emphasizing the importance of understanding cell cycle dynamics for therapeutic interventions (Sherr and Bartek 2017).

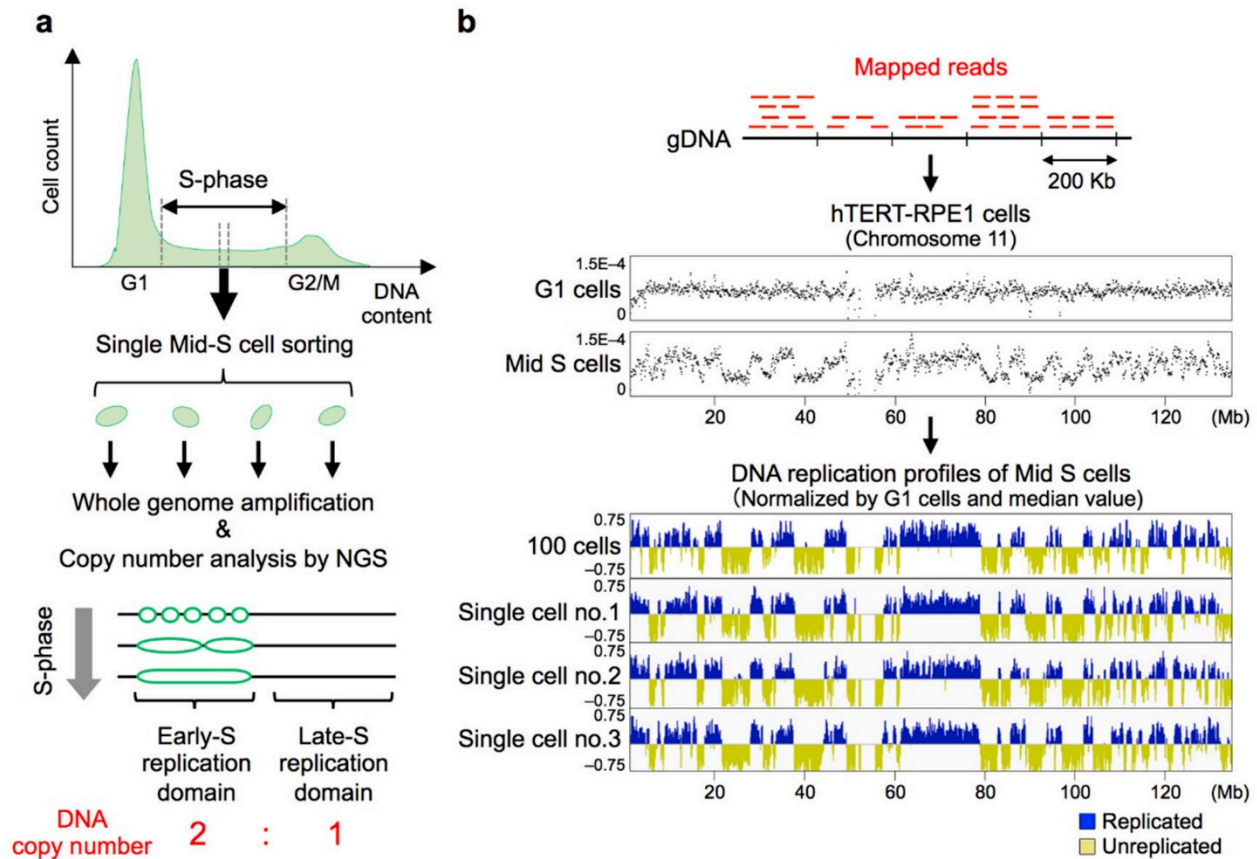


Figure 1.2.2.3.1: The basic principles of genome-wide single-cell DNA replication profiling

(a) An experimental overview of scRepli-seq. A typical cell cycle profile of mammalian cells stained with propidium iodide during flow cytometer analysis is shown, along with the mid-S sorting gate used. Genomic DNA samples isolated from single or 100 mid-S cells were subject to copy number analysis by next generation sequencing (NGS) to detect early and late replication domains throughout the genome; (b) Replication profiling by copy number analysis. Mapped NGS reads of mid-S cells were counted in sliding windows of 200 Kb (Kilobases) at 40-Kb intervals to generate tag density plots (i.e., counts per window normalized by total read counts). Shown are human chromosome 11 (chr11) tag density data from 100 hTERT-RPE1 cells in G1 and mid S-phase. Mappability was corrected using G1 samples, and the numbers were further divided by the median read count (i.e., median centering) to generate $\text{Log}_2[(\text{corrected mid-S})/\text{median}]$ replication profiles of 100 cells and three individual mid-S cells (Hiratani and Takahashi 2019).

1.4.3 Computational tools to study replication in single cell

One of the key advancements in this realm in the last one and half years was the development of tools that were designed for single-cell Replication Timing (scRT) analysis. New methods like Probabilistic Estimation of single-cell Replication Timing (PERT) and Kronos scRT were developed to jointly infer single-cell copy number and replication states from scWGS data (Weiner et al. 2023) (Gnan et al. 2022). Unlike traditional methods, Kronos scRT does not require cell sorting or a specific platform, thus enabling the study of large datasets from asynchronous cell populations. It provides robust predictions of cell cycle phases, quantifies replication timing variability, and estimates relative proliferation rates among tumor subclones.

1.5 Quantifying genetic changes in cancer evolution

One challenge for theorists lies in adapting theoretical models to data that are often noisy and complex. Generative Bayesian modeling is particularly suited for this task, as it allows for direct incorporation of data-generating aspects into the analysis. Other Bayesian techniques, such as approximate Bayesian computation (Gillis and Roth 2020), enable the fitting of simulation-based models to data in a structured manner. This approach includes accounting for sequencing noise, which can be modeled using specific distributions (like the beta-binomial for overdispersed coverage) or by comparing synthetic data generated within a simulation-based framework against real data.

Approaches to measure cancer evolution and its implications for patient outcomes generally fall into two categories. The first involves analyzing similarities and differences across large cohorts of cancer genomes. Recurrent patterns in these studies may indicate convergent evolution towards specific phenotypes, while the absence of certain features might signal negative selection. The second approach focuses on intratumor heterogeneity, providing insights into the evolutionary dynamics of individual malignancies.

1.5.1 Leveraging cohort-level information to study cancer evolution

The clonal evolution model posits that tumors originate from a single cell, which then undergoes somatic alterations. Central to this model is the concept of clonal expansions, primarily driven by 'driver' mutations. These mutations confer an increased fitness, defined as the enhanced capability of the cell to produce surviving offspring within its specific microenvironment. By tracing these mutations, much like tracking species evolution, we can gain insights into the cancer's evolutionary trajectory (Körber et al. 2023). Körber et al. implemented this approach by conducting deep whole-genome sequencing on samples from two neuroblastoma patient cohorts, encompassing a total of 186 patients and covering all clinical stages of the disease. A primary obstacle in such analyses is the inability to directly observe the entire evolutionary process of the tumor. Instead, researchers are limited to analyzing a random sample of the process's outputs. This sampling limitation presents a significant challenge in fully reconstructing and understanding the complex evolutionary pathways of cancer. In the study of neuroblastomas, the sequenced tumors exhibited a range of point mutations along with a plethora of copy-number alterations (CNAs), which include both gains and losses of genetic material. These CNAs play a significant role in altering gene expression and contribute to the genetic instability of the tumor cells. The researchers observed that CNAs

were universally present across the neuroblastoma cells, suggesting that these alterations were acquired early in the tumor's development.

To analyze the mutational landscapes of each neuroblastoma sample, the authors employed a combination of mathematical modeling and statistical inference. A key part of their analysis involved determining the timing of segmental copy-number gains in certain chromosome regions (specifically, 1q and 17q) and whole-chromosome gains (in chromosomes 2, 7, and 17). They utilized neutral mutations, which accumulate at a consistent rate over time, to establish a molecular clock. This approach enabled them to differentiate neuroblastomas into two evolutionary groups. The first group comprised tumors in which the gains occurred temporally close to the most recent common ancestor (MRCA) the latest cell lineage from which all the sequenced neuroblasts are descended. The second group included tumors in which an earlier common ancestor (ECA), preceding the MRCA, had acquired most of the gains. This distinction between groups based on the timing of genetic gains sheds light on the evolutionary pathways leading to the malignant state of the MRCA. In Körber et al.'s study, the latter group of neuroblastomas, which evolved from a premalignant early common ancestor (ECA), showed additional genetic evolution into the malignant state. This was evidenced by the timing of gains in regions 9 and 20q, which occurred after the ECA phase. The research team utilized molecular clocks to draw comparisons between primary tumors and their relapsed counterparts, delineating prototypical models that differentiated early and late most recent common ancestors (MRCAs). This distinction suggests varying disease trajectories for relapsed neuroblastomas.

The study found that 95% of neuroblastomas with early-MRCAs predominantly exhibited aneuploidy, with minimal contribution from point mutations. In contrast, 55% of late-MRCA tumors displayed extended genetic evolution characterized by multiple point mutations and other somatic changes. Further insights were gained from analyzing MRCAs of metastases that were relapsed following initial diagnosis.

The evolutionary history of a tumor can be a powerful predictor of disease outcomes. Utilizing a discovery cohort, Körber et al. demonstrated that the timing of the MRCA is a reliable predictor of both event-free survival and overall survival, offering improvements over traditional clinical variables and neuroblastoma-specific features like telomere maintenance and RAS/p53-pathway mutations. In particular, the study highlighted the role of aneuploidy, especially in late-MRCA tumors, which exhibited broader instability and were associated with poorer prognoses. This finding suggests a correlation between the duration of a tumor's evolution and its eventual outcome. Additionally, the acquisition of telomere maintenance mechanisms was identified as a

significant factor in neuroblastomas, although the exact timing of *TERT* mutations acquisition remained unclear.

1.6 Methods for evaluating genetic tumor evolution

Accurately reconstructing a tumor's evolutionary history from a genomic snapshot, which is typically obtained from a single time point and often from just one section of the tumor, presents significant challenges. Despite a consensus acknowledging the diversity in tumor evolution, inconsistencies in terminology and analytical methods continue to impede efforts to categorize tumors based on their evolutionary patterns.

Many studies, including a notable analysis of 21 breast cancer genomes by Campbell and colleagues (Nik-Zainal et al. 2012), rely on bulk sequencing of a single tumor sample to deduce the tumor's evolutionary trajectory. These studies primarily interpret evolutionary history based on the variant allelic frequencies (VAFs) of somatic mutations found in the tumor. However, such approaches must navigate a range of complicating factors. These include the proportion of non-tumor tissue included in the sample, the specific somatic copy number alteration (SCNA) profile of the tumor, and the precision and depth of the sequencing process. Each of these factors can significantly skew the interpretation of the genomic data.

Another key issue in these studies is the potential for sampling bias. A mutation might appear to be present in all the cancer cells in the sampled tissue, but it might not actually be representative of every cancer cell within the entire tumor. This discrepancy can lead to misinterpretations or oversimplified conclusions about the tumor's evolution, underscoring the need for more comprehensive sampling and sophisticated analytical approaches in studying tumor genomics.

Moreover, differentiating the variant allelic frequencies (VAFs) of mutations in subclones resulting from selective pressures from those in the tail-like distribution characteristic of neutral evolution can be challenging. This tail distribution, representing passenger mutations not subjected to selection, mirrors the random mutagenesis occurring at each cell division. It also reflects the expected correlation between the number of mutational events and their clonal frequency over time. Even when clonal sweeps are precisely identified, distinguishing between a sweep that occurred recently and one present at the tumor's inception is often not feasible. This limitation highlights the inherent complexities in interpreting the evolutionary history of tumors and underscores the need for advanced methodologies and approaches in cancer genomics research.

Multi-sampling at various time points during clonal evolution, while not completely addressing the issue of neutral tails, can improve the precision in identifying and classifying subclones. This approach also helps reduce sampling bias and can provide a higher-resolution view of the phylogenetic relationships among subclones. This enhanced resolution is particularly valuable for subclones with low cancer cell fractions (CCFs), as it allows for the observation of coordinated patterns of CCF fluctuations over time, offering deeper insights into the evolutionary dynamics of the tumor. With increased frequency in sampling time points, it becomes possible to distinctly identify individual subclones based on their cancer cell fractions (CCFs), particularly when these subclones exhibit unique growth dynamics (Caravagna, Sanguinetti, et al. 2020). Serial sequencing not only aids in the more accurate decomposition of clonal populations but also facilitates the measurement of clone-specific fitness. This approach is particularly effective in tracking the growth kinetics of clonal populations in response to therapeutic interventions, as evidenced in studies of both circulating leukemia cells and circulating tumor DNA in solid malignancies. Using mathematical modeling, researchers can assess how clonal growth kinetics are influenced by treatment.

An illustrative example of this is seen in chronic lymphocytic leukemia (CLL) cases that relapse post-chemotherapy. In such instances, an increase in the CCF of subclones with *TP53* mutations compared to the pre-treatment tumor has been observed, suggesting that these *TP53*-mutated subclones gain a significant fitness advantage under the selective pressure of therapy. This finding underscores the potential of dense temporal sampling in enhancing our understanding of clonal evolution, particularly in the context of therapeutic response and resistance. However, this approach is not a complete solution. Typically, only a very small portion of the total tumor is analyzed, and sequencing every single tumor cell, which would be ideal for comprehensive analysis, remains an unattainable goal.

Multi-sampling can also involve analyzing multiple regions within a tumor to evaluate its intratumoral clonal spatial composition. Similar to temporal sampling, multiregional sampling has proven beneficial, particularly in the study of non-small-cell lung carcinomas, where it has helped refine our understanding of clonal relationships and thus enhance clinical stratification.

A notable example of how multiregional sampling can reveal evolutionary selection in tumors involves the study of driver copy number alterations (CNAs). In some cases, multiregional sampling has uncovered the convergent evolution of these driver CNAs, where the same CNA impacts different parental alleles in various regions of the tumor. This type of sampling also allows for the comparison of primary tumors with their metastatic counterparts. For instance, in lung

adenocarcinomas that have metastasized to the brain, multiregional sampling has identified specific drivers of metastasis, such as CNAs in genes like *MYC*, *YAP1*, and *MMP13*, that are distinct from those in the primary tumors. This approach is instrumental in understanding the complex dynamics of tumor evolution and metastasis, offering vital insights for the development of targeted cancer therapies. Hua et al. explicitly compared the ITH of point mutations, SCNAs and DNA methylation in a multi-region study of 84 lung adenocarcinomas⁷⁴. They found that tumour evolutionary trees inferred from SCNAs and DNA methylation were highly similar, demonstrating that patterns of cancer evolution may be agnostic of the variant mechanism (Hua et al. 2020).

Negative selection, which can lead to the elimination of a clone before sampling, remains undetectable using variant allelic frequency (VAF)-based approaches. Consequently, alternative methods are necessary to unravel the sequence of events and differentiate between functional events, which are subject to selection, and non-functional events, which are not.

One strategy to time somatic alterations and distinguish between early and late events in tumor evolution involves analyzing copy number gains and whole-genome duplication events. Such an approach was employed in the Pan Cancer Analysis of Whole Genomes (PCAWG) study. This analysis enabled the inference of clonal architecture trends across various tumor types. For instance, in colorectal adenocarcinoma, mutations in genes like *APC*, *KRAS*, and *TP53* were predominantly early events, whereas certain copy number alterations, including losses on chromosomes 15q, 21q, and 22q, tended to occur later.

However, describing tumor evolution using binary labels such as neutral or branched, and punctuated or gradual, remains challenging. Different analyses of the same data can yield contrasting conclusions. For example, a study by Williams et al., which analyzed VAFs in 904 cancers across 14 types, suggested that 36% (323 out of 904) of these tumors exhibited subclonal VAF distributions consistent with neutral evolution. Yet, a subsequent study using a different method (dN/dS) identified evidence of subclonal selection within cancer genes in these 323 tumors (Williams et al. 2016).

Single cell sequencing platforms are increasingly recognized for their potential to offer deeper insights into tumor evolution. These platforms enable the analysis of cellular genomes alongside other cellular components like epigenomes and transcriptomes. Lineage reconstruction at the single-cell level could significantly enhance our understanding of tumor phylogeny. An example of this is Direct Library Preparation single-cell whole-genome sequencing, which allows for the identification of clonal populations of single cells and their unique genomic characteristics. This

technique enables the aggregation of data into ‘pseudo-bulk’ samples, from which clonal phylogenies can be inferred, offering a more nuanced view of tumor evolution.

1.6.1 The role of Single cell sequencing

Single-cell approaches (SCA) have revolutionized the analysis of individual cancer cell genomes, transcriptomes, epigenomes, proteomes, and even metabolomes, as well as the tumor microenvironment (TME), offering unprecedented resolution (Navin et al. 2011) (Hiley et al. 2014) (Y. Wang et al. 2014) (Kim et al. 2018). The advent of single-cell RNA sequencing (scRNA-seq) has been pivotal, providing a valuable tool for dissecting tumor evolutionary dynamics (Nam, Chaligne, and Landau 2021). Its high resolution allows for the detection of gene expression in minor sub-clones, which might otherwise be missed by bulk RNA-seq, thus enabling the identification of small, treatment-resistant cell populations that could lead to therapeutic failure (Rambow et al. 2018).

While scRNA-seq can infer copy number variation and loss of heterozygosity in individual cells, it typically lacks the resolution required for discovering novel SCNAs (Fan et al. 2018). Innovatively, scRNA-seq generates transcriptomic data from individual cells at various evolutionary stages, facilitating the organization of cells into pseudotime and evolutionary trajectories, which sheds light on tumor evolutionary history and dynamics without indicating the directionality of these trajectories. RNA velocity analyses augment this by measuring the ratio of intronic to exonic reads to infer changes in transcript abundance, offering predictions of future transcriptional states and enhancing our grasp of cellular transcriptional dynamics (La Manno et al. 2018, la m)

Furthermore, single-cell DNA sequencing (scDNA-seq) has also surfaced, providing individual cell genomic profiles to deduce genetic phylogenies without deconvolution. However, first generation of scDNA-seq confronts significant limitations, notably its low coverage due to the minimal DNA content in a single cell, a stark contrast to the nanogram quantities needed for NGS library preparation. Whole-genome amplification, a necessary yet error-prone step, is employed to produce sufficient DNA for sequencing (Y. Wang et al. 2014).

1.6.1.1 First-generation single cell DNA sequencing

A single diploid cell contains approximately six pg of DNA. In tumors, this value can fluctuate up to approximately 12 pg. Extracting these minute amounts and performing whole genome sequence (WGS) creates technical challenges including non-uniformity of the genome coverage and

sequence sparsity. In addition, financial challenges are common due to the prohibitive cost of sequencing the entire genome of hundreds to thousands of cells.

The first generation scDNA-Seq technologies applied DOP-PCR⁵¹. DOP-PCR takes advantage of degenerate primers that randomly align to the genome and cover large fractions of the DNA. While this technique presents high sensitivity and captures small amounts of DNA uniformly, the physical coverage remains limited to approximately 10% of a single cell genome. Moreover, despite the uniform coverage, the random priming nature of the technique combined with WGS prior to fragmentation results in regions of pileups of DNA, which increases coverage bias in copy number studies. Besides, DOP-PCR is a laborious protocol that can take up to 3 days. An alternative method for WGS is multiple displacement amplification (MDA) using the phi29 enzyme. MDA increases the genome physical coverage up to 90%, but it is unsuitable for copy number analysis due to its high non-uniformity coverage. To address this issue multiple annealing and looping based amplification cycles (MALBAC) use short cycles of amplification and denaturation in an attempt to reduce technical bias. However, it remains limited due to its quasi-linear amplification nature and introduces high false positive single nucleotide errors given the lack of proofreading activity of the Bst enzyme. Collectively, first-generation scDNA-Seq allowed the first single cell copy number profiles. Nevertheless, they were limited by technical biases to the breadth of coverage and the uniformity. In addition, the laborious workflows and scale of reagents with expensive enzymes resulted in cost-prohibitive experiments not suitable for high-throughput analysis of scDNA-Seq.

Additional challenges include a heightened false negative rate from allelic dropout and noise from sequencing doublets (Pugh et al. 2013) (Leighton et al. 2021) (Mallory et al. 2020). Interestingly, the integration of scDNA-seq with bulk DNA sequencing can compensate for each method's limitations, leading to more precise phylogenetic inferences and a more comprehensive capture of intratumor heterogeneity (ITH) and evolutionary dynamics.

1.6.1.2 Single Cell Copy-Number Profiling of tumors

Single cell copy number data is a powerful tool to understand tumor evolution. Copy number data, possess singular characteristics that facilitate the comprehension of how tumors develop over time. Aneuploidy events from cell division defects are inherited across all daughter cells, and we can use this property to resolve tumor lineages. The first usage of scDNA-Seq for the inference of tumor evolution was in 2011 when Navin et al. applied single-nucleus sequencing to analyze hundreds of cells from two breast tumors. By examining the copy number profiles and the relationship among cells, the authors detected aneuploid cells without intermediate populations. Which, in turn, is suggestive of 'gaps' during the evolutionary process. Moreover, the same work revealed the

existence of pseudodiploid cells, i. e, cells that present an overall diploid profile with limited copy number events (Navin et al. 2011). This work was pivotal and the first demonstration of high resolution (~54 kb) copy number profiles from single cell genomes using sequencing.

1.6.1.3 High-throughput single-cell DNA Sequencing methods

Modern methods for single cell DNA sequencing introduced technical advances and scaled up sample sizes from dozens to thousands of single cells. Adey et al. introduced a PCR-free library construction methodology by using a transposase to simultaneously fragment DNA and incorporate sequencing adapters (Adey et al. 2010). This method was adapted for microfluidic devices and named Direct Library Preparation (DLP) to sequence hundreds of tumor cells (Laks et al. 2019). The PCR-free approach significantly reduces bias compared to first-generation single cell sequencing due to the absence of DNA amplifications steps, by performing the PCR step after fragmentation. Despite the advance in cell throughput, the DLP method required custom equipment, preventing wider adoption. DLP was later adapted to nano wells once again scaling-up of the number of sequenced cells and labeled Direct Library Preparation+ (Zahn et al. 2017). However, it still required custom equipment for its use. Combinatorial indexing approaches exponentially increased the number of cells sequenced. Instead of single cell chemistry reactions, pooled cells are barcoded using a transposase. Later, they are redistributed across different pools and go through the second round of barcoding through PCR. The combination of barcodes is used as a unique cell identifier and distinguished bioinformatically during sample demultiplexing (Vitag et al. 2017). Despite the robust increase in throughput, the lack of accessibility of the transposase to regions containing nucleosomes limits its genomic resolution. Lastly, with the advent of high-throughput droplet based microfluidics platforms, a commercial method from 10X Genomic platform using MDA chemistry allowed copy number profiling of up to 3000 cells in a single run (Velazquez-Villarreal et al. 2020). Nevertheless, it presented coverage bias related to MDA chemistry, and it is currently discontinued. Hence, despite the throughput increase by different protocols, the need for cost-efficient and easy workflow techniques for single cell copy number sequencing remains.

1.6.1.4 Copy Number Inference from high-throughput scDNA-seq

Several methods can infer copy number data from sequencing reads. Allele-specific copy number methods leverage imbalances across single nucleotide polymorphisms (SNPs) to estimate copy numbers across segments. Notable methods applying this inference are ASCAT, SEQUENZA and FACETS. Allele-specific methods can estimate haplotype-specific copy number calls but require extensive coverage of SNPs regions. Counting methods bin sequencing reads into genomic regions

of different resolutions to obtain total copy number information. Though the latter cannot perform haplotype-specific calls, it can be applied to sparse sequenced samples. Thus, reducing the number of reads needed and increasing cell multiplexing.

1.6.2 Interpretation complexities of tumor evolutionary models

Classifying tumor samples into specific evolutionary models serves to identify common patterns in their evolutionary dynamics, facilitating the stratification of tumors based on these dynamics. However, it is crucial to acknowledge that evolution, including tumor evolution, is fundamentally a continuous and dynamic process. The models used to understand tumor evolution are often constructed based on data from single time point samples and may involve limited spatial sampling and the adoption of simplifying assumptions (Williams, Sottoriva, and Graham 2019).

In reality, the evolutionary trajectory of a tumor could encompass different models at various stages of its development. For instance, tumor cells might initially undergo clonal sweeps under strong selective pressures, aligning with the linear evolution model. As the tumor progresses, particularly during its pre-malignant stages, the conditions might shift to favor a scenario more in line with the branching evolution model, characterized by less pronounced selective pressures and more extensive branching.

Therefore, while these models are invaluable tools in deciphering the complex nature of tumor evolution, they come with inherent limitations. It is important to approach the interpretation of these models with caution and an understanding that they may represent only a part of the tumor's evolutionary story. This nuanced understanding acknowledges the complexity of tumor evolution and the need for a flexible and multifaceted approach to studying and classifying it.

1.7 Integrative scDNA-seq and scRNA-seq data analysis deciphering complex relationships between genotype and phenotype.

Age-related, environmental, and genetic factors significantly influence mutational processes, shaping the acquisition of somatic mutations throughout an individual's lifespan (Brady et al. 2020) (Thatikonda et al. 2023) (Gerstung et al. 2020) (Alexandrov et al. 2020). These mutations contribute to the development of clonal structures in both healthy and diseased tissues. Techniques such as targeted, whole-genome, and whole-exome DNA sequencing of bulk cell populations have been instrumental in delineating these mutational processes and reconstructing clonal trees (Nik-

Zainal et al. 2012) (Ludmil B. Alexandrov et al. 2013) (Gundem et al. 2023). The advent of single-cell DNA sequencing (scDNA-seq) methods, coupled with novel computational approaches, has further refined our ability to reconstruct these clonal populations (Navin et al. 2010) (Darlan C. Minussi et al. 2021) (Laks et al. 2019). However, there remains a substantial gap in our understanding of the functional differences between clones and their molecular phenotypes. Comprehensive characterization of clone's phenotypic properties is crucial for unraveling the mechanisms driving the transition from normal to malignant behavior in tissues and quantify the extent of intra-tumoral heterogeneity and its clinical implications. Recent studies have made strides in mapping single-cell RNA sequencing (scRNA-seq) profiles to clones with distinct copy number states in cancers, providing initial insights into gene expression differences among clones (Tirosch et al. 2016) (Fan et al. 2018). Yet, universally applicable methods for deciphering the origins of single cells and studying genotype-transcriptome relationships are still under development. The theoretical integration of assays to sequence both RNA and DNA from same single cells holds the promise of yielding comprehensive genotype-phenotype correlations, thereby capturing the critical aspects of clonal expansion, proliferation, and metastasis (Macaulay et al. 2015) (Dey et al. 2015). This approach is vital for understanding the mechanisms of drug response in cancer, where responses are often dictated by both positive and negative evolutionary selection of mutation-induced phenotypes, as well as by genome-independent, epigenetically driven alterations in transcriptional programs (Shaffer et al. 2017).

At present without scalable high-throughput methods to simultaneously assay DNA and RNA from the same single cell, parallel single-cell DNA and RNA measurements from independent cell populations must be mapped for genome-transcriptome association.

Computational methods such as cardelino exploit single nucleotide variant information in scRNA-seq reads (the alternate read count using a binomial model) to assign full-length single-cell transcriptomes to the clonal substructure (McCarthy et al. 2020). In this method clonal architecture derived from bulk DNA sequencing such as targeted sequencing, exome or whole genome sequencing could be integrated to investigate gene expression differences between clones at the level of individual genes and in pathways, which provides new insights into the dynamics of clones. Another method clonealign - a statistical method to assigns gene expression states to cancer clones using single-cell RNA and DNA sequencing independently sampled from a heterogeneous population (Campbell et al. 2019). In order to relate the independent measurements, clonealign assumes that an increase in the copy number of a gene will result in a corresponding increase in that gene's expression and vice versa (a copy-number dosage effect on transcript abundance).

Single-cell RNA sequencing methods offer a unique opportunity to link genetic heterogeneity with the overall cellular state. CNVs can be inferred from both transcript abundance and allelic imbalance in heterozygous single-nucleotide polymorphisms (SNPs) (Patel et al. 2014) (R. Gao et al. 2021) (Serin Harmanci, Harmanci, and Zhou 2020). However, reliably determining copy number states remains challenging due to the sparse and noisy nature of single-cell measurements. Allele-based approaches to CNV detection, which are less influenced by sample or cell type variations, face limitations due to data sparsity and allele-specific transcriptional variability in single cells. Existing methods for CNV detection from scRNA-seq do not typically incorporate prior knowledge of haplotypes, which can significantly enhance the detection of allelic imbalances. Population-based phasing, used to computationally phase variants of an individual, has been successfully applied in characterizing chromosomal aberrations in both germline polymorphisms and cancer evolution (Zaccaria and Raphael 2021) (Funnell et al. 2022). This phasing method, however, has not been widely explored in the context of scRNA-seq. Single-cell sequencing provides an unparalleled opportunity to dissect genetically heterogeneous subpopulations that are obscured in bulk measurements. Methods leveraging allele information typically rely on aggregating data across cells to form ‘pseudobulk’ profiles for confident aberration definition. Nevertheless, this approach's effectiveness depends on the accurate inference of clonal cell populations.

A recently published computational method Numbat, integrates expression, allele, and haplotype information obtained from population-based phasing to comprehensively characterize the CNV landscape in single-cell transcriptomes (T. Gao et al. 2022). Numbat employs an iterative approach to jointly reconstruct the subclonal phylogeny and single cell copy number profile of tumor samples. Applying this method to a diverse array of tumor samples, Numbat has demonstrated its capability to reconstruct high-fidelity copy number profiles from scRNA data alone, accurately distinguishing cancer cells from normal cells in the tumor microenvironment and identifying distinct subclonal lineages with allele-specific alterations. Numbat's versatility is evident in its applicability to a wide range of experimental settings and cancer types, without the need for sample-matched DNA data or a priori genotyping. This innovative approach has the potential to unravel previously inaccessible differences in molecular phenotypes between cells from the same individual, significantly advancing our understanding of cancer heterogeneity, evolution, and clone specific response to chemotherapy.

2 THESIS OBJECTIVES

Intratumoral heterogeneity (ITH) is a fundamental characteristic of many cancers, including neuroblastoma (NB), and it plays a crucial role in tumor evolution and treatment resistance.

ITH can begin early in tumor development, often as a consequence of genetic instability in the initial cancer cells. This instability leads to a variety of mutations and chromosomal alterations. As the tumor proliferates, different subclones and different populations of cells with distinct genetic profiles emerge. Environmental pressures, such as limited nutrient supply, tumor microenvironment and immune surveillance, further select for subclones with advantageous traits. ITH becomes particularly evident during and after cancer treatment. Therapies may effectively target the dominant clone but can leave behind or even promote the growth of resistant subclones, leading to relapse.

In neuroblastoma, previous studies have identified the importance of both spatial and temporal genetic heterogeneity, and both are thought to play an important role in neuroblastoma oncogenesis, tumor progression and treatment resistance. The evolutionary genetic trajectory of neuroblastoma remains unclear. Moreover, DNA sequencing of bulk samples confirm pervasive heterogeneity in neuroblastoma but remain limited in resolution and confounded by normal cells contaminant from the stroma. The main aim of the thesis was to perform integrative bioinformatics analysis to investigate ITH and clonal evolution at single cell level in neuroblastoma.

Part A: Manuscript aims and objective:

Aim 1. Dissecting the Clonal and Subclonal Genetic Landscape of Neuroblastoma

To analyze the clonal and subclonal architecture of neuroblastoma, this study focuses on investigating the genetic variations present at diagnosis, progression, and relapse phases. Utilizing single-cell DNA sequencing, we aim to dissect the intricate clonal and subclonal structures within neuroblastoma tumors.

Aim 2. Analysis of Evolutionary Trajectories and Underlying Molecular Mechanisms

The study focuses on a comprehensive examination of tumor evolution in neuroblastoma, covering various aspects from the modalities of tumor evolution to the underlying molecular mechanisms. First, it investigates the different modes of tumor evolution, aiming to understand their roles in driving intratumoral heterogeneity and influencing the course of tumor progression. This aspect is crucial for comprehending how these evolutionary patterns contribute to the dynamic and complex

nature of cancer growth. Additionally, the study delves into the temporal evolutionary patterns of neuroblastoma, tracing its progression from diagnosis onwards. This analysis will shed light on the temporal dynamics and transformation of genetic alterations over time. Finally, the research seeks to unravel the molecular drivers responsible for subclonal evolution, focusing on the factors that lead to tumor progression and relapse.

Aim 3. Investigating Replication Timing and Its Impact on Tumor Progression

The study explores replication timing (RT) variations within neuroblastoma at both the single-cell and pseudobulk levels, focusing on early and late RT. It aims to assess how these RT variations contribute to intratumoral heterogeneity (ITH) by analyzing the differential RT in distinct subclones within the same tumor. Additionally, it investigates the behavior of these subclones, specifically their response to therapeutic interventions, based on their RT profiles. This approach provides an innovative perspective on the relationship between RT variability and the dynamic response of tumor subclones to treatment, offering new insights into tumor progression and potential therapeutic strategies.

Aim 4. Subclonal Heterogeneity: Phenotypic Consequences and Transcriptomic Impact

This aspect of the study aims to elucidate the phenotypical impact of subclonal heterogeneity at the copy number level on the transcriptome. By examining the transcriptomic profiles of neuroblastoma subclones, we seek to understand how variations in replication timing, specifically between early and late replicating subclones, manifest at the transcriptomic level. Additionally, the research delves into the characteristics and transcriptomic expression potential of rare clones in neuroblastoma.

Aim 5. Temporal Intratumoral Heterogeneity through sequential bulk Whole-Exome Sequencing under treatment pressure

Subclonal Dynamics and Targeted Treatment: Study subclonal population's evolution and response under therapeutic pressure, integrating whole-exome sequencing with single-cell DNA sequencing.

These aspects are presented in part A of the results section.

Part B: We aimed to provide further insights into NB ITH by investigating details of genes of interest on one hand and large patients on other hand.

Aim 1 Analyzing ITH for a subset of genes involved in chromatin remodeling and epigenetic modifier genes.

Aim 2 Studying ITH for *ALK*, the gene most frequently altered at a gene level in high-risk (HR) patients on the HR-NBL1/SIOPEN trial to determine their frequency.

Role in part B projects: Bioinformatics data analysis

The aspect of this work is presented in 2 published manuscripts presented in part B of the results section.

3 RESULTS

3.1 Part A. Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing suggest clone specific therapeutic vulnerabilities in neuroblastoma.

Neuroblastoma, a predominant pediatric malignancy, is distinguished by its diverse clinical behavior and genetic heterogeneity. Poor clinical outcomes are still observed in high-risk tumors irrespective of prognostic molecular markers suggesting the importance of inter and intra-tumor heterogeneity (ITH), describing diversity within individual tumors. Whether ITH is due to stochastic accumulation of mutations following the acquisition of early driver events or due to continuous clonal evolution and selection through time and space, remains an open scientific question.

This thesis aims to delve into the molecular intricacies of neuroblastoma by leveraging advanced genomic and transcriptomic techniques such as single cell whole genome sequencing and single cell RNA sequencing to dissect the complexities of intratumoral heterogeneity (ITH), understand replication timing (RT) variances, and explore the clonal evolutionary trajectories under therapeutical pressure.

By elucidating clonal and subclonal dynamics in Neuroblastoma through employing scDNA-seq on 14 patient-derived xenografts (PDXs) and 4 patient tumors, we aim to map the clonal architecture of neuroblastoma. This analysis will provide insights into the evolution of clonal populations, their competitive dynamics, and the emergence of subclones, particularly focusing on how these aspects contribute to therapy resistance and disease progression.

Investigating replication timing (RT) and its Impact on ITH: A critical aspect of this research involves studying the RT profiles in neuroblastoma cells. Our objective is to examine how variations in early and late replication timing within the same tumor contribute to intratumor heterogeneity. By studying the RT variability among subclones, we intend to unravel the distinct replication timing profiles of each subclone, elucidating their roles in tumor progression and response to therapy. This understanding is pivotal in exploring the relationship between DNA replication, genomic instability, and cancer development.

Integrating Single-Cell RNA Sequencing Data: By incorporating scRNA-seq data, this research seeks to link genetic alterations at the DNA level with changes in the transcriptomic landscape at

a single cell level. This integration enables a comprehensive understanding of how genomic variations manifest in functional changes within tumor cells. The study extensively utilizes scDNA-seq to profile individual cells within neuroblastoma tumors. This approach allows for an in-depth examination of genetic diversity at the single-cell level, facilitating a detailed understanding of tumor evolution. In parallel, scRNA-seq data from 16 of the 18 tumor samples offer insights into the gene expression patterns, complementing the genomic data. This transcriptomic analysis helps to establish a direct link between genetic alterations and their functional implications in tumor biology.

Temporal Intratumoral Heterogeneity through Whole-Exome Sequencing: We conduct WES on a six of the established PDX models to capture the temporal dynamics of ITH. This approach is instrumental in identifying novel mutations potentially responsible for therapeutic resistance. This analysis provides a broader view of the genetic landscape across different stages of tumor progression, aiding in the identification of evolving mutational patterns.

Challenges and Limitations: Despite the comprehensive nature of these technologies, challenges persist in fully unraveling the complexity of neuroblastoma. For instance, the limitations of bulk sequencing in capturing the full extent of tumor heterogeneity necessitate continual advancements in sequencing and computational methodologies.

The resulting manuscript is currently being finalized, and submission for publication is eminent.

3.1.1 Manuscript: Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing suggest clone specific therapeutic vulnerabilities in neuroblastoma.

Jaydutt Bhalshankar^{1,2*}, Angela Bellini^{1,2*}, Irene Jimenez^{1,2}, Elnaz Saberi-Ansari^{1,2}, Yasmine Iddir^{1,2}, Alexandra Saint-Charles^{1,2}, Mylène Bohec⁴, Sonia Lameiras⁴, Sylvain Baulande⁴, Didier Surdez⁶, Sakina Zaidi⁶, Isabelle Janoueix-Lerosey⁷, Cécile Thirant⁷, Gaelle Pierron⁸, Coralie Guerin⁹, Angel Montero Carcaboso¹⁰, Maria Eugenia Marques da Costa¹¹, Birgit Geoger¹¹, Andrei Zinovyev¹², Olivier Delattre^{1,8}, Gudrun Schleiermacher^{1,2}

**Contributed equally to this study*

¹INSERM U830, Laboratoire de Génétique et Biologie des Cancers, Research Center, PSL Research University, Institut Curie, Paris, Ile de France, France

²SiRIC RTOP « Recherche Translationnelle en Oncologie Pédiatrique », Translational Research Department, Research Center, PSL Research University, Paris, Ile de France, France

³Unité de Génétique Somatique, Institut Curie, Paris, France

⁴Institut Curie Genomics of Excellence (ICGex) Platform, Paris, France

⁵Department of Biopathology, Institut Curie, PSL Research University, Paris, France

⁶INSERM U830, Équipe Labellisé LNCC, PSL Université, SIREDO Oncology Centre, Institut Curie, Institute Curie Paris, Ile de France, France

⁷Institut Curie, PSL Research University, Inserm U830, Equipe Labellisée Ligue contre le Cancer, F-75005, Paris, France

⁸Department of Somatic Genetics, Curie Institute, Paris, France

⁹Institut Curie, Flow Cytometry Department, Paris, France

¹⁰Preclinical Therapeutics and Drug Delivery Research Program, Department of Oncology, Hospital Sant Joan de Déu, Barcelona, Spain

¹¹Gustave Roussy Cancer Campus, Department of Paediatric and Adolescent Oncology, Université Paris-Saclay, Villejuif, France

¹²INSERM U900, Cancer et génome: Bioinformatique, biostatistiques et épidémiologie des systèmes complexes, Research Center, PSL Research University, Institut Curie, Paris, Ile de France, France

This work was supported by the Ligue contre le Cancer (AO2018), as well as EraCoSysMed (Project ID29 INFER-NB). This work was also supported by the Annenberg Foundation, the Association Hubert Guoin Enfance et Cancer, the Fédération Enfants Cancers Santé, the Société Française de lutte contre les Cancers et les leucémies de l'Enfant et l'adolescent (SFCE), Les Bagouz à Manon, Les amis de Claire and the Fondation ARC pour la Recherche contre le Cancer (ARC). Funding was also obtained from SiRIC/INCa (Grant INCa-DGOS-4654), and PHRC IC2007-09 grant. High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche ("Investissements d'Avenir" program), by the Canceropole Ile-de-France and by the SiRIC-Curie program - SiRIC Grant "INCa-DGOS- 4654". Tumor sequencing data and clinical data have been provided under the MAPPYACTS protocol (clinicaltrials.gov: NCT02613962).

3.1.1.1 Abstract

Background: How cell-to-cell allele specific copy number alterations underpin genetic Intratumor heterogeneity (ITH), drive genomic and phenotypic variation, and consequently the evolution of neuroblastoma (NB), remains understudied

Aims: Here we investigate ITH, timing of specific genomic aberrations, single-cell replication timing and the co-evolution of the genome and transcriptome in NB tumors at single-cell resolution. Further we study subclonal dynamics and clone specific response or resistance under targeted therapeutic pressure.

Methods: In addition to germline/tumor bulk whole exome sequencing (WES), ultra-low depth (0.25x) single-cell whole-genome DNA sequencing (scDNAseq) was performed using 10x genomics Chromium single-cell CNV (scCNV) kit and 9410 tumor cells were characterized from 14 patient-derived xenografts (PDX) NB-models and 4 tumor biopsies from NB-patients, either at diagnosis (n=7), progression (n=3) or relapse (n=8). Single-cell RNA sequencing (scRNAseq) data was obtained from the same PDX and patient tumor samples (Thirant et al. 2023). 6/14 PDX models were subjected to different treatment combinations (targeted treatment with/without chemotherapy) and bulk WES was performed at two time-points, pre- and post-treatment.

Results: Polyclonal (n=11) and monoclonal (n=7) genomes were determined by allele and haplotype specific copy number (CN) alteration using both scDNAseq and scRNAseq data analysis. 2 to 11 clones were observed per polyclonal NB tumor. Whole genome duplication events (n=7) were observed in both polyclonal and monoclonal genomes. Known driver CN (segmental loss in chr1p and chr11q and gain at chr17q, or *MYCN/ALK* amplification) or somatic mutations

(*ALK/ATRX/TP53/NF1*) were early clonal events. scDNAseq analysis showed parallel copy number evolution of two distinct clones, subclone s1/s2, in one PDX model. Data integration of clonal mutational profiles with pre and post targeted therapy (Lorlatinib) reveals clone specific treatment response. Subclone s2 was partially responding with extinction of a sub-set of somatic alterations, whereas no change was observed in subclone s1. The replication timing (RT) profile of these two clones, subclone s1 (early-RT) and s2 (late-RT) were mutually exclusive. Genotype to phenotype analysis revealed subclone s1 genotype was preferentially expressed at transcriptomic level.

Conclusion: Together, these results determine the evolutionary trajectories of NB tumors, linked to distinct replication timing and highlight opportunities for targetable early clonal alteration detection.

3.1.1.2 Introduction

Neuroblastoma (NB) is the most frequent childhood extracranial solid tumor of the peripheral sympathetic nervous system. Clinically, neuroblastoma manifests salient heterogeneity, ranging from spontaneous regression to incurable progression despite intensive therapy (Maris et al. 2007) (Cheung and Dyer 2013).

The clinical heterogeneity observed in cancer can in some instances be explained by a complex interplay between genetic and non-genetic determinants of somatic evolution. Genetic variations including somatic single nucleotide variants (SNVs), somatic copy number alterations (SCNAs) and whole genome duplication (WGD) can act as a primary substrate for tumor evolution (Gerstung et al. 2020).

In neuroblastoma, at diagnosis, genetic alterations mainly concern copy number alterations. *MYCN* amplification observed in approximately 20 – 25% of all cases is associated with a poor prognosis, with other genomic amplifications occurring less frequently, including chromosome regions encompassing the *ALK* gene, or *CDK4/MDM2*. Segmental chromosome alterations, including deletions of chromosome arms 1p, 3p, 4p and 11q, or gains of chromosome arms 1q, 2p, or 17q, are associated with poor outcome. Only few genes have been shown to be altered recurrently in NB, with activating mutations of *ALK* being observed in 8 – 10% of NB at diagnosis, and other SNVs or single gene events targeting genes such as *ATRX*, genes involved in chromatin remodeling, or *PTPN11*. Telomere maintenance either through activation of telomerase activity or by ALT (alternative mechanisms of telomerase) play an important role in the oncogenesis of NB

and are associated with a poor prognosis in particular when occurring together with mutations in the RAS/MAPK pathway (Eleveld et al. 2015) (Ackermann et al. 2018).

Both spatial and temporal genetic heterogeneity play an important role in NB, and some genetic events considered as tumor “driver” events can occur at a subclonal level in only a subset of tumor cells. Furthermore, at relapse new genetic alterations are frequently observed, with apparition of either new CNA, or emergence of mutations in the *ALK* or RAS/MAPK genes, underlining the importance of clonal evolution in the process of tumor progression and treatment resistance (Braekeveldt et al. 2018) (Schmelz et al. 2021).

Study of changes in the mutational landscape before and after treatment can lead to the identification of new mutations that confer resistance to therapy, possibly leading to an understanding of the genetic mechanisms underlying treatment failure (Ben-David, Beroukhim, and Golub 2019).

Accumulation of genetic alterations might reflect distinct evolutionary trajectories of tumor cells, and mapping their appearance over time might reflect an underlying molecular clock. Specific mutational signatures can be observed when somatic single base substitution mutations (such as SBS1 or SBS5) accumulate over time because of specific extrinsic and intrinsic factors such as DNA replication errors or selection pressures (Brady et al. 2020) (McGranahan and Swanton 2017).

Multi-sampling for interrogating multiple regions within a tumor to assess for intra-tumoral spatial heterogeneity or study of different time points during clonal evolution to assess for intra-tumoral temporal heterogeneity can provide higher-resolution phylogenetic relationships even for subclones with low Cancer Cell Fraction (CCF) (Schmelz et al. 2021) (Körber et al. 2023). However, these bulk sequencing data can only resolve clonal and subclonal relationships to a limited extent. (Nam, Chaligne, and Landau 2021)

The emergence of single-cell genomic DNA sequencing technologies permits scalable whole genome single-cell DNA sequencing of hundreds of individual cells in parallel, providing an ideal framework for analyzing genetic ITH (Baslan et al. 2016) (Y. Wang et al. 2014) (Laks et al. 2019). Complementing these technical developments, recent computational advances enable highly accurate ploidy estimates and the inference of allele- and haplotype-specific SCNAs in individual cells and sub-populations of cells from low-coverage single-cell DNA sequencing (Zaccaria and Raphael 2021). This allows cell-by-cell assessment of intra-tumoral SCNA heterogeneity, identification of allele- specific alterations, and reconstruction of the evolutionary history of a tumor from hundreds of individual cancer cells obtained at a single or multiple time points during tumor progression (Funnell et al. 2022).

To better understand neuroblastoma ITH, here we performed whole-genome single-cell DNA sequencing in combination with both whole-exome (bulk), and transcriptomic single-cell RNA sequencing in both primary tumor samples and established patient derived xenografts (PDX). To determine whether the widespread SCNAs in neuroblastoma result from ongoing genomic instability, the study aimed at characterizing in detail the composite clonal and subclonal architectures along with replication timing (RT) profiles associated with neuroblastoma tumorigenesis. Integration of single-cell RNA sequencing, we define the genotype to phenotype relationship. An independent WES (bulk) on 6 NB PDX models before and after (targeted) therapies was performed to study temporal genetic heterogeneity and identify clones and new mutations that might confer resistance to therapy.

3.1.1.3 Material and Methods

3.1.1.3.1 Preparation of the single-cell suspension for DNA sequencing

Samples were obtained from patients treated for neuroblastoma at Institut Curie, following informed consent. Patients were treated according to national or international treatment protocols. Patient derived xenograft (PDX) were established as reported previously (Marques Da Costa et al. 2023). PDX samples listed in Table 3.1.1.4.1 were included in this study. After reaching ethical size and following resection, PDX samples were immediately transferred for dissociation. On arrival, each sample was placed in a sterile Petri dish at room temperature. Tissue was subsequently minced to smaller pieces in a sterile less than 3 mm and digested by trypsin (10 mg/ml, Sigma Merck) during 1h at 37°C followed by inactivation by Soybean trypsin inhibitor (10 mg/ml, Sigma). Collagenase type II (100 mg/ml, serlabo) and DNase I (2 mg/ml, Sigma Merck) were added to cell solution to complete the tissue digestion. Subsequently, the sample was filtered using a 70-mm nylon mesh and centrifugated (300 × g in 1× PBS with 0.04% BSA) to discard supernatant. The cell pellet was resuspended in 1ml of medium, counted and resuspended in recommended media at accurate concentration.

For patients' biopsies, collected with informed consent from patients enrolled in this study, the same protocol is applied by adjusting the enzymes' concentrations according to the size of tumoral sample received.

3.1.1.3.2 Single-cell DNA library generation

The single-cell suspension was processed using chromium single-cell CNV solution (10× Genomics) as described in the user guide to generate a barcoded single-cell DNA library. Briefly,

single-cells were partitioned in a hydrogel matrix by combining with a CB polymer to form cell beads (CBs) using a microfluidic chip. After the first encapsulation, CBs were treated to lyse the encapsulated cells and denature the genomic DNA (gDNA). The denatured gDNA in the CB is then accessible to amplification and barcoding. A second microfluidic encapsulation step is required to partition the CB with 10× barcode gel beads (GBs) to generate an emulsion called GEMs. Immediately after barcoding and amplification, 10× barcoded fragments were pooled and attached to standard Illumina adaptors. Finally, sequencing libraries were quantified by qPCR before sequencing on the Illumina platform (Hi-Seq) to obtain a depth of coverage of 1X.

3.1.1.3.3 Single cell RNA sequencing library generation

Single-cell RNA sequencing (scRNA-seq) data was accessible for 16 out of these 18 tumor samples from Isabelle Janoueix-Lerosey's team at Institut Curie (Thirant et al. 2023). Single cell RNA sequencing was performed on using 10X genomic platform. The scRNA library preparation and sequencing details were described previously (Thirant et al. 2023). While scRNA-seq was conducted on the same patient-derived xenograft (PDX) passage or patient tumor, it was not performed on the exact same single cells.

3.1.1.3.4 In-vivo experiments for the study of clonal evolution

To investigate clonal evolution under targeted treatments, six different PDX models were treated in independent *in vivo* experiments. In each experiment, Swiss Nude mice were engrafted with tumors in their interscapular fat pad. The study planned for eight mice per group, aiming for a minimum of six treated mice in each experimental arm to ensure statistical significance. Randomization into six treatment groups occurred when tumor volumes reached 150-200 mm³. The study team monitored mouse weight and tumor size three times weekly. Mice were euthanized when tumors reached a predetermined ethical size limit of approximately 1,500-2,000 mm³. At the experiment's conclusion, DNA from the end of treatment, PDX tumor was subjected to whole-exome sequencing (WES). These details are summarized in Table 3.1.1.6.3. All procedures were approved by the Institutional Review Board of the Institut Curie. All animal experiments complied with current European/French legislation (articles R.214-87 to R.214-126 of the Decree n°2013-118 of February 1st) and were carried out in accredited animal facilities of the Institut Curie. Treatment schedules is detailed in Supplementary data section, 3.1.1.6.1.2.

3.1.1.3.5 Tumor DNA sampling, processing, and sequencing

PDX tumor tissue was flash-frozen after dissection and stored at -80°C. PDX tumor DNA was extracted by using the QIAamp DNA Mini Kit. DNA concentration after extraction was measured

by Qubit fluorometric assay with dsDNA BR Assay Kit. Tumor DNA was fragmented and then libraries constructed using Kapa Library Preparation Kit Illumina platforms (Kapa Biosystems) with Indexed Adapters included in SeqCap EZ Human Exome Kit v3.0 (Nimblegen Roche Sequencing). The manufacturer's protocol was modified as reported previously (Chicard et al. 2018). For exome capture SeqCap EZ Exome Enrichment Kit v3.0 (Nimblegen Roche Sequencing) was used according to the manufacturer's protocol, using Illumina Hi-seq2500 leading to paired-ends 100x100 bp (expected coverage: 100X). Paired patient's tumor DNA and germline DNA was available for all PDXs and was whole exome sequenced by Agilent SureSelect Human All Exon v5. Tumor DNA from patients was obtained from the same tumor specimen that was used to establish the corresponding PDX excepting for GR-NB10. In this particular case enrolled in the MAPPYACTS program (NCT02613962), the patient's tumor DNA sequenced corresponded to a later timepoint with respect to the establishment of the PDX.

In the experiments performed on IC-pPDX-17, GR-NB4 and GR-NB10, two tumors were sequenced per treatment arm, whereas in the experiments on HSJD-NB-005, IC-pPDX-75 and IC-pPDX-109 only one tumor per experimental arm was sequenced.

3.1.1.3.6 Single-cell DNA sequence data possessing

Chromium single cell DNA sequencing paired-end reads were processed using CellRanger-dna CNV pipeline v1.1 (10x Genomics, Inc). It consists of barcode aware read alignment, post-alignment processing as well as downstream copy number variation (CNV) identification. Reads were aligned over reference genome build GRCh37/hg19 using BWA-MEM mapping tool (Li and Durbin 2009). Post alignment preprocessing includes the step of Mark-duplicate per barcoded cell. Read count was obtained at fixed 20kb genome bin and then normalized according to mappability and GC content. Genomic regions with no reads or fewer reads (mappability < 70%) were removed from downstream analysis. Finally total copy numbers calls per 20kb bin length were calculated. CellRanger-dna CNV calling details can be browsed at https://support.10xgenomics.com/single-cell-dna/software/pipelines/latest/algorithms/cnv_calling

In our custom bioinformatics workflow, normalized count data per bin length was used to call CNVs using CopyKit pipeline. The summary statistics produced by CellRanger-dna pipeline includes data points for bin-to-bin ratio variation and ploidy inference per cell. Bin-to-bin ratio variation and number of breakpoints per cell inferred by CopyKit pipeline (Darlan Conterno Minussi et al. 2022) were then used to classify cells into G1/G2-phase and S-phase cells. The thresholds varied per tumor sample. Non tumoral cells were detected according to Gini coefficient less than 0.12 (R. Wang, Lin, and Jiang 2020) as well as curated manually (visualization). Cells

with complete silent copy number profile (no CN breakpoints) were classified as non-tumor cells. Cells exhibiting extensive dropout events, such as whole-chromosome or multiple chromosome arm level CN alterations, were categorized as noisy cells.

3.1.1.3.7 CNV calling at single cell level and inference of genomic architecture.

We integrated two comprehensive CNV analysis pipelines. First, CopyKit pipeline was used which performs CNV analysis at 250kb resolution and then estimates clonal substructure based on either segment ratio or integer copy number analysis. To delineate the genomic architecture of the tumors, Uniform Manifold Approximation and Projection (UMAP) was implemented (McInnes et al. 2018). Then HDBSCAN clustering was performed using the UMAP embedding (McInnes, Healy, and Astels 2017). Second, the CHISEL pipeline was used which is designed to perform allele- and haplotype specific copy number analysis (Zaccaria and Raphael 2021). Before running CHISEL, we prepared heterozygous single nucleotide polymorphism (SNPs) lists derived from variant analysis of pseudobulk (pooled single cells per tumor) WGS and germline bulk whole exome analysis. All these SNPs were identified as common in the population with gnomAD population database allele frequency, i.e. $\text{popmax} > 0.1\%$ (Q. Wang et al. 2020). These heterozygous SNPs were phased using Eagle v2.4 (Loh et al. 2016) and reference panel TOPMed r2 from TopMed imputation server (Taliun et al. 2021). Genomic coordinates were transformed from hg19 to hg38 or vice versa using CrossMap tool (Zhao et al. 2014). The phased SNP list was used by CHISEL to compute B-Allele frequency (BAF) across each bin in each cell and then to infer allele-specific copy number (ASCN) and haplotype specific copy number (HSCN). We used variable resolution ranging from 2Mb to 5 Mb. CHISEL applies hierarchical clustering to infer clones and subclones. The comprehensive bioinformatics workflow was outlined in Figure 3.1.1.6.22

3.1.1.3.8 Integer copy number and consensus profiles

The ploidy for each cell of a tumor sample was estimated by CellRanger-dna pipeline. Then CopyKit was used to constructs single cell integer copy number profiles by multiplying the segment ratios by ploidy and rounding the values to their nearest integer. To determine cluster stability, CopyKit evaluates the Jaccard Similarity using R package ‘fpc’ (v2.2-9) (Hennig, C. 2020) over a range from $k=10$ to the square root of the number of cells. The value that maximizes the Jaccard Similarity (default = ‘median’) is returned as the suggested value to be passed on for sample clustering algorithms. Subclonal consensus integer profiles were calculated by taking the median of every integer copy number of all the single cells that were assigned to the same subclone and rounded to the nearest integer.

3.1.1.3.9 Phylogenetic reconstruction of clonal lineages

Phylogenetic inference for subclonal consensus trees were performed using MEDICC2 (Kaufmann et al. 2022) based on the minimum event distances using total copy number parameter as an input. A diploid cell with copy number = 2 was added to the tree and designed as the root node, to root the tree. Common ancestors were estimated during tree construction as described in MEDICC2. The trees were plotted using R package ggtree v3.2.1 (Yu 2020)

3.1.1.3.10 Inference of most recent common ancestor (MRCA) profile

Inference of most recent common ancestor (MRCA) were inferred from the consensus copy number profiles (mean integer values of segments) of all the subclones from each sample during MEDICC2 tree construction. The consensus profile of each subclone was used to derive the most recent common ancestor (MRCA). For every segment, we selected the copy number (CN) value among the consensus CN values from each superclone to the average tumor ploidy as the ancestral segment.

3.1.1.3.11 Bulk whole exome and pseudobulk WGS data analysis

Joint variant calling to identify single nucleotide variants (SNVs) and small insertions and deletions was conducted on pseudobulk data (comprising pooled G1/G2-phase single cells WGSs from each tumor sample), whole exome sequencing (WES) of tumor samples, and germline data using Mutect2 in GATK v4.2.6.1 (DePristo et al. 2011). Variants deemed noisy were filtered out using the FilterMutectCalls tool from GATK. Additionally, allele-specific copy number analysis on both pseudobulk and bulk tumor samples was carried out utilizing the Sequenza tool. (Favero et al. 2015). Variants annotation was executed using the snpeff tool (Cingolani et al. 2012). Custom R scripts were used for annotated variant processing and analysis. Genome-wide variants exhibiting a CADD score greater than 10 and a gnomAD population maximum allele frequency below 0.1% were selected for further downstream analysis.

3.1.1.3.12 Mutation timing or Mutation multiplicity analysis

Genome-wide filtered single nucleotide variants (SNVs) were utilized to infer mutation timing. Initially, the number and size of clones/clusters were determined using the Mobster tool. (Caravagna, Heide, et al. 2020). MutationTimeR, an R package, was used to determine the timing of SNVs relative to clonal and subclonal copy number states. This approach measures the timing as a fraction of point mutations, following the methodology outlined by (Gerstung et al. 2020). The timing of clonal mutations, as well as the occurrence of whole chromosome gains or arm-level

gains (specifically on chromosomes 7 and 17/17q), and whole genome duplication events, were ascertained using pseudobulk whole genome sequencing (WGS) data.

3.1.1.3.13 Replication timing analysis using S-phase cell population identified from scWGS

To investigate replication timing at both single cell and pseudobulk levels, we delved into the S-phase cells obtained from single-cell whole genome sequencing (scWGS). For this purpose, we used the Kronos.scRT R package, which specializes in determining Replication Timing (RT) (Gnan et al. 2022). The normalized read counts of single-cell DNA sequencing were calculated per 250kb non-overlapping bin length. This particular bin size was chosen and adapted in accordance with the mean genome-wide coverage observed in our data.

3.1.1.3.14 Single-Cell RNA Sequencing: Genotyping and Copy Number Variation Analysis

In case of patient-derived xenografts (PDX) sequencing, mouse-derived reads were identified using the XenoCell software (Cheloni et al. 2021) and subsequently excluded from further analysis. Single-cell RNA sequencing (scRNA-seq) of tumor-specific cells underwent processing using the CellRanger pipeline, with paired-end reads mapped to the reference genome build GRCh37/hg19. The obtained filtered barcodes and count matrices were subsequently analyzed using the Seurat 4.3 tool (Hao et al. 2021). Basic quality control steps for single-cell RNA sequencing data were implemented to ensure the retention of high-quality cells. Criteria for selection included cells with a minimum of 1000 unique molecular identifier (UMI) counts, over 500 feature counts, and gene expression in at least 5% of cells. Additionally, cells with more than 25% mitochondrial gene expression were excluded. Doublets were identified and removed using the Scrublet tool (Wolock, Lopez, and Klein 2019). These good quality cells were then genotyped using cellsnp-lite v.1.2.2 tool (Huang and Huang 2021). Phased heterozygous single-nucleotide polymorphisms (SNPs) obtained from scWGS were used for genotyping. Allele- and Haplotype specific copy number analysis and clonal substructure was determined using numbat tool (T. Gao et al. 2022).

3.1.1.3.15 Bulk whole exome sequencing (WES) data analysis to study subclonal dynamics under targeted treatment

Exome sequencing paired-end reads were aligned to the human reference genome GRCh37/hg19. For PDX sequencing, a hybrid approach was used, aligning reads to both human GRCh37 and mouse mm38 genome builds. This alignment was executed using the BWA-MEM algorithm, ensuring that only reads uniquely mapped to the human genome with a mapping quality exceeding 35 were retained. Post-alignment, the average read depth observed from bulk exome sequencing of six PDX samples was approximately 110x per sample. However, two PDX samples, which demonstrated poor sequencing quality (mean read depth less than 30x), were excluded from the

post-treatment group of the GR-NB04 sample. Exome sequencing performed at the time of pre and post treatment for each treatment arm per tumor sample and paired germline were jointly analysed using Mutect2 algorithms from GATK v4.2.6.1. The resulting Mutect2 variant calls underwent rigorous filtering using the FilterMutectCalls tool. Any variant calls found in genomic regions of low complexity were systematically removed. Somatic single nucleotide variant (SNVs) with variant allele fraction (VAF) > 0.05 and the SNV position covered by total read depth > 20x in both pre and post treatment sample independently were included for the clonality inference. ASCN and ploidy analysis were performed using sequenza v3.08.

3.1.1.3.16 Inference of pre and post treatment subclonal dynamics

For a given PDX model, pairwise pre and post-treatment clonal and subclonal population structures were inferred by running beta-binomial distribution from PyClone-VI, defining parameters such as a maximum of 10 clusters and 1000 random restarts (Gillis and Roth 2020). Figures were plotted using R statistical environment and packages such as ggplot2, FishPlot and ClonEvol.

3.1.1.3.17 Gene set overlapping analysis

Clonal and subclonal genes were investigated for their significant overlap with gene sets in Molecular Signatures Database (MSigDB). e.g. gene ontology (go) database, canonical pathway database (e.g., Reactome or Kegg) or cancer hallmark database. Overlaps were computed with false discovery rate (FDR) q-value < 0.05.

3.1.1.3.18 Mutational Signature analysis

Mutational signature analysis was performed using deconstructSigs tool (Rosenthal et al. 2016)

3.1.1.4 Results

3.1.1.4.1 Study overview

Single-cell whole genome sequencing (scWGS) was performed on 10,469 cells from 4 patient tumor and 14 PDX models. All 4 patient tumor samples were directly analysed at the time of diagnosis. Samples for establishment of PDX were derived either at the time of diagnosis (n=3), disease progression (n=3) or relapse (n=8), and PDX tumors were then used for downstream analyses (Figure 3.1.1.4.1.A). All patients had INRG stage L2 or M/Ms disease, with a median age at diagnosis of 41 months (range, 17 to 216 months). The clinical information regarding patients is detailed in Supplementary data, Table 3.1.1.4.1. The scWGS runs summarizing total number cells obtained after sequencing, number reads mapped to reference genome (hg19), ploidy inference and other details were detailed in Table 3.1.1.6.1.

Out of the 10,467 single cells sequenced using scWGS, the cells were not sorted based on their cell cycle phase. After quality control (QC) checks (refer Methods section: Single-cell DNA sequence data possessing), single cells were classified into G1/G2-phase (n = 7804), S-phase (n = 1606), non-tumor (n = 817), and noisy cells (n = 240) (pie-chart Figure 3.1.1.4.1.B). Classification was based on two criteria: bin-to-bin segmentation ratio variability and the number of copy number (CN) breakpoints. Non-tumoral cells were exclusively identified in patient tumor samples (P1 = 57%, P2 = 4%, P3 = 62%, P4 = 20%), indicating contamination by non-tumoral cells concordant with histological reports. Both non-tumor and noisy cells were excluded from further analysis. On average, 17.2% of cells in the S-phase and 78.5% in the G1/G2 phase were retained for downstream analysis, as depicted in the stacked bar chart in Figure 3.1.1.4.1.B. Each tumor sample yielded an average of 434 (range, 4-963) G1/G2 cells, with a mean genome-wide read depth coverage of 0.25X, as illustrated in the accompanying box plot in Figure 3.1.1.4.1.B. The observed coverage was 5-fold higher than earlier reported 0.02x-0.05x in scDNA-seq data from 10x genomics platform (Andor et al. 2020) (Wu et al. 2021). Total number of cells per tumor sample according to qc and phase wise classification were detailed in Table 3.1.1.6.2. Ploidy estimations from scWGS were cross validated with bulk exome sequencing of the same tumor samples, demonstrating a strong correlation (R=0.94) as illustrated in Figure 3.1.1.4.1.C. For downstream analysis of clonal estimation, noise-free copy number (CN) profiles with accurate breakpoints from the G1/G2 cell population were selected, as presented in the upper CN profile of Figure 3.1.1.4.1.D. Unlike cells in G1 or G2 phases, which have stable copy numbers (two copies for diploid genomes), S-phase cells display variable copy number ratios due to the ongoing process of DNA replication. This makes them distinct in their CN profiles and characterized by high bin-to-bin segmentation

ratio variation, as depicted in the lower CN profile was essential for DNA replication timing studies, as depicted in the lower CN profile in Figure 3.1.1.4.1.D.

Table 3.1.1.4.1 Clinical information of the 18 patients

Sample id	Samples analysed by scWGS	PDX passage	Sampling time	Disease stage (INSS)	Age at diagnosis (months)	Patient gender
GR_NB10	PDX	p4	Relapse	4	66	M
GR_NB4	PDX	p5	Relapse	4	22	F
GR_NB5	PDX	p6	Relapse	4	48	M
GR_NB7	PDX	p6	Progression	4	17	M
HSJD-NB-003	PDX	p1	Diagnosis	4	32	F
HSJD-NB-004*	PDX	p5	Diagnosis	4	25	F
HSJD-NB-005*	PDX	p6	Progression (5 months after diagnosis)	4	25	F
HSJD-NB-009	PDX	p8	Relapse	4	41	M
HSJD-NB-011	PDX	p9	Relapse	4	19	M
IC-pPDX-109	PDX	p2	Progression	4	97	M
IC-pPDX-112	PDX	p4	Diagnosis	4	42	M
IC-pPDX-17	PDX	p10	Relapse	3	30	F
IC-pPDX-63	PDX	p22	Relapse	4	108	M
IC-pPDX-75	PDX	p6	Relapse	4	216	F
P1	Patient tumor	na	Diagnosis	4	78	M
P2	Patient tumor	na	Diagnosis	3	49	F
P3	Patient tumor	na	Diagnosis	4	40	M
P4	Patient tumor	na	Diagnosis	Localized	3	F

M=Male; F=Female; * PDX derived from same patient

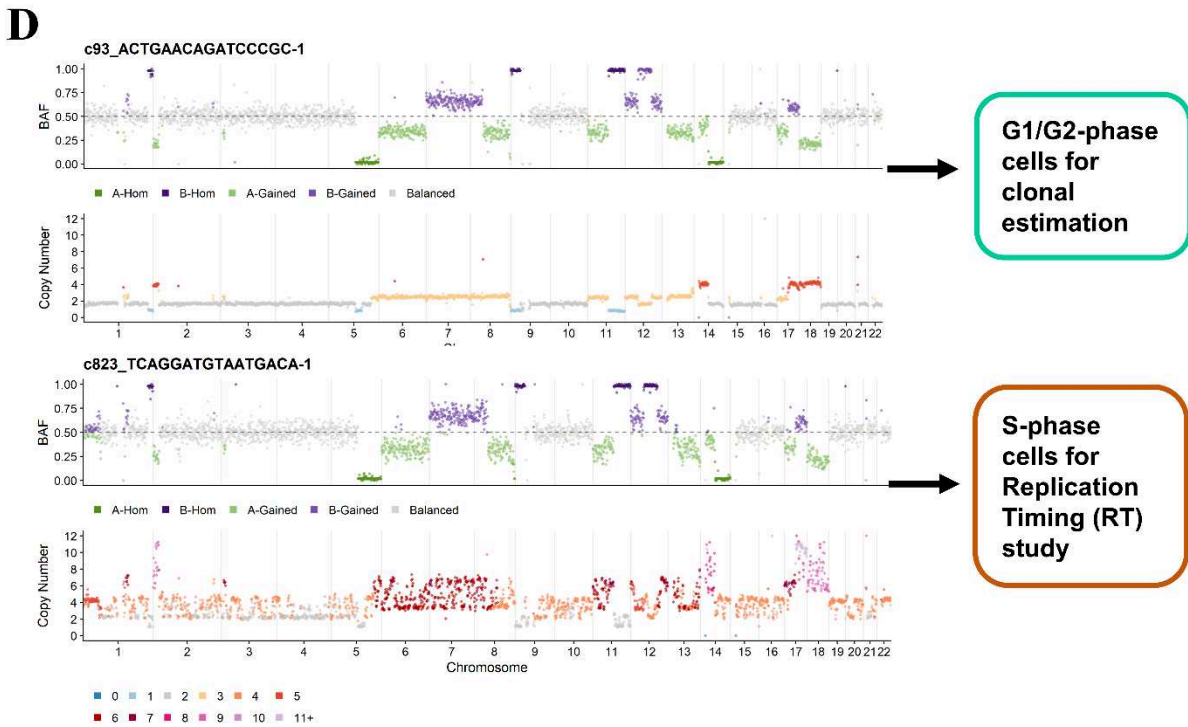
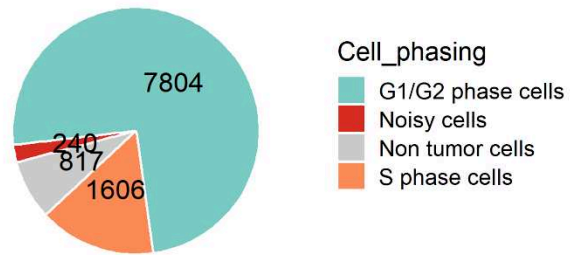
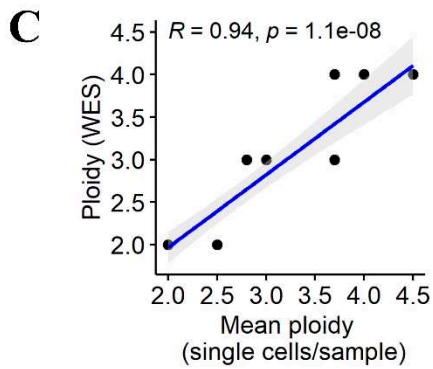
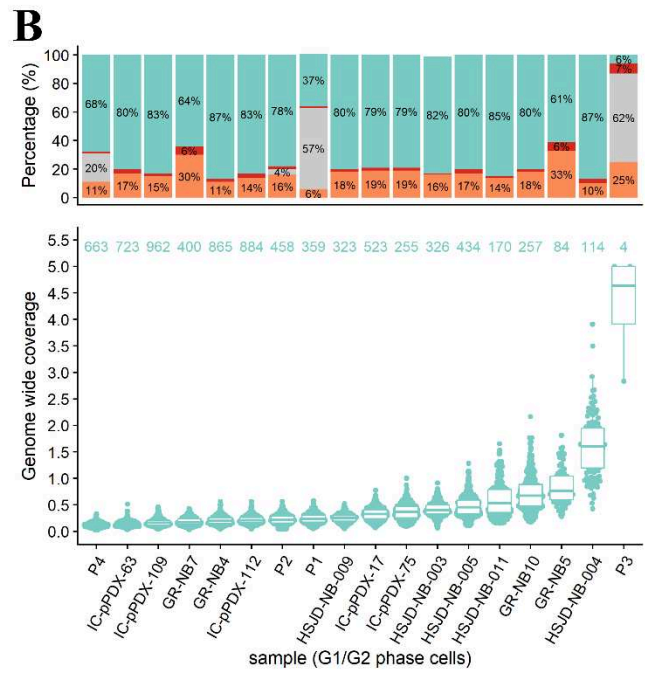
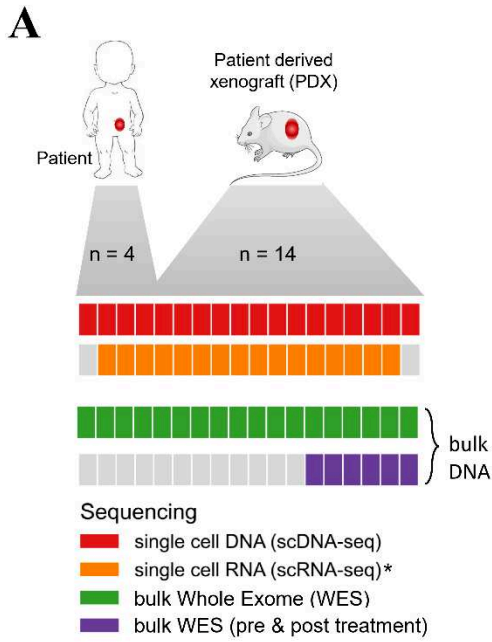


Figure 3.1.1.4.1 Comprehensive analysis of neuroblastoma (NB) genetic heterogeneity using single cell whole genome sequencing (scWGS)

A) Overview and Study Design, which applies single-cell whole genome sequencing (scWGS) to investigate the genetic heterogeneity in neuroblastoma. The study encompasses a total of 18 neuroblastoma tumor samples, of which 4 are freshly obtained from patients and the remaining 14 are patient-derived xenografts (PDXs). The genetic analysis of these samples is performed using a combination of techniques: single-cell DNA sequencing (scDNAseq) denoted in red, single-cell RNA sequencing (scRNAseq) indicated in orange, and bulk whole exome sequencing (WES) represented in green. A subset of the PDX samples, 6 out of 14, has been subjected to both pre-treatment and post-treatment WES to observe the genomic changes following therapeutic intervention, which is highlighted in violet. B) Data quality control outcomes for scDNAseq of 18 neuroblastoma samples. The stacked bar graph categorizes cell counts into G1/G2 population (light green), non-tumor cells (grey), noisy cells (red), and S-phase cells (orange). A box plot details the genome-wide average read depth in G1/G2 cells. A pie chart summarizes the total cell distribution across all samples, split by the same four categories. C) Comparative analysis of ploidy estimations, a robust positive correlation ($R=0.94$) between mean ploidy of single-cells per tumor with those derived from bulk whole exome sequencing (WES) for each tumor analysed. D) The upper panel displays allele-specific and total copy number profile of a single cell representing the G1/G2 cell population, which were used for clonal inference. The lower panel displays allele-specific and total copy number profile of a single cell representing the S-phase cell population, which were used for assessing DNA replication timing across the samples.

3.1.1.4.2 Genetic alterations defined at bulk level and summary of single cell DNA sequencing data analysis

In a first step, the genetic features of the 18 samples were determined by bulk WES sequencing, and then compared to the scWGS pseudobulk (pooled single cells) for each sample (Figure 3.1.1.4.1). *MYCN* amplification was observed in 10 cases, in 1 patient tumor (NB1576) and 9 PDX models, with *ALK* co-amplification observed in 2 of the PDX models (IC-pPDX-17 and GR-NB4). Two PDX models (IC-pPDX-17 and IC-pPDX-75) harbored CDK4 amplification co-occurred with *ATRX* missense and nonsense mutations respectively and co-amplification of MDM2 in IC-pPDX-17. Three PDX models harbored *ALK* pathogenic mutations, p.F1174L in IC-pPDX-75, p.E1419K in HSJD-NB-009 and p.I1171N in HSJD-NB-011. Other genetic alterations observed recurrently in neuroblastoma concerned an *NF1* mutation associated with LoH in HDJD-NB-004 and HDJD-NB-005, missense in HDJD-NB-011 and nonsense in GR-NB10. TP53 mutation was associated with LoH in HDJD-NB-004 and HDJD-NB-005. ARID1A mutation in HDJD-NB-005. All these amplifications and driver/targetable alterations were observed at a clonal level, with a cancer cell fraction > 90%. In 2 cases, a high tumor mutational burden was observed, with a TMB > 10 mutations/Mb in GR-NB10, and a TMB approximately 5 mutations/Mb in IC-pPDX-112. All other cases presented <1 mutation/Mb (Figure 3.1.1.4.1). Single base substitution (SBS), small insertion and deletions (ID) signatures (Ludmil B. Alexandrov et al. 2013) and Copy number (CN) signatures (Steele et al. 2022) were inferred based on whole exome and whole genome (pooled single cells) sequencing. The SBS5 signature frequently observed in NB (Brady et al. 2020) was observed in

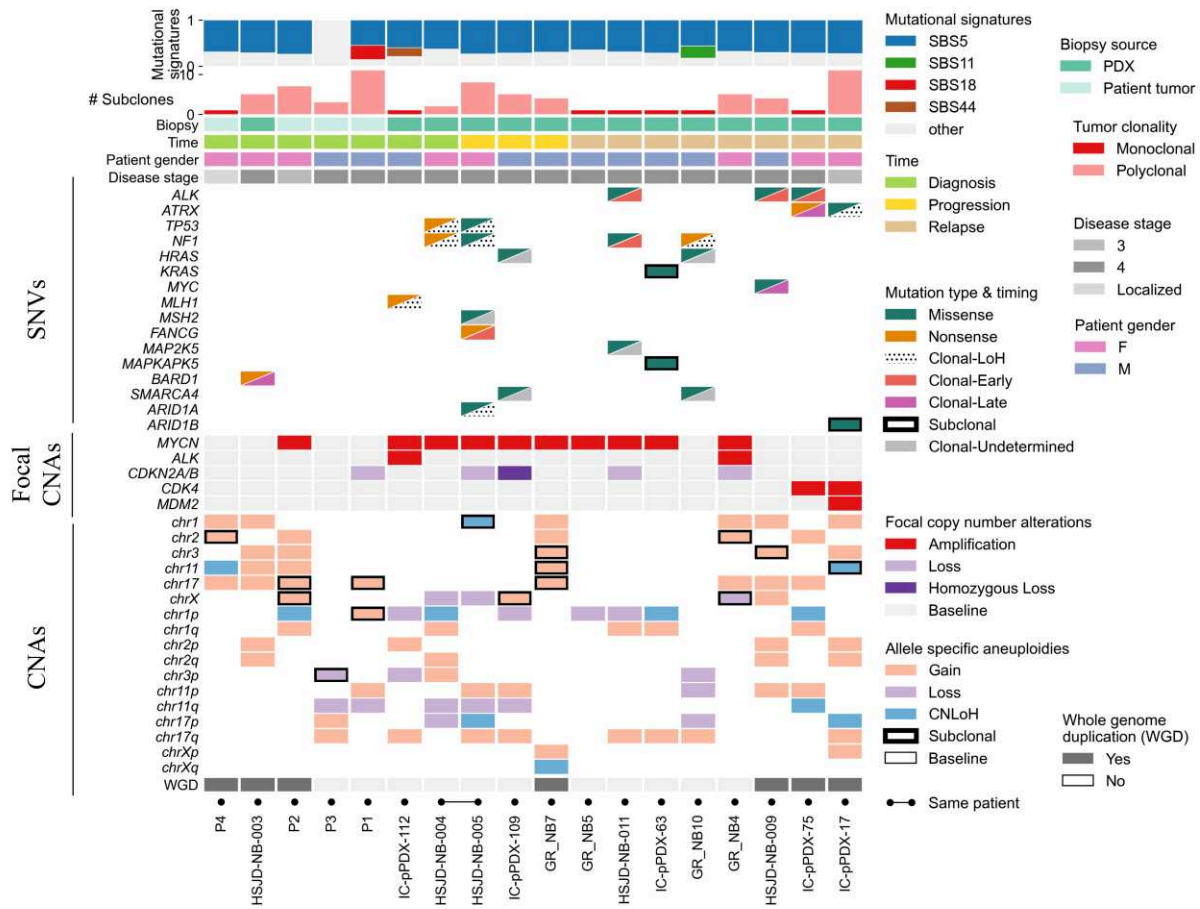
13/18 cases. SBS44 and ID5 signatures associated with defective DNA mismatch repair was observed in the IC-pPDX-112 model, and SNV and Indels analysis from IC-pPDX-112 model confirms microsatellite instability.

Whole genome duplication (WGD) events were inferred based on allele specific copy number analysis in 2 patient tumors (NB1576, NB1583) and 5 PDX models (MAP199, IC-pPDX-17, IC-pPDX-75, HSJD-NB-003 and HSJD-NB-009).

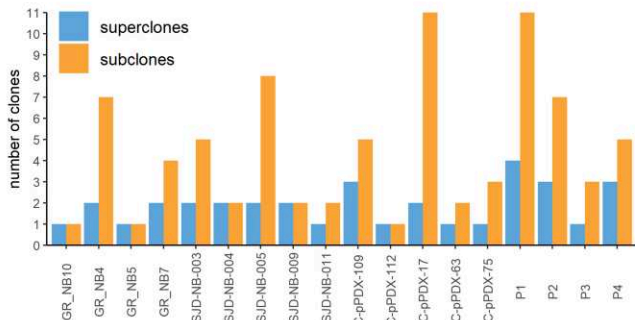
Allele- and haplotype-specific SCNAs established from the low-coverage WGS of the G1/G2 phase cells enabled to cluster cells into clones according to the inferred copy number profile. In total 1 to 11 subclones organized into 1 to 4 superclones were inferred per tumor sample (Figure 3.1.1.4.1.B). Subclones were defined as a minimum of 5 cells that share which highly similar copy number profiles and were < 90% in total cell population, representing a clonal expansion from a single genotype whereas superclones were defined as higher-order organization of subclone groups that share a subset of CNA events. Among the 18 samples, 7 were considered as monoclonal, and in the other 11 samples, from 1 - 4 super clones were identified, composed of up to 11 subclones. No correlation between the inferred clonal composition and clinical or genetic parameters was observed.

Interestingly a majority of subclones harbor copy number alterations known to occur recurrently in NB, such as deletion of chromosome 1p, 1q gain, 17q gain, whereas deletion of 11q was nearly always observed at a clonal level. Numerical chromosome alterations, concerning most frequently chromosome 7 gains, were observed at a clonal level.

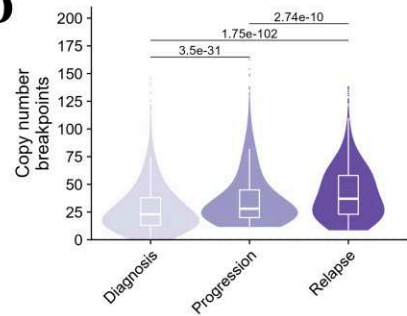
A



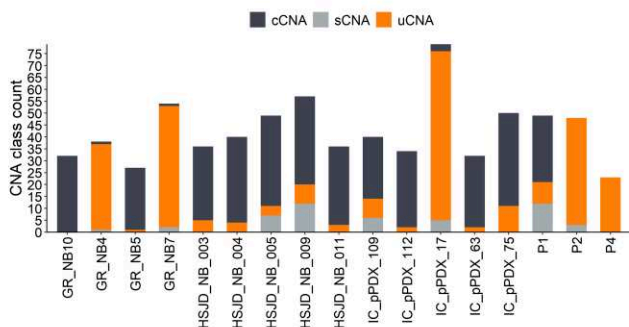
B



D



C



E

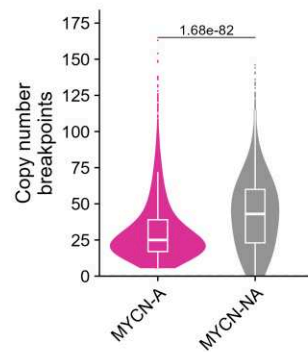


Figure 3.1.1.4.2: Summary of genetic alterations

A. Oncoplot summarizing genetic alterations which included somatic SNVs, focal amplifications and segmental alterations observed at bulk/pseudobulk analysis of 18 tumors. Top bar plot represents for mutational signature and number of subclones/clones observed per tumor sample. B. Bar chart summarizing number of subclones reported per tumor samples and their higher order of organization i.e. superclones. C. Bar chart summarizing clonal, subclonal and unique alterations across the single clones. D comparison of number of breakpoints observed at diagnosis, progression, and relapse. E. Copy number breakpoints comparison between MYCN amplification and tumor with no MYCN amplification.

3.1.1.4.3 Genomic Copy Number Architecture in Neuroblastoma Tumors

For each analysed sample, two independent tools were implemented into the analytical workflow to derive single cell copy number (CN) profiles and estimate clonal structures from the G1/G2 phase cell population. Initially, the clonal substructure was determined using CopyKit tool (Darlan Conterno Minussi et al. 2022) which facilitated the primary resolution of clonal composition. On an average, 1,550 reads per megabase (Mb) genomic window were reported, with a range extending from 195 to 9,211 reads per Mb. In our analysis, two distinct genomic window sizes were utilized, absolute copy number profiles were determined using a 220kb genomic window length using CopyKit tool, while allele- and haplotype-specific copy number profiles were obtained using a 2 to 5 Mb window length with the CHISEL tool (Zaccaria and Raphael 2021). The window length varied due to variable total number of G1/G2 cells per tumor sample and single nucleotide polymorphism (SNP) density which was essential for phasing and haplotype analysis. This approach ensured the cross-validation and robustness of our results.

Across all 18 tumors, clustering of the G1/G2 cell population based on total CN data identified 2-11 subclones (c1-c11) that were organized into 1-4 superclones (s1-s4) across the 18 tumors. Seven tumors IC-pPDX-112, IC-pPDX-63, IC-pPDX-75, GR_NB5, GR_NB10, HSJD-NB-011 and patient tumor P4 displayed minimal subclonal diversity, characterized as monoclonal genomes with over 90% of single cells exhibiting homogeneous copy number profiles in Supplementary data, Figure 3.1.1.6.1. In contrast, the remaining tumors exhibited a greater number of subclones, indicating higher genomic heterogeneity (Figure 3.1.1.4.3). In this study, 'subclones' were identified as groups of cells exhibiting closely similar copy number profiles, indicative of clonal expansion from a single genetic origin. 'Superclones' represent a more complex classification, encompassing groups of subclones that share a subset of copy number alteration (CNA) events (Darlan C. Minussi et al. 2021). This analysis revealed distinct spatial segregation of superclones in the reduced dimensionality space (UMAP), with discernible subclonal structures within several superclones across the 18 tumor samples as depicted in Supplementary data, Figure 3.1.1.6.2

Copy Number Aberrations (CNAs) were defined as genomic segments where a change in integer copy number values, either an increase or decrease, occurs in comparison to the tumor's baseline or 'neutral' DNA ploidy. Through CNA analysis, three primary categories were identified based on their prevalence within the tumor cell population: 1) Clonal CNAs (cCNAs): These CNAs are consistent across all subclones, indicating their presence in the entire tumor cell population. 2) Subclonal CNAs (sCNAs): These were present in a subset of tumor cells and identified in at least two subclones, suggesting a more restricted distribution within the tumor. 3) Unique CNAs (uCNAs): This subclass of sCNAs is characterized by distinct copy number states or breakpoints exclusively found in a single subclone, underlining their specificity. Notably, uCNAs underscore the diversity within sCNAs by representing unique copy number alterations restricted to individual subclones.

In patient P1, diagnosed at 6.5 years old, clustering of the single cell data revealed 11 subclones that were organized into 4 superclones (Figure 3.1.1.4.3.B. We observed clonal LoH across chromosomes 5p, 9p and 11q, as well as clonal gain on chromosomes 2p, 14q, 17q and 18, traced back to the most recent common ancestor (MRCA). We detected genomic events characterizing subclones across chromosomes 1p, 2p, 6, 8q, 12q, 17p and 20p. Many genomic regions with high copy gains and LoH events harbored genes known to be important in neuroblastoma oncogenesis, including but not limited to *MYCN*, *CDKNA*, *TP53*, *NF1*, *ARIDIA*, *ATM* and *MYC*. The genes with somatic mutations were highlighted according to their genomic location in the heatmap presented in Figure 3.1.1.4.3.B.

In patient P2 diagnosed at 4.1 years old, clustering of the single cell data revealed 7 subclones that were organized into 3 superclones (Figure 3.1.1.4.3.C. The observation of clonal amplifications of *MYCN* and loss of heterozygosity (LoH) regions across chromosomes 1p and 8q, traced back to the most recent common ancestor (MRCA), indicates early and genomic events in the tumor's evolution. Subclones c2, c3, and c4 were characterized by initial subclonal variations on chromosomes 14q and 17q. These subclones diverged early in the tumor's development but exhibited limited expansion, suggesting that they might have experienced weak selective pressure. The presence of distinct ploidy levels in subclones c1, c5, c6, and c7 suggests a complex evolutionary pattern. Subclone c1 displayed haploidy, c5 was triploid, while c6 and c7 both exhibited whole genome duplication (WGD) but originated from a common ancestor characterized by 22 numeric copy number alterations. This scenario is indicative of punctuated evolution, where some cells, potentially triggered by additional oncogenic events, underwent WGD. This leads to the emergence of these distinct subclones, each with its unique genomic profile, reflecting a sudden and significant evolutionary shift within the tumor's development.

In the analysis of patient tumor P3, diagnosed at 3.4 years old, an exceptional case was noted regarding the proportion of tumor cells. Out of 71 cells sequenced, only four were identified as tumor cells. Intriguingly, three out of four tumor cells exhibited distinct copy number profiles with clonal 17q gain and subclonal 11q loss, indicating a subclonal diversity within this small population of tumor cells (Figure 1.1.2.5.1). This finding suggests significant heterogeneity even within this limited number of tumor cells from patient tumor P3.

In patient P4, diagnosed at 3 months old, the clustering of single cell copy number data identified five distinct subclones, which were grouped into two superclones (Figure 3.1.1.6.1). top left panel. Notably, this patient's tumor displayed numeric segmental alterations accompanied by a whole genome duplication (WGD). We observed clonal copy neutral loss of heterozygosity (CN-LoH) on chromosomes 8, 11, and 19. Subclones c1 to c4 contained very few numbers of cells ranging from 14 to 19. These subclones were characterized by subclonal alterations at chromosome 2 and 17. However subclone c5, consisting of 585 cells which accounted for 90% of the tumor cell population, suggested a predominantly monoclonal nature. This dominant subclone's significant presence implies a linear evolutionary pattern within the tumor, contrasting with the diverse subclonal landscape exhibited by the smaller subclones c1 to c4.

In IC-pPDX-109 PDX derived at relapse, the clustering of single-cell copy number data revealed eleven distinct subclones, categorized into five subclones grouped three main superclones (Figure 3.1.1.4.1.A). Subclone c5 and c2 clearly stand out due to their unique genetic alterations as subclone c5 exhibited chromosome X duplication, while c2 exhibited whole chromosome 4 and 11p gain without 11q loss. Phylogeny analysis present a complex branching divergence due to chromosome X alterations. *MYCN* amplification, gains such as 3q, 16q and 17q and LoH regions such as 1p, 11q and homozygous deletion at 9p arm were clonal and early events mapped as MRCA.

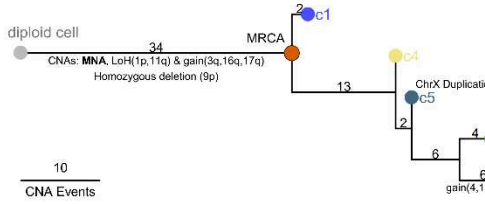
In IC-pPDX-17 PDX derived at relapse, the clustering of single-cell copy number data revealed eleven distinct subclones, categorized into two main superclones (Figure 3.1.1.4.1.D). Subclone c1, comprising 12 cells, diverged early with limited expansion due to weak selective pressure. It appears that some cells, potentially triggered by additional oncogenic events as MRCA, leading to a whole genome duplication (WGD) and significant burst of subclonal diversification from c2 to c11. This suggests that a single or limited number of catastrophic genomic events early in tumorigenesis are responsible for the emergence of these intermixed subclones (Vendramin, Litchfield, and Swanton 2021). We observed clonal co-amplification of *CDK4* and *MDM2*, along with pathogenic clonal mutations in *ATRX* and *CDK2* genes. Pathway analysis indicates the likely

activation of the PI3K-AKT-MTOR signaling pathway (p-value = 0.019), and/or MYC target V1 pathway (p-value = 0.01), during the early stages of tumor evolution. Other genes with potentially pathogenic mutations, such as *ETV1* and *MAFB*, which activate the KRAS signaling pathway (p-value = 0.01), and *GLI*, *YBX3*, *TGFB111*, *PU2F1*, and *KDM5A*, associated with transcriptional regulatory activity (p-value = 0.0031), appear to have played roles later in tumor evolution.

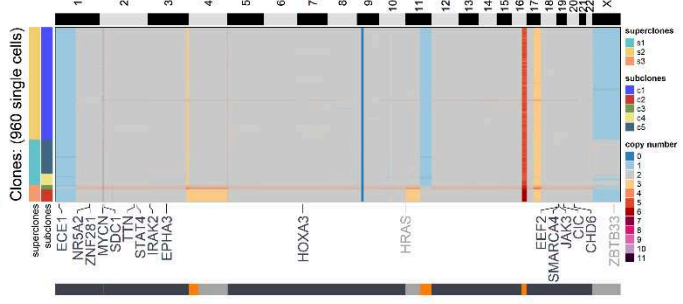
In GR-NB7 PDX derived at progression, the clustering of single-cell copy number data identified four distinct subclones, which were further organized into two overarching superclones. Ploidy analysis revealed that superclone s1 exhibits a triploid genome, while superclone s2 demonstrates hexaploidy (Figure 3.1.1.4.1.E). Phylogenetic analysis indicates that, following the most recent common ancestor (MRCA), these two superclones likely evolved in parallel, with superclone S2 / subclone C1 exhibiting 21 copy number breakpoints. Superclone s1, comprising subclones c2, c3, and c4, appears to have a shared ancestral origin, with shared 13 common copy number breakpoints observed among these subclones. Allele-specific copy number analysis revealed complex loss of heterozygosity (LoH) regions on chromosomes 1p, 11p, 11q, 14, and 17p. Notably, *MYCN* amplification was identified as a clonal event across the tumor cell population. Variant analysis did not reveal any mutation in known driver oncogene. However, protein serine/threonine kinases activity was observed after gene set overlap analysis (p-value=0.007).

To investigate the branching phylogenies post-MRCA in high-risk neuroblastoma, we conducted a comprehensive analysis by computing consensus Copy Number Alteration (CNA) profiles of subclones to construct balanced minimum evolution trees. We also integrated single-cell data by superclone groups to compute allele-specific copy number. This analysis revealed that the majority of Loss of Heterozygosity (LoH) regions were in alignment with the bulk exome sequencing data, with a median region overlap of 96.1%, indicating these events likely occurred prior to the MRCA. Across the 18 tumor samples, an average of 15.21% of the genome (ranging from 1.1% to 25.8%) exhibited LoH events. These findings collectively demonstrate that a significant number of subclonal CNAs (sCNAs) and unique CNAs (uCNAs) emerged post-MRCA, contributing to the diversification of clonal genotypes as the primary tumor mass expanded.

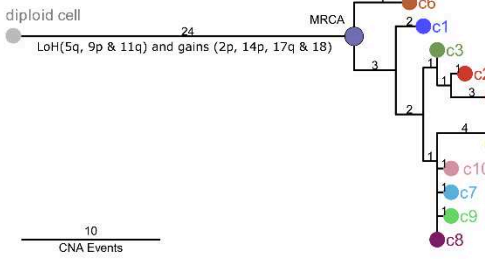
IC-pPDX-109



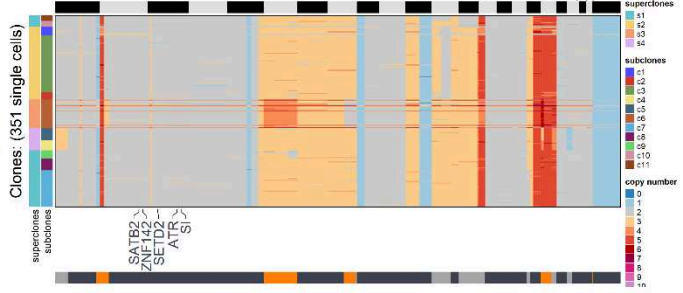
A



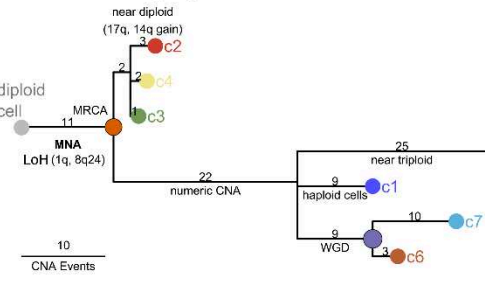
Patient tumor (P1)



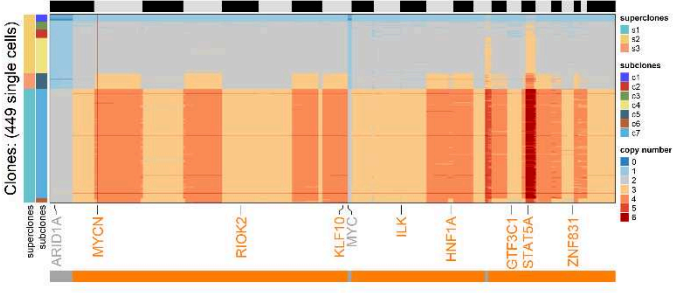
B



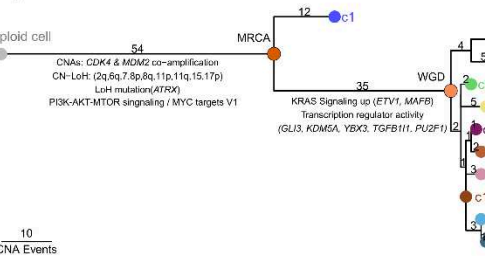
Patient tumor (P2)



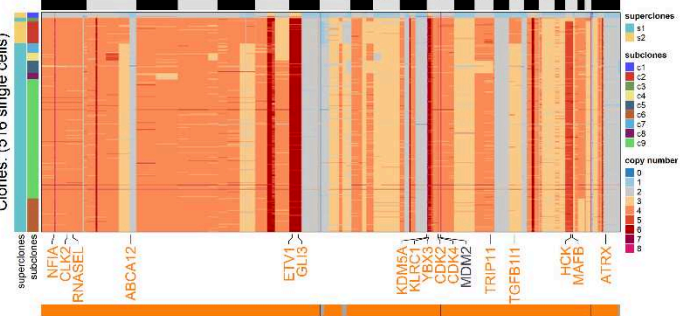
C



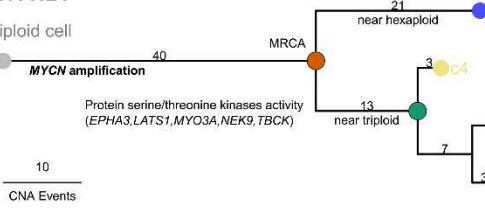
IC-pPDX-17



D



GR-NB7



E

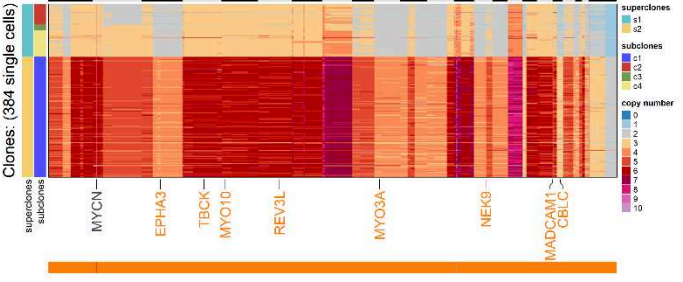


Figure 3.1.1.4.3 Clonal substructure and evolutionary analysis of clonal lineages in NB patient tumors and PDX models

Left panels show the event-based evolutionary trees of subclonal consensus integer copy-number profiles that are rooted by a diploid profile and annotated for MRCA, CNA, WGD events based on ploidy inference and mutations mapped according to their mutation timing during the evolution for two patients (P1 and P2) at diagnosis and three PDX models (IC-pPDX-109, IC-pPDX17 at relapse and GR-NB7 at progression). The first internal node in orange color on the tree represents the most recent common ancestor (MRCA). The right panels show clustered heatmaps of single-cell total copy-number profiles showing the profiles of subclonal clusters grouped into superclones, as represented in left-most header columns. Bottom annotation panels indicate clonal and subclonal classification of CNAs (clonal = black, subclonal = gray and unique subclonal CN = orange) and selected cancer gene annotations. Selected genes are either affected by SNV or CNA.

3.1.1.4.4 Longitudinal single cell sequencing shows distinct evolutionary patterns of SCNA from diagnosis to progression.

Two independent patient-derived xenografts (PDXs), labeled HSJD-NB-004 and HSJD-NB-005, were established from the same patient tumor, with HSJD-NB-004 derived at diagnosis and HSJD-NB-005 at the time of disease progression. Both PDX models underwent comprehensive genomic and transcriptomic analysis, including single-cell whole genome sequencing (scWGS), single-cell RNA sequencing (scRNA-seq), and bulk whole exome sequencing (WES).

The scWGS data analysis of the PDX models NB-004 and NB-005 provided a more intricate view of the clonal architecture, particularly through copy number variation analysis. We examined 114 single cells from the model obtained at diagnosis and 431 single cells from the model obtained at progression. The clustering of these cells indicated the evolution of the tumor's complexity: At diagnosis, two distinct subclones were identified, which then expanded into eight subclones by the time of progression. These subclones were organized into two superclones at both stages (Figure 3.1.1.4.4.A). A clonal amplification of *MYCN* was consistently observed throughout the disease's progression. Extensive and shared regions of loss of heterozygosity (LoH) at chromosomes 1p, 11q, 16q, and 17p were persistent features, suggesting their importance in the tumor's evolution. Notably, major chromosomal gains identified at diagnosis were largely absent at the progression stage, except for a gain on chromosome 4q. The presence of clonal mutations in genes like *TP53*, *NF1*, *NEK8*, and *KMT2A* hinted at the early activation of the MAPK signaling pathway during the onset of the disease. These critical genomic events were all mapped to the most recent common ancestor (MRCA), highlighting their early establishment during tumor evolution (phylogeny tree Figure 3.1.1.4.4.E).

The phylogeny analysis reveals branched evolution highlighting unique alterations at diagnosis and progression. This implies that the tumor cells present at the time of progression are not the same as

those at diagnosis; they have acquired additional changes that distinguish them from their predecessors. At diagnosis, subclones were characterized by unique alterations on chromosome 7. Subclone c1 exhibited copy-neutral loss of heterozygosity (CN-LoH) on chromosome 7, while subclone c2 was characterized by triploidy with common ancestor (internal node in green color) with shared 17 CN breakpoints (phylogeny tree Figure 3.1.1.4.4.E). Mutation in coding sequences of genes such as *AKAP9*, *PDGFRA*, *GSK3A* were unique at diagnosis. Gene set overlap analysis of genes altered uniquely at diagnosis suggest their enrichment for the P13-AKT signaling. Figure 3.1.1.4.4.B demonstrates allele specific copy number profile of both c1 and c2 subclones at diagnosis.

At progression, the subclones c1 to c8 were characterized by genetic alterations on chromosomes 1, 4, and 17. Figure 3.1.1.4.4.B demonstrates allele specific copy number profile of subclones c7, c7 and c8 subclones at progression with the red arrow pointing alterations specific to subclone with very few cells. Notably, subclones c1 to c8 shared a common ancestor, as indicated by an internal node (in purple), with 20 shared copy number CN breakpoints (Figure 3.1.1.4.4.E). This commonality was further defined by loss of heterozygosity (LoH) regions at chromosomes 5q, 9p, and 19q, and gains at chromosomes 11p and 19p. A key mutation in the DNA mismatch repair gene *FANCG* emerged at the progression stage, suggesting a role in the evolving genomic instability of the tumor. The mutation in *MSH2*, observed at both diagnosis and progression, implies its ongoing contribution to the tumorigenesis process. Additionally, mutations unique at progression, including in genes such as *TP73*, *GBX2*, *ETV6*, and *HNF4G*, underscore their potential importance in DNA binding and transcriptional regulatory activities. The sudden genomic changes from the diagnosis to progression stage, especially with the emergence of new mutations and alterations in genomic structure, suggest a punctuated evolution pattern.

Using allele-specific copy number analysis on scRNA-seq data, two unique genotypes (genotype 2 and 3) were discerned, differentiated primarily by a loss of heterozygosity (LoH) alteration on chromosome 13 (Figure 3.1.1.4.4.C). Alteration on chromosome 13 was notably absent at the DNA level, with the exception of three cells identified at progression, which were represented in subclone c5. Single-cell allele-specific copy number (ASCN) profiles derived from scWGS clearly demonstrated the loss on chromosome 13 in these three cells. Figure 3.1.1.4.4.B. In contrast, B-allele frequency (BAF) profiles derived from genotypes 2 and 3 in scRNA-seq data distinctly show that genotype 2 exhibits no loss on chromosome 13, while genotype 3 displays a clear loss of chromosome 13 (Figure 3.1.1.4.4.D). The combined analysis of scRNA-seq and scWGS indicates that even a subclone with a small number of cells, or rare cells, may possess the fitness and selective advantage to exhibit robust expression at the transcriptomic level.

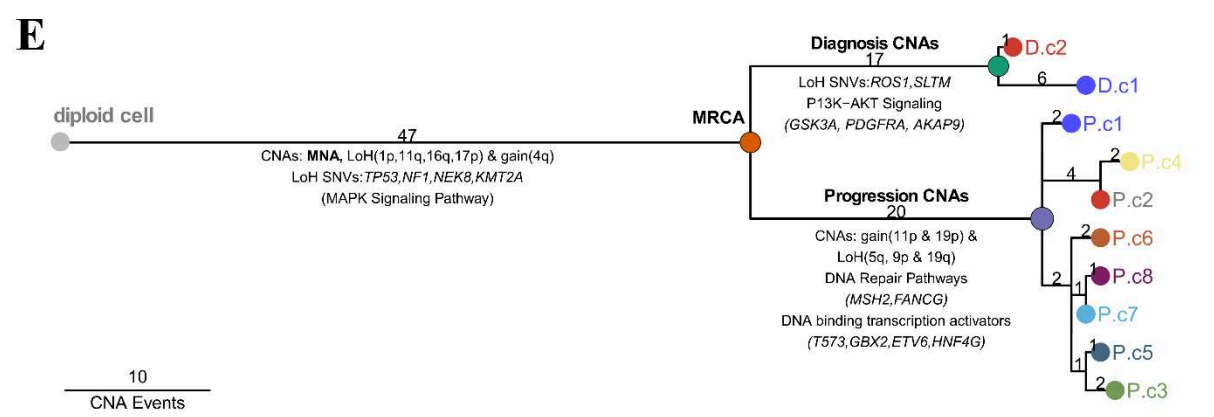
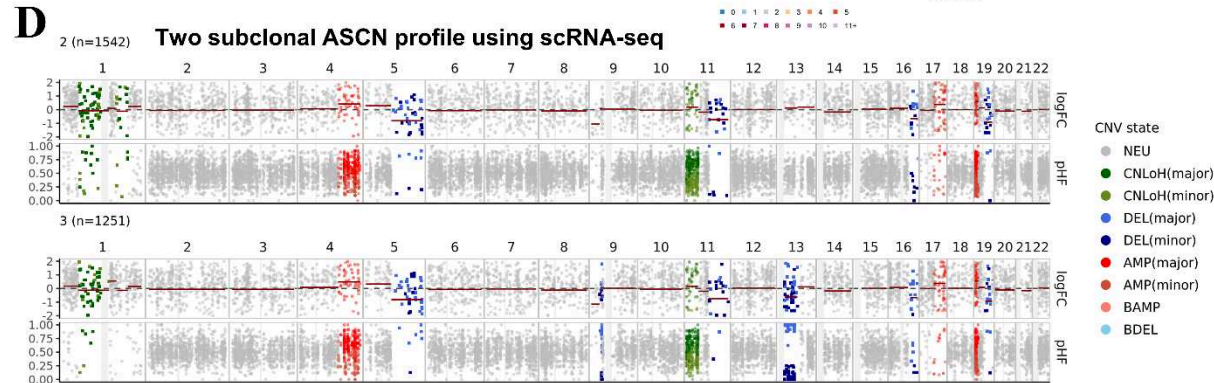
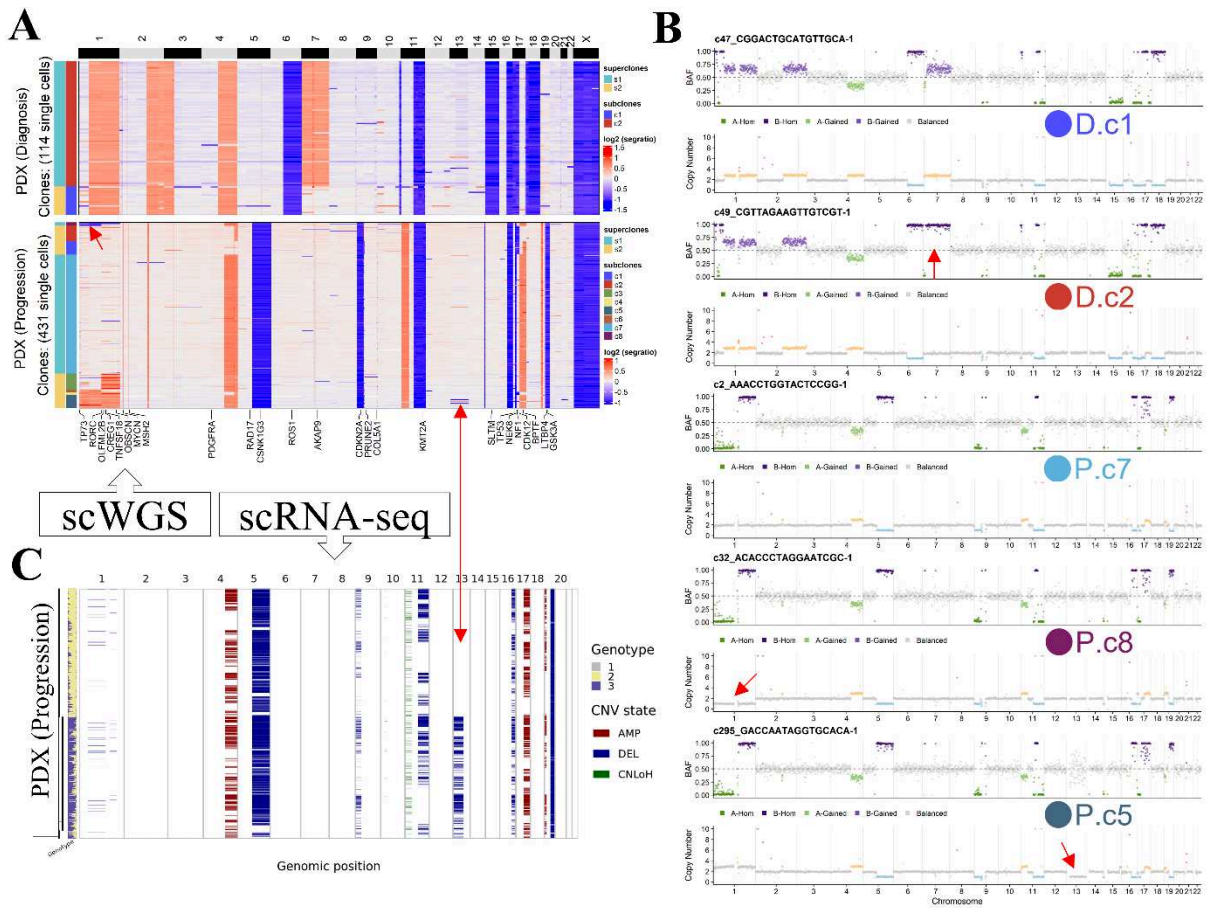


Figure 3.1.1.4.4 Investigating temporal heterogeneity by comparing genetic alterations at diagnosis and progression.

A. Heatmap of segmentation ratio data derived from single cell whole genome sequencing (scWGS) at diagnosis (upper panel) and Progression (lower panel). B. Individual single cell allele specific copy number profile representing subclones at diagnosis (c1 and c2) and progression (c5, c7 and c7). Red arrow highlights some alterations observed on subclone/rare cells. C. Allele- and haplotype specific copy number profile derived at progression using single cell at RNA sequencing. The red arrow which links scWGS CN profile and scRNA-seq CN profile indicates rare cell with unique alteration at chromosome 13 and its abundant expression in scRNA-seq. D. Pseudo bulk allele specific copy number profile of two genotype/clones inferred from scRNA-seq data. E. Phylogeny tree unfolding evolutionary history NB showing divergence of branching post-MRCA for diagnosis and progression.

3.1.1.4.5 Replication timing study using single cell DNA S-phase cell population in neuroblastoma.

We aimed to analyse replication timing as a possible source of ITH and driver of clonal and subclonal composition. To differentiate cells in the G1/G2 phase from those in the S phase, three assumptions were made: i) Majority of cells were from the G1/G2 phase. ii) Bin-to-bin variability is minimal in G1/G2 cells due to similar copy numbers across bins and maximal in S-phase cells due to asynchronous replication. iii) G1/G2 cells have fewer copy number breakpoints than S-phase cells (Gnan et al. 2022). Based on the cell ploidy, the intracellular bin-to-bin CN ratio variability and number of CN breakpoints single cells were used to classify cells into G1/G2-phase cells and S-phase cells. Copy number profiles from S-phase cells were normalized against G1/G2 cells from the same tumor to establish RT in S-phase cells. Quality control involves comparing tumor-specific (pseudobulk) RT profiles with the SK-N-SH neuroblastoma cell line, with observed high Spearman correlation values (>0.65) indicating accuracy and relevance. Three patient tumor (P1, P2 and P3) and two PDX samples (HSJD-NB-004 and HSJD-NB-011) were excluded from the replication timing (RT) analysis because they either had < 30 S-phase cells or showed a Spearman correlation coefficient < 0.65 when compared to the SK-N-SH cell line.

3.1.1.4.5.1 Dissecting replication timing variations between *MYCN*-amplified and non-amplified neuroblastomas

We conducted an analysis of replication timing (RT) in a series of 12 neuroblastoma patient-derived xenografts (PDX), utilizing both single-cell and pseudobulk level. Pseudobulk approach allowed us to average RT across populations of cells according to 250kb genomic windows or bin lengths and to determine early versus late replicating genomic regions. We then compared the early versus late replicating genomic regions between different tumor groups, based on key molecular characteristics, including the presence or absence of *MYCN* amplification and the occurrence of

whole genome duplication (WGD). As illustrated in Figure 7A, significant differences in RT were observed between the *MYCN* amplified (n=7) and non-amplified groups (n=5), as depicted in the density plots of pseudobulk RT profiles. Our findings suggest that neuroblastoma tumors without *MYCN* amplification (no MNA) are characterized by a predominance of late replicating domains, in contrast to *MYCN* amplified (MNA) tumors, which are enriched in early replicating domains (Figure 3.1.1.4.5.A). To quantitatively assess the differences in the distribution of these replication domains between the two groups, we employed the Kolmogorov-Smirnov (KS) test. The results, yielding a p-value of < 0.001 , indicate that the replication timing distributions of these two groups are significantly different.

In conducting dimensionality reduction analysis using Uniform Manifold Approximation and Projection (UMAP), we analyzed replication timing (scRT) profiles of S-phase cells in neuroblastoma. Our findings revealed two distinct trajectories of S-phase progression in the *MYCN*-amplified group, with cells arrayed from late to early replication stages. Conversely, in the *MYCN* non-amplified group, there was a clear bifurcation into two separate groups, indicative of early and late replication domains (Figure 3.1.1.4.5.C).

This variation in replication timing across the two groups underscores the impact of *MYCN* amplification on the cellular replication process. Specifically, the *MYCN*-amplified group exhibited a more linear replication progression, while the non-amplified group showed a dichotomous pattern, suggesting different regulatory mechanisms at play in these two contexts.

Furthermore, the analysis highlighted considerable variability in the extent of genome replication among different tumors, with the median percentage of the replicating genome ranging from 23% to 57% (Figure 3.1.1.4.5.B). This variability underscores the heterogeneity in replication timing across neuroblastoma tumors, which could be influenced by factors like genetic makeup, *MYCN* status, and possibly other underlying molecular mechanisms. Such insights can be crucial for understanding the biology of neuroblastoma and may have implications for therapeutic strategies targeting the replication machinery in these cancers.

3.1.1.4.5.2 Clone specific replication timing (RT) and its consequences towards ITH

Our investigation focused on the dynamics of DNA replication in the GR-NB4 NB PDX model derived at relapse. By examining the replication timing (RT) in single cells, we sought to understand whether variations in RT are observed in subclones which in turn might be linked to

subclone-specific proliferation. We specifically analyzed two distinct superclones, s1 and s2, which comprised 53 and 40 S-phase cells respectively. Superclones s1 and s2 exhibited distinct allele-specific copy number alterations, (Figure 3.1.1.4.5.D). For instance, s2 uniquely exhibited whole chromosome 2, 3 and 8 gains, while s1 was characterized by copy neutral loss of heterozygosity (CN-LoH) of chromosome 8 and a loss of heterozygosity (LoH) on chromosome X and 14q arm. Haplotype analysis suggested parallel evolution of these two superclones, marked by chromosome 17 allele-specific gains. Pseudobulk RT analysis of these two superclones revealed clone-specific distinct replication patterns, superclone s2 (in blue) showing late replicating domains and s1 (in maroon) demonstrating early replicating domains (Figure 3.1.1.4.5.E), which underlines differential replication programs within same tumor.

To quantify the replication variability within each superclone, we employed the concept of Twidth, defined as the time needed for a genomic region to be replicated in 25% to 75% of cells during a 10-hour S-phase (Dileep and Gilbert 2018). Our findings showed a Twidth range of 1.7 to 1.28 hours in s1 versus 1.65 to 1.95 hours in s2 for early and late-replicating regions, respectively (Figure 3.1.1.4.5.E - middle panel). Statistical analysis confirmed a significant difference between early and late Twidth ($p < 0.001$) in both clones. Further analysis of differential replication timing domains with distinct RT profiles between s1 and s2 revealed key neuroblastoma associated genes such as *ALK*, *MYC*, and *EGFR* were affected by these differential replication timings (Figure 3.1.1.4.5.E - heatmap). Figure 3.1.1.4.5.G highlights the chromosome 2p arm region (10 Mb to 50Mb), where the genomic loci of *MYCN* and *ALK* are indicated. The *ALK* gene locus exhibited significantly late replication in s2 versus early replication in s1. Genes in the late-replicating regions are often silenced or expressed at low levels. Therefore, we hypothesized that superclone s1 would be more transcriptionally active compared to late replicating s2 superclone.

Interestingly, when conducting a study to assess the impact of lorlatinib treatment, alone or in combination with chemotherapy, in this model (Figure 3.1.1.4.7), post-treatment bulk exome sequencing revealed a copy number profile in GR-NB4 that closely matched the pseudobulk CN profile of superclone s1, indicating a higher sensitivity of this model to lorlatinib treatment might come from superclone s2 which was late replicating.

In the IC-pPDX-109 PDX model, derived at relapse from a male neuroblastoma patient, we explored the intercellular replication variability by focusing on two superclones, s1 and s2, comprising 73 and 47 S-phase cells respectively. Notably, superclone s2, represented in sky-blue, displayed a duplication of chromosome X, in contrast to the expected single copy in superclone s1, depicted in orange (Figure 3.1.1.4.5.F). Pseudobulk replication timing (RT) analysis of these

superclones uncovered distinct replication patterns: s2 exhibited mid-replicating domains, while s1 showed predominantly early replicating domains, indicating divergent replication programs within the same tumor. Our analysis revealed that the Twidth, a measure of replication timing variability, ranged from 1.99 to 1.64 hours in s1 and from 1.53 to 2.13 hours in s2 for early and late-replicating regions, respectively (Figure 3.1.1.4.5.F - mid panel). Statistical testing confirmed significant differences between early and late Twidth values in both clones ($p < 0.001$), underscoring the distinct replication dynamics within each superclone. Further scrutiny of these differential replication timings between s1 and s2 unveiled specific replicating domains with unique RT profiles. This analysis highlighted the impact of RT variations on key genes, particularly CDK8 (Cyclin Dependent Kinase 8) and CD88 (Complement C5a Receptor 1) (Figure 3.1.1.4.5.F - heatmap). Specifically, the delayed replication timing in s2, potentially influenced by the duplication of chromosome X, suggests a unique genomic and epigenetic landscape that could contribute to the tumor's behavior.

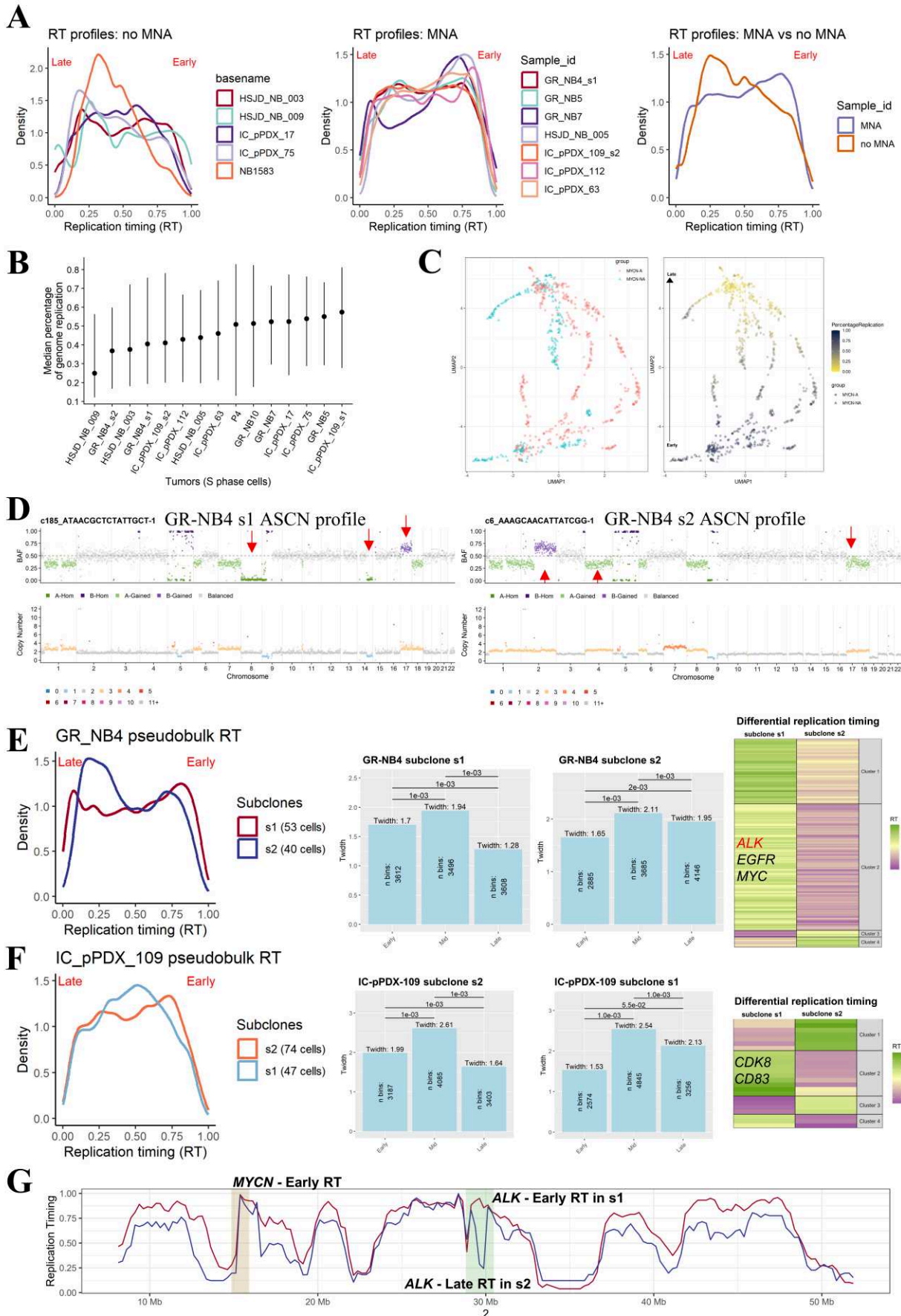


Figure 3.1.1.4.5 Single cell DNA replication timing (RT) profiling

A. Replication timing profiles determined at pseudobulk of each 12 tumor samples grouped according *MYCN* amplification status and compared their RT. These density plots, with X-axis values ranging from 0 to 1, demonstrated that replication timing peaks near zero indicated late replicating domains, while those closer to one represented early replicating domains. B. Inferred median percentage of genome replication per tumor sample with inter quartile range. C. Dimensionality reduction (UMAP) analysis of single cell RT data for the indicated two groups MNA (amplification) and no MNA (no amplification). The second panel of UMAP color-coded based on the on the genome replication percentage at single cell level. D. Two single cells copy number profile representing GR-NB4 subclone s1 and s2 with clone specific CN alterations. (B, F). Twidths calculated for the indicated three RT categories (Early, Mid and Late) based on the pseudo-bulk RT values in the two subclones (s1 and s2) identified in GR-NB4 and IC-pPDX-109. P-values were calculated using the Kronos scRT R package. G. Comparison between the GR-NB4 subclone s1 (marron solid line) and subclone s2 (blue solid line) RT at chromosome 2, zooming 10 Mb to 50 Mb genomic region which encompasses *MYCN* and *ALK* genes.

3.1.1.4.6 Parallel clonal evolution and subclone specific expression

In the GR-NB4 neuroblastoma (NB) patient-derived xenograft (PDX) model, we observed a unique genetic profile characterized by co-amplification of *MYCN* and *ALK* but no driver mutations in genes typically associated with NB. Intriguingly, differential replication timing was noted between two superclones, s1 and s2, leading to the hypothesis that the early replicating superclone s1 might exhibit higher transcriptional activity compared to the late replicating superclone s2.

To delve deeper into the characteristics of these two superclones and their respective subclones, we analyzed a copy number heatmap derived from single-cell whole-genome sequencing (scWGS) and single cell RNA sequencing data (Figure 3.1.1.4.6.A & B) respectively. The heatmap from panel A, based on absolute copy number analysis, displayed 865 single cells clustered into seven distinct subclones. Within this framework, subclones c6 and c7 were categorized under superclone s1, while subclones c1 to c5 were assigned to superclone s2. Allele-specific and haplotype-specific analyses revealed an interesting pattern of allelic gains across the genome, particularly on chromosome 17, where a combination of alleles ABB in superclone s1 and AAB in superclone s2 was observed, suggesting a parallel evolution of these superclones (Supplementary data, Figure 3.1.1.6.4). This is supported by phylogenetic analysis, which indicates a branched evolutionary pattern in this PDX model, in line with the allele-specific gains on chromosome 17. The most recent common ancestor (MRCA) was characterized by amplifications in *MYCN* and *ALK*, gains on chromosomes 1, 6p, 7, 17, and losses of heterozygosity (LoH) on chromosomes 5q and 9p (Figure 3.1.1.4.6.C). Diverging from this, the recent ancestor s1 defining the evolution of superclone s1 featured additional genomic alterations such as gain on chromosome 11p, loss on 14q, CN-LoH on 8 chromosomes, and a whole chromosome X deletion. Unique to subclone c7 (comprising 9 cells)

was the gain of the entire chromosome 12. Conversely, the recent ancestor s2, which delineates the evolutionary path of superclone s2, was marked by gains on chromosomes 2, 3, 8, and 17 (with major allele-specific gain) noted.

We then sought to correlate genomic data derived from single-cell whole-genome sequencing (scWGS) with single-cell RNA sequencing (scRNA-seq) data. This integration was achieved by utilizing phased single nucleotide polymorphisms (SNPs) identified in the scWGS to genotype the scRNA-seq data. The numbat tool was used to infer allele and haplotype specific copy numbers from the scRNA-seq data, aligning closely with the allele-specific copy number profile of superclone s1 observed in the scDNA-seq (Figure 3.1.1.4.6.A & B) (Figure 3.1.1.6.20 & Figure 3.1.1.6.21).

A striking observation from the scRNA-seq heatmap was the preferential expression of subclone c7, which is characterized by a gain of the entire chromosomes 2 and 12. This genomic alteration was prominently expressed at the transcriptomic level, particularly in cells of genotypes 2 and 3. Furthermore, other subclonal alterations specific to superclone s1, such as the loss of chromosome 14q, were abundantly expressed in genotype 2 cell population. The pseudobulk profile highlights all three genotype specific alterations inferred from scRNA-seq data (Figure 3.1.1.4.6.D).

Conversely, genomic alterations unique to superclone s2, like the gain on chromosome 3, were noticeably absent from the transcriptomic profile. However, gain at chromosome 2 observed was abundantly expressed and raises intriguing questions about the differential expression patterns and potential mechanisms driving these variations. One hypothesis to explain these findings is the potential overexpression of *MYCN* and *ALK* due to the amplification and resultant dosage effect on the copy number. This hypothesis is supported by the observation that superclone s1, characterized by such amplifications in addition to LoH regions (CH-LoH at chromosome 8 and LoH at 14q) is more prevalently expressed in the transcriptome compared to superclone s2.

These insights gleaned from integrating scRNA-seq and scWGS data not only enhance our understanding of the clonal dynamics within neuroblastoma but also reveal the complex interplay between subclones (defined by unique set of genetic alteration and differential replication timing) and their transcriptional outcomes. Next, we sought to investigate subclonal s1 and s2 behavior under targeted treatment or chemotherapy as GR-NB4 was included among six PDX included in treatment analysis.

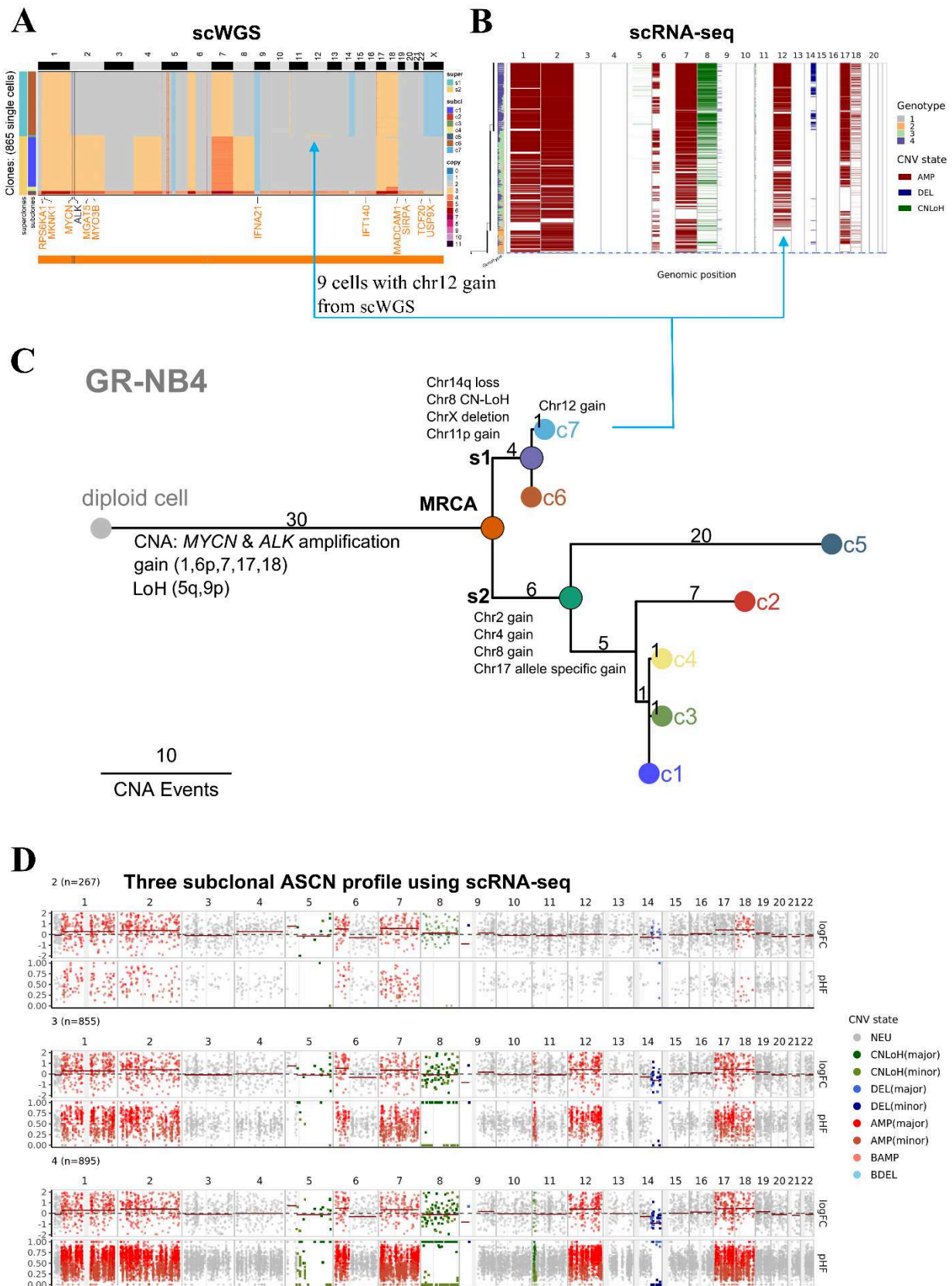


Figure 3.1.1.4.6 GR-NB4 PDX model showing parallel clonal evolution.

A. Heatmap of total copy number data derived from 865 single cell whole genome sequencing (scWGS) using GR-NB4 model. B. Heatmap of copy number data derived from single cell RNA sequencing using GR-NB4 model. C. Phylogeny tree showing tumor evolutionary trajectories. C. Allele specific copy number profile derived from the three genotypes inferred from scRNA sequencing data.

3.1.1.4.7 In Vivo Investigations of Clonal Evolution During Treatment

3.1.1.4.7.1 Analysis of treatment efficacy

We conducted deep whole-exome sequencing (mean coverage $\geq 100x$) pre and post treatment. In our investigation of clonal evolution, we administered targeted therapies to align with the distinct molecular characteristics of each patient-derived xenograft (PDX) model. The post-treatment genetic profiles were then compared with the pre-treatment variations, to discern the molecular changes induced by therapeutic intervention. The comprehensive workflow of these experiments, from engraftment to sequencing, is detailed in Figure 3.1.1.4.7.A.

For instance, lorlatinib, a potent *ALK* inhibitor, demonstrated remarkable antitumor activity (91% tumor growth inhibition (TGI) in the *ALK*-amplified GR-NB4 model, signifying its efficacy in this genetic context. Conversely, the same drug failed to show effectiveness in the IC-pPDX-75 model, which harbored an *ALK* F1174L mutation, with a much lower TGI of 28%. This stark difference in drug response highlights the nuanced nature of targeted therapy, where the presence of specific mutations can dramatically influence the outcome.

In another model, IC-pPDX-109, characterized by an *HRAS* mutation, trametinib a MEK inhibitor demonstrated significant antitumor efficacy with a TGI of 60%. However, in the GR-NB10 model, which presented with an *NFI* mutation coupled with loss of heterozygosity (LoH), trametinib's efficacy was notably diminished, with a TGI of 49%. Additionally, ribociclib, an inhibitor targeting *CDK4*, exhibited substantial efficacy against the *CDK4*-amplified IC-pPDX-17 model, achieving a TGI of 61% and trametinib, an inhibitor of *ARID1A* exhibited substantial efficacy against the *ARID1A* mutated HSJD-NB-005 PDX model, achieving a TGI of 61%

Concerning chemotherapies, etoposide-cisplatin (Chemotherapy-B) were efficient in models GR-NB4 and IC-pPDX-109 (TGI 93.7% and 67% respectively) and doxorubicin-cyclophosphamide (Chemotherapy-A) against models IC-pPDX-17 and GR-NB4 (TGI 70% and 76% respectively). The addition of a targeted treatment to chemotherapy was significantly more efficient than chemotherapy alone in models GR-NB4 (when lorlatinib was added to both chemotherapies), GR-NB10 (when trametinib was added to doxorubicin cyclophosphamide) and IC-pPDX-109 (when lorlatinib was added to doxorubicin-cyclophosphamide). In models IC-

pPDX-17 and IC-pPDX-75, the addition of targeted therapy (ribociclib and lorlatinib, respectively) did not add any significant efficacy to any of the two chemotherapies. Finally, we could find that GR-NB4 was globally sensitive to all treatments, whereas IC-pPDX-75 was globally very resistant. These results are summarized in Table 3.1.1.6.4

3.1.1.4.7.2 Subclonal evolutionary dynamic under treatment

Clonality in six PDXs models were assessed based on the variant allele fraction (VAF) of single nucleotide variants (SNVs) and allele-specific copy number analysis performed pre and post treatment. Using PyClone-VI, we identified distinct clusters per treatment arm, each representing a different subpopulation within the tumor (Figure 1.1.2.5.1.C). The clusters highlighted the dynamic evolution of the tumor: Parental clone in blue, with a cancer cell fraction (CCF) of 100% was persistent before and after treatment. We defined clonal events as those present in a high proportion of cancer cells. Conversely, we categorized events with a CCF of less than 90% as subclonal (McGranahan et al. 2015).

To assess the variance in Variant Allele Frequency (VAF) for each clone, we compared all variants within the clones between the pre- and post-treatment groups. This comparison was conducted using a t-test to statistically evaluate the differences in VAF before and after treatment. Clones observed in the study were categorized into three distinct groups based on their modification from pre to post treatment: 1. Persistent variants/clones, which were consistent across pre- and post-treatment groups with minimal variation in Variant Allele Frequency (VAF), 2. Emerging variants/clones, characterized by a significant increase in VAF from 0% or less than 5% before treatment to more than 5% after treatment, and 3. Diminishing variants/clones, identified by a significant decrease in VAF from greater than 5% before treatment to a lesser extent after the treatment. Detailed comparison from each treatment arm for all 6 models are represented in Supplementary data, Figure 3.1.1.6.25 to Figure 3.1.1.6.60

Delving deeper into GR-NB4 model, Figure 1C showcases the subclonal dynamics in terms of VAF and CCF after lorlatinib alone or in combination with chemotherapy-A targeting *ALK* amplification. Through clonal evolution analysis, seven clones were inferred. Post lorlatinib treatment, cluster 1, 2 and 3 were persistent, showing no significant deviation between pre- and post-treatment. Subclonal cluster 4 (13/133) and 6 (26/133) were significantly diminishing and a new cluster 7 (26/133) emerged. Similarly, when lorlatinib was combined with chemotherapy-A, cluster 1 and 2 were clonal and persistent, cluster 3 (26/139) was significantly diminishing and the other three clusters, 5, 6 and 7 significantly emerged (43/139).

In the GR-NB4 using single-cell DNA sequencing parallel clonal evolution among two subclones s1 (#cells, 411/865) and s2 (#cells, 454/865) were demonstrated (see also section 3.1.1.4.6). Pseudobulk allele specific copy number profiles of subclones s1 and s2 are presented in Figure 3.1.1.4.5.

Furthermore, the SNVs were analysed in pseudobulk of subclones s1 and s2 independently. Then the SNVs from the bulk exome analysis of pre ad post-treatment were mapped to SNVs obtained from pseudobulk of subclones s1 and s2. We observed that the diminishing variants were particularly concentrated in subclone s2 as depicted in Figure 3.1.1.6.30, suggesting its heightened sensitivity to the treatment. Further, CNA exposed segmental level ploidy rectification post-treatment, mirroring the copy number profile of subclone s1 (Figure 3.1.1.4.7.D), and indicating subclone specific sensitivity to lorlatinib. With *ALK* amplification remaining unchanged in all instances, other molecular mechanisms are most likely at the origin of the subclone specific treatment sensitivity.

After analyzing Single Nucleotide Variants (SNVs) from all six treated mice, we observed the majority, approximately 9% on average, of SNVs observed post-treatment were unique to the specific tumor and its treatment regimen. In contrast, diminishing variants and genes, comprising about 4% of the observed variants, were not exclusive to a single treatment but shared across at least two different treatment arms. This indicates a level of commonality in the genetic response of tumors to various treatments. Remarkably, the largest proportion of SNVs, accounting for approximately 75%, remained persistent even after treatment (Figure 3.1.1.4.7.E). This persistence suggests a significant level of genetic stability or resistance to change in these tumors, despite therapeutic interventions.

A detailed review of the remaining PDX models showed numerous SNVs either diminishing or emerging post various treatments. These dynamics can be further explored in Supplementary Figure 3.1.1.6.25 to Figure 3.1.1.6.60. However, the focal amplification status of *ALK* and *CDK4* and the clonal attributes of the SNVs (*ALK*, *ARID1A*, *NF1* and *HRAS*) showed no deviation post-treatment.

Discussion about the main driver genomic events i.e. amplification is really at the origin of the oncogenesis and not variable. or the SNVs modifications which we observe might also be stochastically "drifts" which might or might not be of biological relevance.

SNVs merged from all six PDX models, were classified into three distinct groups based on their response dynamics: Group A (persistent genes), Group B (diminishing genes), and Group C (emerging genes) for gene set overlap analysis. Interestingly, unique gene set enrichments were

evident in Group C post-treatments. For instance, genes regulating immune effector processes stood out post chemotherapy-A treatment, while transcription regulator binding activity emerged post chemotherapy-B. Furthermore, genes associated with transmembrane transport activity specifically neurotransmitter receptors linked to the postsynaptic specialization membrane came to the fore after targeted treatment. Genes pivotal to mitotic spindle and neurogenesis were enriched post the targeted treatment combined with chemotherapy-B (Figure 3.1.1.4.7.D)

Although this analysis indicated shifts in subclonal composition following targeted treatment and chemotherapy. As expected, no common SNVs, targeting genes recurrently altered after a given treatment were identified, but an enrichment in pathways recurrently altered as observed in emerging and diminishing clones.

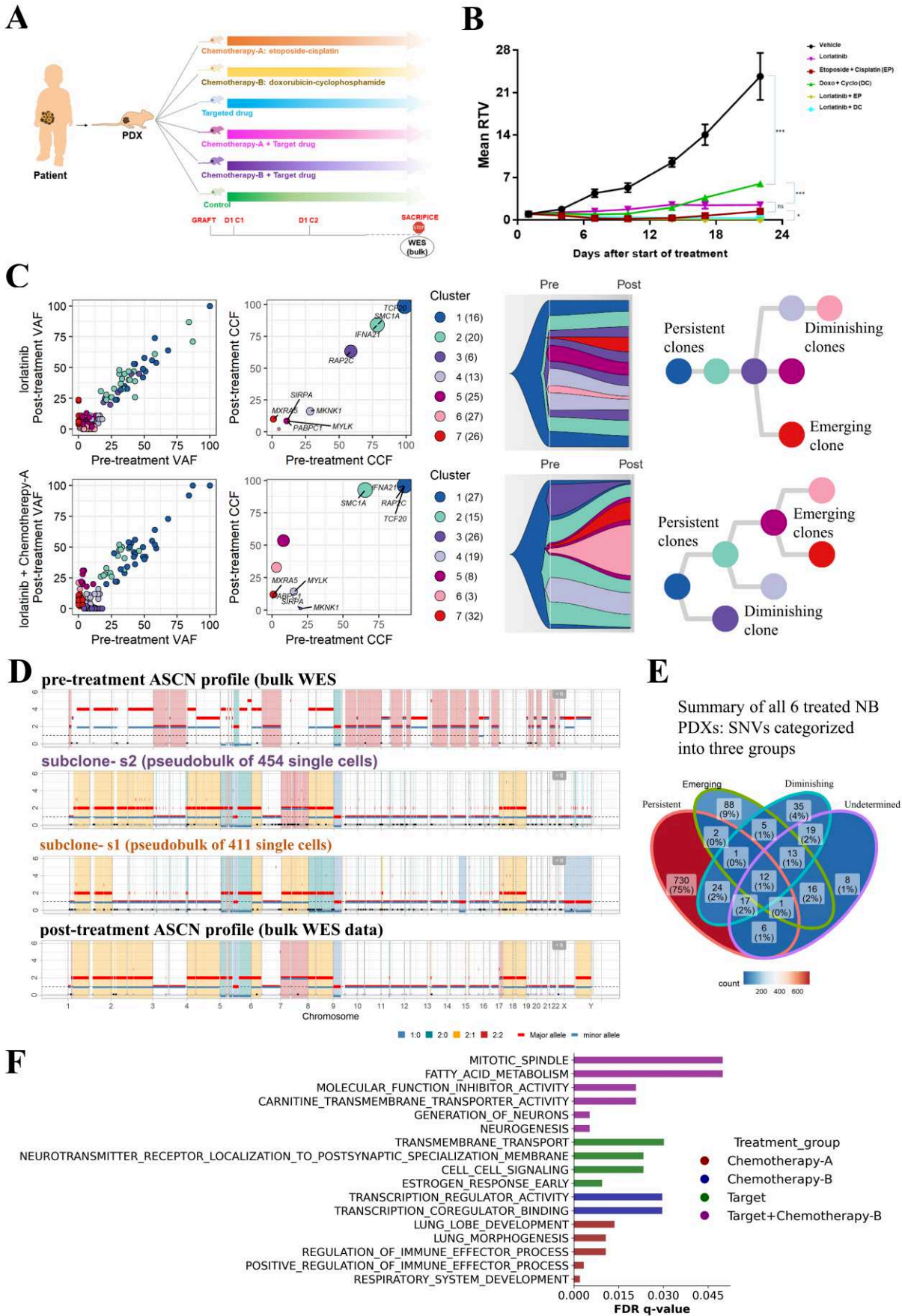


Figure 3.1.1.4.7 Clonal Evolution During Treatment

(A) In vivo experimental design for six high-risk neuroblastoma PDX models. Post-engraftment, mice were distributed among six experimental treatment arms, receiving targeted therapy, chemotherapy, or a combination of both. (B) relative tumor volume curves for the GR-NB4 xenograft model. A striking reduction in tumor volume was observed post-treatment, especially within the groups treated with lorlatinib alone and in combination with doxorubicin (chemotherapy-A). In contrast, the control group exhibited uninterrupted tumor growth. (C) Genetic clonal dynamics observed both at pre and post-treatment. Clusters were derived from the binomial distribution of the cancer cell fraction, and the resultant maximum parsimony trees elucidate the clonal trajectories. (D) Allele-specific copy number alterations, identified via bulk exome sequencing, are analyzed. Post-treatment, we observed a noteworthy ploidy correction. Additionally, the single-cell DNA sequencing of the GR-NB4 PDX model spotlighted two distinct subclones, A and B. The copy number profile of Clone A aligned with the post-treatment bulk copy number profile. (E) a Venn diagram, which consolidates the observed genetic variants (SNVs) from all six treated PDX models into three distinct groups: Persistent, Diminishing, and Emerging. This classification is based on their dynamics pre and post-treatment. (F). Gene set overlap analysis for genes that emerged (surfaced) post-treatment. It highlights common processes, pathways, and underlying biological themes emerged after each treatment regime across all six models.

3.1.1.5 Discussion

This study delves into genetic heterogeneity at a single cell level in neuroblastoma, a cancer characterized by varying genomic complexities. By distinguishing between monoclonal (7/18 cases) and polyclonal (11/18 cases) tumors, we gain vital insights into the diverse evolutionary paths. 5/7 these monoclonal tumors were presented by patients over age of ≥ 42 months. Among these 5, 2 were “hypermutated” cases showing high tumor mutation burden (GR-NB10 > 10 mutations/Mb and IC-pPDX-112 > 5 mutation/Mb) suggesting a shift from CN alterations to the accumulation of SNVs. Indeed in these samples a tumor mutational burden (TMB) significantly higher than the mean in NB was observed (Gröbner et al. 2018).

In monoclonal tumors, the linear evolution (LE) model in tumor growth suggests that tumor progression is driven by selective sweep following the acquisition of key driver mutations. The resulting phylogenetic tree under this model is characterized by the predominance of a dominant clone. Intermediate clones from earlier in the tumor's history are rare, typically persisting only as remnants of previous selective sweeps (Davis, Gao, and Navin 2017) (Black and McGranahan 2021). 6/7 these monoclonal tumors presented at least one early clonal driver gene mutation which includes known tumor suppressor genes (*NF1*, *MLH1*, *PTEN* and *CHEK2*) and oncogenes (*ALK* and *KMT2A*) known to be of importance in NB oncogenesis.

Conversely, polyclonal neuroblastoma exhibits more dynamic evolutionary models, including branched and punctuated evolution. 8/11 these polyclonal tumors were observed in patients over age of < 42 months. GR-NB4, GR-NB7 and P2 showed branched evolution while HSJD-NB-003, HSJD-NB-005, IC_pPDX_17 and P1 showed punctuated evolution. In contrast to LE selective

sweeps are uncommon in BE, and multiple clones expand simultaneously because they all have increased fitness. In punctuated evolution, a large number of genomic aberrations may occur in short bursts of time, at the earliest stages of tumor progression (R. Gao et al. 2016) (Vendramin, Litchfield, and Swanton 2021). These patterns suggest early seeding or parallel evolution, where distinct subclones evolve simultaneously but independently, driven by unique molecular mechanisms. In neuroblastoma both linear and branched evolutionary model have been reported previously based on bulk sequencing studies (Karlsson et al. 2018) (Andersson et al. 2020) (Schmelz et al. 2021) (Gundem et al. 2023). Evidence of punctuated evolution has been observed in a variety of cancer types, including breast cancer (R. Gao et al. 2016; Navin et al. 2011). prostate cancer (Baca et al. 2013), uveal melanoma (Field et al. 2018), and colorectal cancer (Cross et al. 2018). The scWGS analyses now documents punctuated evolution as a prevalent evolutionary pattern in neuroblastoma as well, expanding our understanding of its genetic development.

In our analysis of 18 neuroblastoma tumors, 7 exhibited whole genome doubling (WGD) events, a significant class of copy number alterations (Steele et al. 2022) (Sanz-Gómez et al. 2023). WGD typically arises from mechanisms such as failed cytokinesis, endoreplication, and mitotic slippage (Gemble et al. 2022) (Krupina, Goginashvili, and Cleveland 2021). These events, prevalent in over 30% of solid tumors across various types, are often linked to the inactivation of the TP53 gene (Zack et al. 2013). WGD is a critical event in cancer evolution, leading to increased genomic instability and complexity (TRACERx Consortium et al. 2020) (Frankell et al. 2023). Among the 7 cases with WGD, 2 (GR-NB7 and P2) were *MYCN* amplified. The phylogeny analysis indicates that in both instances, *MYCN* amplification and loss of heterozygosity (LoH) regions, such as 1p and 8q, were events preceding the most recent common ancestor (MRCA), while WGD occurred post-MRCA. Another 2 cases (IC-pPDX-17 and IC-pPDX-75) showed CDK4 amplification. Specifically, in IC-pPDX-17, CDK4 amplification coupled with an *ATRX* mutation was identified as a pre-MRCA event, followed by a post-MRCA WGD. However, in IC-pPDX-75, which presented as a monoclonal tumor, the timing of CDK4 amplification and WGD could not be determined. All 7 cases primarily exhibited numeric copy number alterations, aligning with previous reports that numeric copy number presentations lead to higher ploidy in neuroblastoma. Although, focal CN-LoH at the 17p arm, which harbors the TP53 gene, was reported in 3 of the 7 cases (GR-NB7, IC-pPDX-17, and IC-pPDX-75), Notably, none of the WGD cases harbored a TP53 mutations.

Chromosome regions altered frequently by LoH e.g., 1p, 9p and 11q in the 18 neuroblastoma cases in this study were clonal and early events. Interestingly, monoclonal tumors exhibit a lower percentage of loss of heterozygosity (LoH) at approximately 12.7%, compared to polyclonal

tumors, which show a higher percentage around 16.5%. This difference indicates a greater likelihood of subclonal expansion in polyclonal tumors, underscoring the significance of LoH alterations in driving tumor evolution. This observation aligns with findings reported in prior studies, such as those by (McGranahan and Swanton 2017) and (Gerlinger et al. 2014), which emphasize the impact of LoH on intratumoral heterogeneity and subclonal dynamics. HSJD-NB-005 PDX derived at progression present complex LoH and CH-LoH copy number alteration at chromosome 1 and 17 suggesting these known most frequent alteration fuel substantially subclonal diversification. Both clonal amplifications and LoH regions, typically irreversible, were foundational in clonal architecture (Alfieri, Caravagna, and Schaefer 2023) (Hwang et al. 2021).

Minor clones or rare cells, often characterized by rare and unique genetic alterations, might demonstrate a remarkable potential for growth and gene expression, challenging the prevailing view that major clones are the primary drivers of tumor progression (Schuh et al. 2020). This observation underscores the importance of understanding the selective pressures that influence tumor evolution. Our study in neuroblastoma (NB) provides evidence of this dynamic process, as exemplified in the cases of HSJD-NB-005 and GR-NB4. In HSJD-NB-005 (Figure 3.1.1.4.4) three cells with chromosome 13 loss of heterozygosity (LoH) and in GR-NB4 (Figure 3.1.1.4.6) nine cells with chromosome 12 gain, though few in number, were significantly expressed at the transcriptomic level. This finding demonstrates how a small subset of cells with distinct genetic alterations can have a profound impact on the tumor's transcriptional landscape. The GR-NB10 patient-derived xenograft (PDX), a hypermutated case, further illustrates the rapid clonal evolution in cancer. A single cell with a chromosome 2q gain, initially a minor component at the first relapse, became dominant in the tumor population at the second relapse (data not shown), highlighting its significant selective advantage. This evolution underscores the dynamic nature of cancer progression, where minor clones can expand and influence disease course. The mechanisms and triggers that govern the proliferation and expression of these minor clones are critical areas for future research, offering insights into the complex evolution of cancer.

A key aspect of our research involved studying replication timing in S-phase cell populations, revealing that subclones might have different cellular organisation with regards to replication. This finding has profound implications for predicting tumor behavior and therapeutic response. We hypothesize that early replication timing might be a predictor of a subclone's propensity to proliferate and express genes abundantly. Previous data has indicated that in NB, chromosome breakpoints might map referentially to early replicating regions (Schleiermacher et al. 2003) (Janoueix-Lerosey et al. 2005). RT analyses at a single cell level now leads to novel findings: **First**, we observed a notable association between *MYCN* amplification and early replication timing

(RT) domains in neuroblastoma tumors. Tumors with *MYCN* amplification primarily demonstrated enrichment of early RT domains (Figure 3.1.1.4.5.A), in contrast to those lacking *MYCN* amplification. This finding is significant, considering *MYCN*'s established role as a key regulator of transcription. The amplification of *MYCN* is hypothesized to intensify transcriptional activity in early replicating regions. This increase could potentially lead to conflicts between transcription and replication processes, thereby inducing replication stress and DNA damage. Moreover, *MYCN* amplification may cause premature S-phase entry, increasing the risk of replication errors, especially in early replicating domains. This observation is consistent with previous studies highlighting the connection between *MYCN* amplification and replication stress in neuroblastoma (Molenaar et al. 2012) (Gogolin et al. 2013) (Briu, Maric, and Cadoret 2021). **Second**, we noticed distinct RT between two subclones evolving parallelly within the same tumor. This is exemplified in the PDX model GR-NB4, where two distinct subclones, s1 and s2, were identified. These clones, evolving in parallel, exhibited contrasting replication patterns: clone s1 was characterized by early replication domains, while clone s2 displayed late replication domains. This dichotomy in RT between the two clones within the same tumor model highlights the intricate and varied replication behavior that can exist within neuroblastoma tumors. Recent studies demonstrated subclone specific distinct RT in breast cancer, using cell line MCF7 (Gnan et al. 2022), PDX and primary cancer tissue (Weiner et al. 2023) highlighting these subclones often exhibit unique proliferation rates, allowing them to progress more rapidly through the cell cycle. This acceleration in cell cycle progression is a key factor in the rapid growth and evolution of tumors, as these advantageous genetic changes enable subclonal populations to thrive and expand within the tumor microenvironment.

We now demonstrate for the first time that in NB differential replication timing can be observed in distinct subclones, associated with differential treatment sensitivity. The differential replication timing (RT) patterns within subclones could potentially dictate their sensitivity or resistance to specific therapies. In the case of the GR-NB4 model treated with lorlatinib targeting *ALK* amplification, we postulate that the treatment responsiveness was aligned with the RT characteristics of the dominant subclone where the *ALK* genomic locus (Figure 3.1.1.4.5) was late replicating. Figure 3.1.1.4.5

Our research has revealed notable differences in chemotherapy response among various PDX models of neuroblastoma, underscoring the complexity of treatment efficacy in relation to specific genetic backgrounds. For instance, in the GR-NB4 model (Figure 3.1.1.4.7), which harbors *ALK* amplification, we observed a marked response to lorlatinib, a targeted *ALK* inhibitor. This response was further enhanced when lorlatinib was used in combination with chemotherapy, as indicated by

the growth response curves and significant variations in variant allele frequency (VAF) in subclonal populations pre- and post-treatment. Specifically, our integrative analysis of subclonal specific single nucleotide variant (SNV) profiles revealed that diminishing variants were predominantly associated with subclone s2 (Figure 3.1.1.6.30), suggesting a targeted treatment response in this subclone. In contrast, the IC-pPDX-75 model, characterized by a pathogenic *ALK* mutation, exhibited no significant response to lorlatinib treatment.

In summary, in our cohort of the neuroblastoma samples (n=18) but analysing ~7800 and ~1600 single cell whole genome from G1/G2 and S-phase cell population sheds light on complex nature of tumor evolution. Driver amplifications (>10 copies) in *MYCN*, *ALK*, *CDK4* and *MDM2* were clonal and extensive LoH, identified as early evolutionary events, occurred before the MRCA. Post-MRCA, an initial phase of instability gives rise to numerous subclones during progression, later there may be stabilizing selection (Brennan et al. 2013) into a continuous rate of copy number evolution at relapse. This indicates that, despite some stabilizing selection, tumor cells persistently adapt and evolve throughout tumor expansion. The investigation into the relationship between DNA replication and subclonal growth and behavior is pivotal for understanding the unique characteristics, evolutionary fitness, and drug sensitivities. This research provides crucial insights into the rare subclones within a tumor, how they adapt and survive during tumor expansion, and their varying responses to therapeutic interventions.

This study faces several limitations that must be considered. Initially, the conclusions are based on a restricted dataset, which may necessitate further experimental validation for robustness. While the genomic and transcriptomic single cell analyses were conducted on samples from the same tumor or patient-derived xenograft (PDX) passage, ideally, such analyses should be performed on the same single cells. This approach would offer a more accurate representation of the genotype-to-phenotype relationships without biases. Secondly, the challenge of accurately detecting mutations from high-throughput single-cell whole genome sequencing is amplified due to the sparse genomic coverage typical of such techniques, complicating the mutation calling process at the single-cell level. Finally, the scope of this study is limited by its small patient cohort, consisting of only 18 tumor samples. To validate and generalize the biological findings and the evolutionary models proposed, it is crucial to extend these studies to larger cohorts of patients.

In conclusion we report on the first extensive scDNA study in NB. We show that in NB both monoclonal and polyclonal structures can occur. Importantly rare single cells can evolve to a major clone upon relapse. Subclones can present distinct replication timing profiles. Furthermore, we

document for the first time clone specific sensitivity to treatment, highlighting the importance of clonal evolution upon tumor progression and treatment resistance in NB.

3.1.1.6 Supplementary data

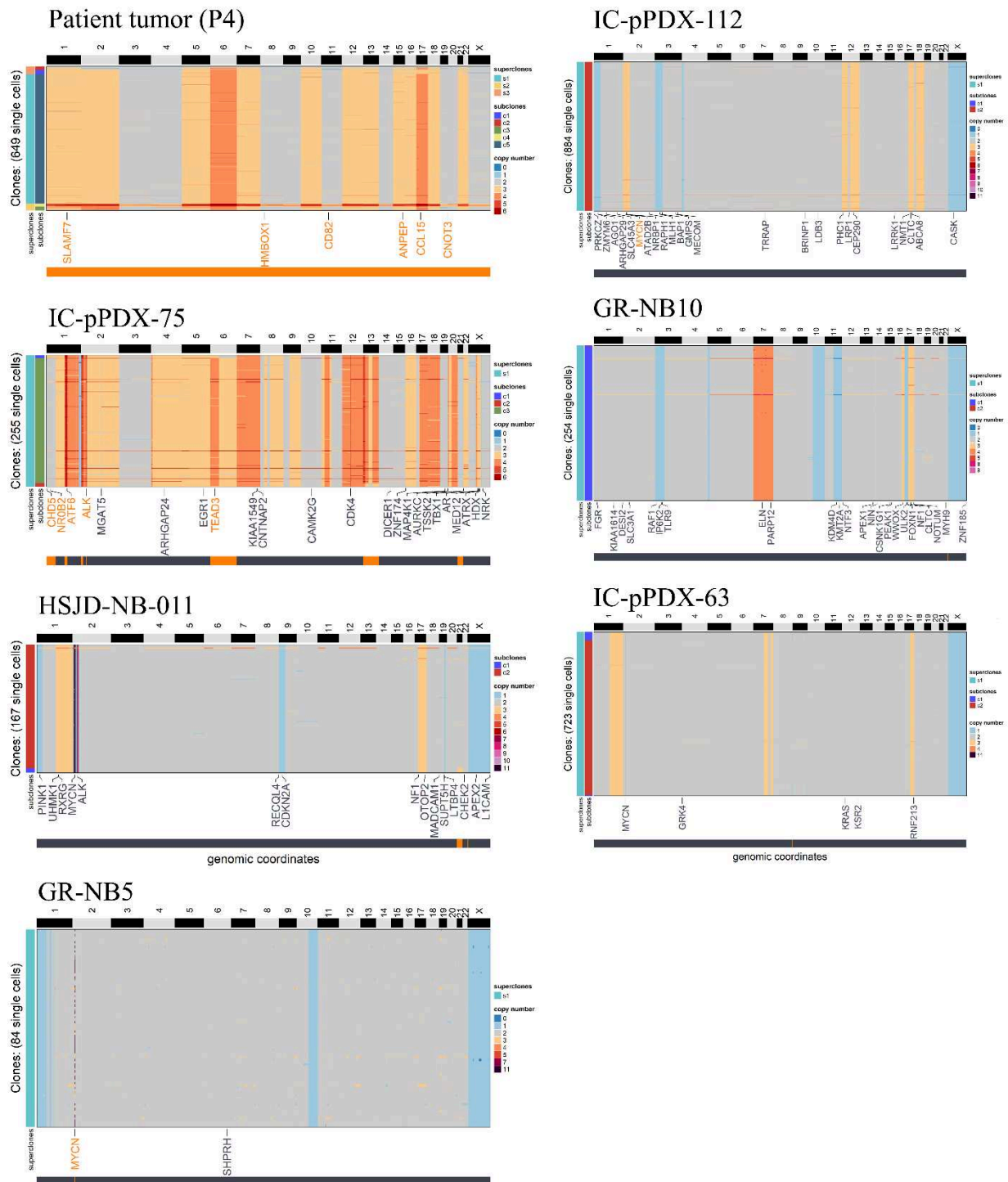


Figure 3.1.1.6.1 Monoclonal genomes with minimal subclonal diversity.

Tumor id is presented on top of each heatmap derived from single cell copy number profile.

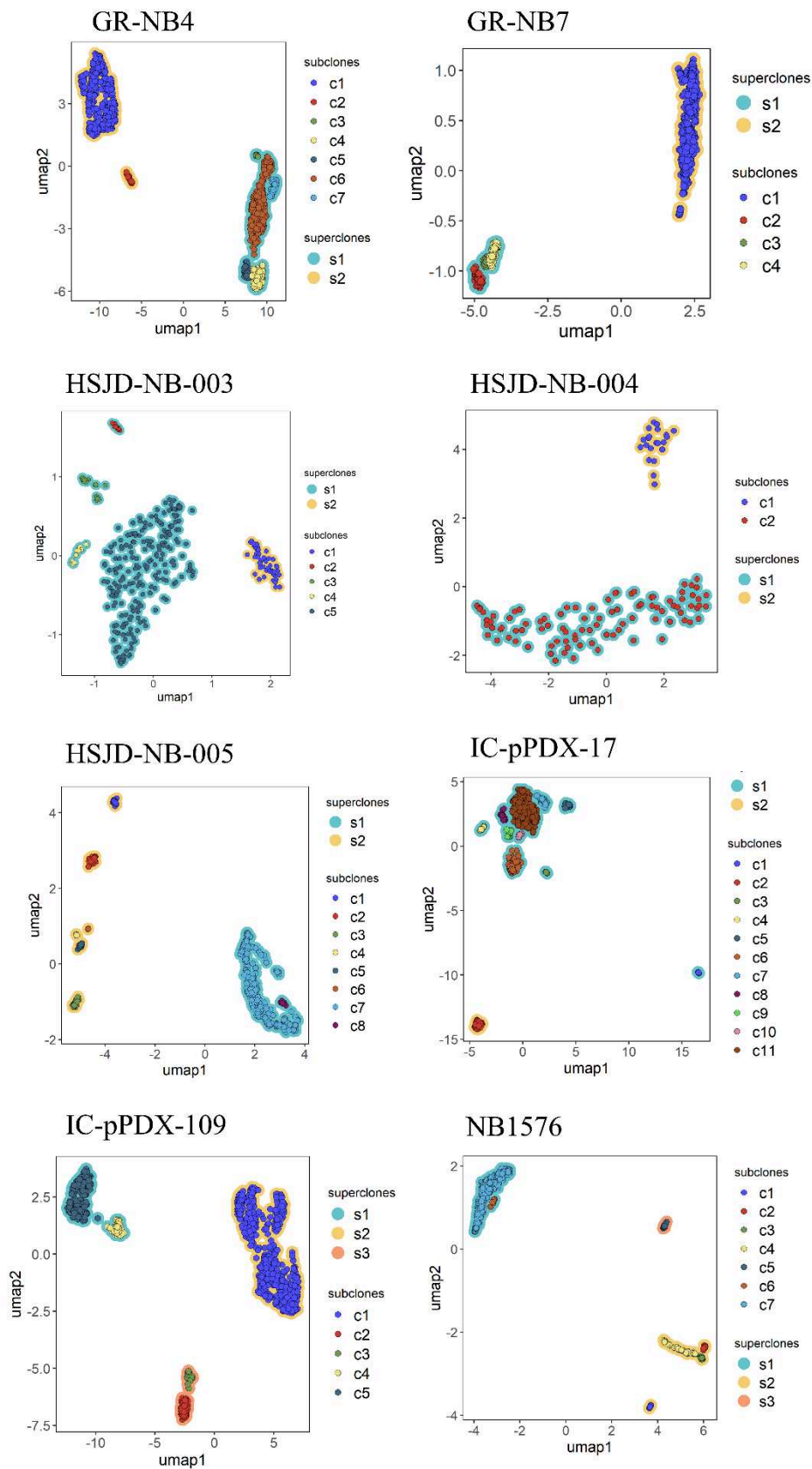


Figure 3.1.1.6.2 Clonal substructure of neuroblastoma Patient tumor (P2) and 7 PDXs samples.

UMAP clustering of single cell copy number data from twelve tumors showcasing superclones (contour colors) and subclones (colored points). This visualization highlights the superclone and subclone distributions within each tumor, illustrating the complex clonal architecture.

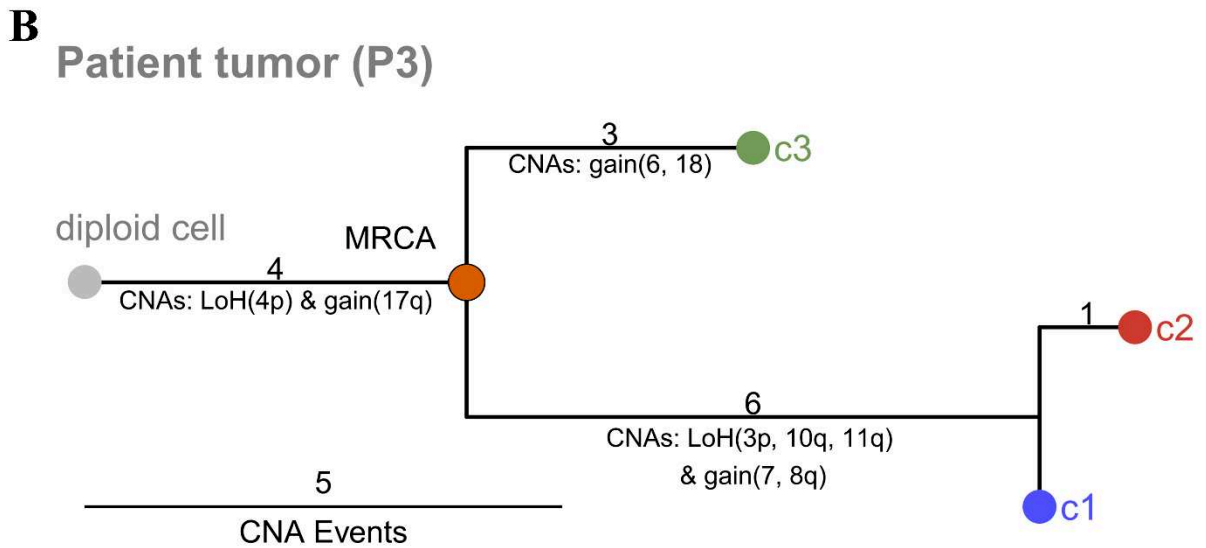
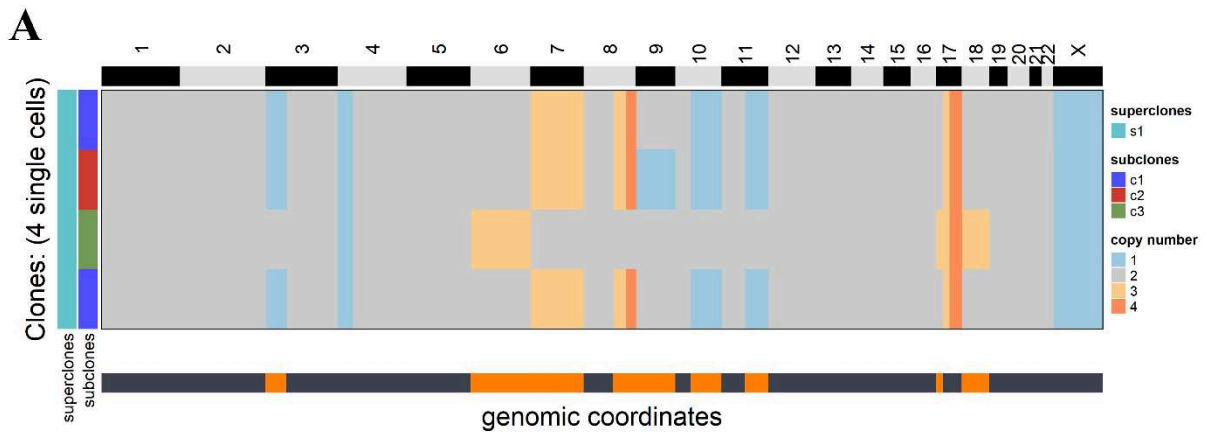


Figure 3.1.1.6.3: Clonal substructure and evolutionary analysis of clonal lineages in NB patient tumors P3.

A) The upper panel is clustered heatmaps of single-cell total copy-number profiles showing subclones in NB patients P3 at diagnosis. Single-cell copy-number clustered heatmaps, with left-most header columns indicating superclone and subclone groups. B) Bottom annotation panels indicate clonal and subclonal classification of CNAs (clonal = black, subclonal = gray and unique subclonal CN = orange). The bottom panel show the inferred evolutionary rooted trees with MRCA profile and different CNA classes.

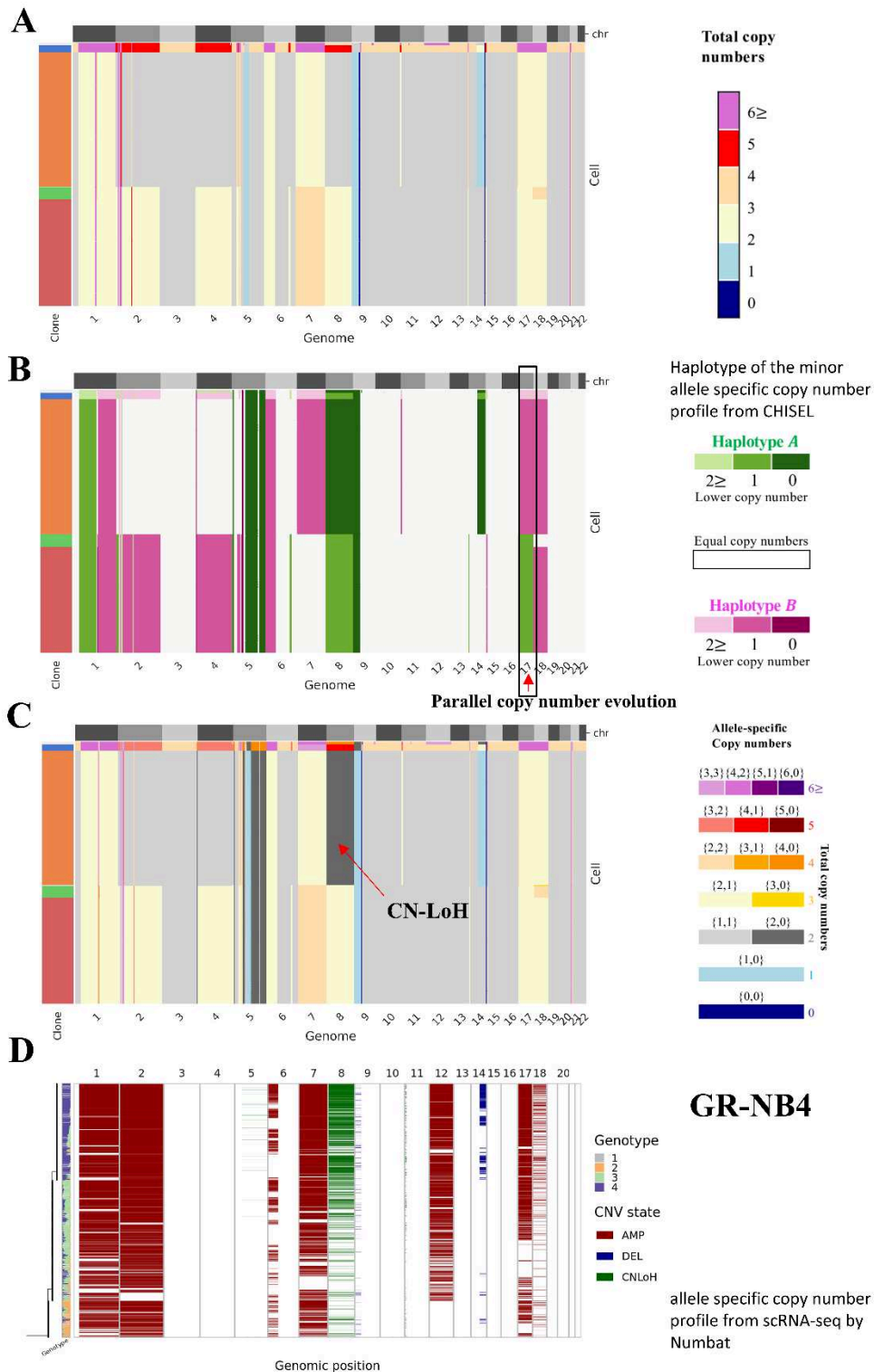


Figure 3.1.1.6.4: Allele- and haplotype-specific copy numbers for 865 cells in GR-NB4 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 865 cells into 4 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

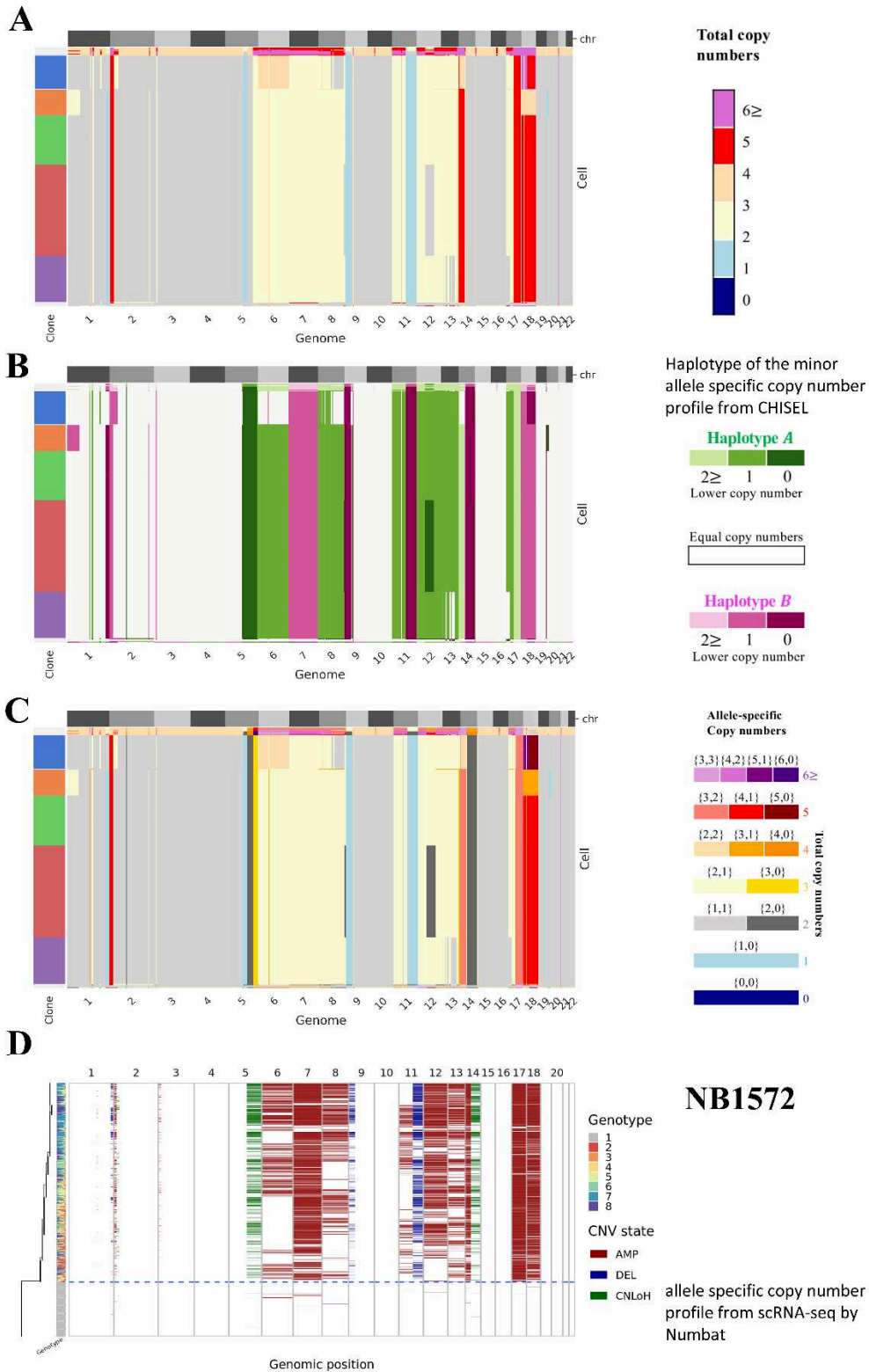


Figure 3.1.1.6.5: Allele- and haplotype-specific copy numbers for 359 cells in GR-NB4 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 359 cells into 5 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

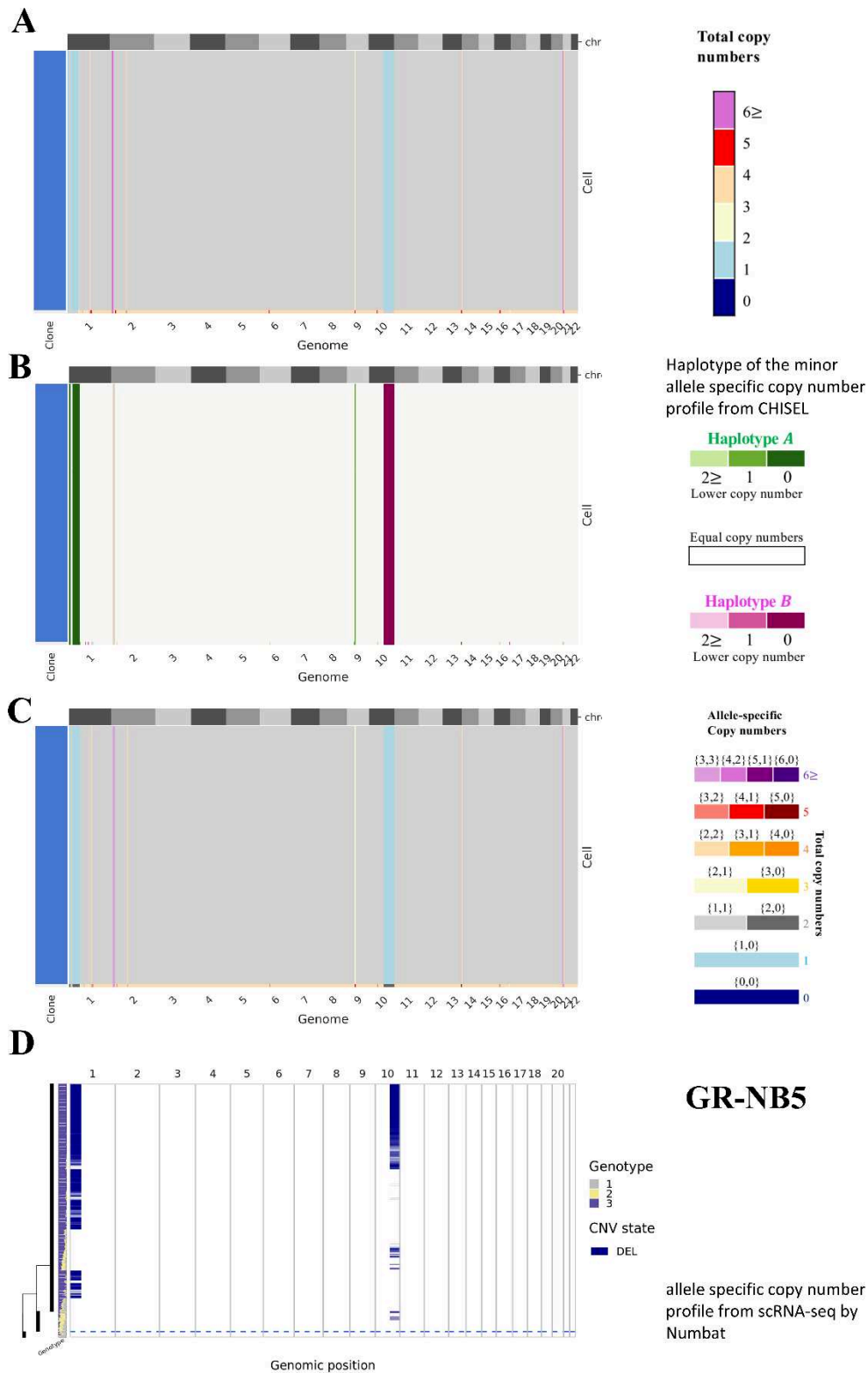


Figure 3.1.1.6.6: Allele- and haplotype-specific copy numbers for 84 cells in GR-NB5 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 84 cells into 1 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

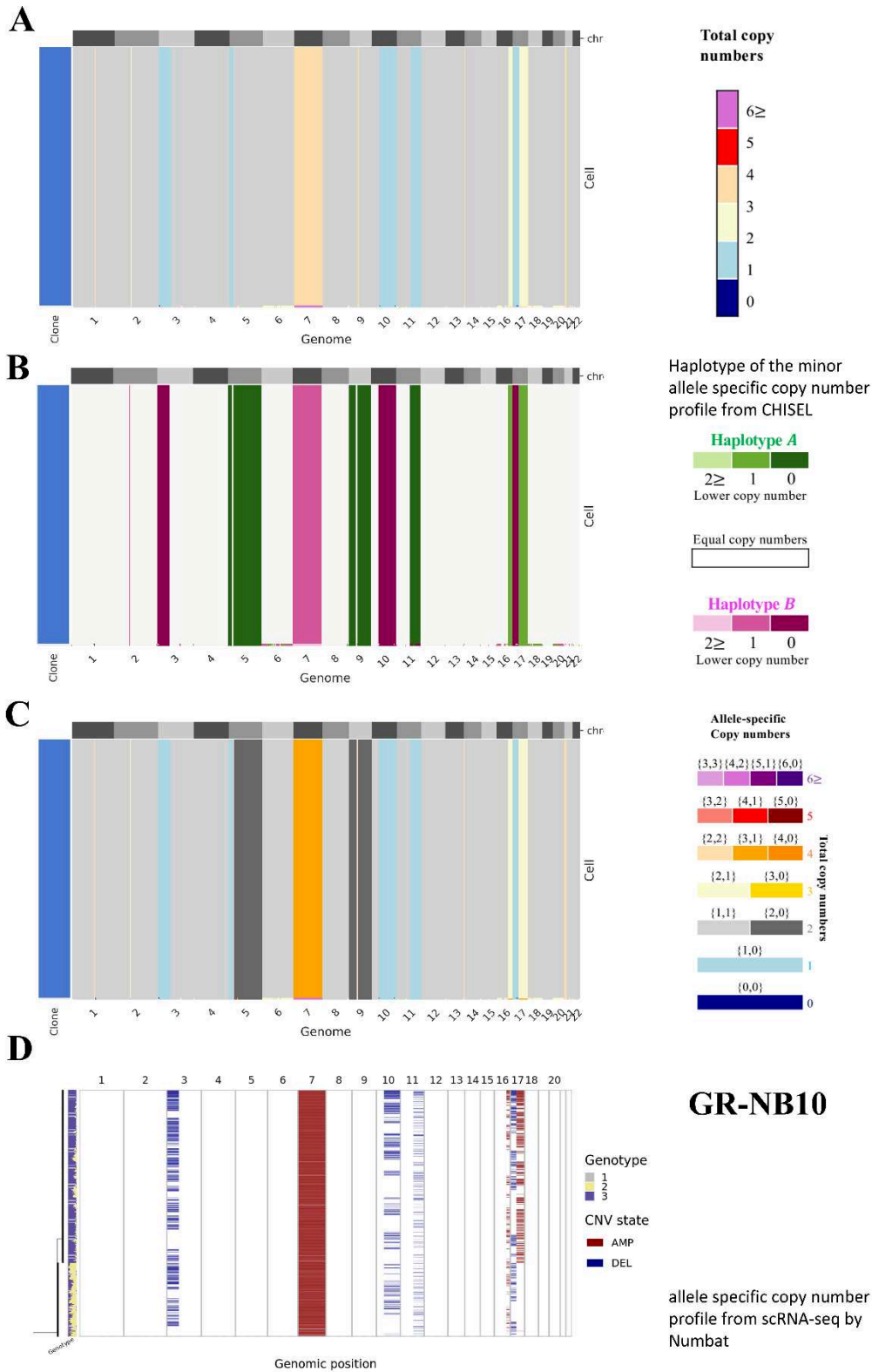


Figure 3.1.1.6.7: Allele- and haplotype-specific copy numbers for 257 cells in GR-NB10 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 257 cells into 1 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

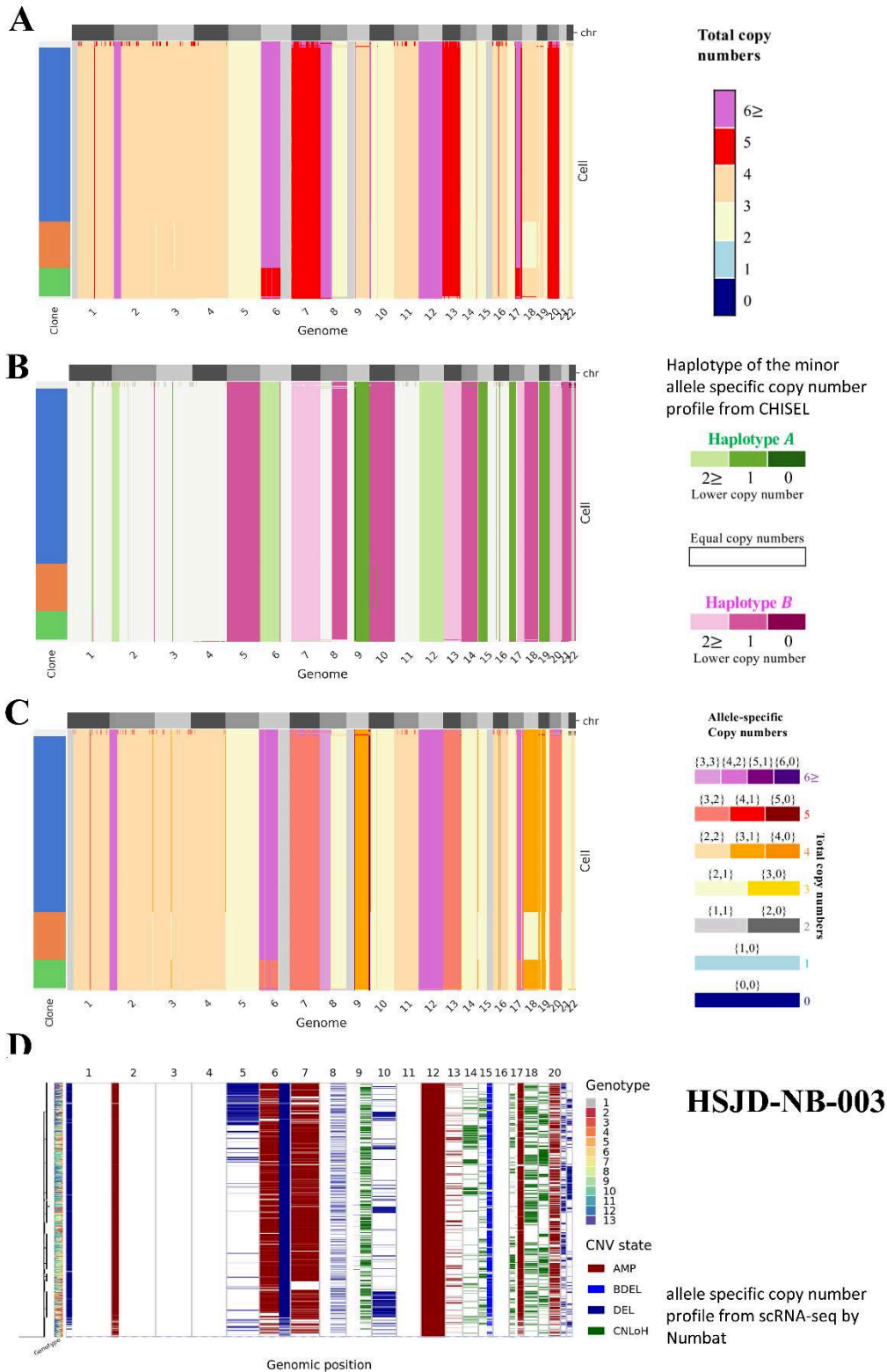


Figure 3.1.1.6.8: Allele- and haplotype-specific copy numbers for 326 cells in HSJD-NB-003 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 326 cells into 3 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

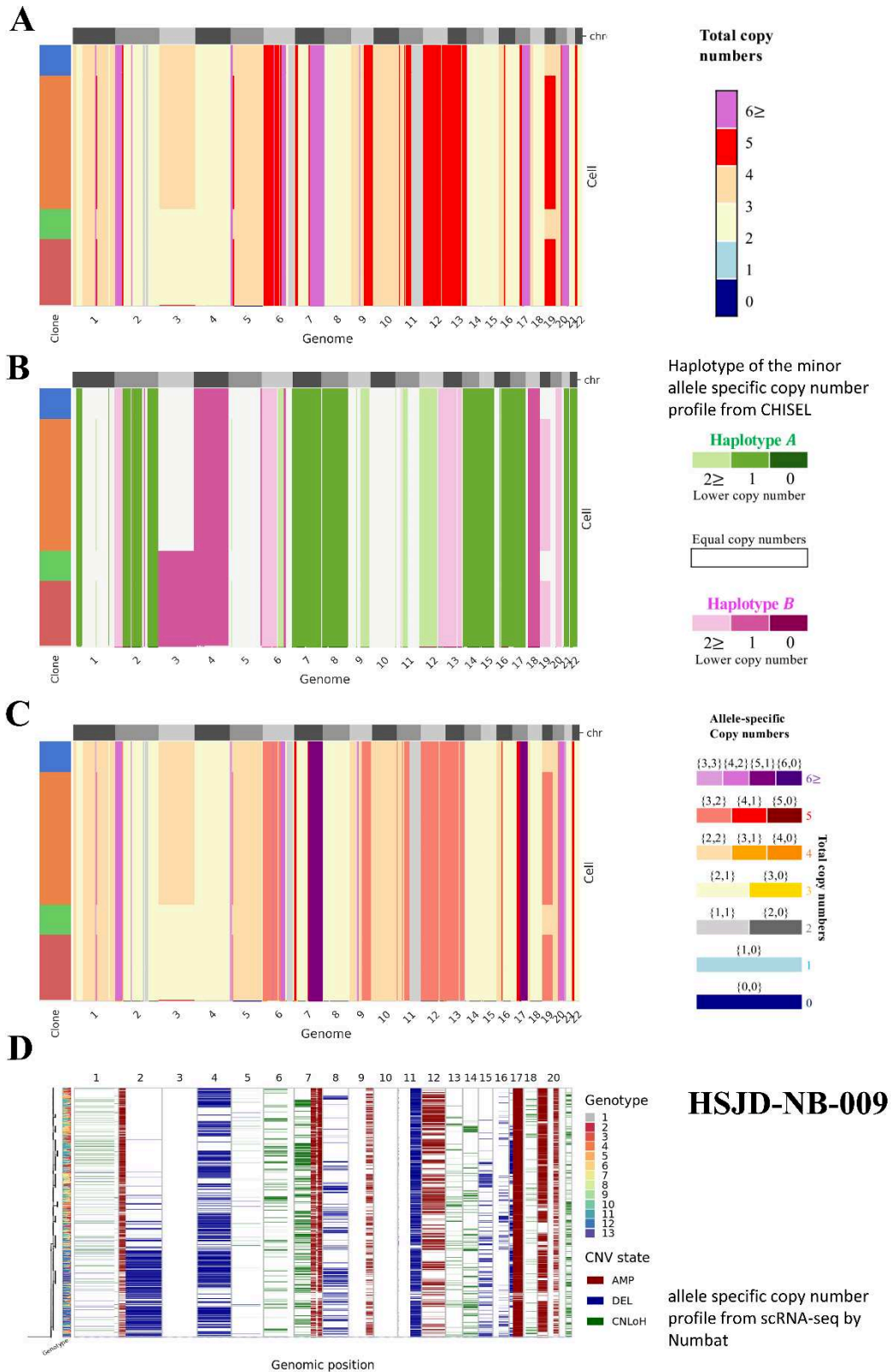


Figure 3.1.1.6.9: Allele- and haplotype-specific copy numbers for 323 cells in HSJD-NB-009 PDX model. The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 323 cells into 4 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

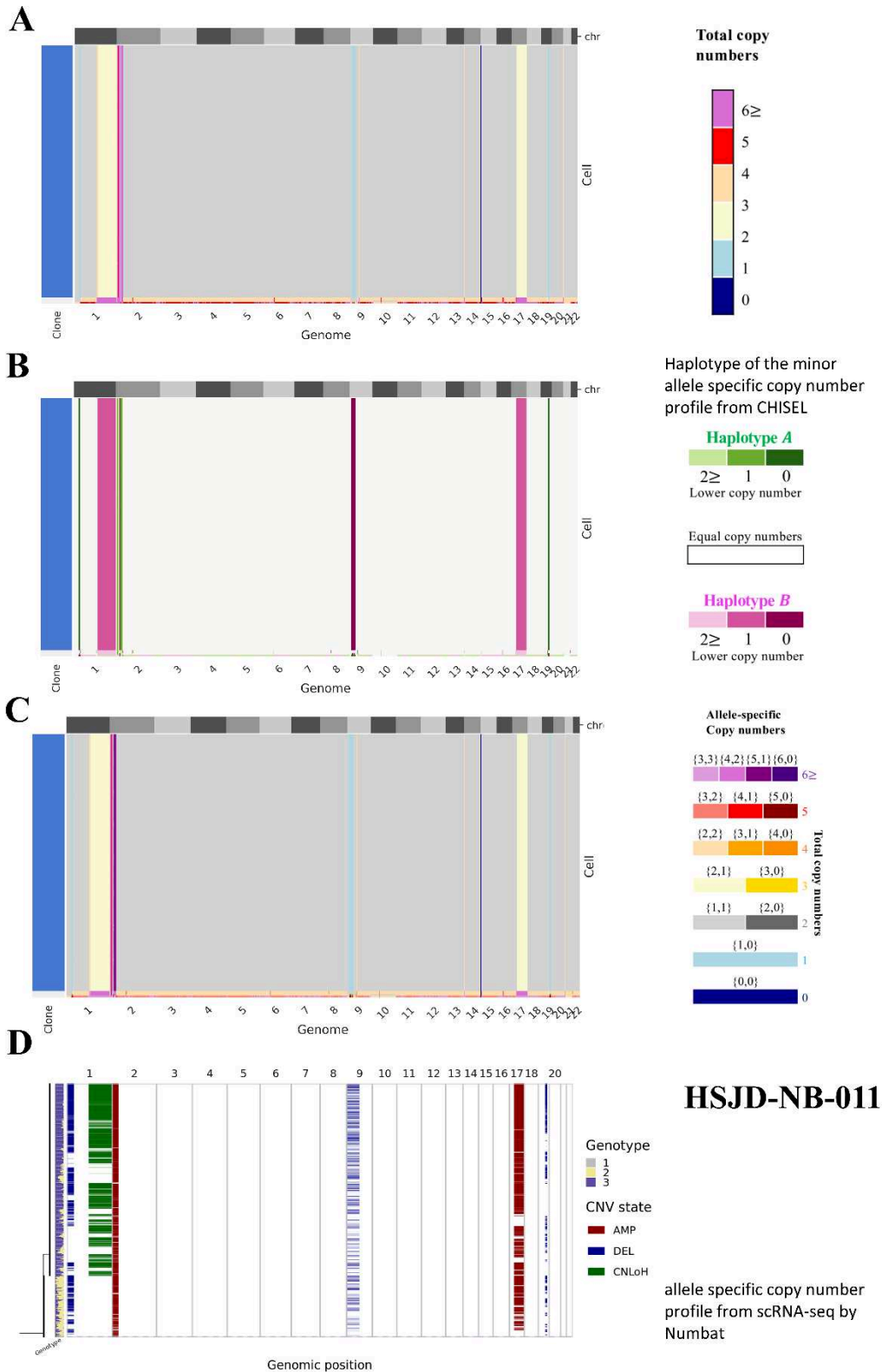
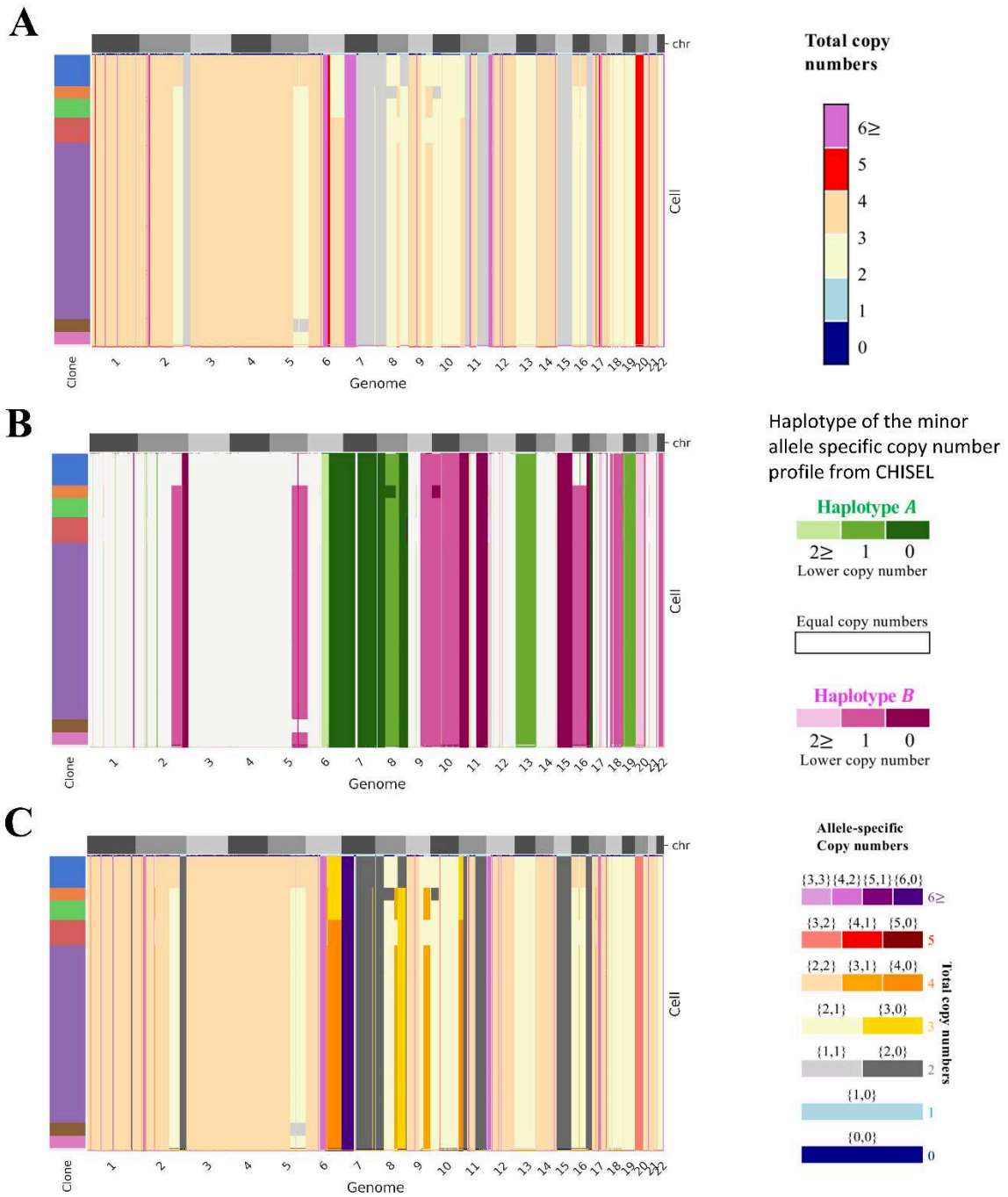


Figure 3.1.1.6.10: Allele- and haplotype-specific copy numbers for 170 cells in HSJD-NB-011 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 170 cells into 1 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.



IC-pPDX-17

No scRNA-sequencing

Figure 3.1.1.6.11: Allele- and haplotype-specific copy numbers for 523 cells in IC-pPDX-17 PDX model. The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 523 cells into 7 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar).

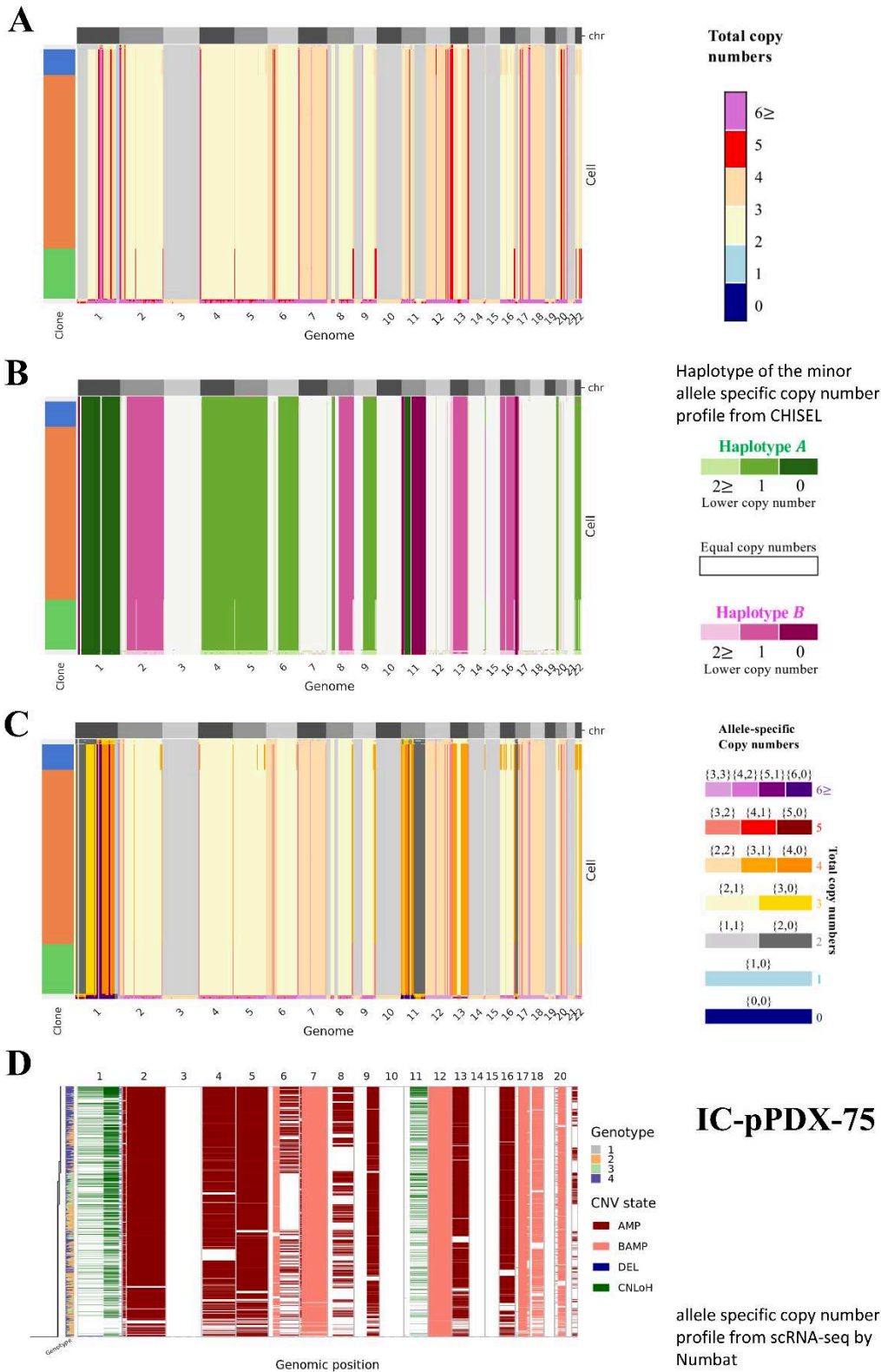


Figure 3.1.1.6.12: Allele- and haplotype-specific copy numbers for 255 cells in IC-pPDX-75 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 255 cells into 3 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

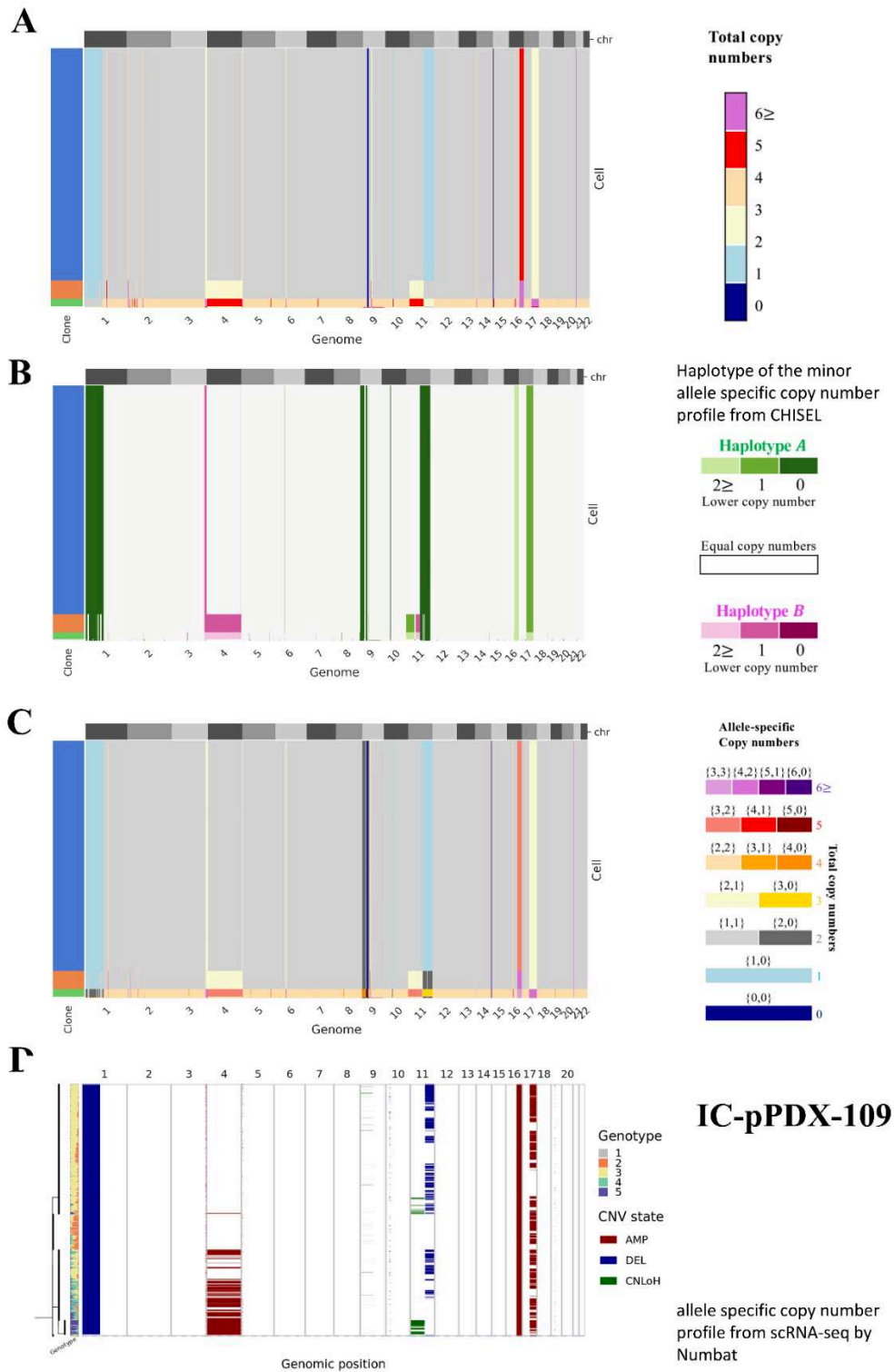


Figure 3.1.1.6.13: Allele- and haplotype-specific copy numbers for 962 cells in IC-pPDX-109 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 962 cells into 3 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

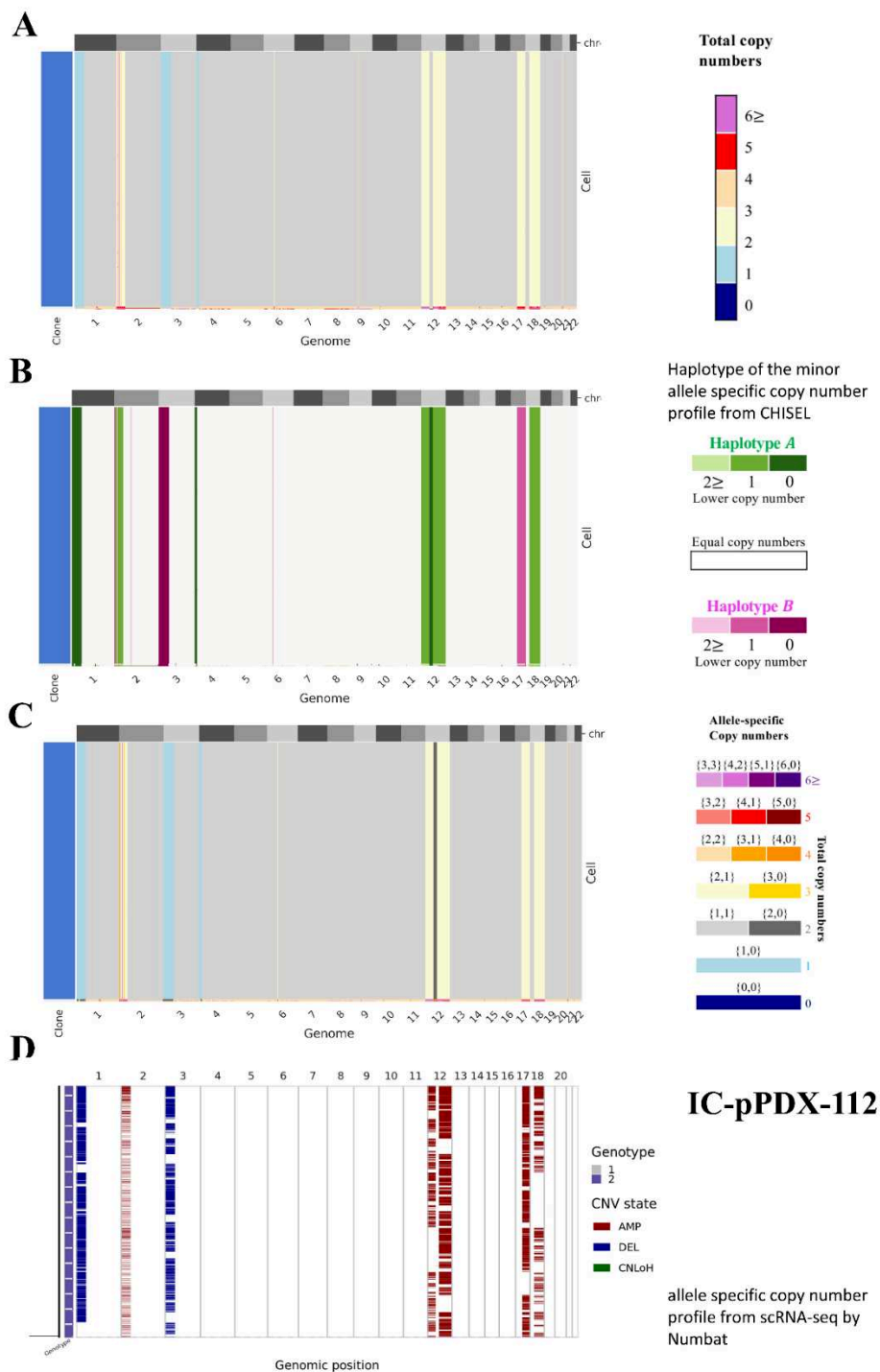


Figure 3.1.1.6.14: Allele- and haplotype-specific copy numbers for 884 cells in IC-pPDX-112 PDX model.

The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 884 cells into 1 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.

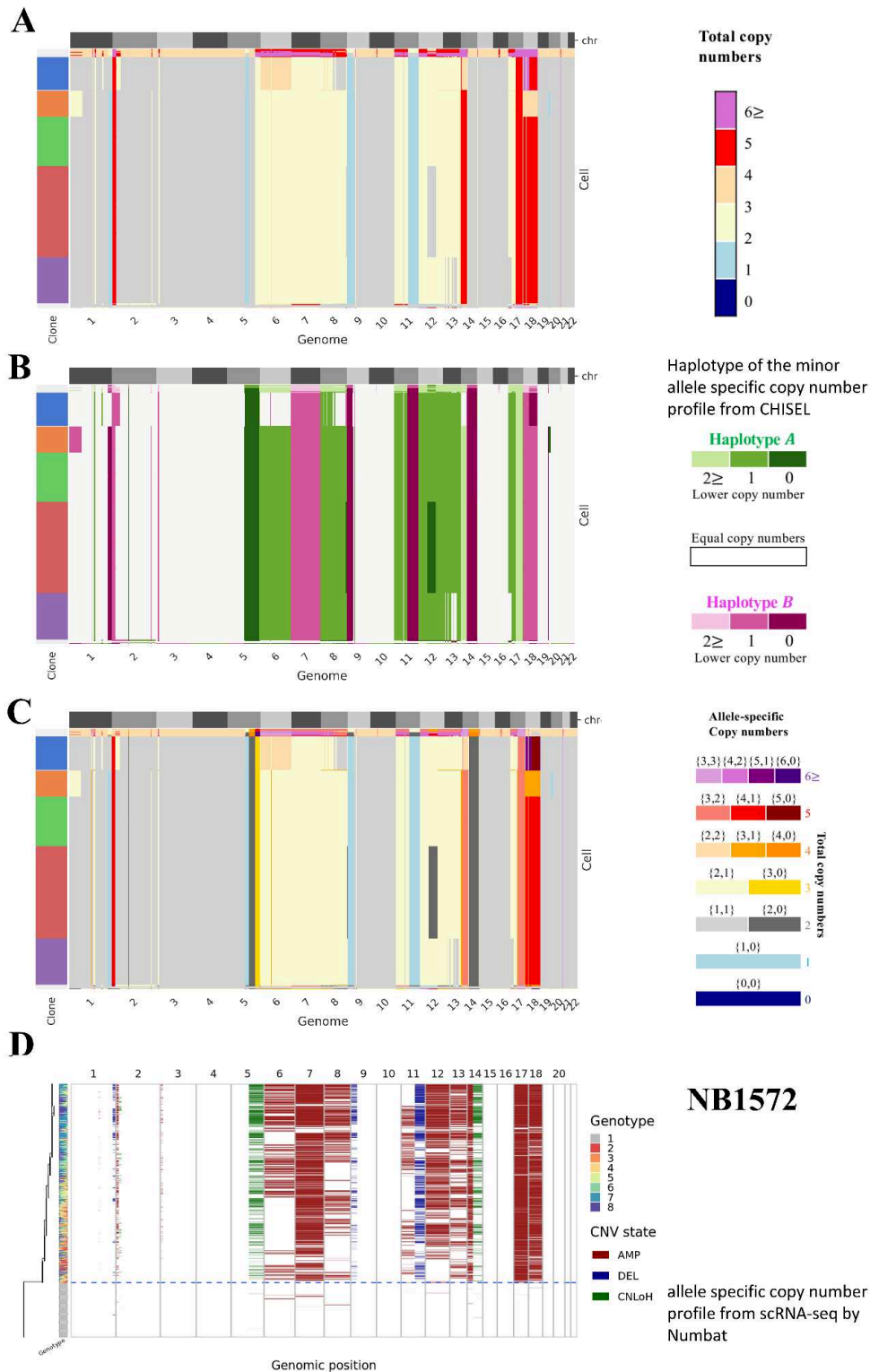
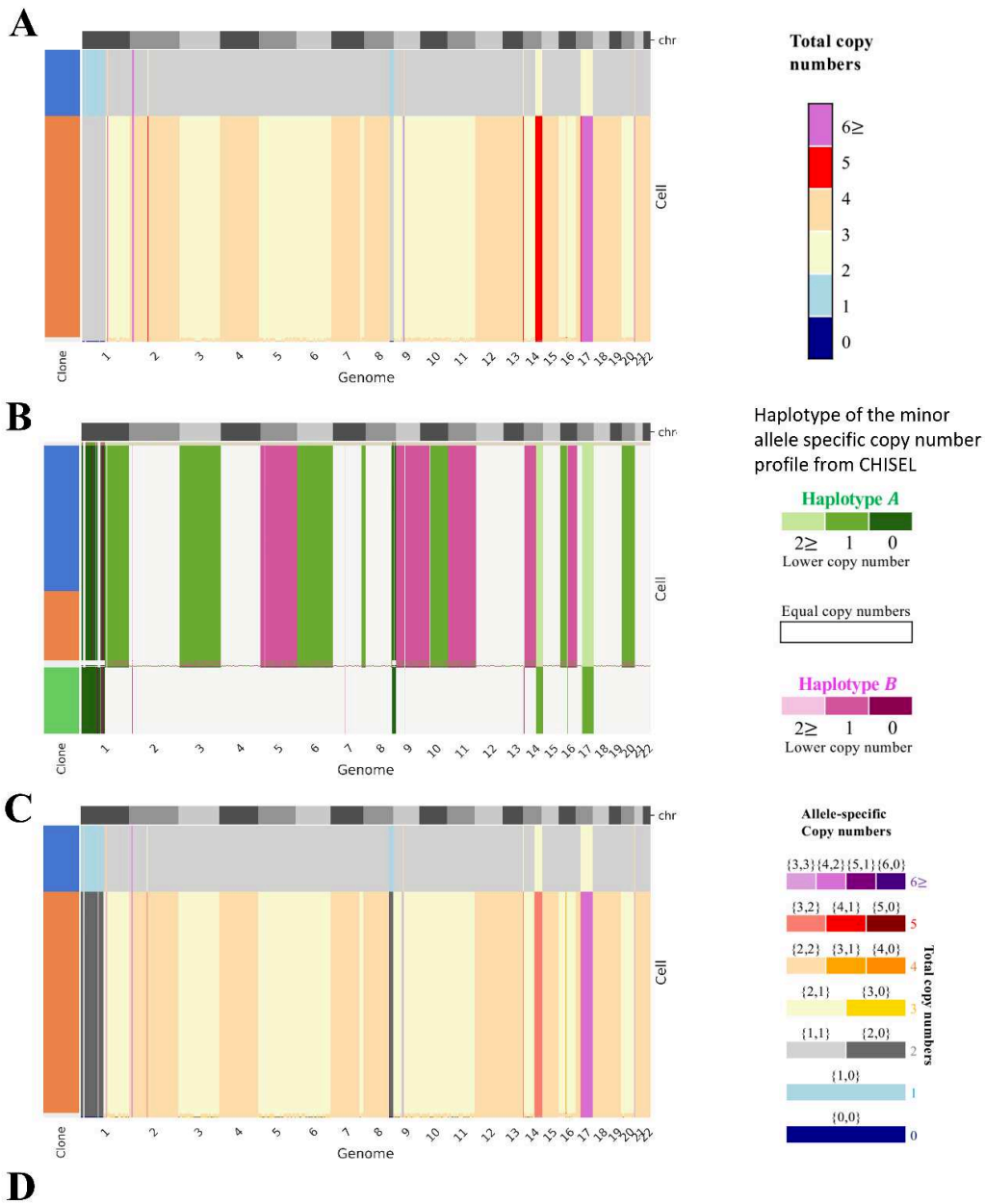


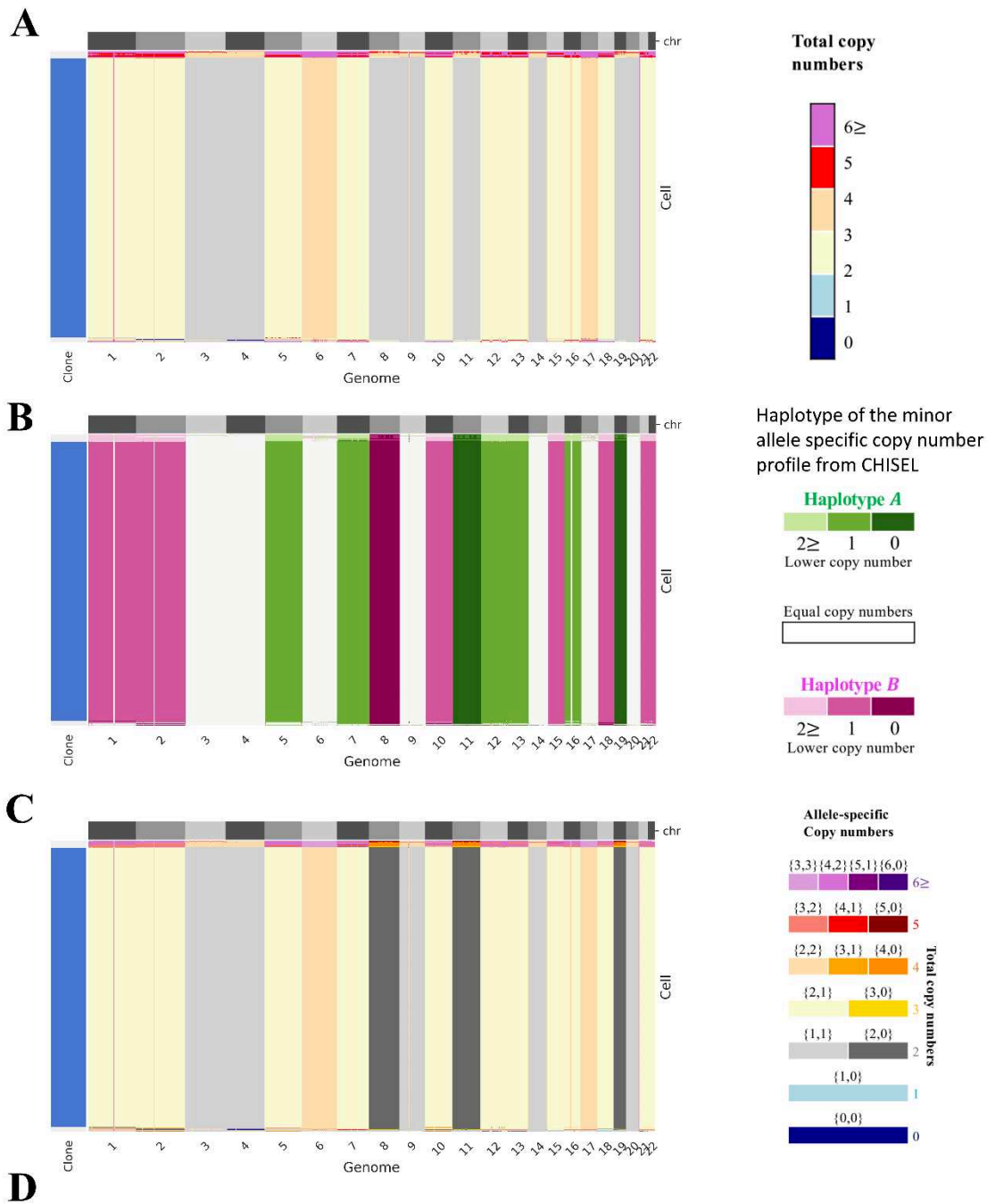
Figure 3.1.1.6.15: Allele- and haplotype-specific copy numbers for 359 cells in NB1572 patient tumor. The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 359 cells into 5 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar). D. Single-cell CNV landscape inferred by Numbat using scRNA-seq data.



NB1576

**scRNA-sequencing
data not analysed**

Figure 3.1.1.6.16: Allele- and haplotype-specific copy numbers for 458 cells in NB1576 patient tumor. The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 458 cells into 2 clones (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar).



NB1583

No scRNA-sequencing

Figure 3.1.1.6.17: Allele- and haplotype-specific copy numbers for 663 cells in NB1583 patient tumor. The allele-specific, haplotype-specific, and total copy numbers are reported in A, B, and C, respectively, across all autosomes (grey rectangles in the first row). CHISEL tool groups 663 cells into 1 clone (colors in the left-side bar) and classifies the remaining cells as noisy (grey in the left-side bar).

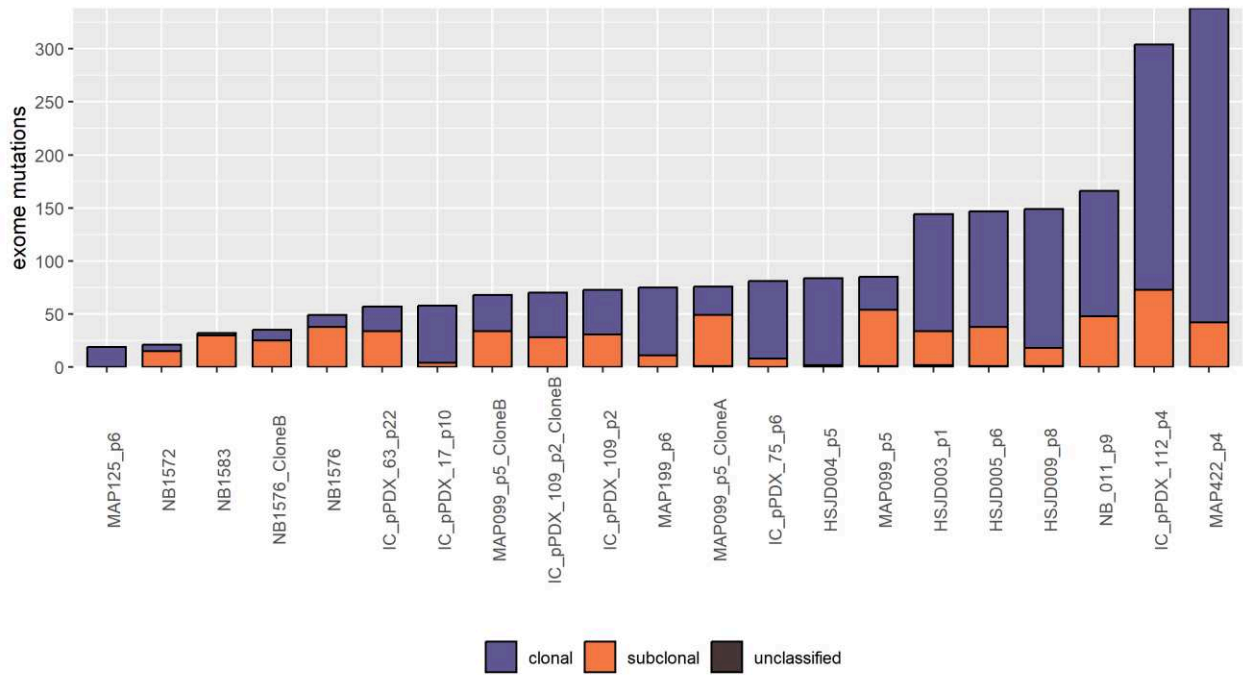


Figure 3.1.1.6.18 Exome mutation counts according to clonality.

Exome mutation counts of each tumor indicating mutations that were classified as clonal or subclonal based on allele-specific copy number frequencies and all these single nucleotide variants in analysis were supported by at least two functional prediction algorithms out of 8 such tools we used in the SNV analysis.

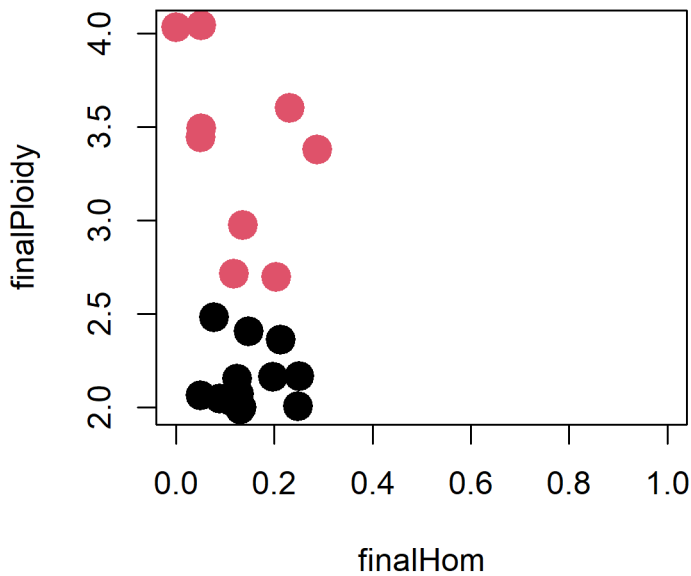


Figure 3.1.1.6.19: Analysis of Loss of Heterozygosity (LoH) regions

Analysis of Loss of Heterozygosity (LoH) and Copy-Neutral Loss of Heterozygosity (CN-LoH) genome wide occurrences in percentage in a set of 20 tumors, including two distinct subclones each from GR-NB4 and P2. And compared with overall ploidy estimates derived from pseudobulk data (aggregated single cells per tumor sample).

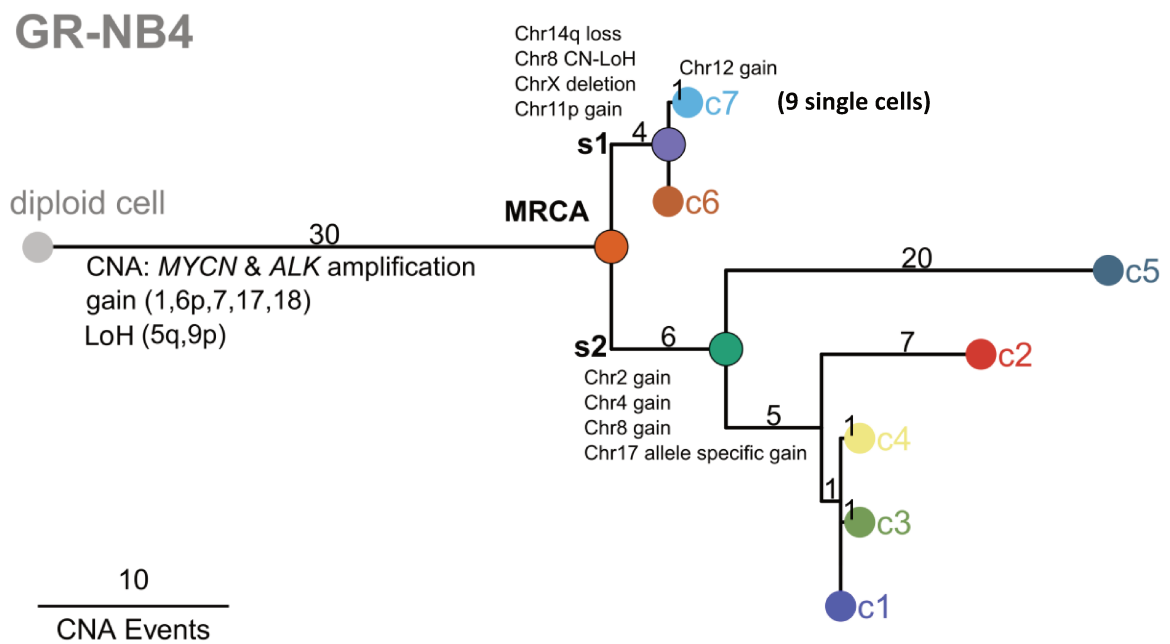


Figure 3.1.1.6.20 GR-NB4, phylogeny tree with seven clones.

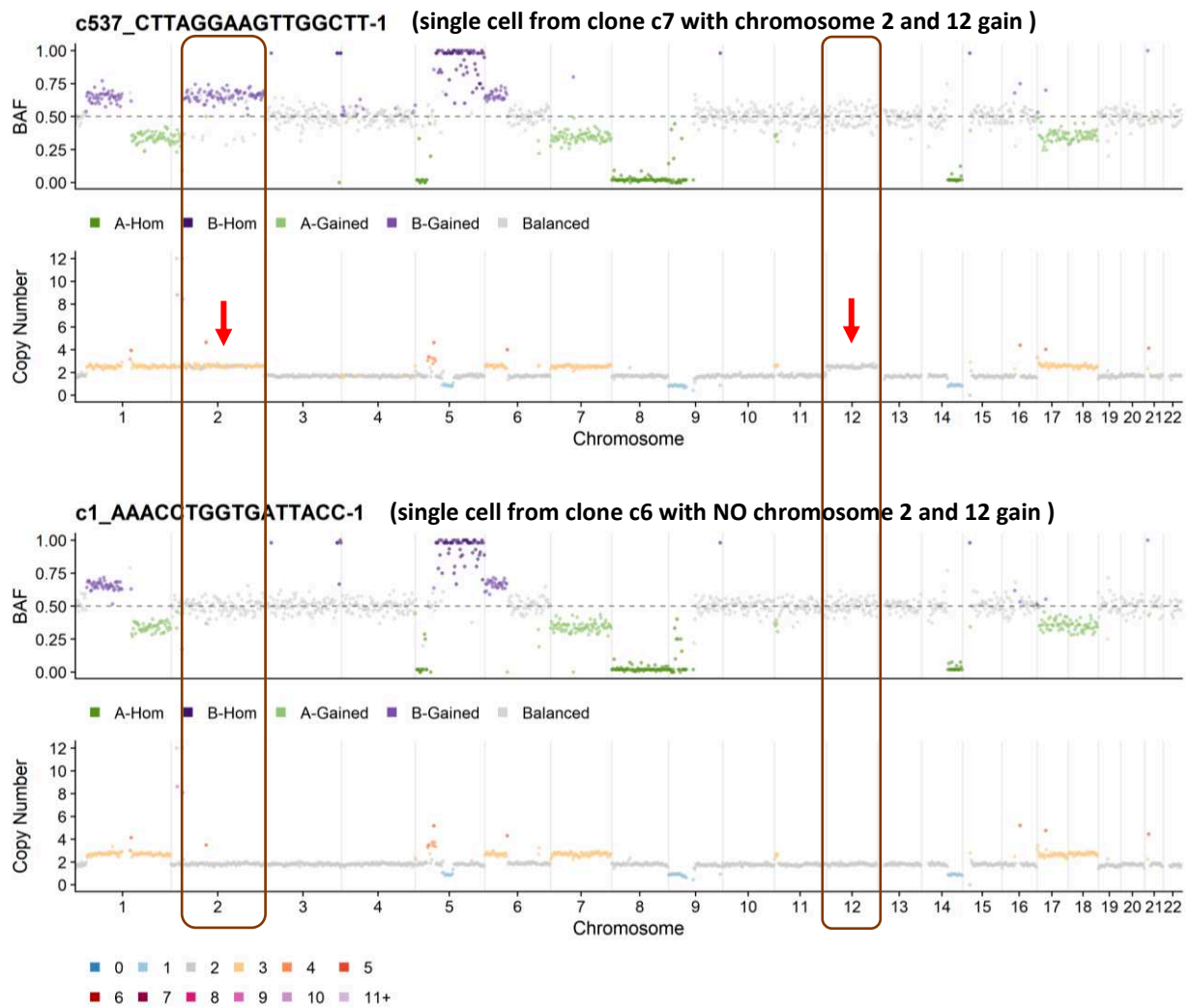


Figure 3.1.1.6.21 Single cell ASCN profiles from GR-NB4 PDX model representing clone 6 and 7 with and without chromosomes 2 and 12 gain.

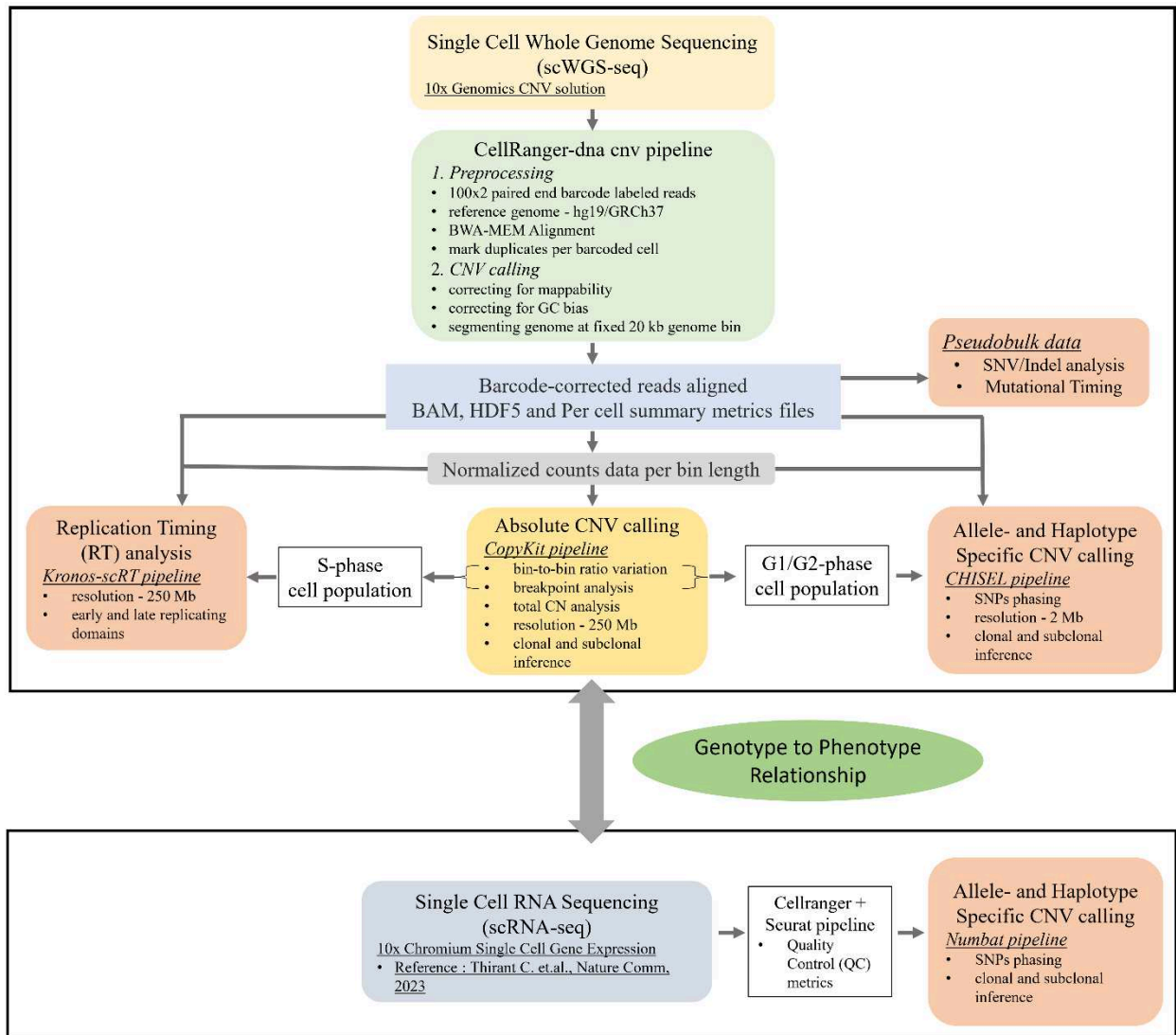


Figure 3.1.1.6.22 Bioinformatics data analysis workflow.

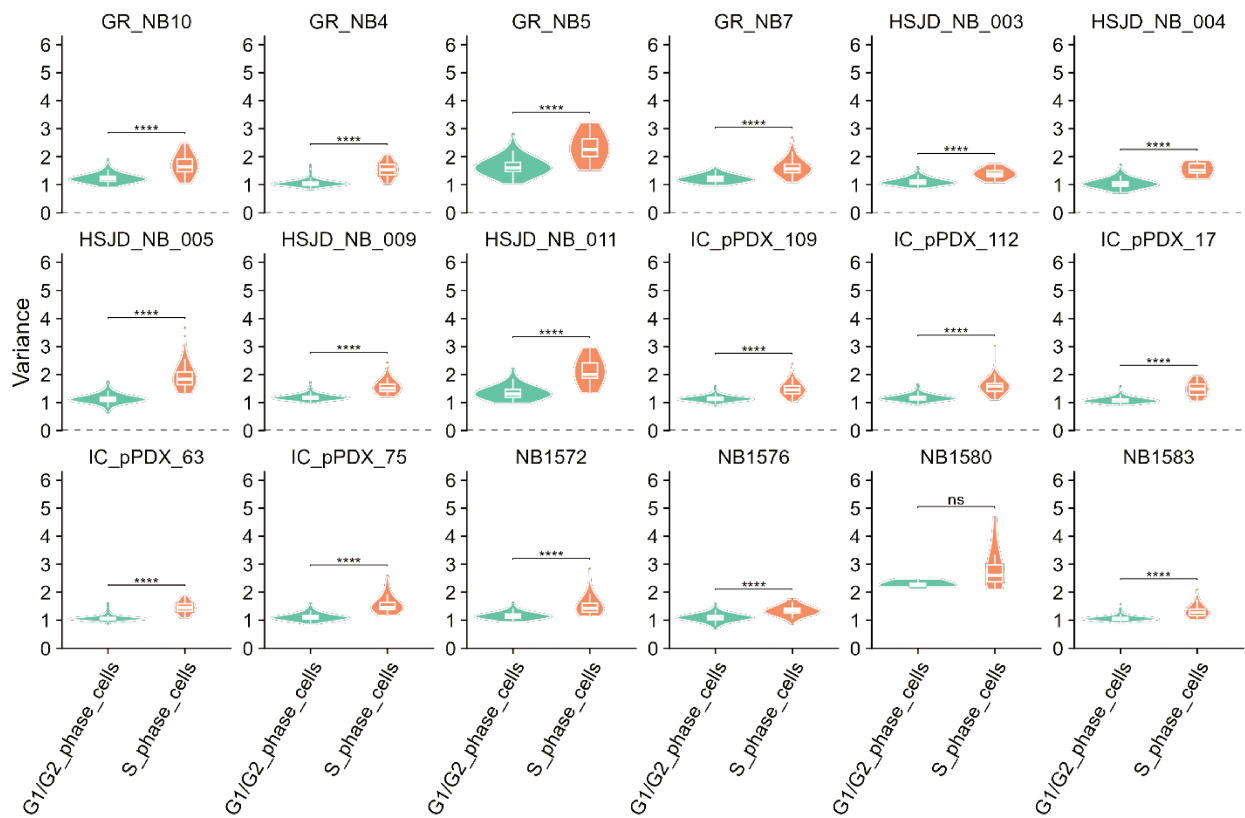


Figure 3.1.1.6.23 Comparative Analysis of bin-to-bin CN ratio variance Between G1/G2 and S-Phase Cell Populations in Neuroblastoma.

The P-value was determined using a Welch t-test.

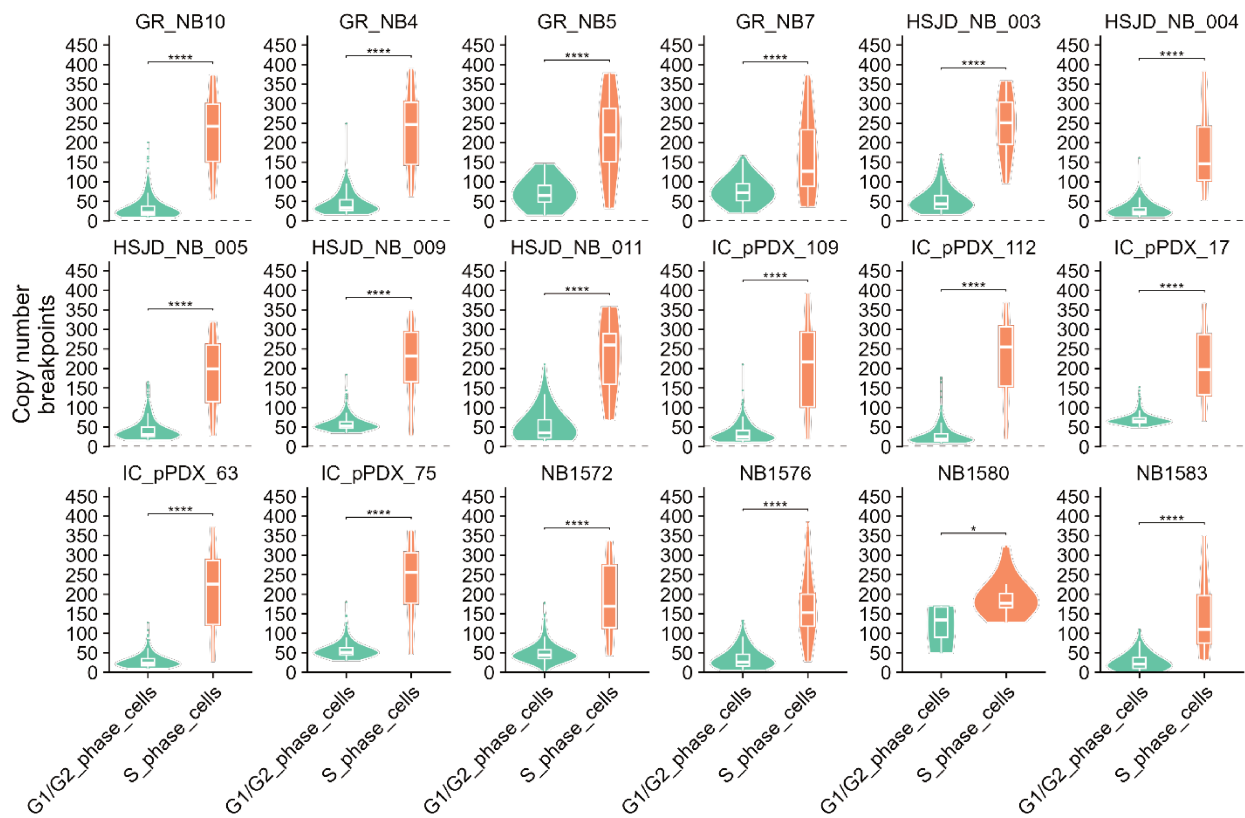


Figure 3.1.1.6.24 Comparative Analysis of Copy Number Breakpoints Between G1/G2 and S-Phase Cell Populations in Neuroblastoma.

The P-value was determined using a Welch t-test.

Table 3.1.1.6.1 Single cell DNA sequencing run summary details.

Sample id	Estimated number of Cells	Total reads	Reads from cells	Total mapped de-duplicated reads in cells	Mean mapped de-duplicated reads per cell	Median effective reads per Mb	Median duplicate fraction per cell	Ploidy (Estimated)	Fraction of noisy cells (Estimated)
GR_NB10	322	11097 M	7973 M (20.1%)	11097 M (71.9%)	6.9 M	2104	63%	2	15%
GR_NB4	994	9928 M	7060 M (41.5%)	9928 M (71.1%)	4.1 M	1292	33%	2	13%
GR_NB5	137	6624 M	3894 M (14.4%)	6624 M (58.8%)	6.99 M	2044	67.30%	2	20.40%
GR_NB7	625	5446 M	3727 M (45.2%)	5446 M (68.4%)	3.94 M	1193	29%	5	53%
HSJD-NB-003	396	6438 M	5525 M (46.1%)	6438 M (85.8%)	7.5 M	2361	38%	4	18%
HSJD-NB-004	131	11661 M	7166 M (7.6%)	11661 M (61.5%)	6.8 M	2181	79%	2	14.50%
HSJD-NB-005	543	11059 M	8665 M (39.5%)	11059 M (78.3%)	8 M	2479	42%	2	19%
HSJD-NB-009	405	5445 M	3400 M (35.5%)	5445 M (62.4%)	4.8 M	1575	34%	4	21.50%
HSJD-NB-011	199	5979 M	4112 M (33.6%)	5979 M (68.8%)	10 M	2862	41%	2	16%
IC_pPDX_17	662	9741 M	7468 M (26.5%)	9741 M (76.7%)	3.9 M	1247	57%	4	24%
IC-pPDX-109	1157	8924 M	6912 M (52.1%)	8924 M (77.5%)	4 M	1198	23%	2	14%
IC-pPDX-112	1059	10064 M	7951 M (47.9%)	10064 M (79.0%)	4.55 M	1427	30%	3	26%
IC-pPDX-63	903	5362 M	4378 M (50.5%)	5362 M (81.7%)	3 M	950	29%	2	20.80%
IC-pPDX-75	322	5949 M	4060 M (35.8%)	5949 M (68.2%)	6.6 M	2060	38%	3	20.50%
NB1572	977	8723 M	7358 M (50.1%)	8723 M (84.4%)	4.47 M	1389	31.60%	2	9.40%
NB1576	584	5107 M	4297 M (52.8%)	5107 M (84.1%)	4.6 M	1468	29%	3	26%
NB1580	71	10964 M	8327 M (5.5%)	10964 M (75.9%)	8.54 M	2623	84%	2	13%
NB1583	980	4761 M	4036 M (56.6%)	4761 M (84.8%)	2.75 M	851	24%	3	11%

M=million

Table 3.1.1.6.2 Single-Cell Phasing and Sample Selection for Replication Timing Analysis in Neuroblastoma Study.

Sample id	G1/G2-phase cells	Noisy cells	Non tumor cells	S-phase cells	Inclusion in Replication Timing (RT) study
GR-NB10	257	6	0	59	Yes
GR-NB4	865	15	0	114	Yes
GR-NB5	84	8	0	45	Yes
GR-NB7	400	36	0	189	Yes
HSJD-NB-003	326	5	0	65	Yes
HSJD-NB-004	114	4	0	13	No
HSJD-NB-005	434	16	0	93	Yes
HSJD-NB-009	323	10	0	72	Yes
HSJD-NB-011	170	2	0	27	No
IC-pPDX-109	962	24	0	171	Yes
IC-pPDX-112	884	29	0	146	Yes
IC-pPDX-17	523	15	0	124	Yes
IC-pPDX-63	723	30	0	150	Yes
IC-pPDX-75	255	5	0	62	Yes
NB1572	359	7	555	56	No
NB1576	458	10	23	93	No
NB1580	4	5	44	18	No
NB1583	663	13	195	109	Yes

Table 3.1.1.6.3. In Vivo Investigation of Clonal Evolution in High-Risk Neuroblastoma Using Patient-Derived Xenograft (PDX) Models and Targeted Therapies.

HR-NB PDX Models	Somatic Genetic Alterations					Chosen target	Targeted drug	Sequenced samples
	<i>MYCN</i>	<i>ALK</i>	<i>TP53</i>	<i>ATRX</i>	Other			
GR-NB4	A	A	WT	WT		<i>ALK</i>	Lorlatinib	Tumor
GR-NB10	WT	WT	c.555+1G>A	WT	<i>NF1</i> mut + LOH	<i>NF1</i>	Trametinib	Tumor
IC-pPDX-17	WT	WT	WT	c.5242G>A	<i>CDK4</i> amp	<i>CDK4</i>	Ribociclib	Tumor
IC-pPDX-75	WT	F1174L	WT	c.6391C>T		<i>ALK</i>	Lorlatinib	Tumor
IC-pPDX-109	A	WT	WT	WT	<i>HRAS</i> mut	<i>HRAS</i>	Trametinib	Tumor
HSJD-NB-005	A	WT	c.517G>T	WT	<i>ARID1A</i> mut	<i>ARID1A</i>	Tazemetostat	Tumor

A/amp=Amplification; WT=Wild type; mut = somatic single nucleotide variant; LOH = Loss of heterozygosity

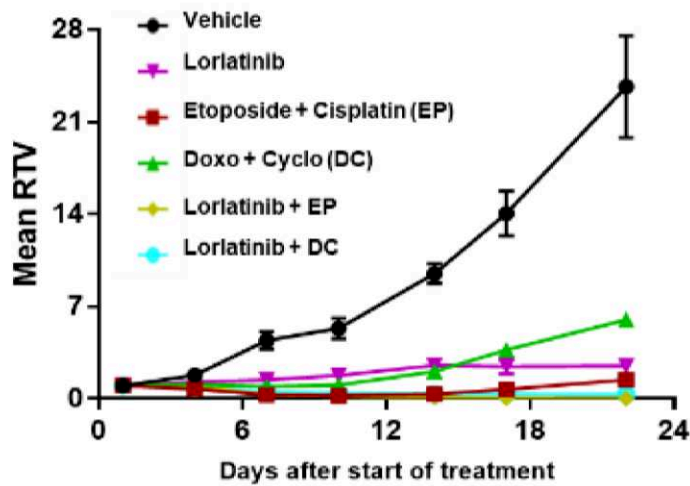


Figure 3.1.1.6.25. GR-NB4, a *ALK*-amplified NB PDX model | Relative tumor volume (mean).

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independently and in combination with lorlatinib. These curves highlighted the individual activity of each chemotherapy, as well as the impact of lorlatinib, a *ALK* inhibitor, when used alone.

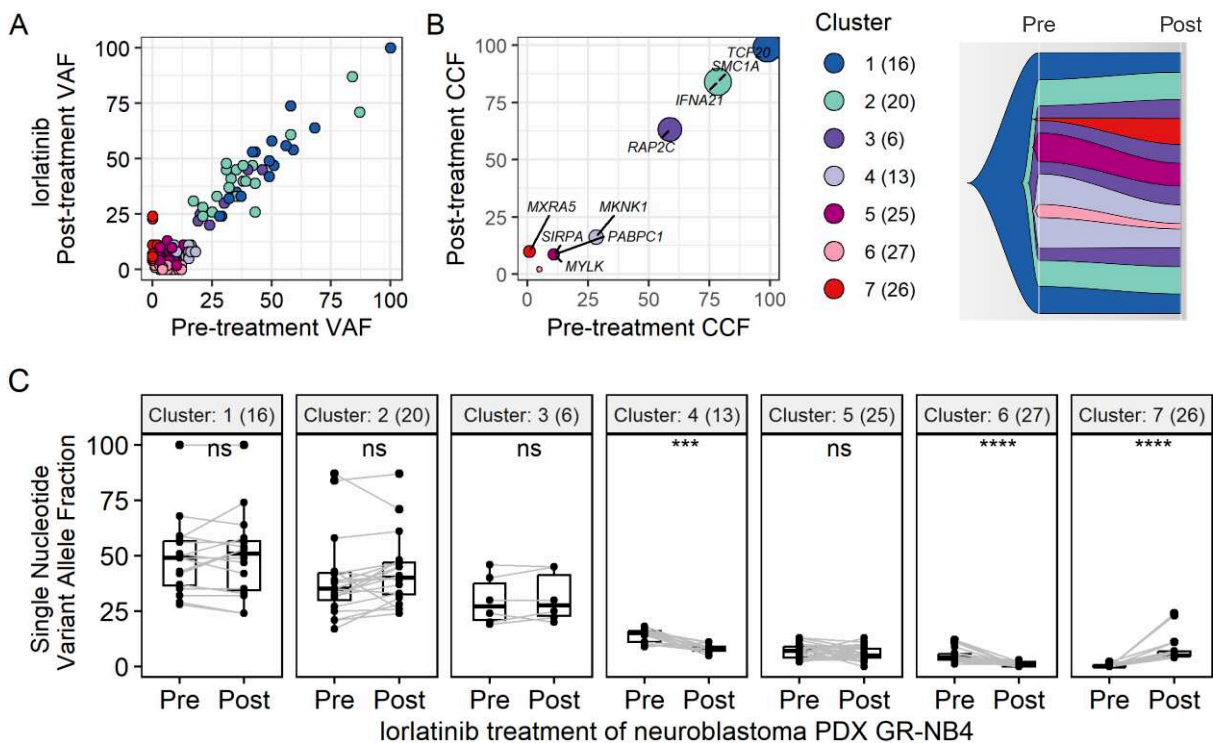


Figure 3.1.1.6.26: GR-NB4, a *ALK* amplified NB PDX model | The impact of targeted treatment (lorlatinib) alone on different clonal populations within a tumor

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-6 and clone-7, displayed dynamic behavior, with clone-4 and clone-6 gradually diminishing and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

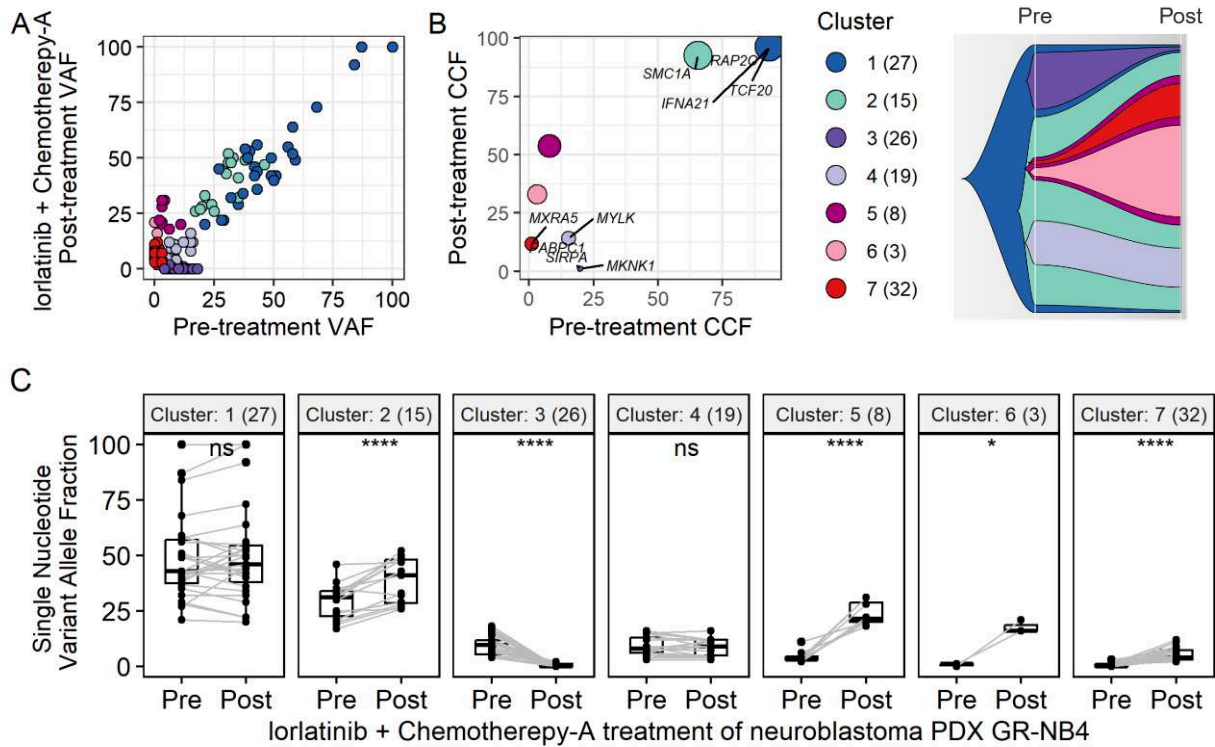


Figure 3.1.1.6.27: GR-NB4, a *ALK* amplified NB PDX model | The impact of targeted treatment (lorlatinib) in combination with chemotherapy-A.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3, clone-5, clone-6 and clone-7, displayed dynamic behavior, with clone-3 gradually diminishing and clone-5, clone-6 and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

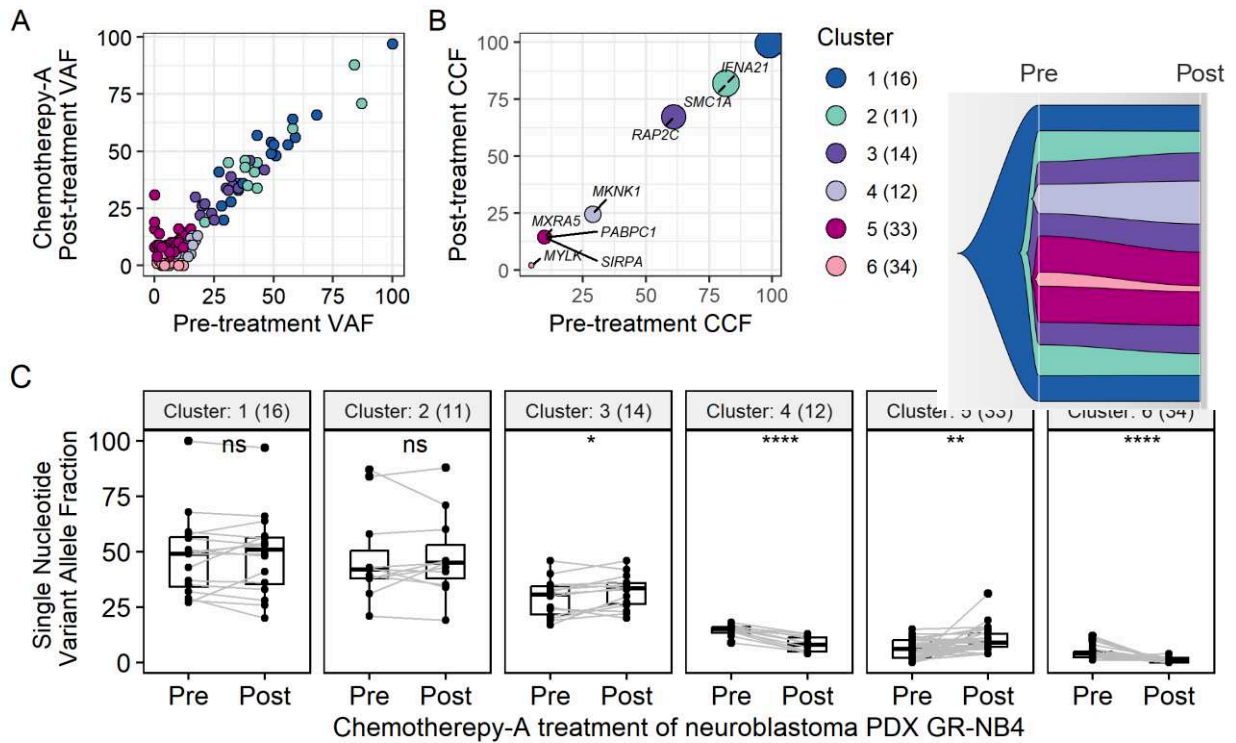


Figure 3.1.1.6.28. GR-NB4, a *ALK* amplified NB PDX model | The impact of chemotherapy-A treatment.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5 and clone-6 displayed dynamic behavior, with clone-4 and clone-6 gradually diminishing while clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***); 0.0001 (****).

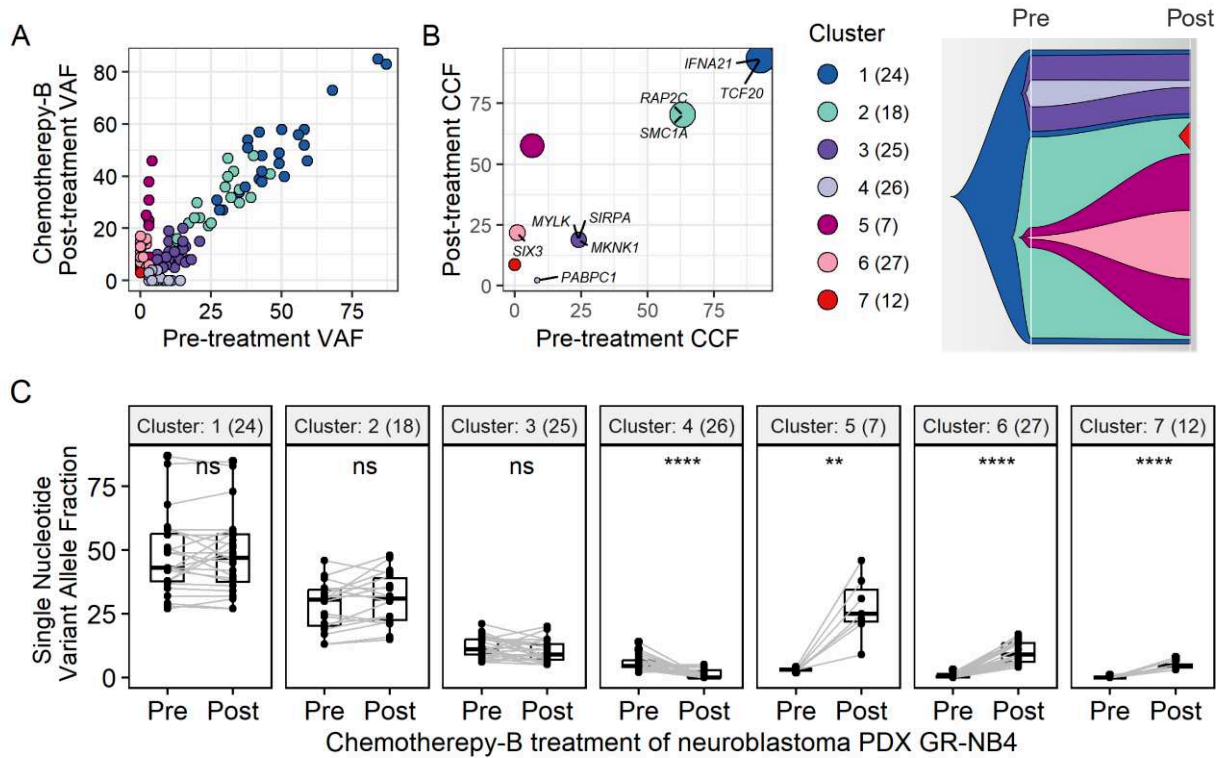


Figure 3.1.1.6.29: GR-NB4, a *ALK* amplified NB PDX model | The impact of chemotherapy-B treatment.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5, clone-6 and clone-7 displayed dynamic behavior, with clone-4 gradually diminishing while clone-5, clone-6 and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

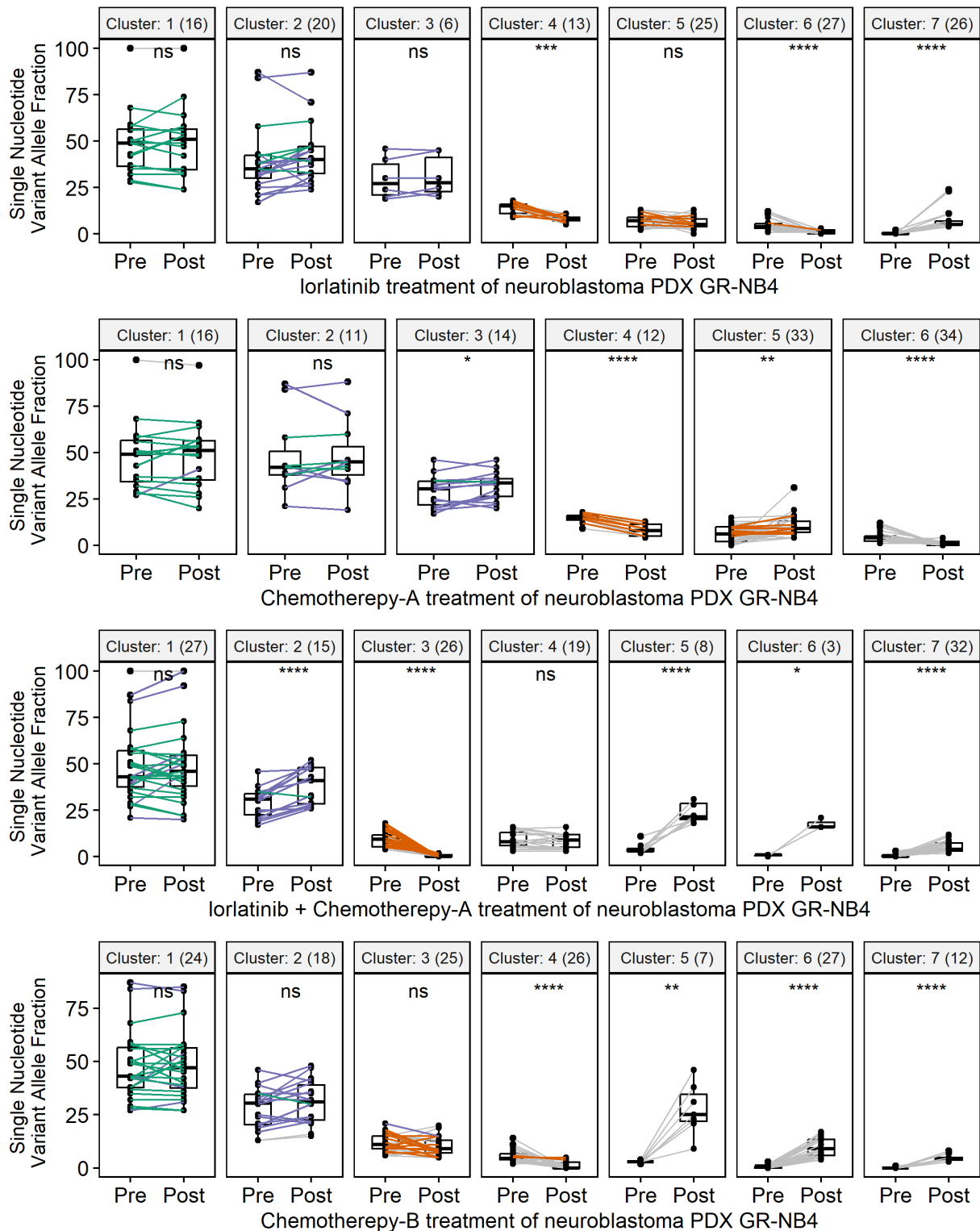


Figure 3.1.1.6.30: GR-NB4 model: SNVs from s1 and s2 subclone mapped to WES derived clones.

In the given data visualization, the green line symbolizes the common or shared Single Nucleotide Variants (SNVs) between superclone/subclone s1 and s2, indicating the genetic overlap between these two subclonal groups. The purple line denotes SNVs unique to superclone/subclone s1, highlighting the distinct genetic features of this subgroup. Notably, the orange line represents SNVs exclusive to subclone s2. The predominance of orange colored SNVs in the diminishing group suggests that superclone/subclone s2 is more responsive to the treatment regimen. These unique variants in the s2 subclone appear to exhibit increased sensitivity to the combined therapy of lorlatinib and chemotherapy-A.

Summary:

The in vivo experiment and bulk exome sequencing data analysis focusing on GR-NB4 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment:

- **Lorlatinib:** Lorlatinib alone demonstrated substantial efficacy when used as a single agent, achieving a Tumor Growth Inhibition (TGI) of 91%. This indicates that lorlatinib is highly effective in targeting and inhibiting the growth of neuroblastoma cells in this specific model.
- No change in clonal status of *ALK* amplification before and after any treatment regimen
- In GR-NB4 *MYCN* and *ALK* amplified model, no single nucleotide variations (SNVs) were identified that are typically associated with or considered as drivers of neuroblastoma.
- Gradually diminishing and emerging subclones were significantly evident after targeted treatment alone or in combination with chemotherapy.
- The pathway analysis of emerging genes suggests that these emerging genes might play a key role in negative regulation of cell differentiation (e.g. *PROX1*, *SIX3*, *CRI*, *DUSP10*, *ANP32B*, *CCL3*, *SIX2*, *RNF10*, *ADAMTS7*) and epithelial to mesenchymal transition (e.g. *MYLK*, *MXRA5*).
- The hypothesis posits that the majority of mutations that gradually diminished after lorlatinib treatment, either as a standalone therapy or in combination with chemotherapy, originated from Clone-B. This clone, which was evolving in parallel, was identified through the analysis of single-cell whole genome sequencing (scWGS) data. This suggests that Clone-B was sensitive to lorlatinib and its susceptibility contributed significantly to the overall reduction in these mutations post-treatment.
- **Combination with Cisplatin-Etoposide:** The combination of lorlatinib with the chemotherapy-B agents cisplatin and etoposide significantly enhanced the therapeutic effect compared to either treatment alone. This regimen resulted in a higher number of complete responses, signifying that lorlatinib may synergize with traditional chemotherapy to combat neuroblastoma more effectively.
- **Combination with Doxorubicin-Cyclophosphamide:** Similarly, combining lorlatinib with doxorubicin and cyclophosphamide also significantly increased the anti-tumor efficacy compared to these chemotherapeutic agents alone. The presence of complete responses in

this combination group further underscores the potential of lorlatinib to improve treatment outcomes when used in conjunction with established chemotherapy protocols.

Overall, these results suggest that lorlatinib, either alone or in combination with standard chemotherapy regimens, shows promising therapeutic potential in the treatment of neuroblastoma, as evidenced by the high TGI and increased complete response rates observed in the GR-NB4 model.

Relative tumor volume (mean)

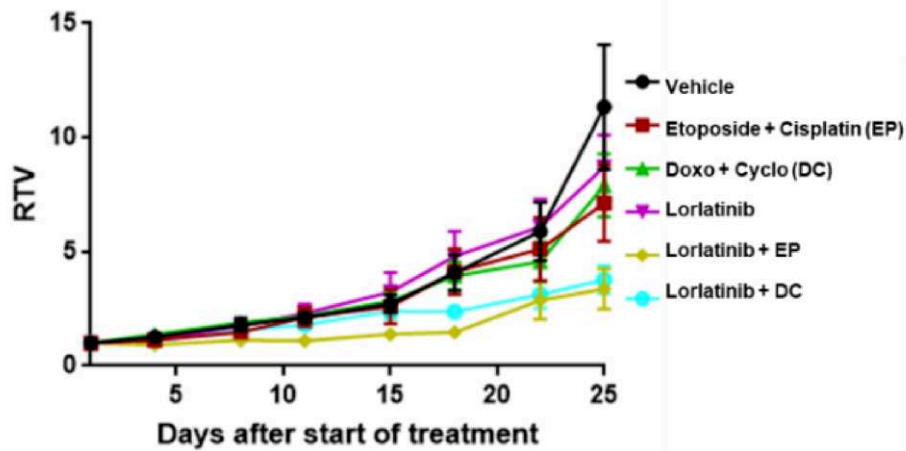


Figure 3.1.1.6.31. IC-pPDX-75, a *ALK*-mutated NB PDX model, growth curves

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independently and in combination with lorlatinib. These curves highlighted the individual activity of each chemotherapy, as well as the impact of lorlatinib, a *ALK* inhibitor, when used alone.

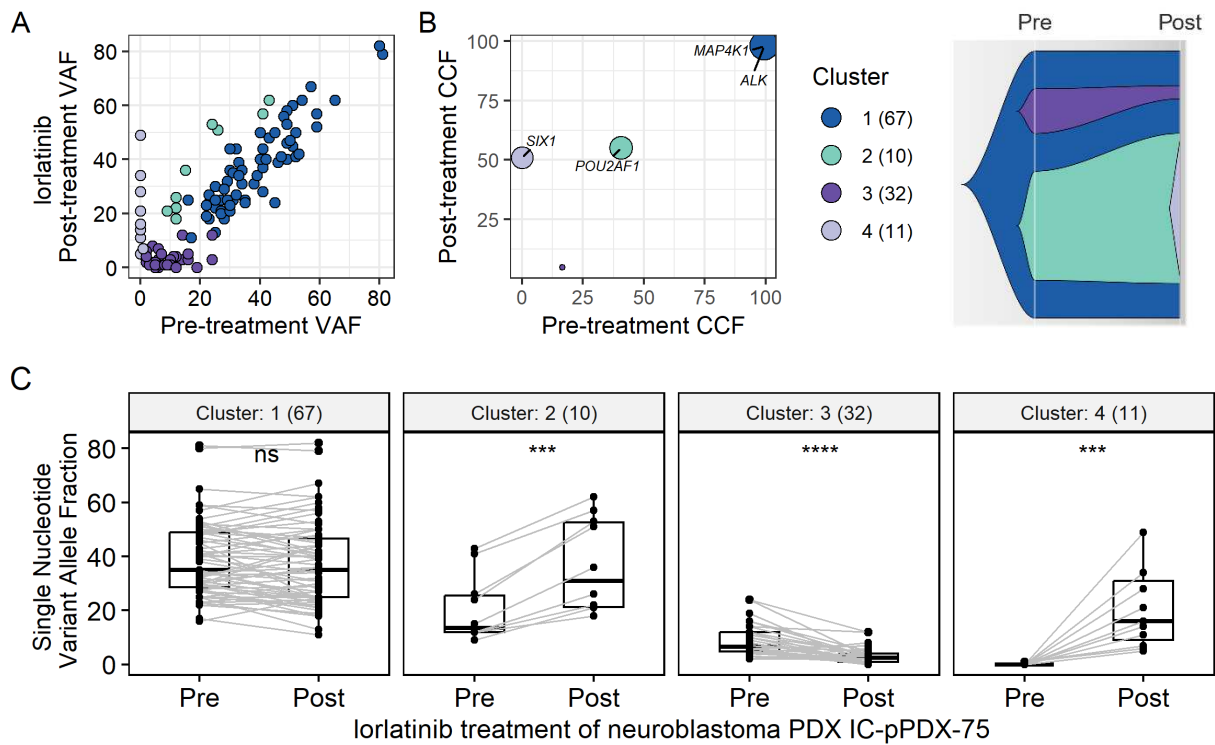


Figure 3.1.1.6.32. IC-pPDX-75, *ALK* mutated NB PDX model | The impact of targeted treatment (lorlatinib) alone.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-4, displayed dynamic behavior, with clone-3 gradually diminishing and clone-4 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

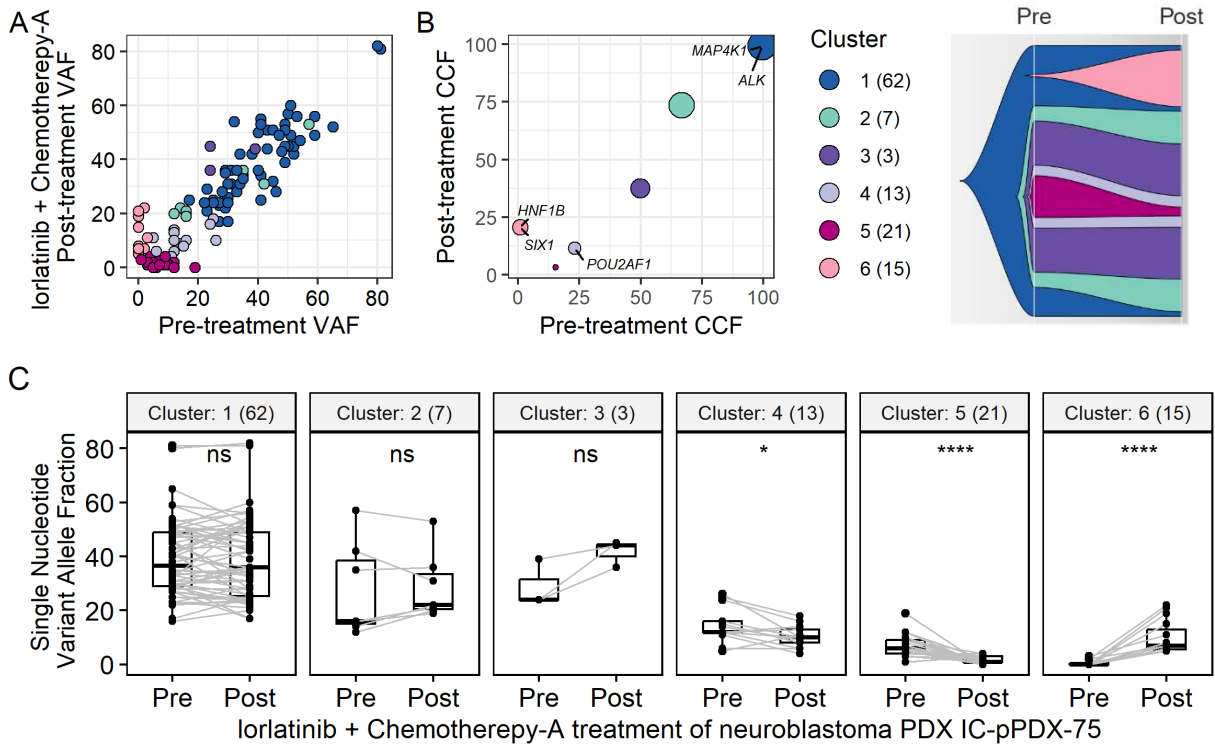


Figure 3.1.1.6.33. IC-pPDX-75, *ALK* mutated NB PDX model | The impact of targeted treatment (lortatinib) in combination with chemotherapy-A

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5 and clone-6, displayed dynamic behavior, with clone-4 and clone-5 gradually diminishing and clone-6 emerging over the course of treatment. This fluctuation in the subclonal population was significant

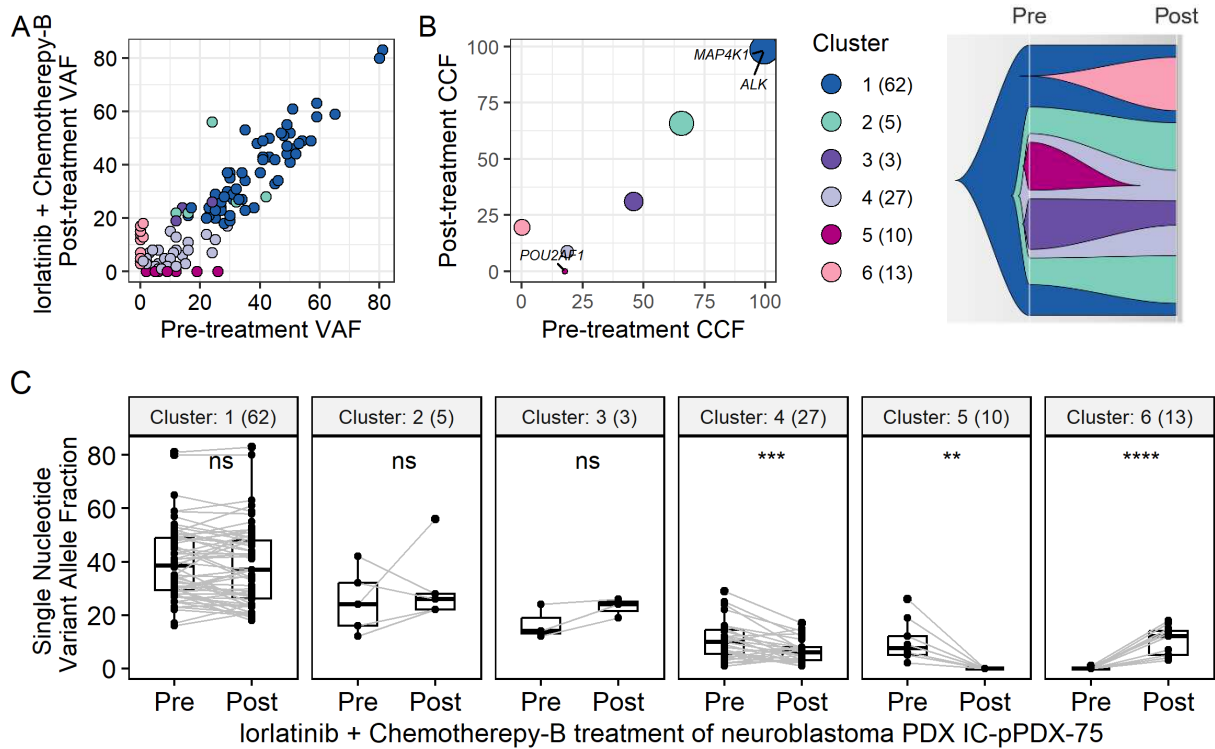


Figure 3.1.1.6.34. IC-pPDX-75, *ALK* mutated NB PDX model | The impact of targeted treatment (lorlatinib) in combination with chemotherapy-B

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5 and clone-6, displayed dynamic behavior, with clone-4 and clone-5 gradually diminishing and clone-6 emerging over the course of treatment. This fluctuation in the subclonal population was significant

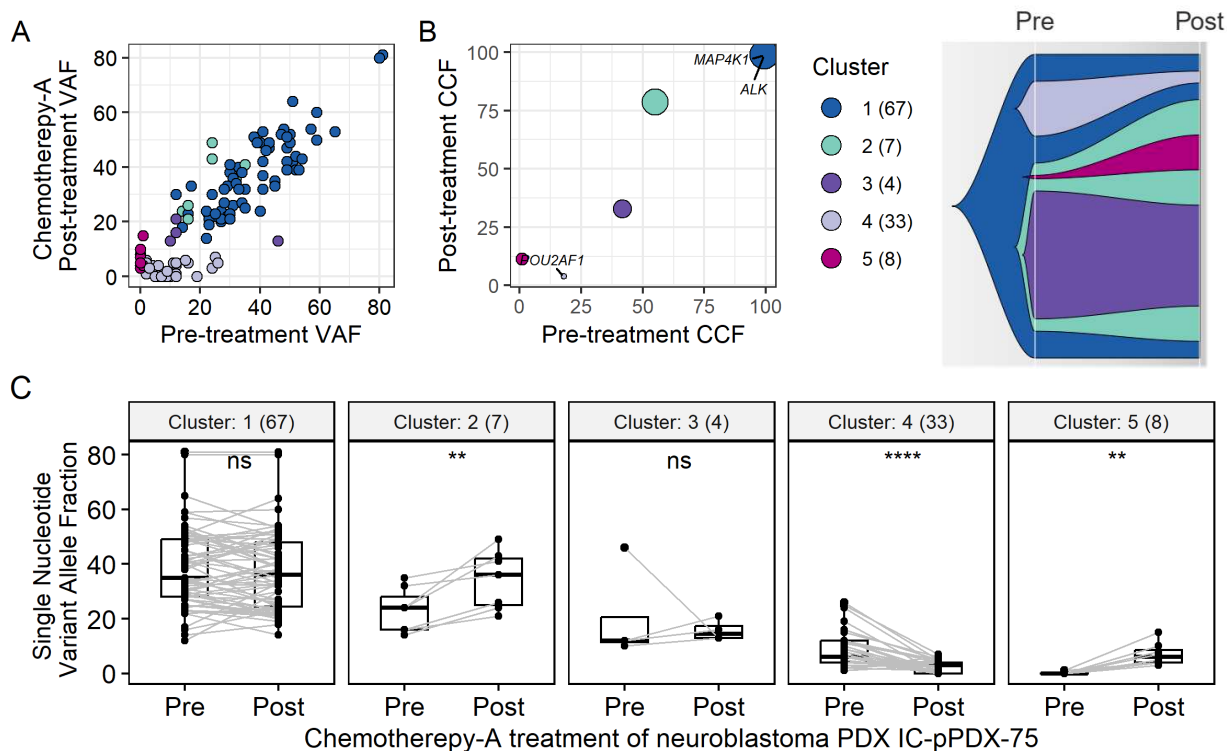


Figure 3.1.1.6.35: IC-pPDX-75, *ALK* mutated NB PDX model | The impact of chemotherapy-A treatment alone

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4 and clone-5, displayed dynamic behavior, with clone-4 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant

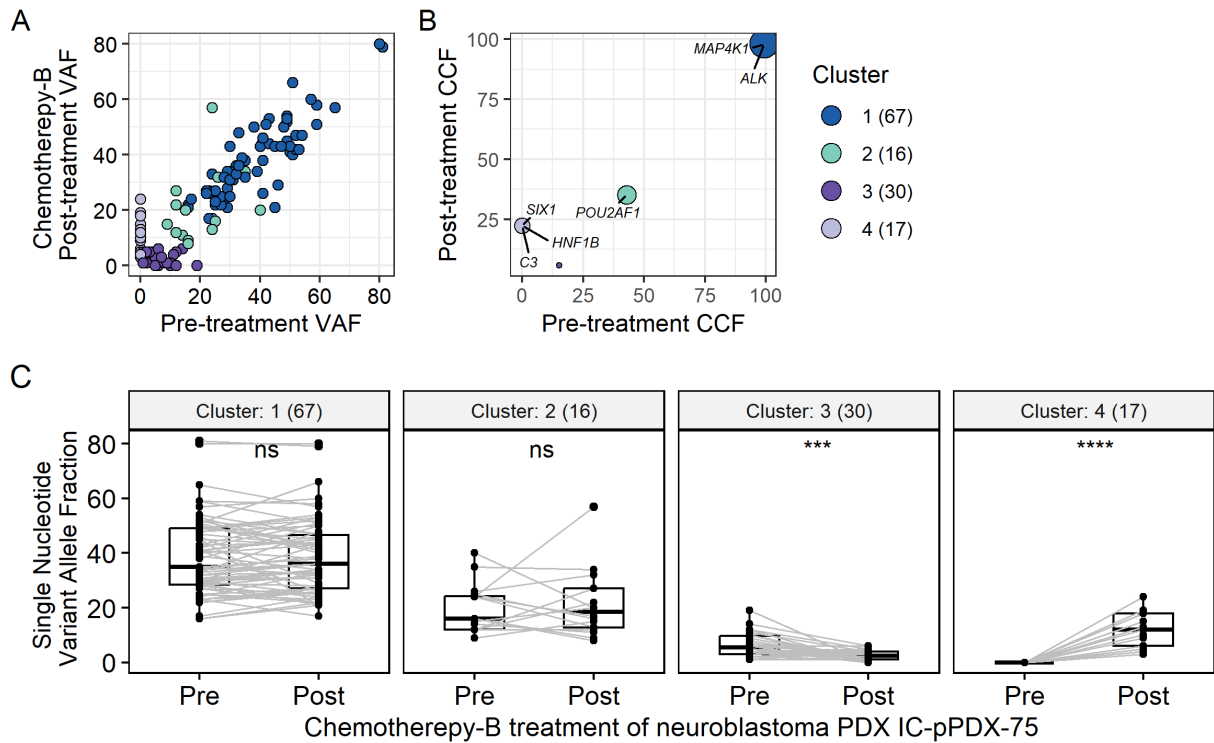


Figure 3.1.1.6.36: IC-pPDX-75, *ALK* mutated NB PDX model | The impact of chemotherapy-B treatment alone

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-4 displayed dynamic behavior, with clone-3 gradually diminishing and clone-4 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

Summary:

The in vivo experiment and bulk exome sequencing data analysis focusing on IC-pPDX-75 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment:

- Lorlatinib: In the IC-pPDX-75 model, lorlatinib alone, when administered as a solitary treatment, showed limited efficacy with a Tumor Growth Inhibition (TGI) rate of only 28%. This suggests that lorlatinib on its own is not sufficiently effective against this neuroblastoma PDX model targeting *ALK* pathogenic mutation.
- No change in *ALK* pathogenic mutation (F1174L) irrespective of treatment regimen.
- Emerging genes might play key role in cell-cell signaling biological process (e.g. *HNF1B*, *ADAM10*, *CACNG2*, *BSN*, *C3*, *PCDHB3*, *RNF207*, *GRID2IP*, *HTRIF* and *AMFR*)

- In combination with Cisplatin + Etoposide (chemotherapy-B): When lorlatinib was combined with the chemotherapy-B drugs cisplatin and etoposide, there was an increase in the efficacy of lorlatinib. However, this combination did not enhance the effectiveness of the chemotherapy agents themselves. This indicates a partial synergy where lorlatinib benefits from the presence of these chemotherapeutic drugs.
- Combination with Doxorubicin + Cyclophosphamide: Notably, the combination of lorlatinib with doxorubicin and cyclophosphamide significantly improved the anti-tumor efficacy compared to either lorlatinib or the chemotherapy agents used alone. This outcome suggests a more pronounced synergistic effect, where the combination therapy notably outperforms individual treatments.

Overall, these results indicate that while lorlatinib alone lacks substantial efficacy in the IC-PDX-75 neuroblastoma model, its combination with certain chemotherapy regimens, particularly doxorubicin and cyclophosphamide (chemotherapy-A), markedly improves anti-tumor activity. This finding underscores the potential of lorlatinib as part of a combination therapy, rather than as a standalone treatment, in certain neuroblastoma cases.

Relative tumor volume (mean)

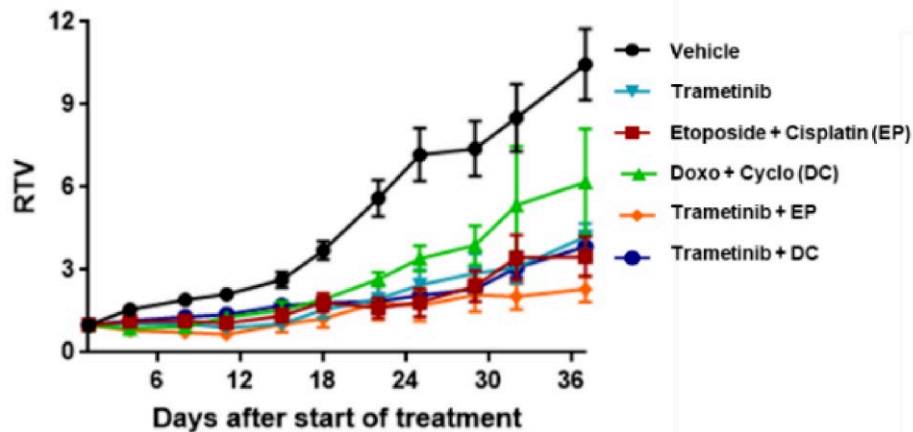


Figure 3.1.1.6.37. IC-pPDX-109, a *HRAS*-mutated NB PDX model, growth curves

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independently and in combination with trametinib. These curves highlighted the individual activity of each chemotherapy, as well as the impact of trametinib, a *HRAS* inhibitor, when used alone.

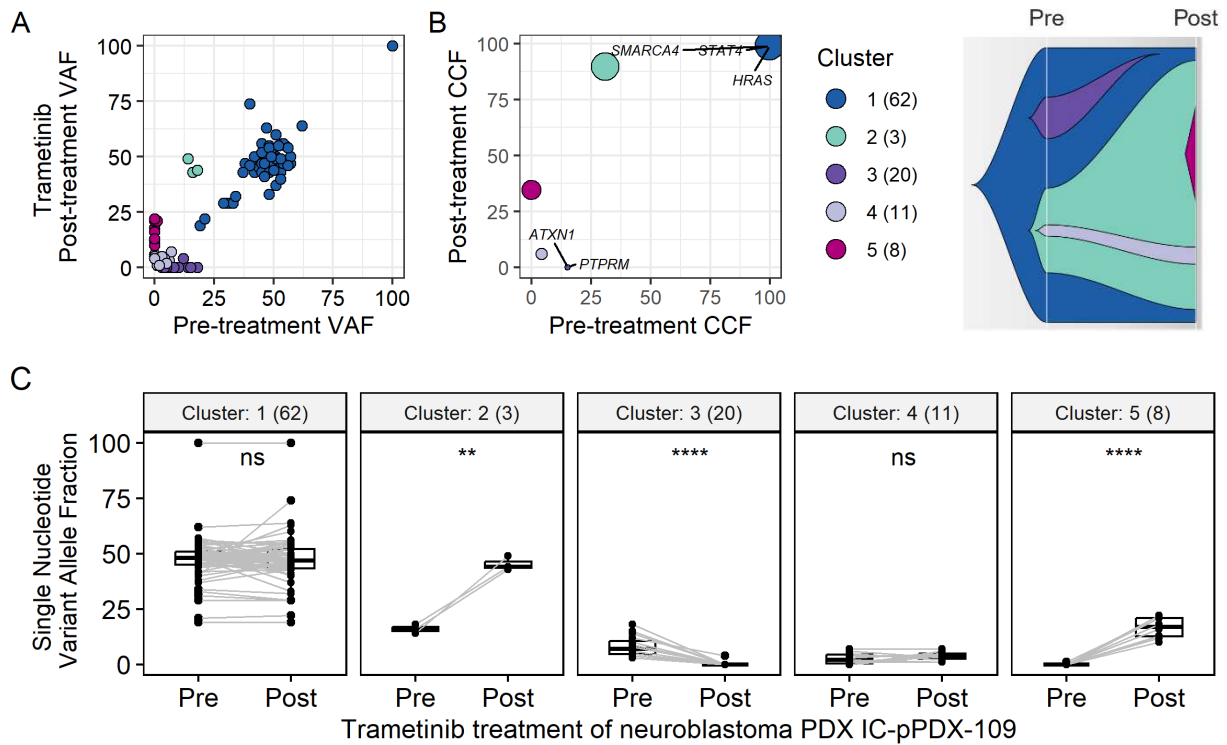


Figure 3.1.1.6.38: IC-pPDX-109, HRAS mutant NB PDX model | The impact of Targeted treatment (trametinib) alone.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-5, displayed dynamic behavior, with clone-3 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant

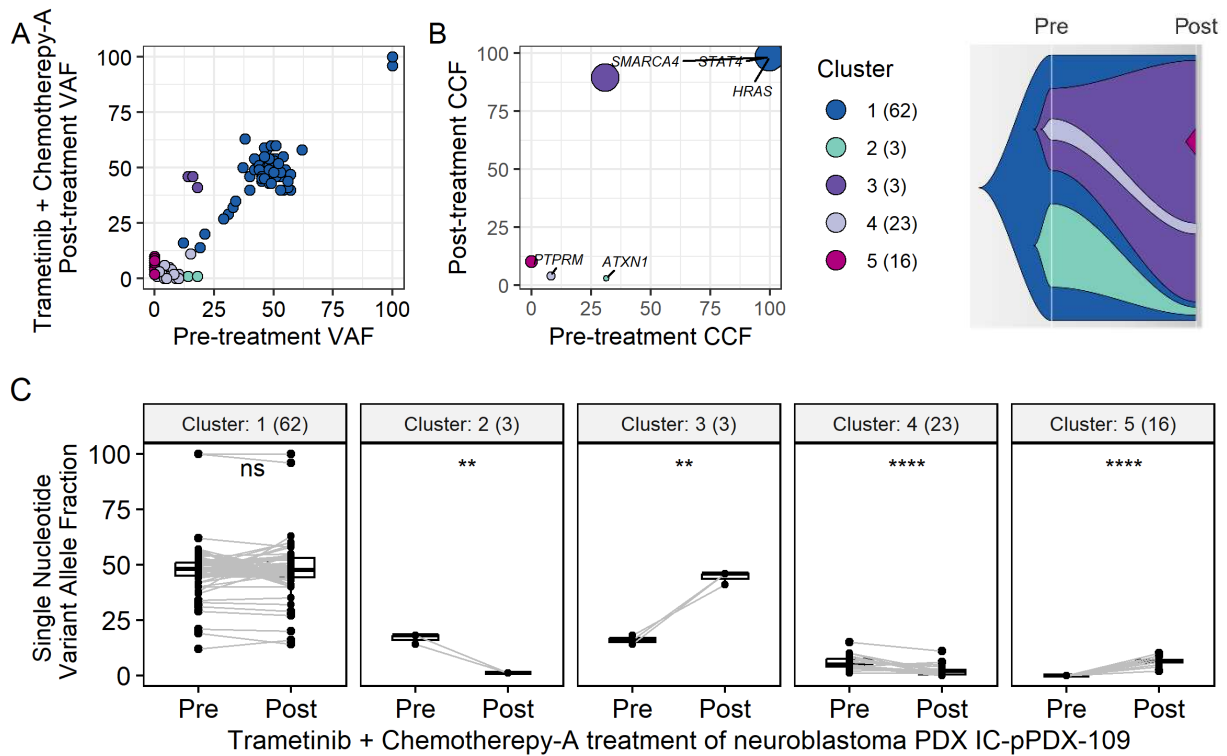


Figure 3.1.1.6.39: IC-pPDX-109, HRAS mutant NB PDX model | The impact of Targeted treatment (trametinib) in combination with chemotherapy-A

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-2, clone-3 and clone-5, displayed dynamic behavior, with clone-2 and clone-4 gradually diminishing while clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant

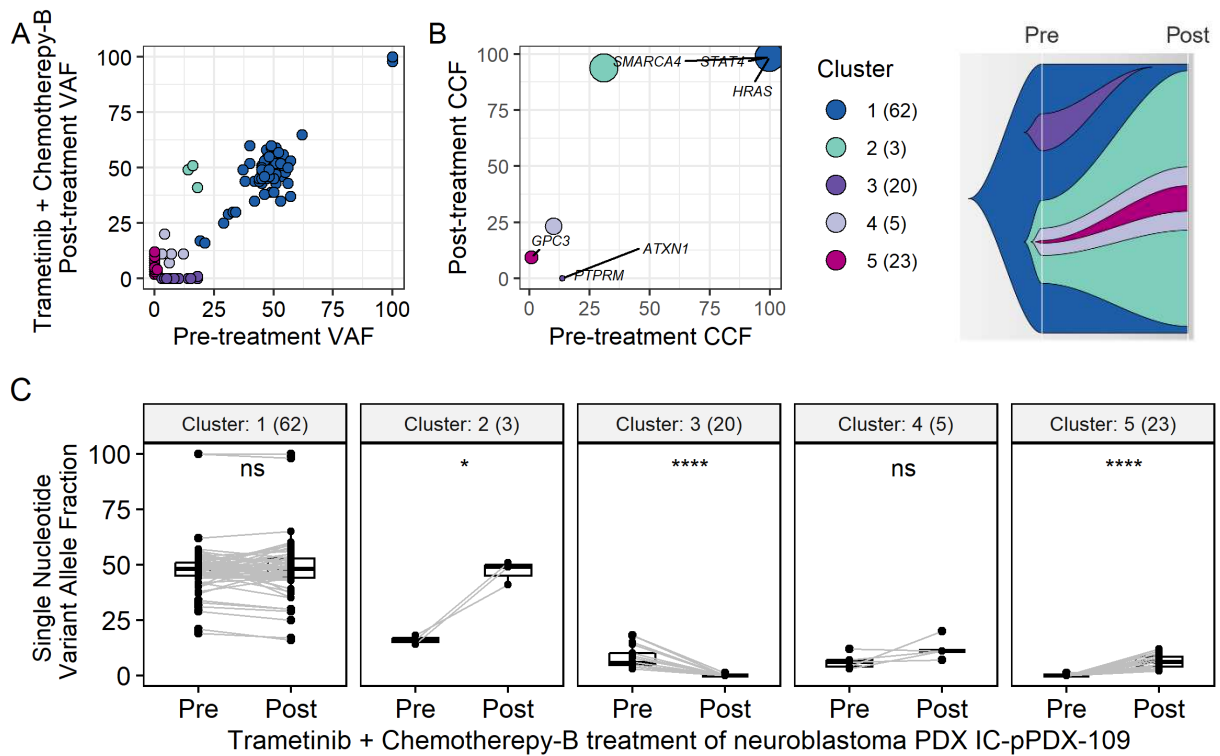


Figure 3.1.1.6.40. IC-pPDX-109, HRAS mutant NB PDX model | The impact of Targeted treatment (trametinib) in combination with chemotherapy-B treatment.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-5, displayed dynamic behavior, with clone-3 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

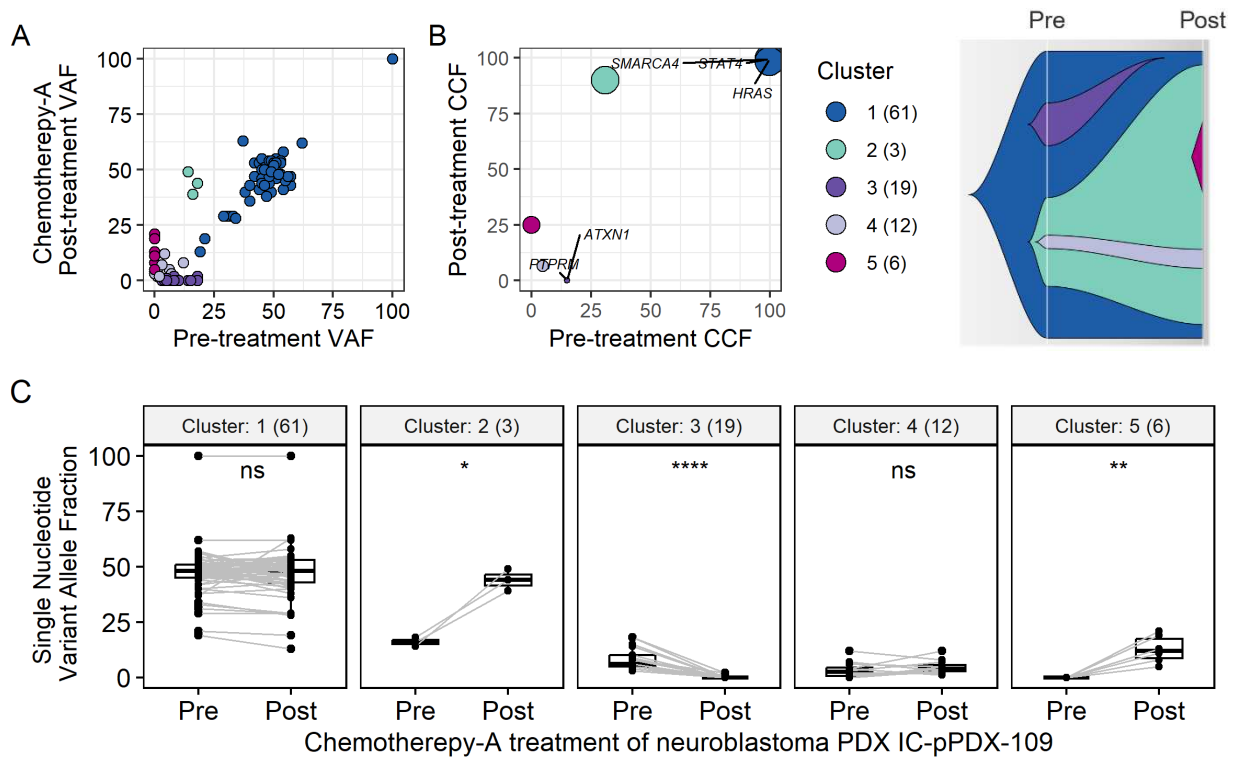


Figure 3.1.1.6.41. IC-pPDX-109, HRAS mutant NB PDX model | The impact of chemotherapy-A treatment alone.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-5, displayed dynamic behavior, with clone-3 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

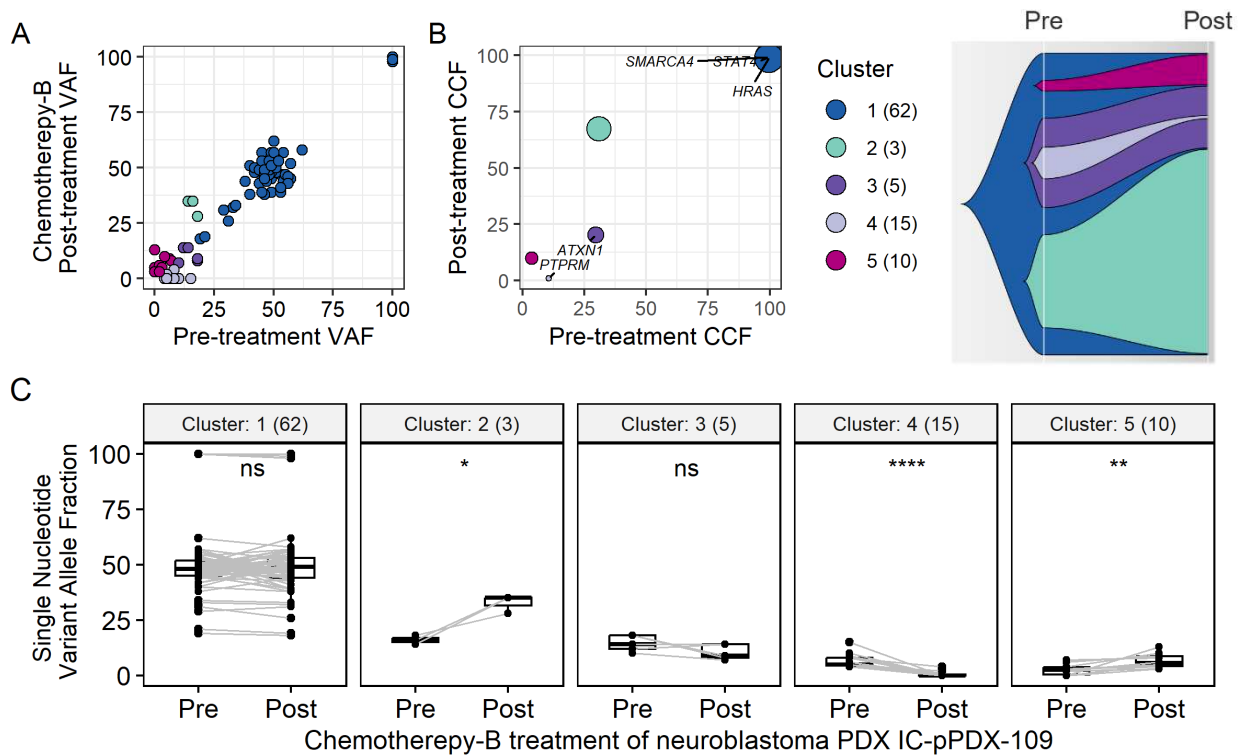


Figure 3.1.1.6.42. IC-pPDX-109, HRAS mutant NB PDX model | The impact of chemotherapy-B treatment alone.

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4 and clone-5, displayed dynamic behavior, with clone-4 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

Summary:

The in vivo experiment and bulk exome sequencing data analysis focusing on IC-pPDX-109 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment:

- Trametinib: Trametinib alone demonstrated significant anti-tumor activity when used as a single agent against the IC-pPDX-109 model, achieving a Tumor Growth Inhibition (TGI) rate of 60%. This indicates that trametinib is effective in reducing tumor growth in this specific neuroblastoma context.
- Molecular target, HRAS mutation and other likely pathogenic genes such as SMARCA4, and STAT4 remain clonal (CCF ~ 100%) irrespective of any treatment regimen.
- Emerging genes might play key role in mesenchymal to epithelial transition (e.g. *GJAI*, *MYLK*, *LAMA1*) or KRAS signaling pathway (e.g. *PCDHBI*, *RYRI*, *UBX2*)

- Combination with Cisplatin + Etoposide: Notably, when trametinib was combined with cisplatin and etoposide, the efficacy was significantly higher than trametinib used alone ($p=0.0317$). However, the combination did not surpass the efficacy of cisplatin and etoposide used without trametinib. This suggests that while trametinib enhances the effectiveness of these chemotherapy drugs, the combination does not yield a superior outcome compared to the chemotherapy duo alone.
- Addition of Doxorubicin + Cyclophosphamide: The inclusion of doxorubicin and cyclophosphamide with trametinib did not result in increased efficacy compared to trametinib used alone. This implies that the combination of trametinib with these chemotherapeutic agents does not confer an additional benefit in terms of anti-tumor activity.

In summary, trametinib shows notable efficiency as a monotherapy in the IC-pPDX-109 neuroblastoma model and enhances the anti-tumor effects of the cisplatin and etoposide combination. However, its combination with doxorubicin and cyclophosphamide does not provide added advantage over trametinib alone.

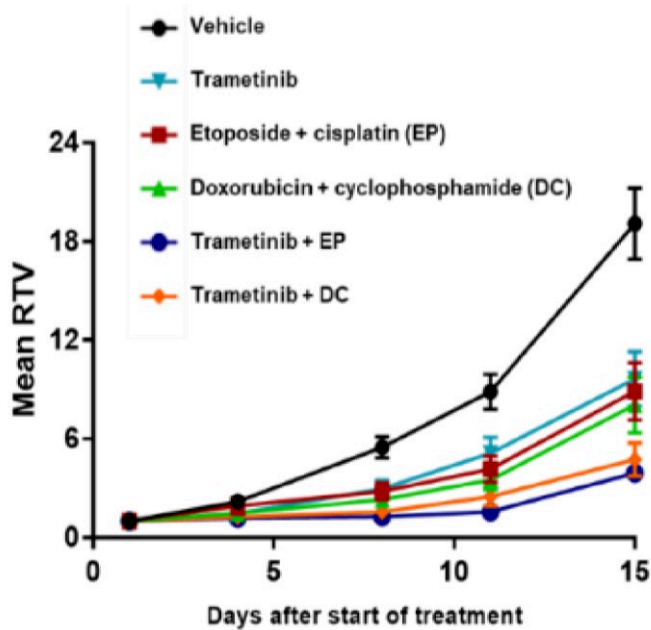


Figure 3.1.1.6.43. GR-NB10, a NF1-mutated NB PDX model, growth curves

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independently and in combination with trametinib. These curves highlighted the individual activity of each chemotherapy, as well as the impact of trametinib, a NF1 inhibitor, when used alone.

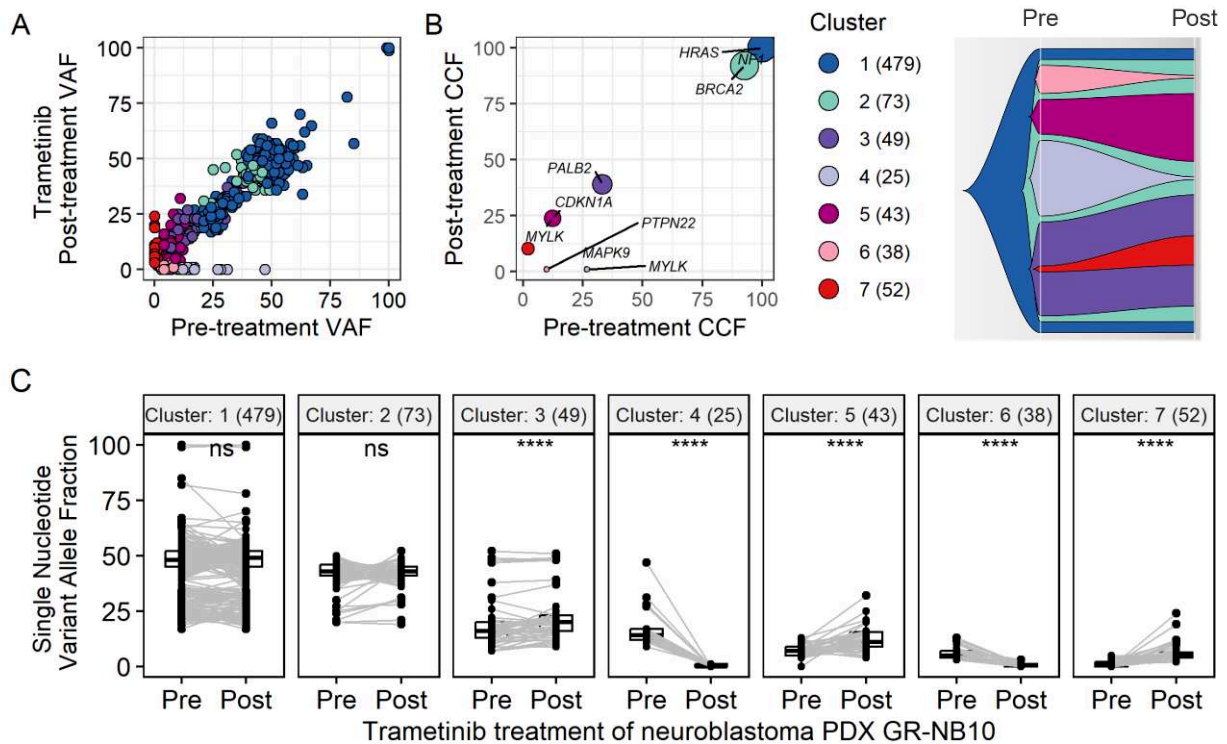


Figure 3.1.1.6.44. GR-NB10, a NF1-mutated NB PDX model | The effect of targeted treatment (trametinib)

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5, clone-6 and clone-7, displayed dynamic behavior, with clone-4, and clone-6 gradually diminishing while clone-5 and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***); 0.0001 (****).

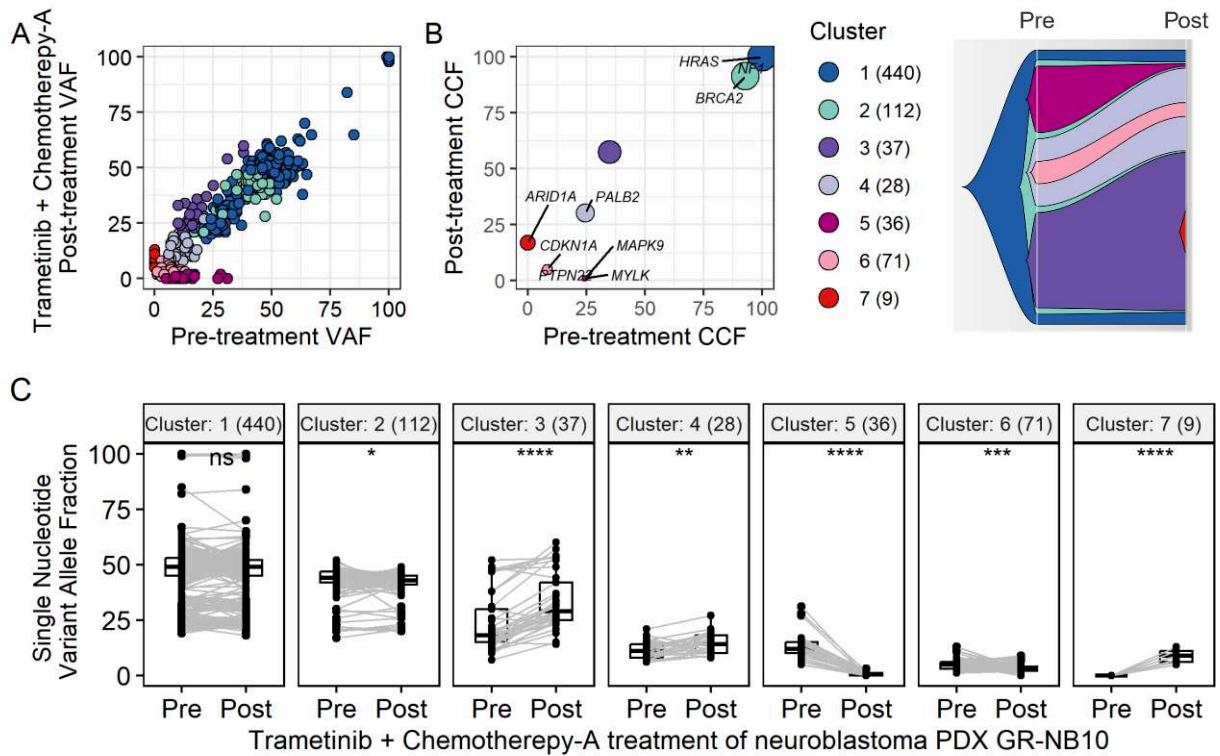


Figure 3.1.1.6.45. GR-NB10, a NF1-mutated NB PDX model | The impact of targeted treatment (trametinib) in combination with chemotherapy-A

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-5, clone-6 and clone-7, displayed dynamic behavior, with clone-5, and clone-6 gradually diminishing and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***); 0.0001 (****).

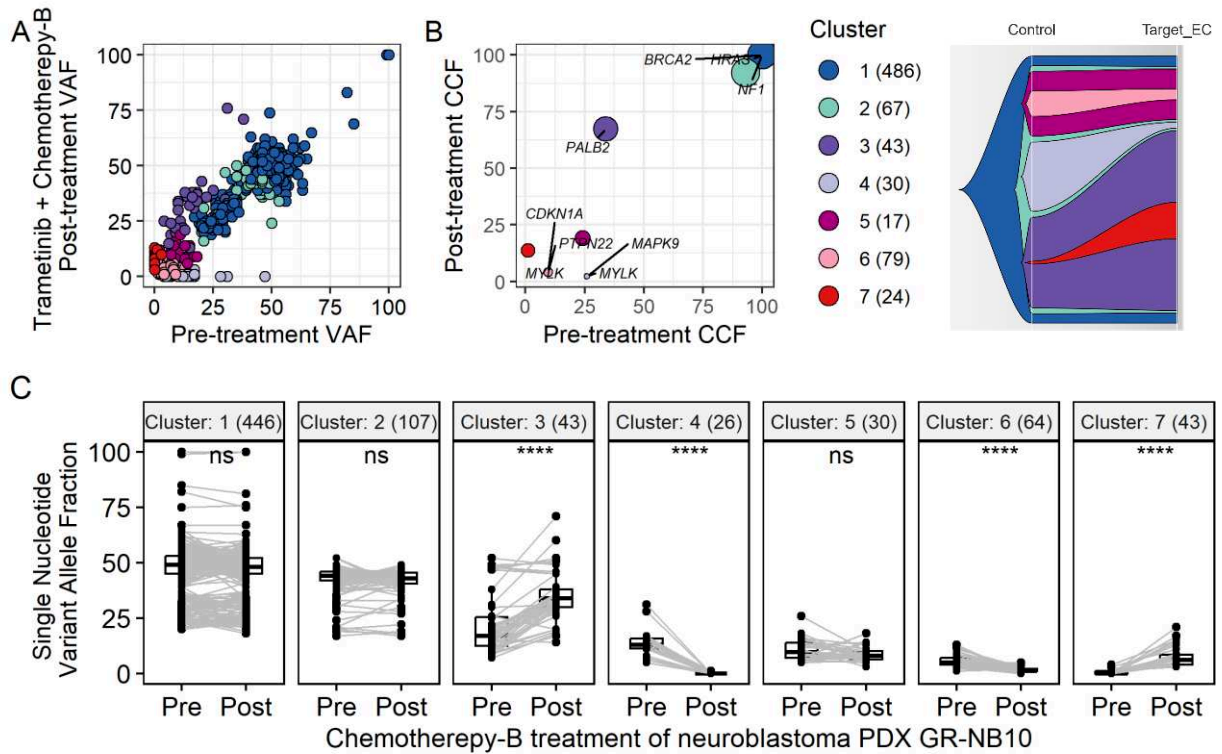


Figure 3.1.1.6.46. GR-NB10, a NF1-mutated NB PDX model | The impact of targeted treatment (trametinib) in combination with chemotherapy-B

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-6 and clone-7, displayed dynamic behavior, with clone-4, and clone-6 gradually diminishing and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) 0.0001 (****).

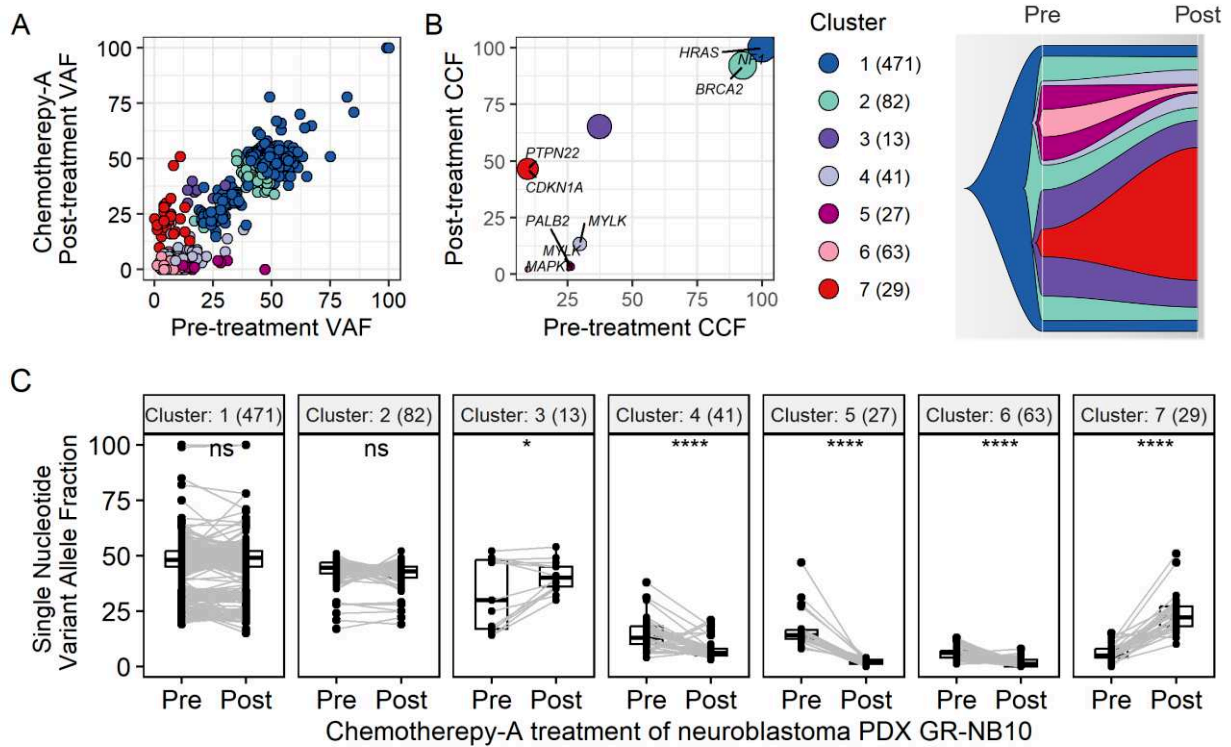


Figure 3.1.1.6.47. GR-NB10, a NF1-mutated NB PDX model | The impact of chemotherapy-A treatment alone

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-5, clone-6 and clone-7, displayed dynamic behavior, with clone-4, clone-5 and clone-6 gradually diminishing and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***); 0.0001 (****).

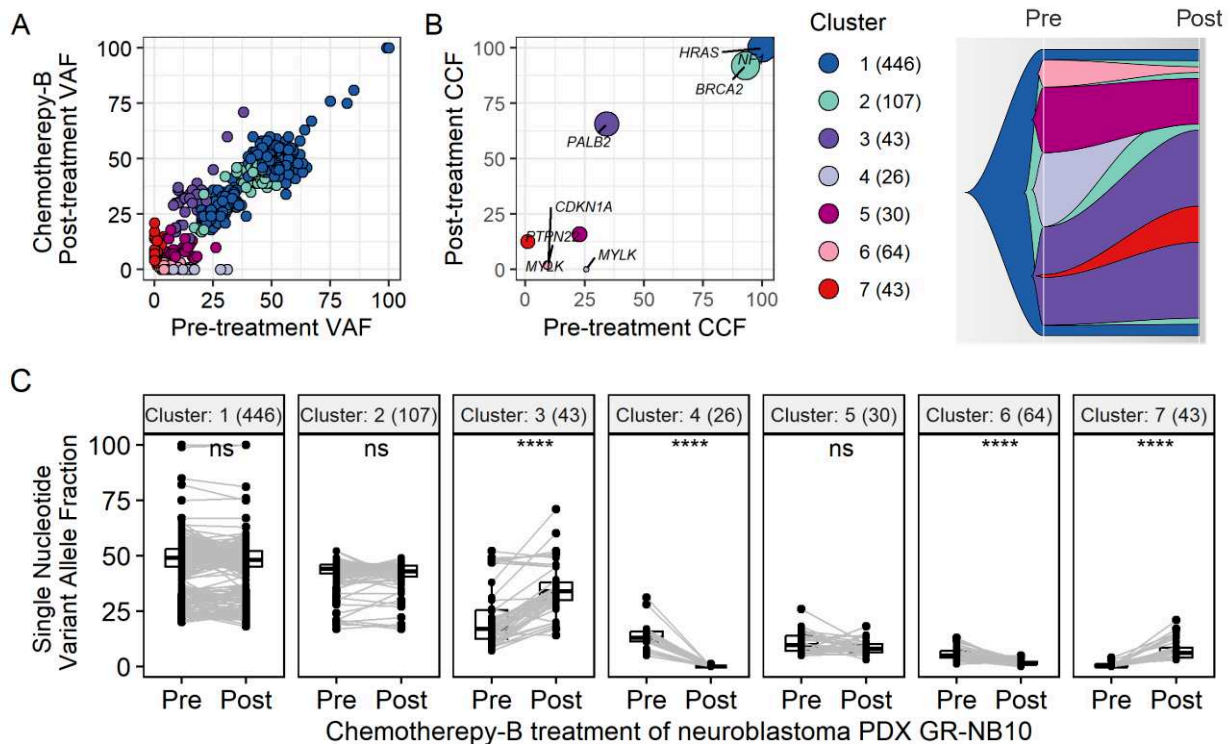


Figure 3.1.1.6.48. GR-NB10, a NF1-mutated NB PDX model | The impact of chemotherapy-B treatment alone

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4, clone-6 and clone-7, displayed dynamic behavior, with clone-4 and clone-6 gradually diminishing and clone-7 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

Summary:

The in vivo experiment and bulk exome sequencing data analysis focusing on GR-NB10 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment.

- In the GR-NB10 neuroblastoma model, trametinib exhibited modest effectiveness as a single-agent therapy, achieving a tumor growth inhibition (TGI) of 49%.
- Molecular target, NF1 mutation and other likely pathogenic genes such as HRAS, and BRCA2 remain clonal (CCF ~ 100%) irrespective of any treatment regimen.
- GR-NB10 case is known hypermutated case (TMB > 10 mutations/Mb). Activities related to Neurogenesis pop-up when emerging genes were studied.

- When trametinib was combined with cisplatin-etoposide chemotherapy, this pairing did not enhance the efficacy beyond that observed with either trametinib alone or the chemotherapy alone.
- However, the combination of trametinib with doxorubicin-cyclophosphamide was notably more effective than either trametinib or the chemotherapy when used independently. This increased efficacy was evident even in the early stages of the experiment.

These findings indicate that while trametinib demonstrates some effectiveness in the GR-NB10 model, its combination with doxorubicin-cyclophosphamide significantly boosts its anti-tumor capabilities, a synergy not observed with the etoposide-cisplatin combination.

Relative tumor volume (mean)

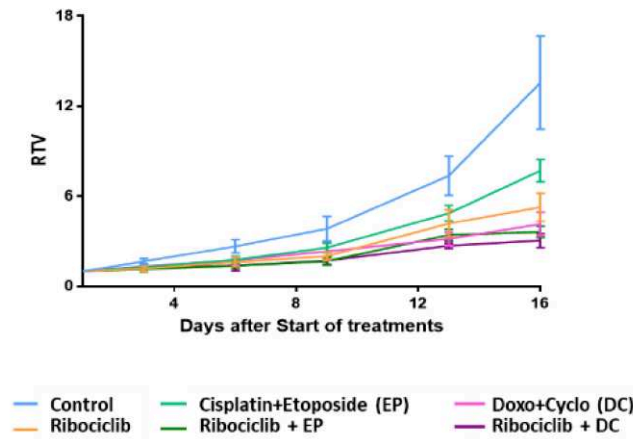


Figure 3.1.1.6.49. IC-pPDX-17, a CDK4-amplified model, growth curves

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independently and in combination with ribociclib. These curves highlighted the individual activity of each chemotherapy, as well as the impact of ribociclib, a CDK4 inhibitor, when used alone.

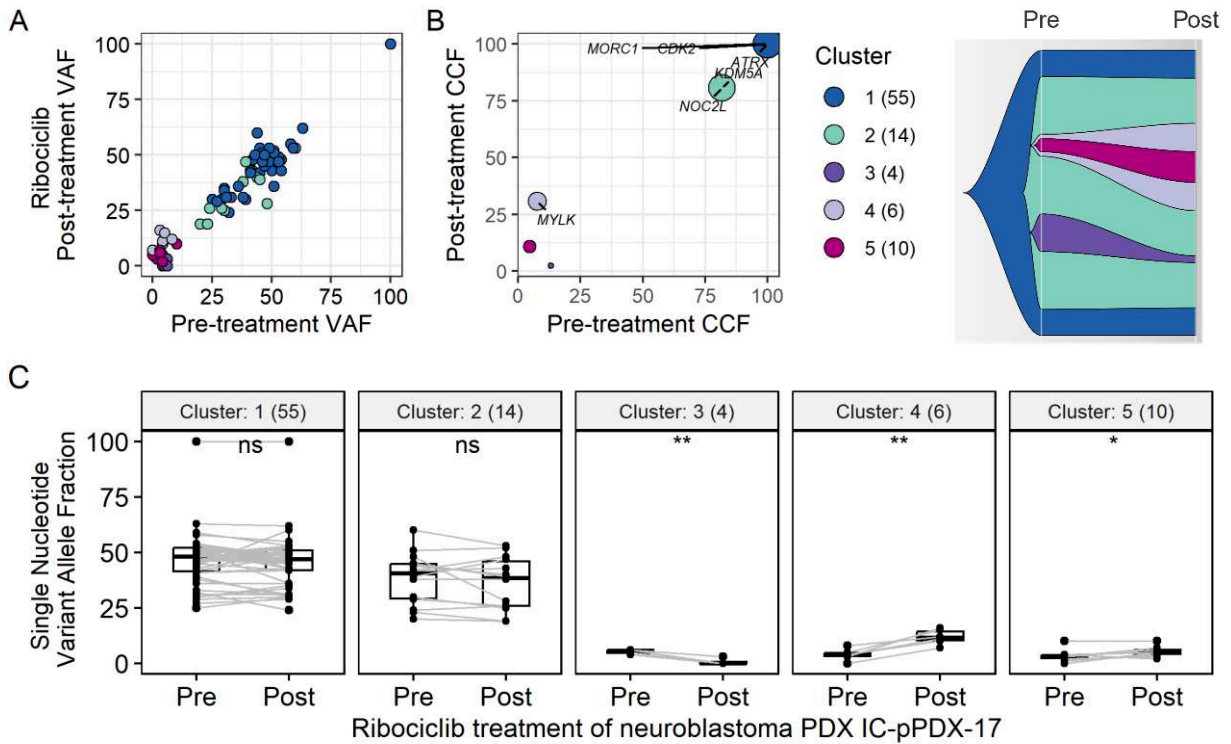


Figure 3.1.1.6.50. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) alone on different clonal populations within a tumor. A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3 and clone-4, displayed dynamic behavior, with clone-3 gradually diminishing and clone-4 and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant suggesting some degree of targeted treatment effect.

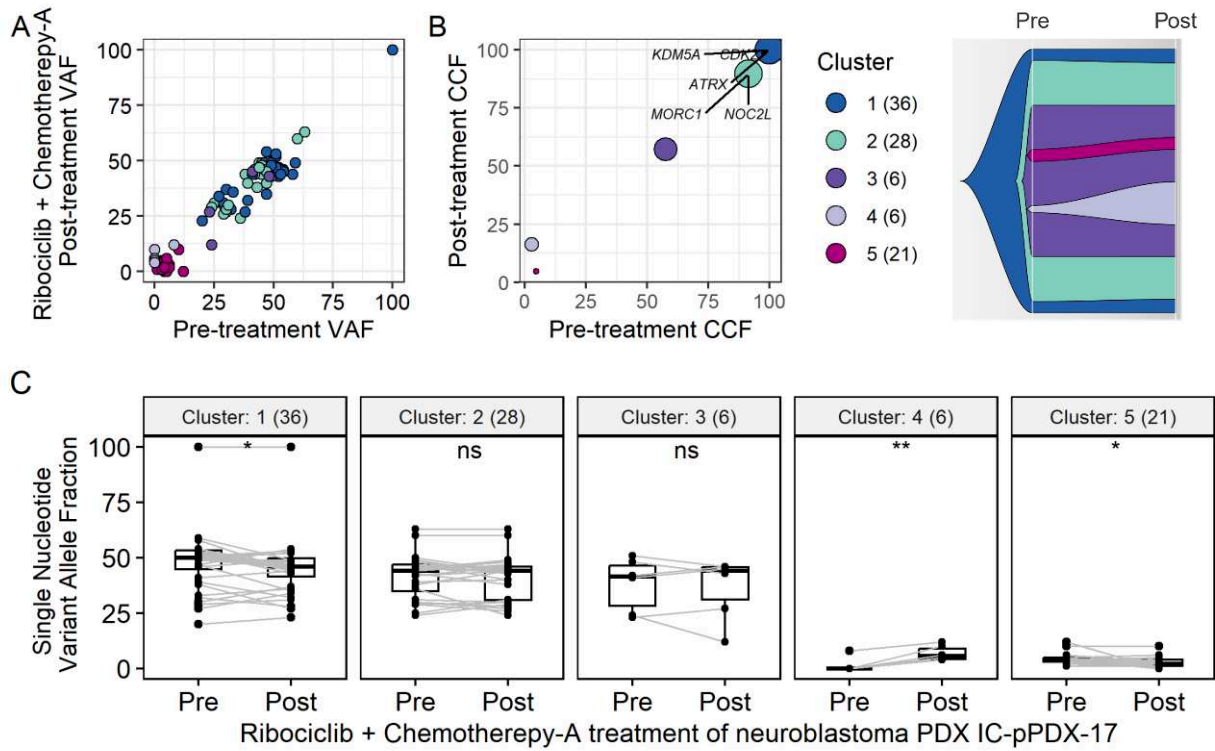


Figure 3.1.1.6.51. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) in combination with chemotherapy-A

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clones from 1 to 3 displayed static behavior. However, clone-4 emerged after the treatment. The fluctuation in the subclonal population was no significant except emerging clone, suggesting no treatment effect.

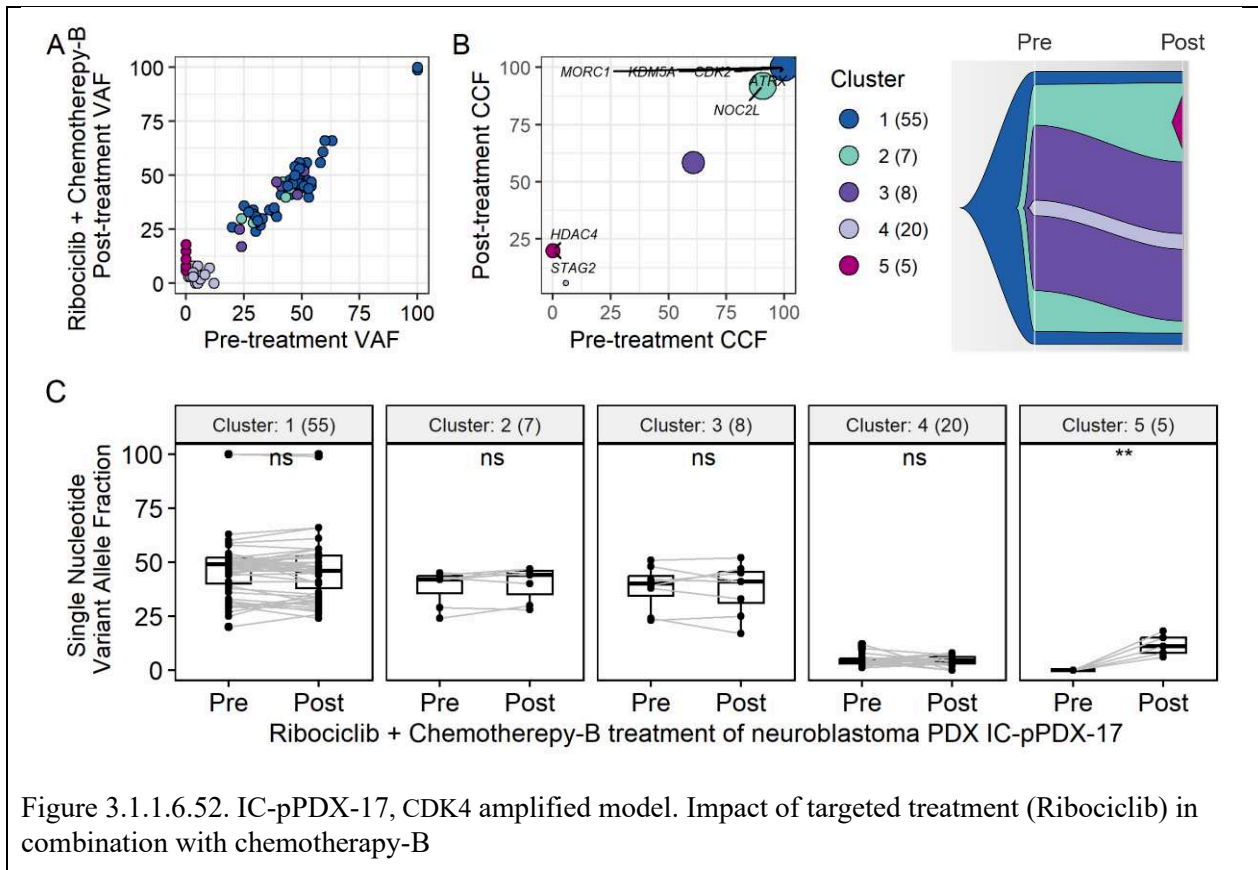
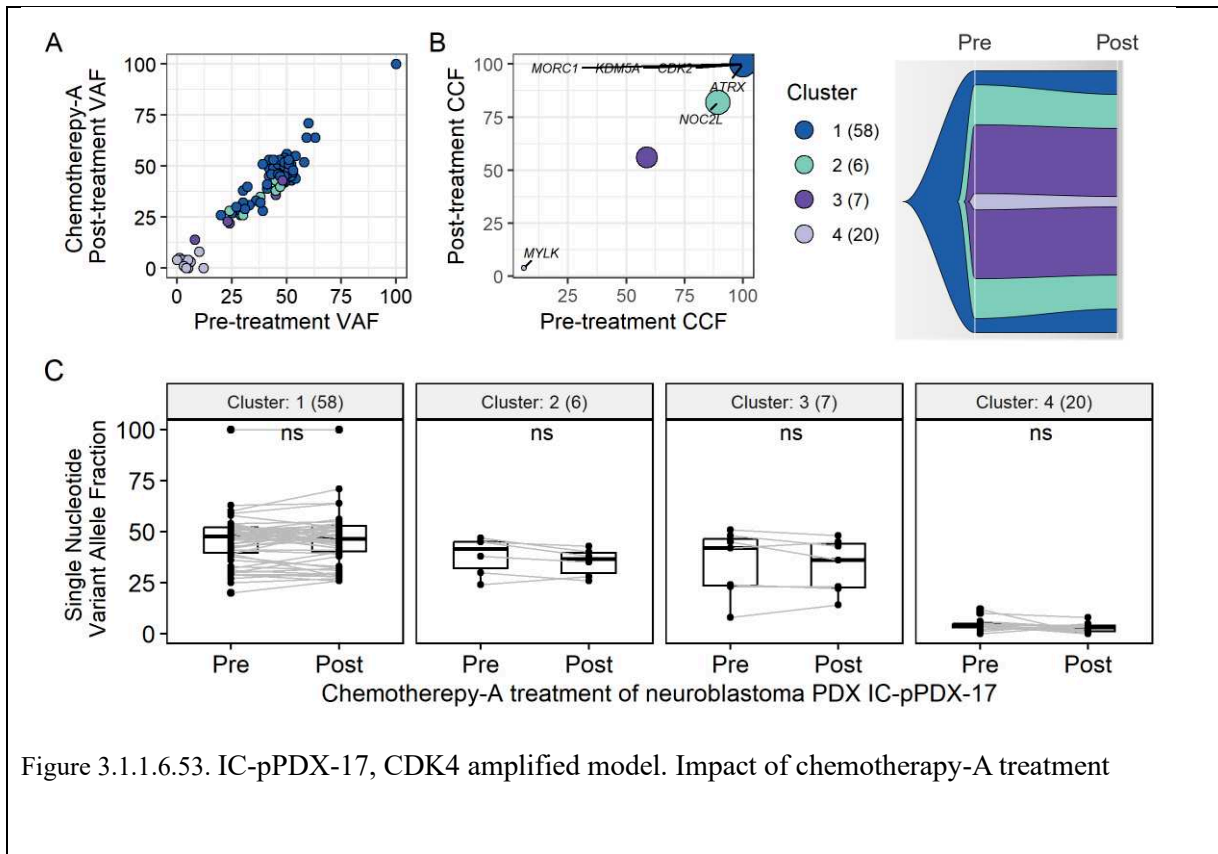
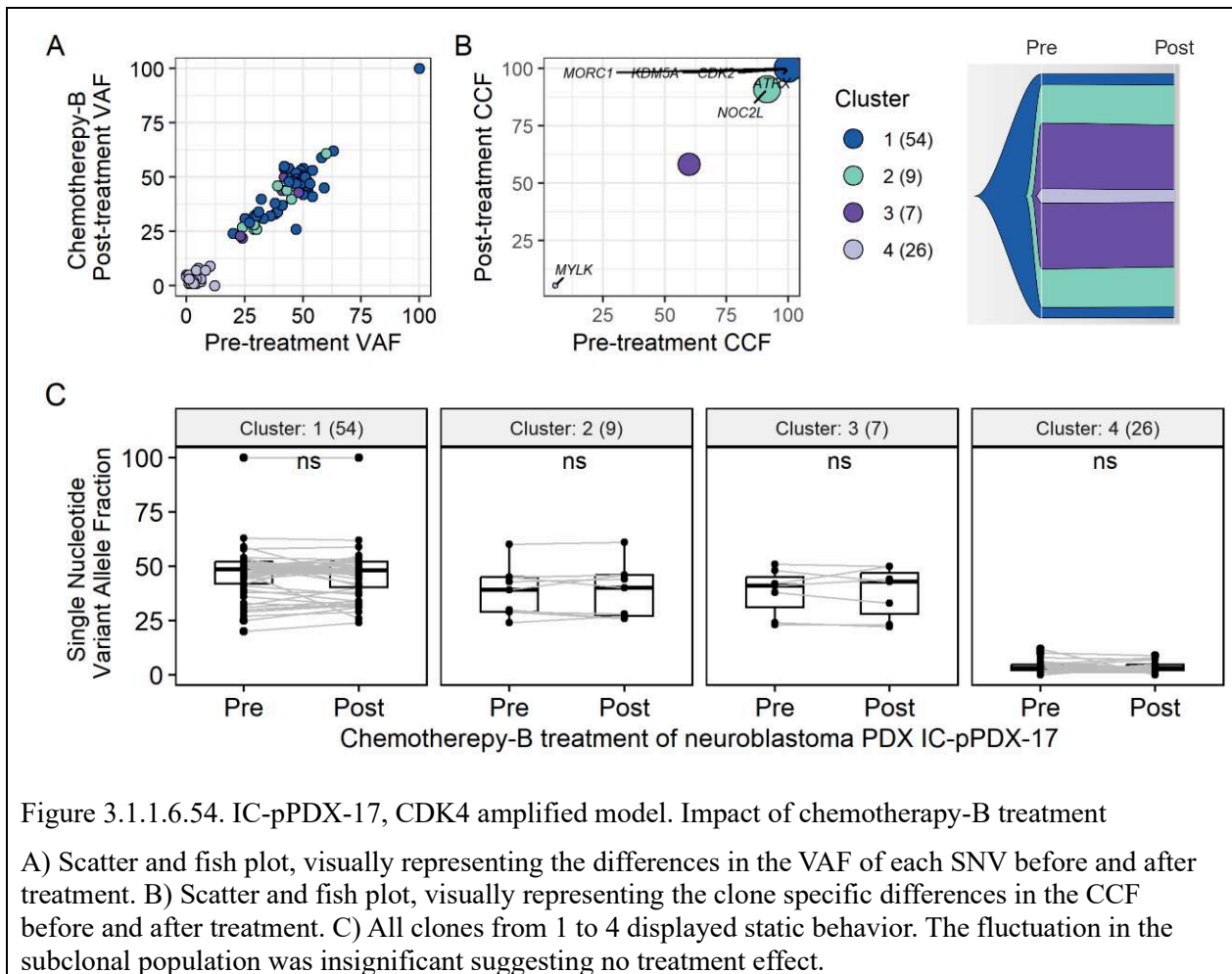


Figure 3.1.1.6.52. IC-pPDX-17, CDK4 amplified model. Impact of targeted treatment (Ribociclib) in combination with chemotherapy-B

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clones from 1 to 4 displayed static behavior. However, clone-5 emerged after the treatment with likely pathogenic mutations in *STAG2* and *HDAC4* genes. The fluctuation in the subclonal population was significant suggesting no treatment effect.



A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) All clones from 1 to 4 displayed static behavior. The fluctuation in the subclonal population was insignificant suggesting no treatment effect.



Summary:

The in vivo experiment and bulk exome sequencing data analysis focusing on IC-pPDX-17 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment:

- Ribociclib, when used alone, demonstrated notable effectiveness, achieving a tumor growth inhibition (TGI) of 61%.
- No change in CDK4 amplification before and after any treatment regimen
- Emerging subclones were evident after targeted treatment alone or in combination with chemotherapy. Mainly involved in post-translation protein modification activity. E.g. *HDAC4*, *C3*, *STAG2* and *TRAPPC9*. The other two emerging genes *XRCC* and *NHEJ1* were involved in DNA repair pathway activity.
- Likely pathogenic genes e.g. *ATRX*, *ETV1*, *GLI3*, *KDM5A* and *CDK2* with clonal mutation did not respond to treatment.

- The addition of doxorubicin and cyclophosphamide to ribociclib did not yield a significant improvement over the use of ribociclib alone or the combined chemotherapy.
- The combination of cisplatin and etoposide with ribociclib did not markedly enhance the overall treatment efficacy compared to using either treatment independently. However, this combination did show a significant increase in TGI at the 20-day mark.

In summary, ribociclib proved to be substantially effective for the IC-pPDX-17 model, and its combination with cisplatin/etoposide showed a notable improvement in chemotherapy efficiency.

Relative tumor volume (mean)

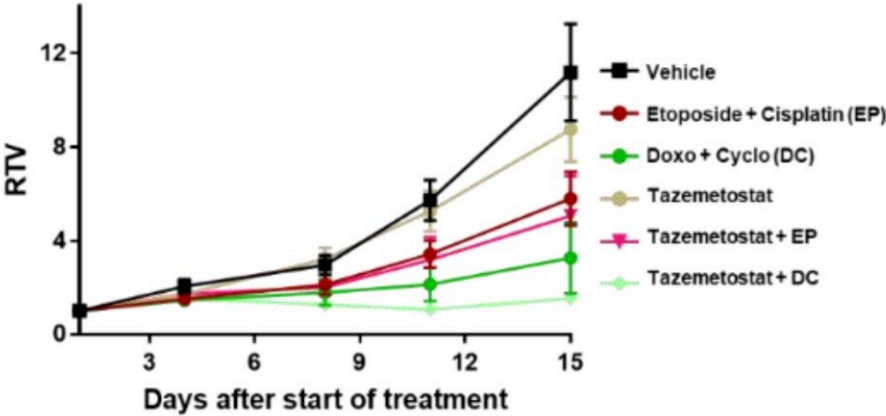


Figure 3.1.1.6.55. HSJD-NB-05, a ARID1A-mutated model, growth curves.

Growth curves were plotted to demonstrate the effects of two different chemotherapy regimens, termed Doxorubicin-Cyclophosphamide (chemotherapy-A) and Etoposide-Cisplatin (chemotherapy-B), both independent and in combination with tazemetostat. These curves highlighted the individual activity of each chemotherapy, as well as the impact of tazemetostat, a *ARID1A* inhibitor, when used alone.

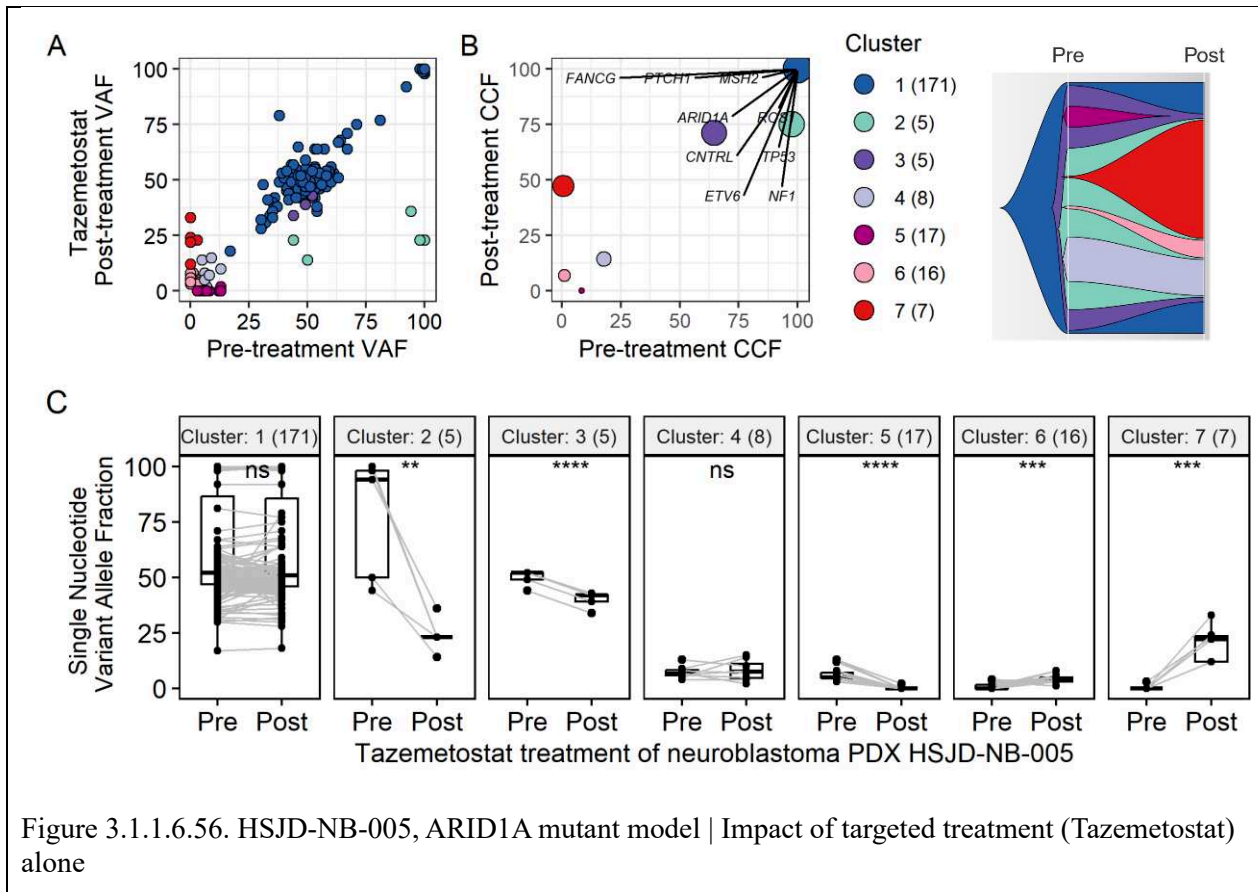
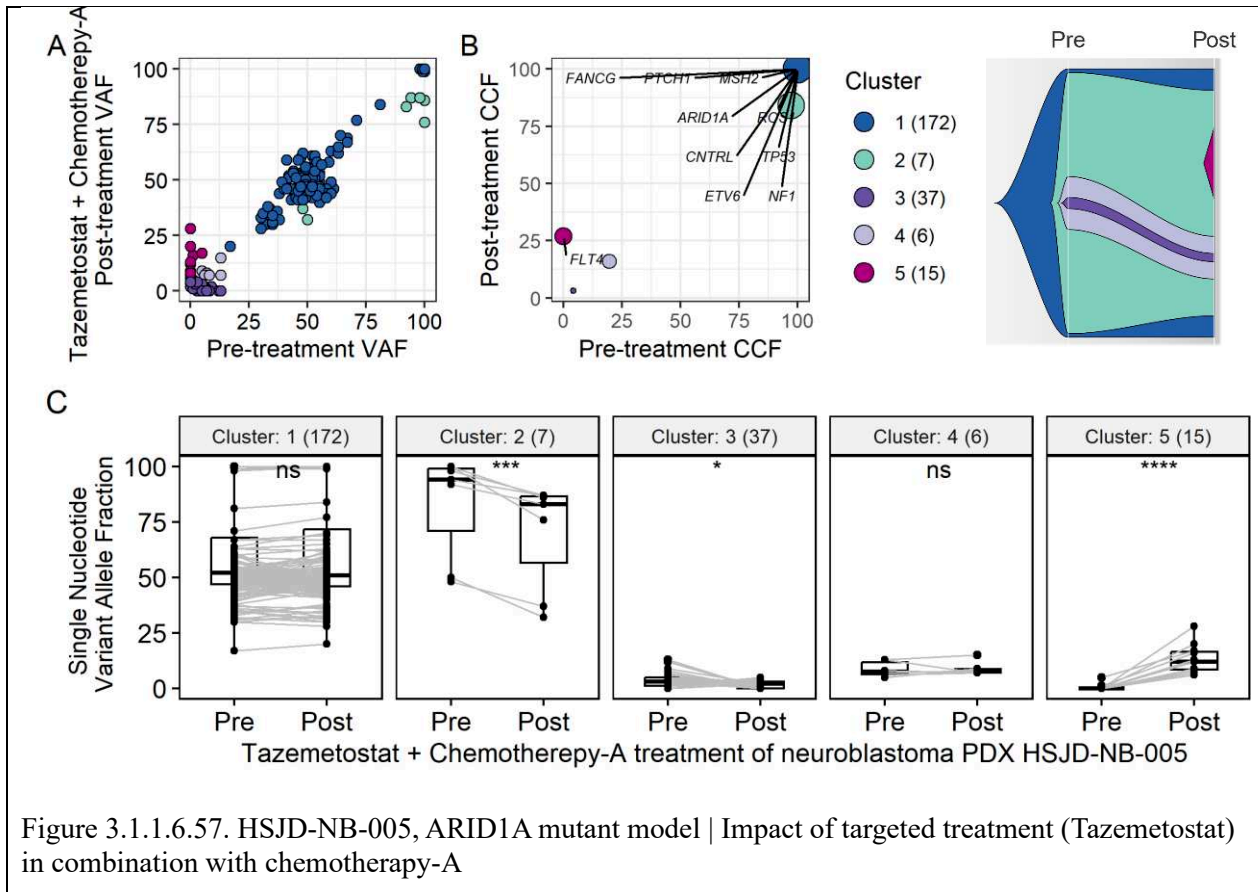


Figure 3.1.1.6.56. HSJD-NB-005, ARID1A mutant model | Impact of targeted treatment (Tazemetostat) alone

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-5 and clone-7, displayed dynamic behavior, with clone-5 gradually diminishing and clone-7 and clone-6 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).



A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4 and clone-5, displayed dynamic behavior, with clone-4 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

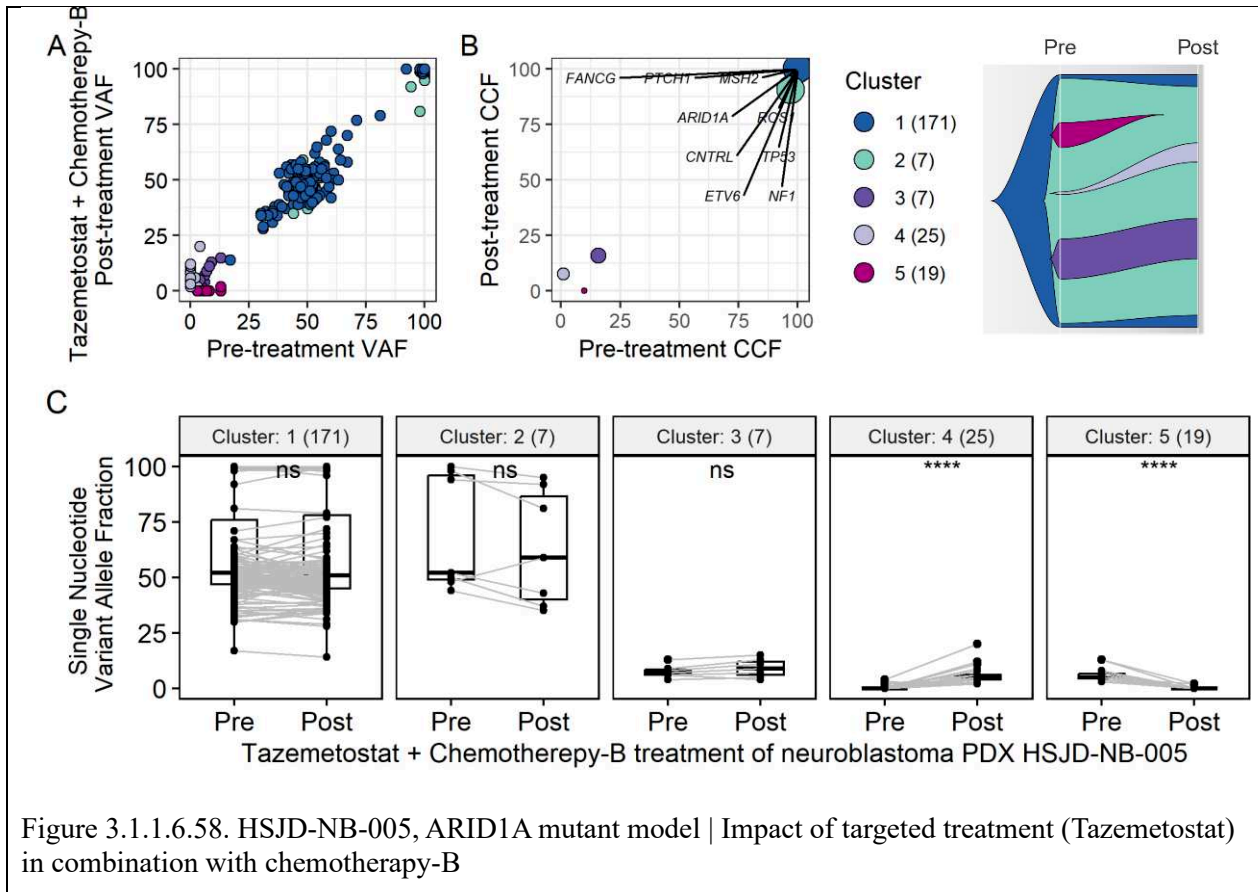


Figure 3.1.1.6.58. HSJD-NB-005, ARID1A mutant model | Impact of targeted treatment (Tazemetostat) in combination with chemotherapy-B

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4 and clone-5, displayed dynamic behavior, with clone-4 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***); 0.0001 (****).

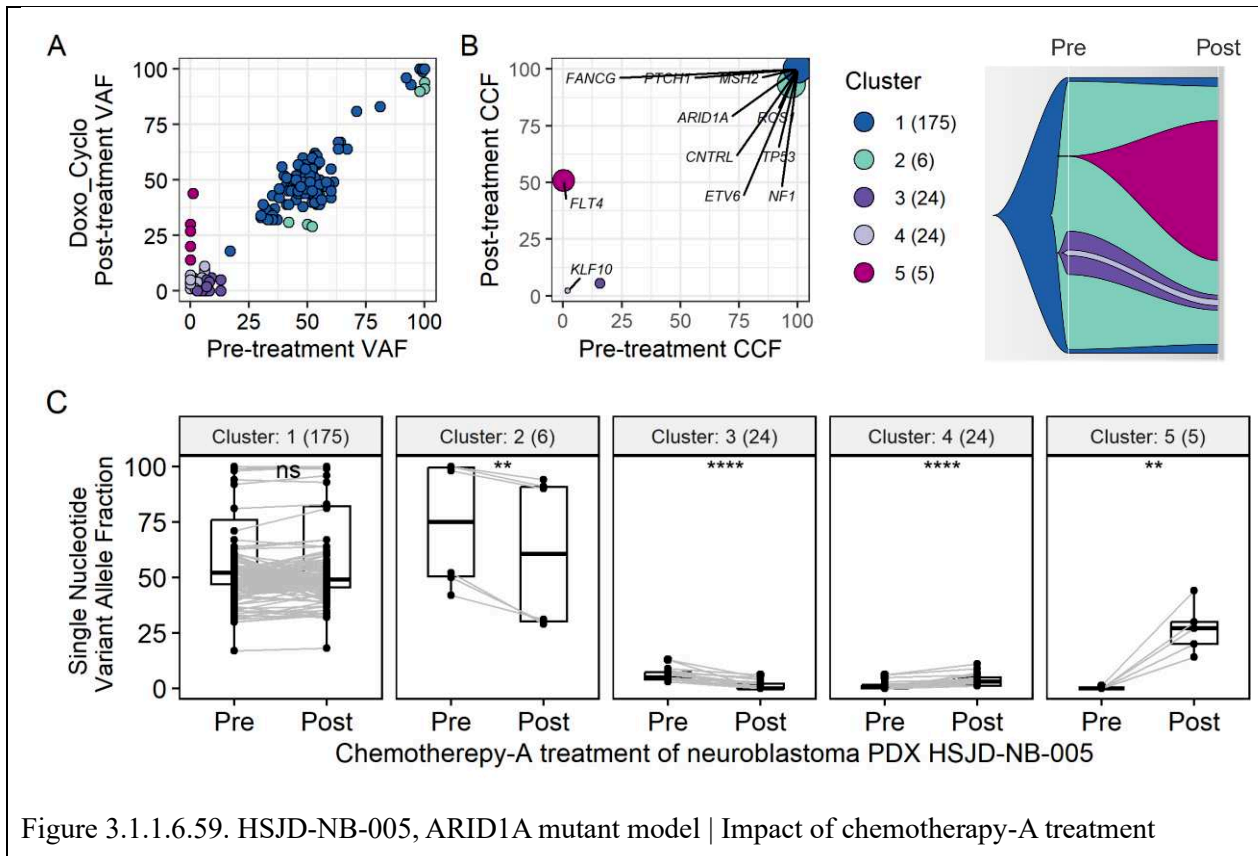


Figure 3.1.1.6.59. HSJD-NB-005, ARID1A mutant model | Impact of chemotherapy-A treatment

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-4 and clone-5, displayed dynamic behavior, with clone-4 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

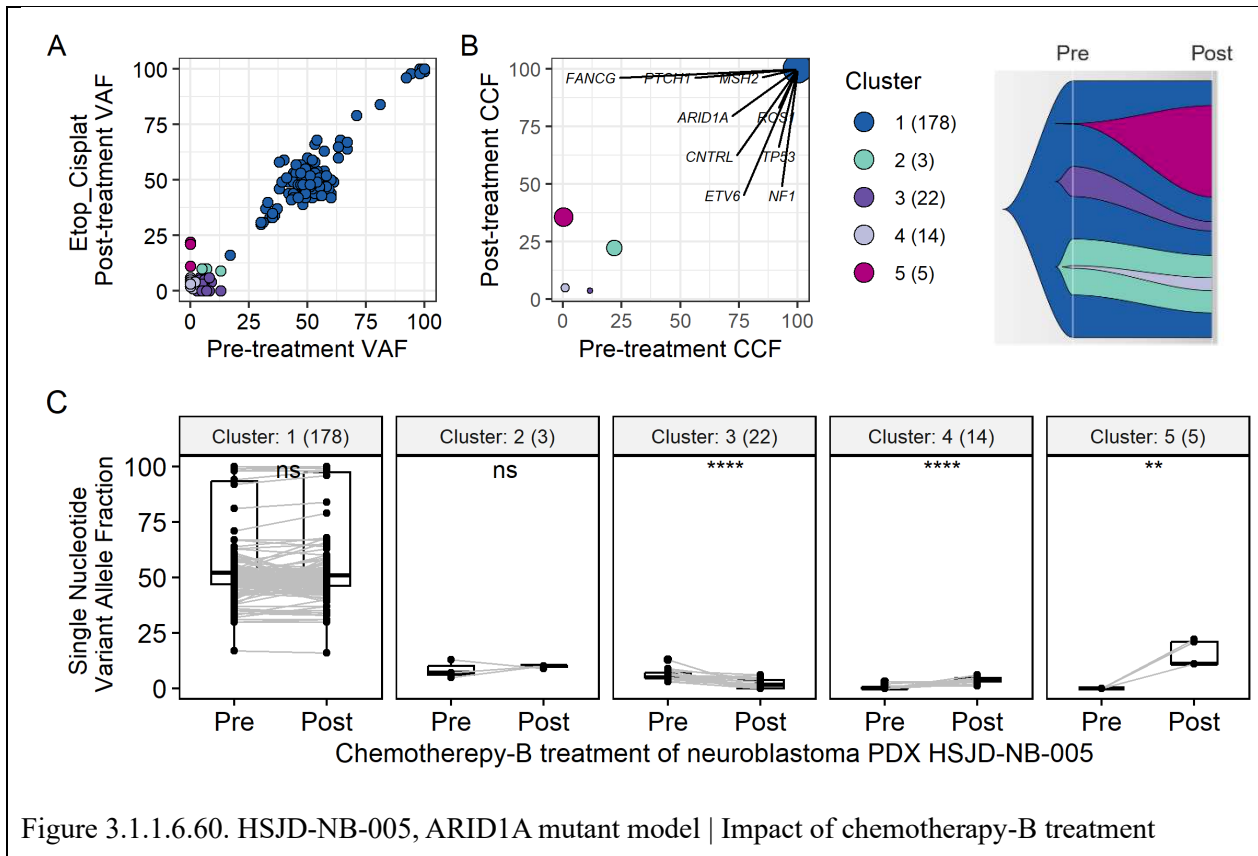


Figure 3.1.1.6.60. HSJD-NB-005, ARID1A mutant model | Impact of chemotherapy-B treatment

A) Scatter and fish plot, visually representing the differences in the VAF of each SNV before and after treatment. B) Scatter and fish plot, visually representing the clone specific differences in the CCF before and after treatment. C) Clone-1 and clone-2 exhibited clonal persistence throughout the treatment course. Subclones, namely clone-3, clone-4 and clone-5, displayed dynamic behavior, with clone-3 gradually diminishing and clone-5 emerging over the course of treatment. This fluctuation in the subclonal population was significant. Statistically significant differences determined by paired, two-tailed t-test. $P \leq 0.05$ (*); 0.01 (**); 0.001 (***) ; 0.0001 (****).

Summary

The in vivo experiment and bulk exome sequencing data analysis focusing on HSJD-NB-005 PDX model provided key insights into treatment efficacy and the changes in subclonal dynamics before and after treatment:

- Targeted agent - tazemetostat, when used alone, demonstrated minimal efficacy against the neuroblastoma PDX, with a tumor growth inhibition (TGI) of only 22%.
- Molecular target, ARID1A mutation and other likely pathogenic genes such as TP53, NF1 and MSH2 remain clonal (CCF ~ 100%) irrespective of any treatment regimen.
- Emerging clones had significant overlap with Epithelial to mesenchymal transition in colorectal cancer pathways. Genes such as *TGFB2*, *NOTCH3*, *WNT3*, *PKD1*, *ZEB2* and *COL4A5* were found significantly overlapped with this pathway.

- Combining tazemetostat with cisplatin and etoposide (Chemotherapy-B) did not enhance the efficacy of either drug regimen on their own. However, this combination did result in a higher number of responders compared to either tazemetostat alone or the cisplatin and etoposide combination alone.
- A significant increase in antitumor efficacy was observed when tazemetostat was combined with doxorubicin and cyclophosphamide (Chemotherapy-A). This suggests that the combination with tazemetostat improves the effectiveness of this chemotherapy regimen. However, there was no significant increase in the antitumor effect when doxorubicin and cyclophosphamide were used alone.

The results clearly demonstrated that tazemetostat alone did not exhibit effectiveness against the NB005 neuroblastoma patient-derived xenograft (PDX) model. Furthermore, combining tazemetostat with either doxorubicin and cyclophosphamide or etoposide and cisplatin did not significantly enhance the anti-tumor efficacy of these chemotherapy regimens. This suggests that tazemetostat, in this specific neuroblastoma model, may not be beneficial as either a standalone treatment or in conjunction with the tested chemotherapy combinations.

3.1.1.6.1 In vivo experiments for the study of clonal evolution under treatment

3.1.1.6.1.1 Ethical considerations of in vivo experimentation

The use of the mouse model remains essential for the outcome of this project: it represents an integrated biological model with a physiological whole context system, in contrast to invitro models. For PDX establishment, tumor tissue was obtained after informed consent provided by the patients' legal guardians. All animal experiments complied with current European/French legislation (articles R.214-87 to R.214-126 of the Decree n°2013-118 of February 1st) and were carried out in accredited animal facilities of the Institut Curie.

The animal's (mice) well-being was be monitored through strict scale endpoints. Following international recommendations, experiments were stopped before any animal suffering or when the tumor volume reached 1500 to 2000 mm³. The number of mice used considered the reduction principle. To reduce its number, whenever possible experiments were pooled to limit the number of control groups. In addition, data obtained from control groups could be used for future projects and leftover tumor material was stored for further projects.

All procedures were approved by the Institutional Review Board (IRB) of the Institut Curie. This project was approved by the Ethics Committee of the Institut Curie and the Ministry of Higher Education, Research and Innovation, with reference number AFAPIS #13980.2018030813227748 v2.

3.1.1.6.1.2 Treatment schedules

The engrafted mice in this study were randomized into six experimental arms, each receiving a distinct treatment regimen over two cycles. The treatments consisted of standard chemotherapy, targeted therapy, or a combination of both. The targeted therapies were chosen based on genetic abnormalities found in the primary tumors and patient-derived xenografts (PDXs), identified through molecular analysis that included whole exome sequencing (WES) at 100X coverage, excluding germline variations, and RNA sequencing. The selection of targetable alterations and corresponding therapeutic strategies was guided by the MAPPYACTS, NCI-MATCH, and INFORM algorithms (Worst et al. 2016). Detailed information on the dosages used is provided below:

1. Chemotherapy-A: Doxorubicin-Cyclophosphamide
 - a. Doxorubicin 2 mg/kg IP at D1 and D22
 - b. Cyclophosphamide 100 mg/kg IP at D1 and D22 3.

2. Chemotherapy-B: Etoposide-Cisplatin
 - a. Etoposide 12 mg/kg IP at D1, D2, D3 (cycle 1) and D22, D23, D24 (cycle 2)
 - b. Cisplatin 6 mg/kg IP at D1 and D22 2.
3. Targeted agents were administered orally 5/7 days from D1 to D42. Depending on the PDX model, mice received:
 - a. IC-pPDX-17: Ribociclib 75 mg/kg
 - b. IC-pPDX-75: Lorlatinib 10 mg/kg
 - c. IC-pPDX-109: Trametinib 0.4 mg/kg from D1 to D5 and 0.2 mg/kg from D8 to D42
 - d. GR-NB4: Lorlatinib 10 mg/kg
 - e. GR-NB10: Trametinib 0.4 mg/kg from D1 to D5 and 0.2 mg/kg from D8 to D42
 - f. HSJD-NB-005 (ARID1A mutation): Tazemetostat 100 mg/kg
4. Chemotherapy-A + Targeted agent
5. Chemotherapy-B + Targeted agent
6. Control group

3.1.1.6.1.3 In vivo treatment efficacy

The efficacy of the different treatments tested was studied in the different PDX models used for the study of clonal evolution.

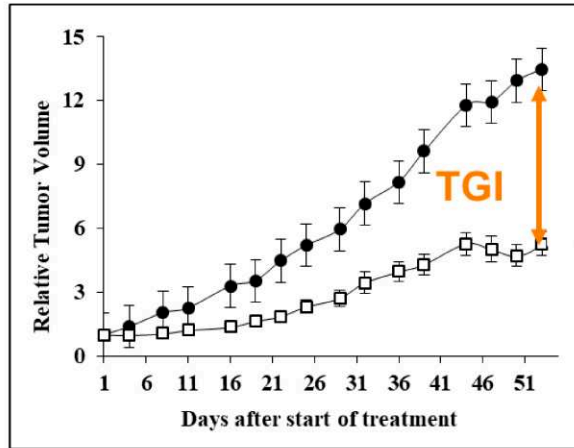
Determination of the treatment efficacy Some concepts are described here for the analysis of the treatment efficacy, such as relative tumor volume, tumor growth inhibition, overall response rate. Statistical significance of differences observed between the individual RTVs corresponding to the treated mice and control groups was calculated by the two-tailed Student's t test (Némati et al. 2010).

3.1.1.6.1.4 Tumor Growth Inhibition (TGI) determination

- Tumor volume (V) was measured as $V = a \times b \times b/2$, where a and b are the largest and smallest perpendicular tumor diameters, respectively.
- Relative tumor volume (RTV) is calculated as $RTV = (V_x/V_1)$, where V_x is the tumor volume on day x and V_1 is the tumor volume at initiation of therapy (day 1).
- Growth curves were obtained by plotting the mean values of RTV on the Y-axis against time.

$$TGI (\%) = 1 - (RTV_t/RTV_c)$$

where RTV_t is the relative tumor volume of the treated mouse and RTV_c the median of RTV of the corresponding control group at a time corresponding to the end of treatment. Fifty percent TGI was the limit for a meaningful biological effect.



$$\Rightarrow \text{TGI (\%)} = 100 - (5.3/13.5) = 61\%$$

3.1.1.6.1.5 Overall response rate (ORR) determination

- The objective is to evaluate the response to treatment according to individual mouse variability.

$$\text{ORR} = [(\text{RTVV}) - 1]$$

- The relative tumor volume variation (RTVV) of each treated mouse is calculated as

$$\text{RTVV} = [(\text{RTV}_t/\text{RTV}_c) - 1]$$

- A tumor was considered as responding to therapy when the ORR was lower than -0.5.

3.1.1.6.1.6 Analysis of treatment efficacy

In the study of clonal evolution, the selection of targeted therapies was tailored according to the unique molecular characteristics of each patient-derived xenograft (PDX) model. The effectiveness of these targeted therapies varied across different models. For instance, lorlatinib demonstrated a high tumor growth inhibition (TGI) rate of 91% in the *ALK*-amplified model GR-NB4, but it showed no significant efficacy (TGI of only 28%) in the *ALK* F1174L mutated IC-pPDX-75 model. In the case of trametinib, a considerable antitumor effect was observed in the *HRAS* mutated IC-

pPDX-109 model, marked by a TGI of 60%. However, this efficacy was not replicated in the GR-NB10 model, which harbored an NF1 mutation accompanied by loss of heterozygosity (LOH), where it only achieved a TGI of 49%. Additionally, ribociclib, a CDK4 inhibitor, was notably effective against the CDK4-amplified IC-pPDX-17 model, achieving a TGI of 61%.

In the study of high-risk neuroblastoma, the efficacy of chemotherapies varied across different patient-derived xenograft (PDX) models. Etoposide-cisplatin demonstrated high efficiency in models GR-NB4 and IC-pPDX-109, with tumor growth inhibition (TGI) rates of 93.7% and 67%, respectively. Similarly, doxorubicin-cyclophosphamide was effective against models IC-pPDX-17 and GR-NB4, achieving TGIs of 70% and 76%, respectively.

Moreover, the combination of targeted therapy with chemotherapy resulted in a significantly greater efficiency compared to chemotherapy alone in certain models. For example, in the GR-NB4 model, the addition of lorlatinib to both etoposide-cisplatin and doxorubicin-cyclophosphamide chemotherapies enhanced their effectiveness. In the GR-NB10 model, the combination of trametinib with doxorubicin-cyclophosphamide proved more effective. Similarly, in the IC-pPDX-109 model, adding lorlatinib to doxorubicin-cyclophosphamide showed increased efficacy. However, in the IC-pPDX-17 and IC-pPDX-75 models, the addition of targeted therapies (ribociclib and lorlatinib, respectively) did not significantly improve the efficacy of either of the two chemotherapy regimens.

Table 3.1.1.6.4. Tumor Growth Inhibition Efficacy Across Various PDX Models Under Distinct Treatment Regimens.

HR-NB PDX Models	Days of analysis	Targeted therapy	Doxorubicin-Cyclophosphamide (Chemotherapy-A)	Etoposide-Cisplatin (Chemotherapy-B)	Targeted therapy + Chemotherapy-A	Targeted therapy + Chemotherapy-B
IC-pPDX-17	16	Ribociclib	61	43	78	73
IC-pPDX-75	25	Lorlatinib	28	35	62	62
IC-pPDX-109	37	Trametinib	61	47	63	78
GR-NB4	22	Lorlatinib	91	76	93.7	99.9
HSJD-NB-005	15	Tazemetostat	22	71	86	55
GR-NB10	15	Trametinib	49	58	79	75

TGI activity values in percentage (%)

3.2 Part B: Further analysis of genetic heterogeneity in neuroblastoma

3.2.1 Paper I: Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma

3.2.1.1 Introduction

It is estimated that mutations in SWI/SNF genes and other epigenetic modifier genes (EMGs) are present in about 20% of human cancers. In neuroblastoma (NB), the most common extracranial solid tumor in children, genetic alterations are primarily copy number changes, including *MYCN* amplification and segmental chromosome alterations linked to poor outcomes. Although recurrent genetic alterations in a few genes like *ALK* have been identified in NB, the extent of CRG/EMG alterations in this cancer type is not well-defined. The study's goal is to ascertain the frequency of CRGs/EMGs variations in NB patients and investigate their association with clinical outcomes.

3.2.1.2 Materials and Methods

To assess their frequency and clinical impact, paired diagnostic tumor and normal DNA from 55 patients were whole-genome (n = 16) and/or whole-exome (n = 39) sequenced, and ultra deep targeted sequencing was performed (TSCA®) on 283 NB patients. We compared the frequency of CRG/EMG variations in NB cases to the population frequency Genome Aggregation Database (gnomAD).

3.2.1.3 Results

Our analysis revealed single nucleotide variations (SNVs), small insertions/deletions (InDels), or focal copy number alterations (CNAs) in CRGs/EMGs in 20% (56/283) of all cases. These variations occurred at a somatic level in 4 (7.2%) cases, at a germline level in 12 (22%) cases, and in remaining cases, only tumor material was available for analysis. The most frequently altered genes were *ATRX* (5%), *SMARCA4* (2.5%), *MLL3* (2.5%), and *ARID1B* (2.5%). Notably, double events, such as SNVs/small InDels/CNAs associated with loss of heterozygosity (LOH), were observed in *SMARCA4* (n = 3), *ATRX* (n = 1), and *PBRM1* (n = 1).

Variations in *SMARCA4* and *ATRX* were more frequent in NB cases compared to the general population in the gnomAD cohort (OR = 4.49, 95%CI: 1.63–9.97, p = 0.038; OR3.44, 95%CI: 1.46–6.91, p = 0.043, respectively). Our case-control study approach suggests a significant role for *SMARCA4* in NB oncogenesis. These findings underscore the crucial role of chromatin

remodeling in neuroblastoma and open new avenues for therapeutic strategies targeting these genetic pathways.

In the study of NB cell lines (n=30), nine variations were detected across seven cell lines, again with SMARCA4, MLL2, and MLL3 being the most frequently targeted genes. Notably, in the CLB-Re cell line, a partial loss of the MLL2 gene was detected, confirmed by SNP6 analysis showing a large region of copy number loss on chromosome 12 encompassing the MLL2 gene. In the SKNFI cell line, a double event in SMARCA4 (SNV and LOH) was observed, leading to the absence of SMARCA4 expression. Conversely, SMARCA4 expression was maintained in the SKNSH cell line, which only had a single SNV event in this gene.

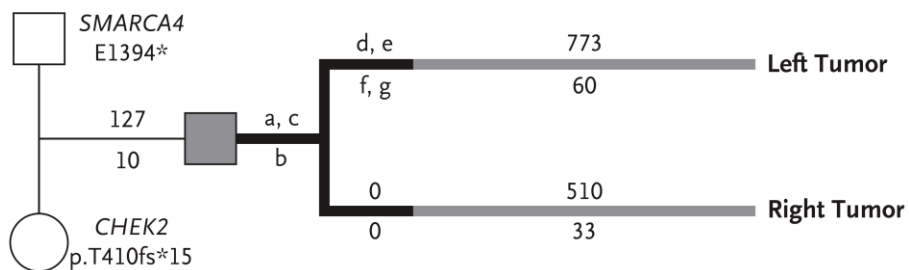
3.2.1.4 Conclusion

The study concludes that CRG/EMG variations with likely functional impact are present in 8.4% of NB patients. These variations correlate with poorer overall survival, suggesting more aggressive disease or differential responses to salvage treatments. Future research should focus on correlations between CRG/EMG variations and other genetic factors linked to poor NB outcomes. Finally, the findings support exploring treatment strategies involving chromatin remodeling processes, such as histone deacetylase inhibitors (HDACi) or DNA methyltransferase inhibitors, especially for patients with high-risk NB and CRG/EMG variations.

Recently Coorens et al. showed in one NB patient, inherited two pathogenic variants: a truncating missense variant in SMARCA4 from the father and a truncating frameshift CHEK2 variant from the mother and the tumors had parallel evolution, including distinct second hits (19p loss) in SMARCA4, a putative predisposition gene for neuroblastoma (Coorens et al. 2020).

A

Phylogenetic Origin of Tumors



B

Timing of Copy-Number Gains from Somatic Mutations

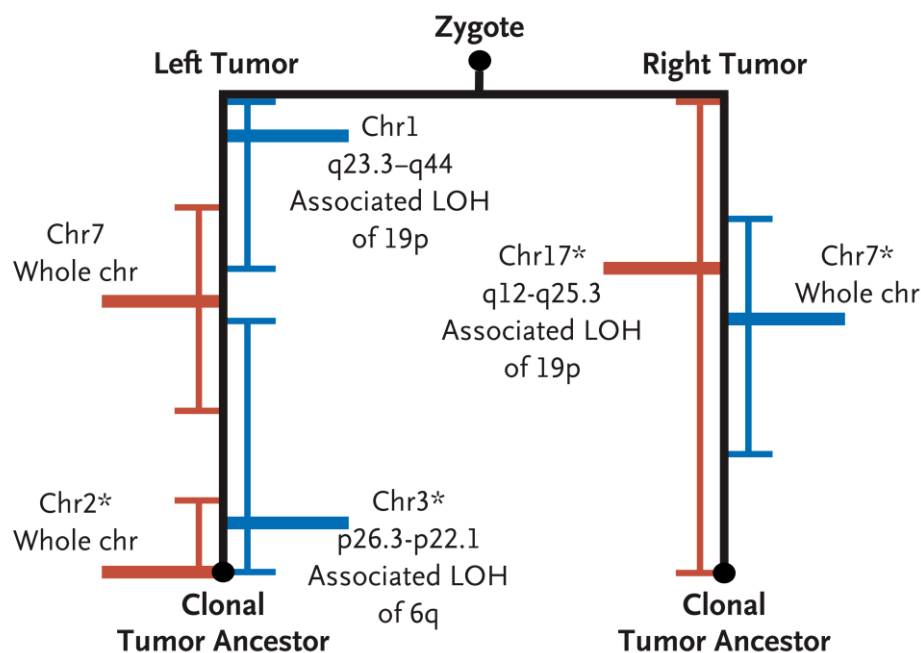


Figure 3.2.1.4.1: Independent Origin and Convergent Parallel Evolution of Tumors in Patient 1 with SMARCA4 mutation.

(A) shows the phylogenetic origins of the right and left tumors in Patient 1. The patient inherited two pathogenic variants: a truncating missense variant in *SMARCA4* from the father and a truncating frameshift *CHEK2* variant from the mother. Branches in black depict variants (labeled a through g) in the tumors that were shared with blood. The numbers of substitutions are annotated above the branch, and insertions or deletions (indels) are annotated below the branch. (B) shows the timing of the acquisition of copy-number gains from somatic mutations before and after duplication. Gains of maternal chromosomes (chr) are indicated in red and gains of paternal chromosomes in blue. Within each branch, asterisks highlight pairs of rearrangements that are unlikely to have occurred at different time points ($P > 0.05$ by the Poisson test). LOH denotes loss of heterozygosity.

Another comprehensive computational meta-genomic analysis of solid tumors from 131,668 cancer patients revealed the presence of SMARCA4 gene alterations in 9434 cases. This study highlighted a notable prevalence of homozygous SMARCA4 mutations, particularly in non-small cell lung cancer (NSCLC), where they were associated with decreased survival rates. The extensive dataset enabled the identification of previously uncharted hotspot missense mutations within the SMARCA4 helicase domain. These mutations were found to significantly impair chromatin remodeling activity. Intriguingly, some of these SMARCA4 missense variants demonstrated the ability to partially or completely compensate for the dependency on its paralog, SMARCA2. This finding emphasizes the necessity for meticulous patient selection based on the nature of SMARCA4 mutations to determine those who might benefit from therapies targeting SMARCA2 (Fernando et al. 2020) (Jiménez et al. 2022).

These studies contribute significantly to our understanding of the role of SMARCA4 mutations in cancer and their potential implications for targeted therapy. This comprehensive analysis underscores the significance of CRG/EMG alterations in neuroblastoma and their potential impact on disease characteristics and progression.

Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma

Angela Bellini^{1,2,3*}, Nadia Bessoltane-Bentahar^{1,2,3*}, Jaydutt Bhalshankar^{1,2,3}, Nathalie Clement^{1,2,3}, Virginie Raynal^{4,5}, Sylvain Baulande⁶, Virginie Bernard⁷, Adrien Danzon^{1,2,3}, Mathieu Chicard^{1,2,3}, Léo Colmet-Daage^{1,2,3}, Gaëlle Pierron⁸, Laura Le Roux⁸, Julien M. Planchon^{3,7}, Valérie Combaret⁹, Eve Lapouble^{3,8}, Nadège Corradini¹⁰, Estelle Thebaud¹¹, Marion Gambart¹², Dominique Valteau-Couanet¹³, Jean Michon³, Caroline Louis-Brennetot^{2,3,4}, Isabelle Janoueix-Lerosey^{1b2,3,4}, Anne-Sophie Defachelles¹⁴, Franck Bourdeaut^{1,2,3}, Olivier Delattre^{2,3,4,8} and Gudrun Schleiermacher^{1b1,2,3}

¹Equipe SiRIC RTOP Recherche Translationnelle en Oncologie Pédiatrique, Institut Curie, Paris, France

²INSERM U830, Laboratoire de Génétique et Biologie des Cancers, Institut Curie, Paris, France

³SIREDO: Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer, Institut Curie, Paris, France

⁴Institut Curie, PSL Research University, Inserm U830, Equipe Labellisée Ligue contre le Cancer, Paris, France

⁵Plateforme de Séquençage ICGex, Institut Curie, Paris, France

⁶Institut Curie, PSL Research University, NGS Platform, Paris, France

⁷Department of Biopathology, Institut Curie, PSL Research University, Paris, France

⁸Unité de Génétique Somatique, Institut Curie, Paris, France

⁹Laboratoire de Recherche Translationnelle, Centre Léon-Bérard, Lyon, France

¹⁰Institut d'Hématologie et Oncologie Pédiatrique, Lyon, France

¹¹Service d'Oncologie Pédiatrique, Hôpital de la Mère et l'enfant, Nantes, France

¹²Unité d'Hémo-Oncologie, Hôpital des Enfants, Toulouse, France

¹³Service d'Oncologie Pédiatrique, Institut Gustave Roussy, Paris, France

¹⁴Pediatric Oncology Unit, AntiCancer Center Oscar Lambret, Lille, France

In neuroblastoma (NB), genetic alterations in chromatin remodeling (CRGs) and epigenetic modifier genes (EMGs) have been described. We sought to determine their frequency and clinical impact. Whole exome (WES)/whole genome sequencing (WGS) data and targeted sequencing (TSCA[®]) of exonic regions of 33 CRGs/EMGs were analyzed in tumor samples from 283 NB patients, with constitutional material available for 55 patients. The frequency of CRG/EMG variations in NB cases was then compared to the Genome Aggregation Database (gnomAD). The sequencing revealed SNVs/small InDels or focal CNAs of CRGs/EMGs in 20% (56/283) of all cases, occurring at a somatic level in 4 (7.2%), at a germline level in 12 (22%) cases, whereas for the remaining cases, only tumor material could be analyzed. The most frequently altered genes were *ATRX* (5%),

Key words: neuroblastoma, chromatin remodeling complex, SWI/SNF, epigenetic modifier, SMARCA4

Abbreviations: aCGH: array-comparative genomic hybridization; CNAs: copy number alterations; CRGs: chromatin remodeling genes; EMGs: epigenetic modifiers genes; OS: overall survival; PFS: progression-free survival; SNVs: single nucleotide variants; TSCA: TruSeq[®] Custom Amplicon; WES: whole exome sequencing; WGS: whole genome sequencing

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare no conflict of interest.

Grant sponsor: Agence Nationale pour la Recherche; **Grant number:** ANR-10-INBS-518 09; **Grant sponsor:** Annenberg Foundation;

Grant sponsor: Equipex; **Grant number:** ANR-10-EQPX-03; **Grant sponsor:** France Génomique Consortium; **Grant number:** ANR-

10-INBS-09-08; **Grant sponsor:** Association Hubert Gouin Enfance et Cancer; **Grant sponsor:** Association Les amis de Claire;

Grant sponsor: Associations Enfants et Santé; **Grant sponsor:** Associations Les Bagouz à Manon; **Grant sponsor:** Canceropole Ile-de-France;

Grant sponsor: CEST of Institut Curie; **Grant sponsor:** Fondation Nelia et Amadeo Barletta; **Grant sponsor:** Programme Hospitalier de

recherche en Cancérologie (PHRC); **Grant number:** IC2007-09; **Grant sponsor:** SiRIC/INCa; **Grant number:** INCa-DGOS-4654;

Grant sponsor: SiRIC-Curie Program-SiRIC

*A.B. and N.B.-B. contributed equally to this work

DOI: 10.1002/ijc.32361

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

History: Received 16 Nov 2018; Accepted 14 Mar 2019; Online 24 Apr 2019

Correspondence to: Gudrun Schleiermacher, Department of Pediatric Oncology, Institut Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France, Tel.: +33-1-44-32-45-50, E-mail: gudrun.schleiermacher@curie.fr

SMARCA4 (2.5%), *MLL3* (2.5%) and *ARID1B* (2.5%). Double events (SNVs/small InDels/CNAs associated with LOH) were observed in *SMARCA4* ($n = 3$), *ATRX* ($n = 1$) and *PBRM1* ($n = 1$). Among the 60 variations, 24 (8.4%) targeted domains of functional importance for chromatin remodeling or highly conserved domains but of unknown function. Variations in *SMARCA4* and *ATRX* occurred more frequently in the NB as compared to the gnomAD control cohort (OR = 4.49, 95%CI: 1.63–9.97, $p = 0.038$; OR 3.44, 95%CI: 1.46–6.91, $p = 0.043$, respectively). Cases with CRG/EMG variations showed a poorer overall survival compared to cases without variations. Genetic variations of CRGs/EMGs with likely functional impact were observed in 8.4% (24/283) of NB. Our case–control approach suggests a role of *SMARCA4* as a player of NB oncogenesis.

What's new?

Mutations that affect chromatin remodeling can lead to cancer. In this paper, the authors investigated the impact of variations in chromatin remodeling genes and epigenetic modifier genes on neuroblastoma patients. They compared the frequency of these variations in NB cases with data from the Genome Aggregation Database (gnomAD). Neuroblastoma cases had a higher frequency of *SMARCA4* and *ATRX* gene variations than the general population. Furthermore, NB patients with CRG/EMG mutations had poorer overall survival than NB cases without such mutations. These findings highlight the importance of chromatin remodeling in neuroblastoma as an avenue for new therapeutics.

Introduction

Alterations in chromatin remodeling genes (CRGs) have emerged as an important cause of cancer; in particular, the identification of inactivating genetic alterations in the *SMARCB1* gene in rhabdoid tumors (RTs) has highlighted the involvement of the SWI/SNF complex in tumor formation.^{1,2} Indeed, SWI/SNF gene mutations and alterations in other epigenetic modifier genes (EMGs) are thought to occur in 20% of human cancers.³

In neuroblastoma (NB), the most common extracranial solid tumor in children, genetic alterations concern predominantly copy number alterations, with *MYCN* amplification and segmental chromosome alterations associated with a poor outcome.⁴ Only few genes have been shown to be altered recurrently by genetic events, including *ALK*.⁴ Genes involved in chromatin remodeling such as *ATRX*, *DAXX*, *ARID1A* and *ARID1B* genes have also been shown to be targeted by mutations or structural rearrangements,^{5–9} but the frequency of CRG/EMG alterations has not been described in detail in NB.

The aim of our study was to determine the frequency of CRGs/EMGs variations in NB patients and to correlate findings with clinical parameters and outcome.

Materials and Methods

Study series

The study consisted of 283 NB patients, constituting a clinically representative cohort (Table 1, Supporting Information Table S1; Supporting Information Fig. S1).

Patients were treated in French pediatric oncology centers according to the relevant national/international protocols. Written informed consent was obtained from parents/guardians according to national law. Our study was authorized by the ethics committees “Comité de Protection des Personnes Sud-Est IV”, references L07–95/L12–171, and “Comité de Protection des Personnes Ile de France”, reference 0811728.

In 55 patients for whom paired normal and tumor tissue was available, whole-exome sequencing (WES)/whole-genome sequencing (WGS) techniques were performed, whereas for 248 diagnostic tumor samples, a TruSeq[®] Custom Amplicon (TSCA) approach was used. Twenty patients are common to the two subsets (Supporting Information Table S1).

Furthermore, 30 NB cell lines (Supporting Information Table S1) and six germline controls from healthy donors were studied.

To analyze the frequency of CRG/EMG variations, 28 major genes involved in chromatin remodeling processes were selected: *ACTB*, *ACTL6A*, *ACTL6B*, *ARID1A*, *ARID1B*, *ARID2*, *BCL11A*, *BCL11B*, *BCL7A*, *BCL7B*, *BCL7C*, *BRD7*, *BRD9*, *DPF1*, *DPF2*, *DPF3*, *PBRM1*, *PHF10*, *SMARCA2*, *SMARCA4*, *SMARCB1*, *SMARCC1*, *SMARCC2*, *SMARCD1*, *SMARCD2*, *SMARCD3*, *SMARCE1*, *SS18*. Five epigenetic modifier genes, *ATRX*, *DAXX*, *CHD7*, *MLL2* (*KMT2D*), *MLL3* (*KMT2C*), were also selected based to their involvement in a wide variety of human malignancies (Supporting Information Table S2).

DNA extraction and aCGH analysis

DNA was extracted from diagnostic tumor samples harboring >20% of tumor cells using standard procedures. Genomic copy number profiles were determined by aCGH, and *MYCN* copy number status was confirmed by FISH.

Whole exome/whole genome sequencing

Paired diagnostic tumor and normal DNA from 55 patients were whole-genome ($n = 16$) and/or whole-exome ($n = 39$) sequenced (Illumina[®] HiSeq2500; average coverage: 80×–100× per sample, respectively¹⁰). WGS for eight cases has been reported previously.¹⁰

TruSeq[®] amplicon panel sequencing

Tumor DNA from 248 NB cases and six healthy donor germline DNA samples (negative controls) were analyzed using a TSCA[®]

panel approach covering 33 CRGs/EMGs, (average coverage >1,500×; TSCA[®]v1.5, Illumina Inc., San Diego, CA; Supporting Information Table S2).

Bioinformatics analysis

Bioinformatics approaches depended on the genetic alteration (single nucleotide variations [SNVs]), small or large InDels; focal or extended copy number alterations (CNAs) and the sequencing technique.

Whole-exome and whole-genome sequencing analyses

The WGS/WES sequencing raw reads were mapped to the reference human genome using BWA (assembly GRCh37/hg19)^{11,12} followed by analysis of the exons of the 33 CRGs/EMGs (Supporting Information Table S2).

SNVs/InDels <30 bp were called using GATK's Haplotype-Caller v3.5 algorithm.^{12,13} The Manta tool (1.0.3) was used to detect structural variations and large InDels.¹⁴ All SNVs/small InDels with a VAF >20% were retained.

Tumor copy number profiles were calculated with FREEC (0.9) using constitutional DNA as reference,^{15–17} with annotations of CRGs/EMGs to highlight focal variations.¹⁸

Genetic variations were termed “somatic” if no evidence of these variations was observed in the constitutional reads.

TSCA[®] analysis

The TSCA[®] sequencing reads were mapped to the reference human genome with Bowtie.¹⁹ To prevent strand or coverage bias due to noncomplete overlap of forward (FW) and reverse (RV) reads of amplicons, the bam files were split to FW and RV read bam(s).

An adapted approach was developed to enable variant calling in TSCA[®] data.²⁰

For SNVs, coverage analysis was performed at each base position (GATK DepthOfCoverage). The background of noise (variability) of each amplicon was then analyzed in the six negative controls.

All possible variants (e.g., ref = A; A > T, A > C, A > G, A > -) and their alternative allele fraction (AF) were calculated for each position, and Fisher's exact two-sided tests with a Bonferroni correction were performed to compare percentages of variant allele fraction (VAF) for a given base between a case and the negative controls. Significant variations were filtered-in in case of an increase in the percentage of variant base (5% significance level). Only positions with total depth of coverage higher than 100× were considered for variant analysis. All SNVs/small InDels with a VAF >5% were retained.

Recurrent variants and variants with VAF >20% detected in more than two samples were filtered-out as they most likely indicate polymorphisms. Finally, FW and RV results were confronted and ambiguous cases were filtered-out.

Amplicon InDels Hunter tool was used to detect large InDels.²¹ Recurrent InDels detected in more than two samples were filtered-out.

For focal CNAs, the depth of coverage of targeted regions was calculated for FW and RV bam (GATK DepthOfCoverage). For each sample, the coverage of each amplicon was normalized by the median of sample depth followed by comparison to negative controls. The FW and RV normalized data were then merged to reconstruct copy number profiles. The copy number profile obtained from TSCA data using our in-house developed pipeline was also confirmed by another technique ONCOCNV, a method that includes a multifactor normalization with respect to library size, CG-content and target length to detection of large copy number changes from amplicon deep sequencing data.²²

Finally, the list of SNVs and InDels was manually curated and visually inspected in IGV (Integrated Genome Viewer).²³

Variant annotation, prioritization and classification

All SNVs/InDels were annotated using SnpEff/SnpSift.²⁴ Annotation was performed using canonical transcripts and publicly available databases dbSNPv150,²⁵ gnomADv2.0.2 (<http://gnomad.broadinstitute.org/>),²⁶ COSMICv74,²⁷ ClinVar(25-02-2018),²⁸ dbNSF Pv2.9.3¹⁸ and known cancer Hotspots.²⁹ Loss of function (LoF) assessment was performed using the snpEff tool. SNVs/InDels were discarded as benign/likely benign/polymorphic based on population minor allele frequency (popmax) ≥0.1% from gnomAD database. However, SNVs/inDels reported as pathogenic/likely pathogenic in ClinVar were retained. In addition to COSMIC annotations, six prediction algorithms VEST3,³⁰ MetaSVM,³¹ REVEL,³² Mutation Taster,³³ M_CAP³⁴ and CADD³⁵ were used to classify all likely protein-altering SNVs/InDels (frameshift, nonsense, splice donor/acceptor, nonsynonymous and in-frame insertions/deletions) as pathogenic, likely pathogenic. The flowchart of variant classification is indicated in Supporting Information Figure S4A and the prediction score thresholds in Supporting Information Figure S4B.

SNVs/small InDels validation

All retained SNVs/small InDels with VAF >20% were confirmed by Sanger sequencing. SNVs with VAF <20% were validated by NGS/targeted sequencing in an independent second experiment.³⁶

Comparison of CRG/EMG variation frequencies between NB and gnomAD reference cohort

We compared the CRG/EMG variation frequencies from the NB cohort (Supporting Information Table S3) with those reported in a large unrelated population (gnomAD database; $n = \sim 123,000$). The gnomAD VCF file was intersected with the CRG/EMG TSCA[®] panel bed file and the corresponding subset of variants from gnomAD were extracted. Using the same approach as in the NB cohort, SNVs/InDels from gnomAD were annotated/classified into pathogenic/likely pathogenic and their frequency was calculated by dividing the sum of the allele count by the median of total alleles. To compare the frequencies of CRG/EMG variations in the NB vs. the control cohort, for each gene an Odds Ratios was calculated: OR = (NB cases with pathogenic/likely pathogenic variant/cases without pathogenic or likely pathogenic variant)/(gnomAD

control with pathogenic or likely pathogenic variant/gnomAD control without pathogenic or likely pathogenic variant) along with 95% confidence intervals. *p*-Value and OR were obtained using two-sided Fisher's test performed in R.³⁷

Statistical analysis

Correlation analyses using chi-squared test was done with MedCalc (Medical Calculator 13.3.0.0). Progression-free survival (PFS) and overall survival (OS) were estimated using the Kaplan–Meier method, and comparisons were made using log-rank tests.

Western blot

Immunoblots were done as reported previously using monoclonal rabbit SMARCA4/BRG1 (EPR3912, #GTX62750, Genetex, Irvine, CA) and HRP-conjugated GAPDH (#HRP-60004, ProteinTech, Rosemont, IL) antibodies.

Results

Genetic variations of CRGs/EMGs detected in 283 NB patients and in 30 NB cell lines

In a cohort of 283 patients studied by WGS/WES ($n = 35$) or by TSCA[®] ($n = 228$), or both ($n = 20$), pathogenic/likely pathogenic genetic variations (focal CNAs, InDels, SNVs) in at least one CRG/EMG were identified in 56/283 cases (20%). Altogether 60 pathogenic/likely pathogenic genetic variations were identified (Table 1, Fig. 1 and Supporting Information Table S3). The most frequently altered genes were *ATRX*, *SMARCA4*, *MLL3* and *ARID1B*.

ATRX was targeted by focal CNAs, small InDels and SNVs in 6, 2 and 6 cases, respectively (Supporting Information Table S3 and Fig. 1). *SMARCA4* was altered by 6 SNVs/InDel, 1 focal CNA and 1 large deletion. *MLL3* was targeted by SNVs in seven cases (Supporting Information Table S3, Fig. 1). *ARID1A* and *ARID1B* genes were altered in altogether 10/283 (3.5%) cases, with two SNVs and one InDel in *ARID1A* and five InDels and two SNVs in *ARID1B*.⁹

Variations in two different CRGs/EMGs were observed in three cases (Supporting Information Table S3, in bold; Fig. 1), targeting *MLL3* and *ATRX*, *MLL3* and *ARID1B*, and *PBRM1* and *ATRX*, respectively.

Double events encompassing both SNV/small InDels, focal CNAs and additional LOH/copy number loss of the same gene were observed in six cases (Fig. 1). The *SMARCA4* gene was targeted by double events in three cases (NBSW80, NBSW251 and NBSW77) with either a SNV and/or InDel (3 bp) in the presence of only 1 copy of the *SMARCA4* gene, corresponding to focal deletions (Figs. 2a and 2b).

Case NBSW165 showed a loss of one copy and a focal CNA covering a region from EX2 to EX9 of the second copy of *ATRX* (Fig. 1; Supporting Information Table S3). A copy number loss involving the *PBRM1* gene is observed in case NBSW23; a somatic SNV is also observed in the same gene (Fig. 1).

In another patient (NBSW26) presenting with speech delay, attention deficit and hyperactivity disorder as well as an adrenal

stage *INSS stage 1*, *INRG L1* NB diagnosed at 4 years of age, a germline deletion of 1.7 Mb at chromosome band 19p13.2 (chr19:10462524–12157782) was detected (Supporting Information Fig. S2). The germline deleted region encompassing among others the *SMARCA4* gene was confirmed by aCGH on peripheral blood lymphocytes (Supporting Information Table S2).

Targeted TSCA[®] sequencing of the 33 CRGs/EMGs was performed in 30 NB cell lines (Supporting Information Table S1). Nine variations were detected in seven NB cell lines (Table 1, Supporting Information Table S3, Fig. 1), most frequently targeting *SMARCA4* (3 cases), *MLL2* (2 cases) and *MLL3* (2 cases) (Fig. 1, Supporting Information Table S3).

Interestingly, a genetic event consisting of partial loss of the *MLL2* gene was detected in the NB cell line CLB-Re (Fig. 3a). SNP6 analysis in the same NB cell line revealed a large region of copy number loss in chr12 encompassing the *MLL2* gene at the position chr12:49449107–58196639, confirming the observation obtained by TSCA sequencing (Fig. 3b).

In the SKNFI cell line, a double event (SNV and LOH) was observed in *SMARCA4* (Figs. 4a and 4b). The absence of expression of *SMARCA4* was confirmed by expression analysis and western blot. On the other hand, expression of *SMARCA4* was maintained in the SKNSH cell line which harbors a single SNV event in this gene (Figs. 4c and 4d).

Germline vs. somatic occurrence of variations

Altogether, among the 60 pathogenic/likely pathogenic CRG/EMG variations, the variation occurred at a somatic level in five cases, targeting *ATRX*, *PBRM1* or *MLL3*, whereas a germline variation could be confirmed in 13 other cases, targeting *ARID1B*, *MLL3*, *MLL2*, *ATRX*, *ARID1A* and *SMARCA4* (Supporting Information Table S3; Fig. 1). For the remaining 42 variations detected in tumor samples without paired germline material available, a distinction between germline and somatic variation could not be made.

We then compared the frequencies of pathogenic/likely pathogenic SNVs/inDels of the 33 studied CRGs/EMGs in the NB cohort to those observed in gnomAD as a control cohort. The analysis revealed a significantly higher frequency of variations ($p < 0.05$ after Bonferroni correction) in two CRGs, *SMARCA4* (OR 4.49, 95% CI: 1.63–9.97, $p = 0.038$) and *ATRX* (OR 3.44, 95% CI: 1.46–6.91, $p = 0.043$) in NB cases when compared to gnomAD, independent of whether the observed SNVs were of known germline, somatic or unknown origin, indicating an enrichment for variations in the genes *SMARCA4* and *ATRX* in NB cases. All other CRGs/EMGs included in our study were not altered with a higher frequency than in the general population (Supporting Information Table S4).

Potential functional impact

The potential functional impact of the observed variations was analyzed *in silico*. For a given variation, within the studied gene, the localization of the amino acid change was studied with regards

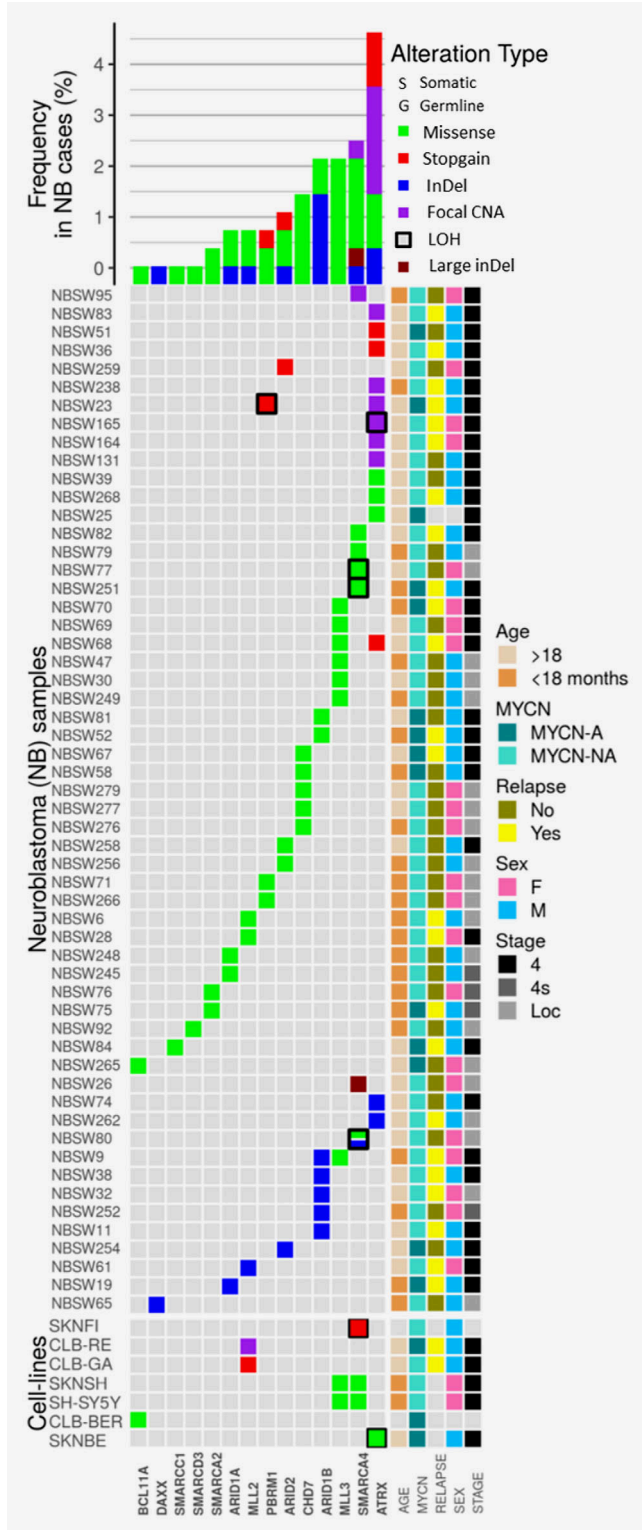
to domains of functional importance for chromatin remodeling or highly conserved domains but with unknown function.

Among the 30 variations detected in CRGs, 12 variations were localized in these domains (9 SNVs, 1 focal CNA, 1 InDel and

1 large inDel; Supporting Information Table S3). Their position was then analyzed with regards to protein-specific domains (DNA binding, bromodomain, histone binding, protein-protein interaction domains). In the *SMARCA4* gene, four identified variations mapped to these functional domains (Fig. 4d), including three missense point mutations in the ATPase domain of *SMARCA4* gene (Fig. 4d).

Among the 30 variations observed in EMGs, 12 variations were localized in domains of functional importance (10 SNVs, 1 focal CNA, 1 InDel; Supporting Information Table S3).

Overall, among all 60 variations observed in CRGs/EMGs, 24 mapped to domains of functional importance at a protein level (Supporting Information Table S3).



Correlation of CRGs/EMGs variations with clinical parameters

There were no statistically significant correlations between the presence of a CRG/EMG variation and the main clinical prognostic parameters of NB (data not shown). Furthermore, among patients whose tumors were analyzed by WGS ($n = 8$),¹⁰ the presence of CRG/EMG variations, was not associated with an overall increase in the tumor mutational burden (data not shown).

A statistically significant poorer overall survival was observed for patients whose tumors harbored CRG/EMG variations in the overall cohort (Supporting Information Fig. S3A). However, there was no statistically significant difference of overall survival between the two patient groups in patient subgroups defined by age, stage or MYCN status, nor among all high-risk patients only ($n = 124$; Supporting Information Fig. S3C). Furthermore, no statistically significant difference in PFS was observed (Supporting Information Fig. S3B).

Discussion

Recent reports have highlighted chromatin remodeling as an important player in oncogenesis, with the main mechanisms of action consisting in tumor suppression, but variations in CRG/EMGs might also play a role in oncogenesis via gain of function.^{38,39}

Figure 1. Genetic variations in chromatin remodeling and epigenetic modifier genes across a series of 283 NB patients and 30 NB cell lines analyzed by TSCA and/or WGS/WES approaches. In the lower part of the figure, genes found to be altered ($n = 14$) are arranged in rows; cases for whom a genetic variation is detected in the studied genes ($n = 56$) are arranged in columns, respectively. The 60 SNVs, InDels, and focal CNAs detected in the 56 cases are represented by colored cases. Double events (SNV/inDels and LOH) are represented by colored cases surrounded by a black square. In the lower part of the graph, the 9 SNVs, InDels and CNAs detected in 7 NB cell lines are represented. The overall frequencies are indicated in the upper half of the figure. Nonsynonymous SNVs are represented in green, stop-gain SNVs are represented in red, InDels are represented in blue, focal CNA in purple and large deletions in brown. The right data grid summarizes clinical information of each neuroblastoma sample.

Table 1. Cases showing SNVs, InDels and focal CNAs in CRGs/EMGs in 283 neuroblastoma patients and 30 neuroblastoma cell lines

Cohort	MYCN status	Stage	Age at diagnosis	Analytical technique	Cases with pathogenic/likely pathogenic variations in CRGs/EMGs	Number of pathogenic/likely pathogenic variations in CRGs/EMGs	Pathogenic/likely pathogenic variations in CRGs/EMGs	Origin of variations
Patients <i>n</i> = 283	MNA <i>n</i> = 64; MN-NA <i>n</i> = 219	Stage 4 <i>n</i> = 128; Localized <i>n</i> = 130; Stage 4s <i>n</i> = 25	<18 months <i>n</i> = 159; ≥18 months <i>n</i> = 124	WES/WGS <i>n</i> = 55	16	18	ATRX <i>n</i> = 5	2/5 germline; 3/5 Somatic
							ARID1B <i>n</i> = 5	Germline
							MLL3 <i>n</i> = 3	2/3 Germline; 1/3 Somatic
							MLL2 <i>n</i> = 2	Germline
							PBRM1 <i>n</i> = 1	Somatic
							ARID1A <i>n</i> = 1	Germline
							SMARCA4 <i>n</i> = 1	Germline
				TSCA <i>n</i> = 228	40	42	ATRX <i>n</i> = 9	Unknown
							SMARCA4 <i>n</i> = 7	
							MLL2 <i>n</i> = 1	
							MLL3 <i>n</i> = 4	
							ARID1A <i>n</i> = 2	
							ARID2 <i>n</i> = 4	
							CHD7 <i>n</i> = 5	
							SMARCC1 <i>n</i> = 1	
							SMARCA2 <i>n</i> = 2	
							SMARCD3 <i>n</i> = 1	
							PBRM1 <i>n</i> = 2	
							DAXX <i>n</i> = 1	
							ARID1B <i>n</i> = 2	
							BCL11A <i>n</i> = 1	
Cell lines <i>n</i> = 30	MNA <i>n</i> = 22 MN-NA <i>n</i> = 8			TSCA <i>n</i> = 30	7	7	ATRX <i>n</i> = 1	
							MLL2 <i>n</i> = 2	
							MLL3 <i>n</i> = 1	
							SMARCA4 <i>n</i> = 2	
							BCL11A <i>n</i> = 1	

In the 55 samples analyzed by WES/WGS, 7 CRG/EMG genes show at least one variation. In the 248 samples analyzed by the TSCA approach, 14 CRG/EMG genes show a variation. Five genes harbor a variation in NB cell lines.

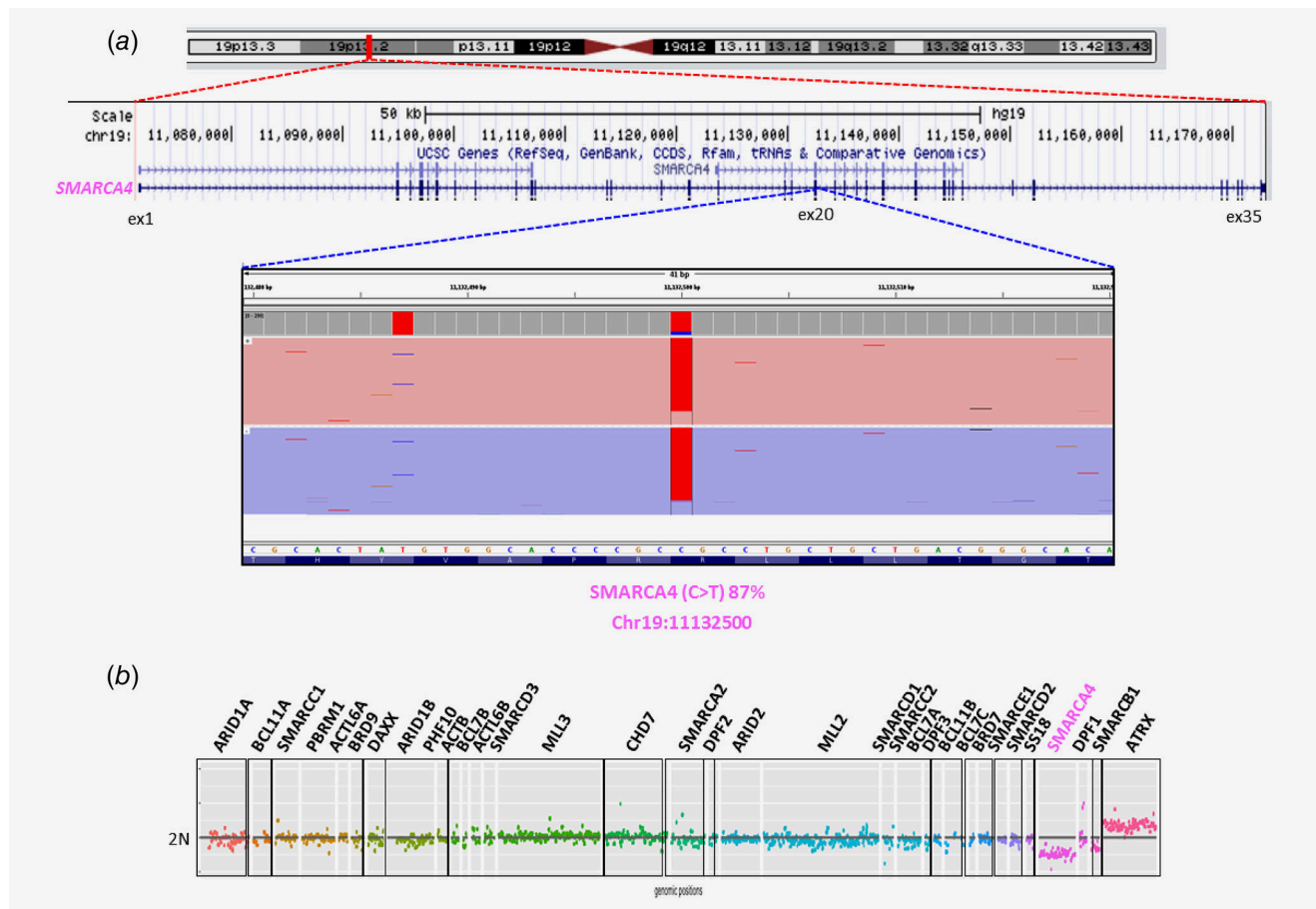


Figure 2. Example of double event detection (SNV and copy number loss in *SMARCA4* gene) by TSCA and aCGH data analysis in NBSW80. (a) The IGV profile from TSCA analysis revealed the presence of one SNV in *SMARCA4* (C>T; chr19: 11132500) gene with variant allele fraction of 87%. (b) The comparison between the copy number profiles obtained by TSCA revealed a second event with copy number loss in the *SMARCA4* gene.

CRG alterations have previously been identified in NB patients, targeting *ARIDs*, *ATRX*, *DAXX* or *SMARCA4*.^{7–9,40,41}

Our study now focused on the genetic status of 33 CRGs/EMGs in tumor samples of 283 NB patients. Altogether, CRG/EMG variations (focal CNA, InDels, SNVs) were identified in 56 cases (20%), mapping to CRG/EMG functional domains in 8.4% (24/283) of cases and most frequently targeting *ATRX*, *SMARCA4*, *MLL3* and *ARID1B*. Other CRGs and many more EMGs, not covered by the panel used in this targeted analysis, exist throughout the genome,⁴² and thus it is possible that variations targeting other CRG/EMGs not taken into account in our study might also be present in the analyzed samples.

In NB, *ATRX* alterations have been associated with activated alternative lengthening of telomere (ALT).⁸ In our study, 5% (14/283) of cases showed an *ATRX* variation (SNV/InDel or focal CNA). Further studies will be required to establish the role of these variations on patient survival.

The *MLL3* gene is rarely altered in NB, with two *MLL3* variations (A293V and P309L) reported outside of annotated protein domains.⁴³ In our cohort, seven missense SNVs were detected, four of them within *MLL3* functional domains.

Alterations in *ARID1A/ARID1B* have been described in 11% of NB patients,⁹ with the presence of variations in a single allele of *ARID1A/ARID1B* possibly corresponding to a dominant tumor suppressor.⁴¹ Ten cases showed *ARID1A/ARID1B* variations, with two SNV/inDels occurring in functional domains.

Altogether, among the 56 cases with CRG/EMG variations, four patients harbored somatic and 12 showed germline variations. For the remaining 40 cases, due to the absence of paired germline DNA, it could not be established whether the observed variations were somatic or might concern private constitutional polymorphisms. Yet constitutional variations of potential functional impact might also be of oncogenic importance.

Indeed, germline events in CRGs such as *SMARCB1*, *SMARCA4* or *SMARCE1* have been described as predisposition syndromes for other cancer types, including RT.¹ In our series, germline events in CRGs/EMGs were observed in 23.6% (13/55) of patients with available germline material.

In RT, the recurrent germline loss of one *SMARCB1* allele followed by somatic loss of the second allele indicates a classic tumor suppressor, with loss of critical subunits of SWI/SNF.² Given the possible suppressor behavior of CRG/EMGs, we searched for

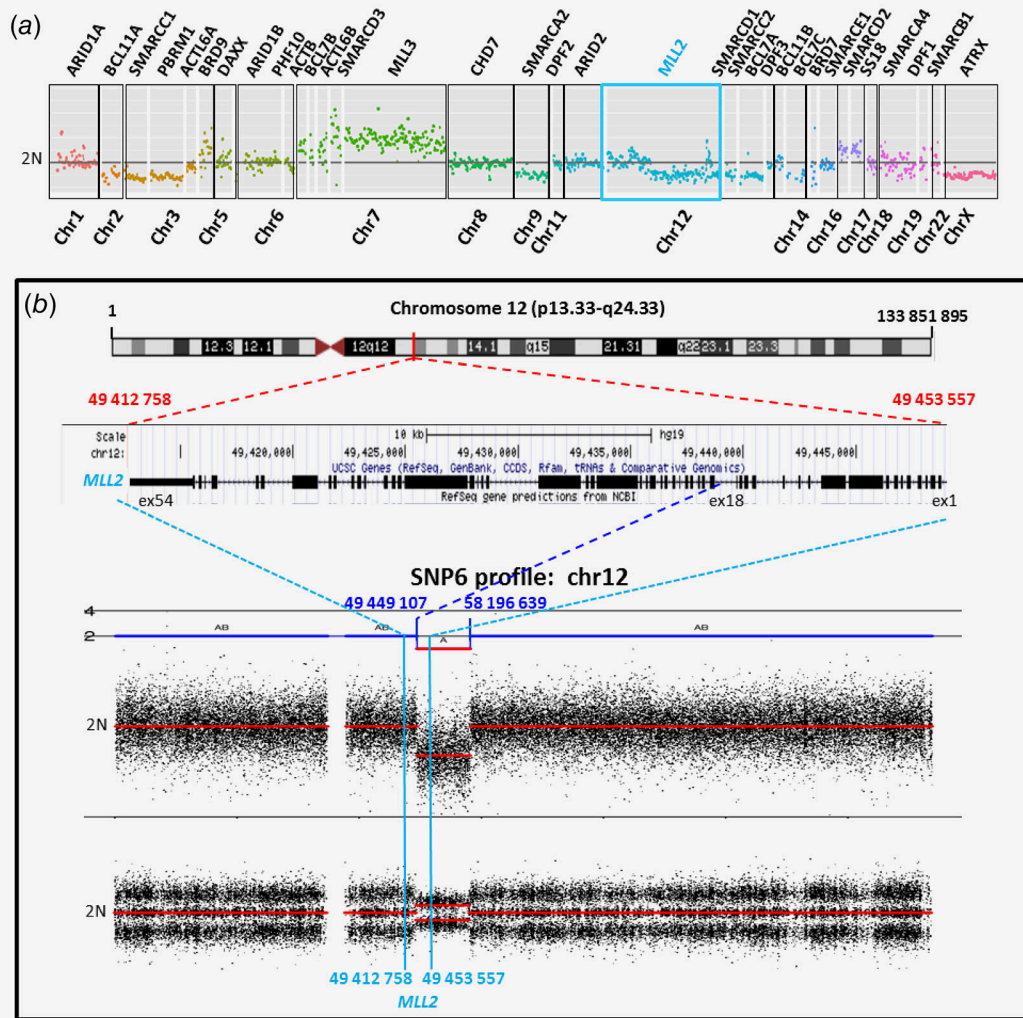


Figure 3. Partial loss event detected in *MLL2* gene in CLB-Re NB cell line. (a) The copy number profile obtained by TSCA shows the presence of a partial loss starting in *MLL2* gene. (b) SNP6.0 array-based copy number encompassing the Chromosome 12 confirms the presence of a large loss with a breakpoint within the *MLL2* gene.

double events possibly causing loss of function, with double events identified in 2.1% (6/283) of patients, most frequently in *SMARCA4* (3/283).

SMARCA4 has been shown to play a role in the oncogenesis of different tumors and thus are not specific to NB. Indeed, germline alterations in *SMARCA4* conferring predisposition to SCOHT and RT.^{39,44,45} At a somatic level, missense point mutations in *SMARCA4* mapping to the ATPase domain might cause loss of direct binding between BAF and PRC1 which could contribute to oncogenesis or epigenetic plasticity during tumor development.⁴⁶ Such *SMARCA4* missense point mutations have been described in medulloblastoma.^{47,48}

In our study, 5/8 *SMARCA4* variations concerned missense point mutations. Although it has been suggested that *SMARCA4* might function as an oncogene in NB, and the overexpression of *SMARCA4* in Stage 4 NB patients is associated with poorer outcomes,⁴⁹ the functional impact of genetic variations can be

difficult to determine without in-depth functional studies which are beyond the scope of this article.

Even though the somatic or germline origin of *SMARCA4* variations in our study remained undetermined, our data suggest that variations in *SMARCA4* might play a role in the oncogenesis of NB, and several arguments now underline its role. First, we describe a case of NB within the context of a germline deletion encompassing *SMARCA4*. Second, our findings highlight a significantly higher frequency of variations in *SMARCA4* in NB cases when compared to the gnomAD reference population. Third, we identified four NB cases showing double events targeting *SMARCA4*. In addition, a double event (stop-gain variation and LOH) of *SMARCA4* was observed in the SKNFI NB cell line causing the absence of *SMARCA4* expression.

Further functional studies are needed for a deeper understanding of *SMARCA4* variations in NB in order to guide the development of more effective therapies.

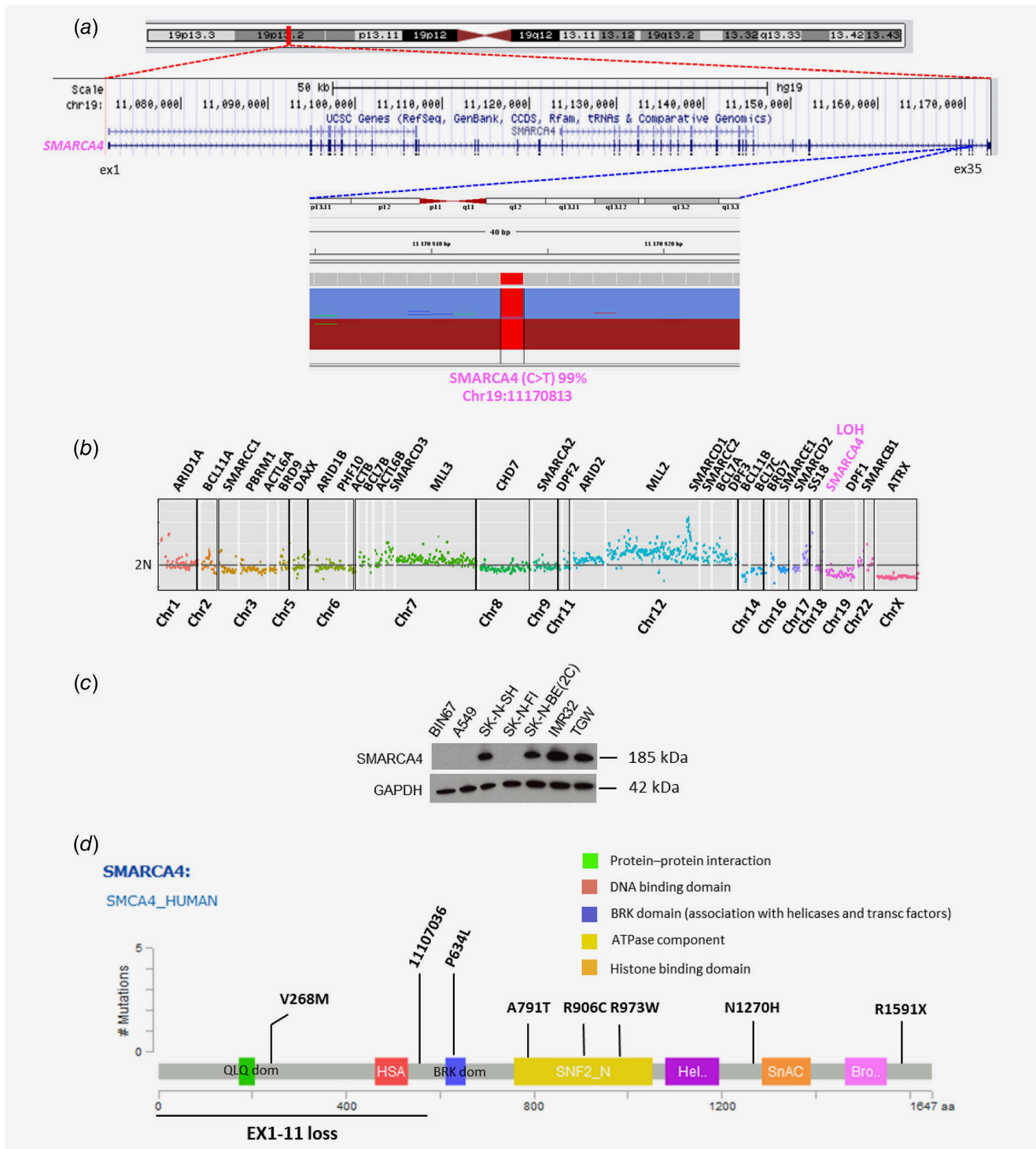


Figure 4. Example of double event (LOH and SNV in *SMARCA4* gene) in SKNF1 NB cell line. (a) The IGV profile from TSCA amplicon-seq analysis revealed the presence of one SNV in *SMARCA4* gene (C>T; Chr19: 11170813) with variant allele fraction of 99%. (b) The copy number profile obtained by TSCA shows the presence of a second event with copy number loss in the same gene. (c) Immunoblot of *SMARCA4* expression in different neuroblastoma and nonneuroblastoma cell lines. BIN67: Small cell carcinoma of the ovary of hypercalcemic type; A549: lung carcinoma; SKNSH, SKNFI, SKNBE(2C), IMR32 and TGW: neuroblastoma. (d) SNVs, inDels and focal CNAs identified in *SMARCA4* gene mapping functional domains in patients enrolled in the study.

Altogether, our data demonstrate CRG/EMG variations with likely functional impact in 8.4% (24/283) of all NB patients.

Although no difference in PFS between cases with or without CRG/EMG variations was observed, a poorer overall survival in cases harboring CRG/EMG variations was observed. This suggests

that cases harboring CRG/EMG variations might present more aggressive disease or that salvage treatment used at relapse might potentially be more efficient in cases showing no CRG/EMG variation. Correlations between the presence of variations in CRG/EMG and other genetic factors associated with poor outcome in NB should be determined in future studies and larger series, such as the overall mutational burden⁵⁰ or factors contributing to a mechanistical classification of NB including mutations of the RAS: MAPK pathway and telomere maintenance mechanisms.⁵¹

New treatment strategies are necessary for patients with high-risk NB, and our data further highlight the potential interest of drugs modulating chromatin remodeling processes. Histone deacetylase inhibitors (HDACi) or DNA methyltransferase inhibitors were the first epigenetic compounds to reach clinical trials with potential benefits in some patients with CRG/EMG variations.⁵² Further preclinical development of therapeutic

approaches involving CRGs/EMGs and in particular *SMARCA4* will be warranted.

Acknowledgements

Our study was supported by the Annenberg Foundation, the Nelia and Amadeo Barletta Foundation (FNAB) and the Association Hubert Gouin Enfance et Cancer. Our study was also funded by the Associations Enfants et Santé, Les Bagouz à Manon, Les amis de Claire. Funding was obtained from SIRIC/INCa (Grant INCa-DGOS-4654), as well as from the CEST of Institute Curie, and Programme Hospitalier de recherche en Cancérologie (PHRC) IC2007-09 grant. High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche ("Investissements d'Avenir" program), by the Cancerpole Ile-de-France and by the SIRIC-Curie program-SIRIC Grant "INCa-DGOS-4654", as well as the grant of collaboration with CEA/IG/CNG financed by France Génomique infrastructure, as part of the 517 program "Investissements d'Avenir" from the Agence Nationale pour la Recherche (contract ANR-10-INBS-518 09).

References

- Masliyah-Planchon J, Bièche I, Guinebretière J-M, et al. SWI/SNF chromatin remodeling and human malignancies. *Annu Rev Pathol Mech Dis* 2015;10:145–71.
- Versteeg I, Sévenet N, Lange J, et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* 1998;394:203–6.
- Kadoch C, Hargreaves DC, Hodges C, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* 2013;45:592–601.
- Matthay KK, Maris JM, Schleiermacher G, et al. Neuroblastoma. *Nat Rev Dis Primer* 2016;2:16078.
- Molenaar JJ, Koster J, Zwijnenburg DA, et al. Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 2012;483:589–93.
- Cheung N-KV, Dyer MA. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat Rev Cancer* 2013;13:397–411.
- Pugh TJ, Morozova O, Attiyeh EF, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* 2013;45:279–84.
- Kurihara S, Hiyama E, Onitake Y, et al. Clinical features of ATRX or DAXX mutated neuroblastoma. *J Pediatr Surg* 2014;49:1835–8.
- Sausen M, Leary RJ, Jones S, et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat Genet* 2013;45:12–7.
- Elefeld TF, Oldridge DA, Bernard V, et al. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* 2015;47:864–71.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;25:1754–60.
- McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
- DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 2016;32:1220–2.
- Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 2012;28:423–5.
- Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568–76.
- Seshan V, Olshen A. *DNACopy: DNA copy number data analysis. R package version 1.46.0*. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- Liu X, Wu C, Li C, et al. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutat* 2016;37:235–41.
- Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods* 2012;9:357–9.
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987–93.
- Kadri S, Zhen CJ, Wurst MN, et al. Amplicon Indel hunter is a novel bioinformatics tool to detect large somatic insertion/deletion mutations in amplicon-based next-generation sequencing data. *J Mol Diagn* 2015;17:635–43.
- Boeva V, Popova T, Lienard M, et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinformatics* 2014;30:3443–50.
- Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–6.
- Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 2012;6:80–92.
- Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308–11.
- Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285–91.
- Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* 2017;45:D777–83.
- Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016;44:D862–8.
- Chang MT, Bhattarai TS, Schram AM, et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov* 2018;8:174–83.
- Carter H, Douville C, Stenson PD, et al. Identifying Mendelian disease genes with the variant effect scoring tool. *BMC Genomics* 2013;14(Suppl 3):S3.
- Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 2015;24:2125–37.
- Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* 2016;99:877–85.
- Schwarz JM, Cooper DN, Schuelke M, et al. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361–2.
- Jagadeesh KA, Wenger AM, Berger MJ, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet* 2016;48:1581–6.
- Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5.
- Bellini A, Bernard V, Leroy Q, et al. Deep sequencing reveals occurrence of subclonal ALK mutations in neuroblastoma at diagnosis. *Clin Cancer Res* 2015;21:4913–21.
- R Core Team (2018). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available from: <https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing> [date last accessed June 18, 2018].
- Biegel JA, Busse TM, Weissman BE. SWI/SNF chromatin remodeling complexes and cancer. *Am J Med Genet C Semin Med Genet* 2014;166C:350–66.

39. Holsten T, Bens S, Oyen F, et al. Germline variants in SMARCB1 and other members of the BAF chromatin-remodeling complex across human disease entities: a meta-analysis. *Eur J Hum Genet* 2018;26:1083–93.
40. Peifer M, Hertwig F, Roels F, et al. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 2015;526:700–4.
41. Cancer Genome Atlas Research Network, Kandoth C, Schultz N, et al. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;497:67–73.
42. Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics. *Sci Adv* 2015;1:e1500447.
43. Chmielecki J, Bailey M, He J, et al. Genomic profiling of a large set of diverse pediatric cancers identifies known and novel mutations across tumor spectra. *Cancer Res* 2017;77:509–19.
44. Witkowski L, Goudie C, Foulkes WD, et al. Small-cell carcinoma of the ovary of hypercalcemic type (malignant Rhabdoid tumor of the ovary): a review with recent developments on pathogenesis. *Surg Pathol Clin* 2016;9:215–26.
45. Witkowski L, Carrot-Zhang J, Albrecht S, et al. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat Genet* 2014;46:438–43.
46. Stanton BZ, Hodges C, Calarco JP, et al. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet* 2017;49:282–8.
47. Robinson G, Parker M, Kranenburg TA, et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* 2012;488:43–8.
48. Pugh TJ, Weeraratne SD, Archer TC, et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 2012;488:106–10.
49. Jubierre L, Soriano A, Planells-Ferrer L, et al. BRG1/SMARCA4 is essential for neuroblastoma cell viability through modulation of cell death and survival pathways. *Oncogene* 2016;35:5179–90.
50. Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers. *Nature* 2018;555:321–7.
51. Ackermann S, Cartolano M, Hero B, et al. A mechanistic classification of clinical phenotypes in neuroblastoma. *Science* 2018;362:1165–70.
52. Gryder BE, Sodji QH, Oyelere AK. Targeted cancer therapy: giving histone deacetylase inhibitors all they need to succeed. *Future Med Chem* 2012;4:505–24.

3.2.2 Paper II: Impact of *ALK* Amplifications and Mutations in the European Neuroblastoma Study Group (SIOPEN) High-Risk Neuroblastoma Trial (HR-NBL1)

3.2.2.1 Introduction

The study focused on high-risk neuroblastoma (HR-NB), a particularly challenging childhood cancer to treat effectively. The primary objective was to determine if the frequency of *ALK* (anaplastic lymphoma kinase) alterations, either amplifications (> 10 copies) or single nucleotide variants (SNVs), in a large cohort of HR-NB patients and to correlate *ALK* alterations with outcome in patients undergoing treatment as part of the European high-risk neuroblastoma (HR-NBL1) /SIOPEN trial.

3.2.2.2 Materials and Methods

Diagnostic tumor samples from 1,092 HR-NBL1/SIOPEN patients were analyzed to determine the status of *ALK* amplification (*ALKa*) and *ALK* mutational profile. The study included 330 cases for *ALKa* determination, 191 for *ALK* mutational profiling, and 571 for both alterations.

3.2.2.3 Results

Ultra deep targeted sequencing techniques resulted in a higher detection rate of *ALK* mutations. These *ALK* activating mutations in neuroblastoma predominantly occur within the kinase domain. The hotspots for these mutations are primarily located at three key positions: F1174, R1275, and F1245. The frequency of these mutations varies within the tumor population, as they can occur at different levels of mutation allele fraction (MAF). The mutations are categorized as either clonal or subclonal based on their prevalence within the tumor cells:

Clonal Mutations: These are defined as mutations with a MAF of more than 20%. In this case, a significant proportion of the tumor cells carry the mutation, suggesting that these mutations are dominant within the tumor population and likely play a significant role in tumor development and progression. Clonal *ALK* mutations (n5 = 65) 33% [95% CI, 21 to 44] were associated with poorer patient outcomes and held independent prognostic significance (see Figure 3.2.2.3.1).

Subclonal Mutations: In this study, subclonal *ALK* mutations defined by a Mutated Allele Fraction (MAF) of less than 20%, to account for contamination with normal tissue, were (n5=22) 48% [95% CI, 26 to 67] cases observed and suggest their presence in a minor subset of tumor cells, reflecting a lower prevalence within the tumor. Subclonal events, did not demonstrate a similar impact on prognosis (see Figure 3.2.2.3.1). These subclonal events accounted for 28% of all *ALK* mutations,

with a very low MAF (less than 5%) in some cases. The MAF cutoff, was defined at > 20% given that tumor cell content was considered at > 50% in all cases., ideally the MAF should consider factors like copy number variation and tumor purity for accurate assessment of clonality which enables accurate calculation of CCF. The low MAF observed is indicative of intratumoral heterogeneity, which has been previously reported for *MYCN* amplification (MNA) and segmental chromosome alterations in NB (Theissen et al. 2009) (López-Carrasco et al. 2021). These subclonal mutations likely emerge later during tumor progression or represent minor tumor cell populations with such pathogenic mutations that remain dormant until selective pressures favor their expansion. This phenomenon is evident in neuroblastoma where subclonal *ALK* mutations detected at diagnosis can undergo clonal expansion, contributing to relapse (Schleiermacher et al. 2014) (Eleveld et al. 2015).

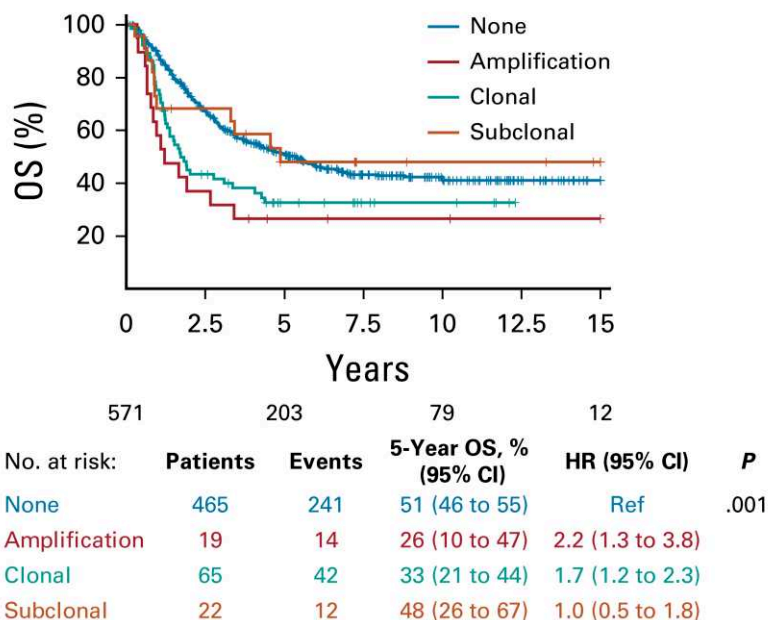


Figure 3.2.2.3.1
OS according to the type of ALK alteration in the cohort of 571 patients with known ALK amplification and ALK mutation status: no alteration (n 5 465), 5-year OS 51% (95% CI, 46 to 55); clonal mutations (n 5 65), 5-year OS 33% (95% CI, 21 to 44); subclonal mutations (n 5 12), 5-year OS 48% (95% CI, 26 to 67); and ALK amplification (n 5 19), 5-year OS 26% (95% CI, 10 to 47), respectively; P 5 .001.

Figure 3.2.2.3.2: OS according to the type of ALK alteration in the cohort of 571 patients with known ALK amplification and ALK mutation status

3.2.2.4 Conclusion

Genetic alterations in *ALK*, specifically clonal mutations and amplifications, are independent predictors of poorer survival in HR-NB. These findings provide a strong rationale for integrating *ALK* inhibitors into the upfront treatment of HR-NB patients with *ALK* alterations, potentially improving outcomes for this high-risk group.

Frequency and Prognostic Impact of *ALK* Amplifications and Mutations in the European Neuroblastoma Study Group (SIOPEN) High-Risk Neuroblastoma Trial (HR-NBL1)

Angela Bellini, PhD^{1,2,3}; Ulrike Pötschger, PhD^{4,5}; Virginie Bernard, PhD⁶; Eve Lapouble, PhD⁷; Sylvain Baulande, PhD⁶; Peter F. Ambros, PhD⁵; Nathalie Auger, PhD⁸; Klaus Beiske, MD, PhD⁹; Marie Bernkopf, PhD⁵; David R. Betts, PhD¹⁰; Jaydutt Bhalshankar, MSc^{1,2,3}; Nick Bown, PhD¹¹; Katleen de Preter, PhD¹²; Nathalie Clément, PhD^{1,2,3}; Valérie Combaret, PhD¹³; Jaime Font de Mora, PhD¹⁴; Sally L. George, MD, PhD¹⁵; Irene Jiménez, MD^{1,2,3}; Marta Jeison, PhD¹⁶; Barbara Marques, PhD¹⁷; Tommy Martinsson, PhD¹⁸; Katia Mazzocco, PhD¹⁹; Martina Morini, PhD²⁰; Annick Mühlethaler-Mottet, PhD²¹; Rosa Noguera, MD²²; Gaelle Pierron, PhD⁷; Maria Rossing, PhD²³; Sabine Taschner-Mandl, PhD⁵; Nadine Van Roy, PhD¹²; Ales Vicha, PhD²⁴; Louis Chesler, MD, PhD²⁵; Walentyna Balwierz, MD²⁶; Victoria Castel, MD, PhD²⁷; Martin Elliott, MD²⁸; Per Kogner, MD, PhD²⁹; Geneviève Laureys, MD, PhD³⁰; Roberto Luksch, MD³¹; Josef Malis, MD²⁴; Maja Popovic-Beck, MD³²; Shifra Ash, MD³³; Olivier Delattre, MD^{2,3,6}; Dominique Valteau-Couanet, MD, PhD³⁴; Deborah A. Tweddle, MD, PhD³⁵; Ruth Ladenstein, MD, PhD^{36,37}; and Gudrun Schleiermacher, MD, PhD^{1,2,3}

PURPOSE In neuroblastoma (NB), the *ALK* receptor tyrosine kinase can be constitutively activated through activating point mutations or genomic amplification. We studied *ALK* genetic alterations in high-risk (HR) patients on the HR-NBL1/SIOPEN trial to determine their frequency, correlation with clinical parameters, and prognostic impact.

MATERIALS AND METHODS Diagnostic tumor samples were available from 1,092 HR-NBL1/SIOPEN patients to determine *ALK* amplification status ($n = 330$), *ALK* mutational profile ($n = 191$), or both ($n = 571$).

RESULTS Genomic *ALK* amplification (*ALKa*) was detected in 4.5% of cases (41 out of 901), all except one with *MYCN* amplification (MNA). *ALKa* was associated with a significantly poorer overall survival (OS) (5-year OS: *ALKa* [$n = 41$] 28% [95% CI, 15 to 42]; no-*ALKa* [$n = 860$] 51% [95% CI, 47 to 54], [$P < .001$]), particularly in cases with metastatic disease. *ALK* mutations (*ALKm*) were detected at a clonal level ($> 20\%$ mutated allele fraction) in 10% of cases (76 out of 762) and at a subclonal level (mutated allele fraction 0.1%-20%) in 3.9% of patients (30 out of 762), with a strong correlation between the presence of *ALKm* and MNA ($P < .001$). Among 571 cases with known *ALKa* and *ALKm* status, a statistically significant difference in OS was observed between cases with *ALKa* or clonal *ALKm* versus subclonal *ALKm* or no *ALK* alterations (5-year OS: *ALKa* [$n = 19$], 26% [95% CI, 10 to 47], clonal *ALKm* [$n = 65$] 33% [95% CI, 21 to 44], subclonal *ALKm* ($n = 22$) 48% [95% CI, 26 to 67], and no alteration [$n = 465$], 51% [95% CI, 46 to 55], respectively; $P = .001$). Importantly, in a multivariate model, involvement of more than one metastatic compartment (hazard ratio [HR], 2.87; $P < .001$), *ALKa* (HR, 2.38; $P = .004$), and clonal *ALKm* (HR, 1.77; $P = .001$) were independent predictors of poor outcome.

CONCLUSION Genetic alterations of *ALK* (clonal mutations and amplifications) in HR-NB are independent predictors of poorer survival. These data provide a rationale for integration of *ALK* inhibitors in upfront treatment of HR-NB with *ALK* alterations.

J Clin Oncol 39:3377-3390. © 2021 by American Society of Clinical Oncology

Licensed under the Creative Commons Attribution 4.0 License 

INTRODUCTION

Neuroblastoma (NB), the most frequent solid, extra-cranial malignancy in children, exhibits wide clinical and genetic heterogeneity. High-risk neuroblastoma (HR-NB), defined as metastatic disease over the age of 12 months or *MYCN*-amplified (MNA) disease at any age, remains associated with long-term survival rates of only 50%.¹ Current treatment approaches consist of

intensive induction chemotherapy, surgical resection of the primary tumor, consolidation with high-dose chemotherapy (HDC), and autologous stem-cell rescue, and for minimal residual disease, isotretinoin in combination with human or mouse chimeric anti-GD₂ antibody, ch14.18.²⁻⁸

In NB, several recurrent genetic alterations have been described. MNA is a strong biomarker associated with

ASSOCIATED CONTENT

Appendix
Data Sharing Statement
Protocol

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on April 21, 2021 and published at ascopubs.org/journal/jco on June 11, 2021; DOI <https://doi.org/10.1200/JCO.21.00086>

CONTEXT

Key Objective

High risk neuroblastoma (HR-NB) is one of the most difficult childhood cancers to cure. This study examined whether the presence of an *ALK* alteration (amplification or mutation) was associated with a poor prognosis in a large patient series treated on the prospective European high-risk neuroblastoma trial (HR-NBL1).

Knowledge Generated

We found that *ALK* amplification or clonal mutation was associated with inferior prognosis in patients with HR-NB and both are independent prognostic variables on multivariate analysis. To our knowledge, this is the first study to report the highly prognostic significance of *ALK* amplification in HR-NB.

Relevance

As *ALK* can be targeted therapeutically, this study convincingly argues for the introduction of *ALK* inhibitors for upfront management of patients with HR-NB with *ALK* aberrations. Importantly, the prognostic significance of *ALK* alterations included a subgroup of trial patients treated with the current standard of care for HR-NB including anti-GD₂ immunotherapy.

rapid tumor growth.⁹ Other copy-number alterations occur over more extensive chromosome regions, with segmental chromosome alterations being associated with a poor outcome.¹⁰ Recurrent mutations have been described in the RAS-MAPK pathway, chromatin remodeling genes (*ATRX* and *ARID1A*), and *TERT* rearrangements.¹¹⁻¹⁴

Activating anaplastic lymphoma kinase (*ALK*) mutations are the most frequent mutations in NB, occurring in both familial and sporadic cases, with somatically acquired *ALK* mutations (*ALK*m) observed in 6%-12% of sporadic NBs in all risk groups.¹⁵⁻¹⁸

These *ALK* activating mutations are localized most frequently within the kinase domain at hotspots identified at the F1174, R1275, and F1245 positions, with mutations occurring both at clonal (> 20% mutated allele fraction [MAF]) or subclonal levels (< 20% MAF).¹⁹⁻²³

ALK can also be activated by genomic focal amplification, described in 1%-2% of NBs, almost exclusively with MNA,^{17,24} or, more rarely, following structural rearrangements.²⁵ Genetic alterations of *ALK* are associated with poorer survival in the overall NB population.^{24,26} However, their prognostic role in HR-NB has been less well studied.^{10,17,24} Altogether, *ALK* alterations are an important molecular target, given the role of *ALK* as a driver oncogene in NB and its actionability with small molecule therapies.²⁷⁻²⁹

To determine the frequency of *ALK* alterations (mutations and amplifications), their correlation with clinical characteristics, and their prognostic impact in HR-NB, we analyzed a large series of 1,092 diagnostic NB samples from patients on the HR-NBL1/SIOPEN trial.

MATERIALS AND METHODS

Patients and Samples

Patients were treated within the HR-NBL1/SIOPEN Protocol (ClinicalTrials.gov: [NCT01704716](https://clinicaltrials.gov/ct2/show/study/NCT01704716), EudraCT: 2006-

001489-17; Protocol [online only]), an international, randomized, multiarm, open-label, phase III trial.^{2-5,30,31} Patients with International Neuroblastoma Staging System stages 2, 3, 4, or 4S with MNA, or International Neuroblastoma Staging System stage 4 without MNA \geq 12 months of age at diagnosis were eligible for the trial up to 20 years of age. Within the trial, several randomized treatment arms were conducted over different periods (Appendix Fig A1, online only). Induction random assignments included the following: R0—random assignment of prophylactic granulocyte colony-stimulating factor during rapid COJEC induction³¹; R3—comparison of two induction regimens, rapid-COJEC versus modified N7.³² HDC was evaluated in the R1 random assignment: busulfan or melphalan versus carboplatin or etoposide or melphalan.³ Anti-GD₂ immunotherapy random assignments during maintenance phase were explored in R2 (2009-2013) and R4 (2014-2017), both comparing dinutuximab beta with oral isotretinoin to dinutuximab beta and subcutaneous interleukin-2 with oral isotretinoin, but with altered schedules.^{5,30} In the interim, dinutuximab beta with oral isotretinoin was the recommended standard.

Patients were enrolled on the HR-NBL1/SIOPEN trial after approval by national regulatory authorities and by national, and institutional, ethical committees or review boards in participating countries. Parents or guardians and patients according to age provided written informed consent for treatment, data collection, and analysis.

The *ALK* analysis cohort consisted of patients for whom a contributive tumor sample obtained at diagnosis was available in a SIOPEN reference laboratory³³ for additional molecular analysis with available follow-up data (Fig 1).

MYCN status and tumor genomic copy-number profiles were determined in SIOPEN reference laboratories as described previously.^{10,33-36} Samples were required to contain at least 20% tumor cells on pathologic examination.

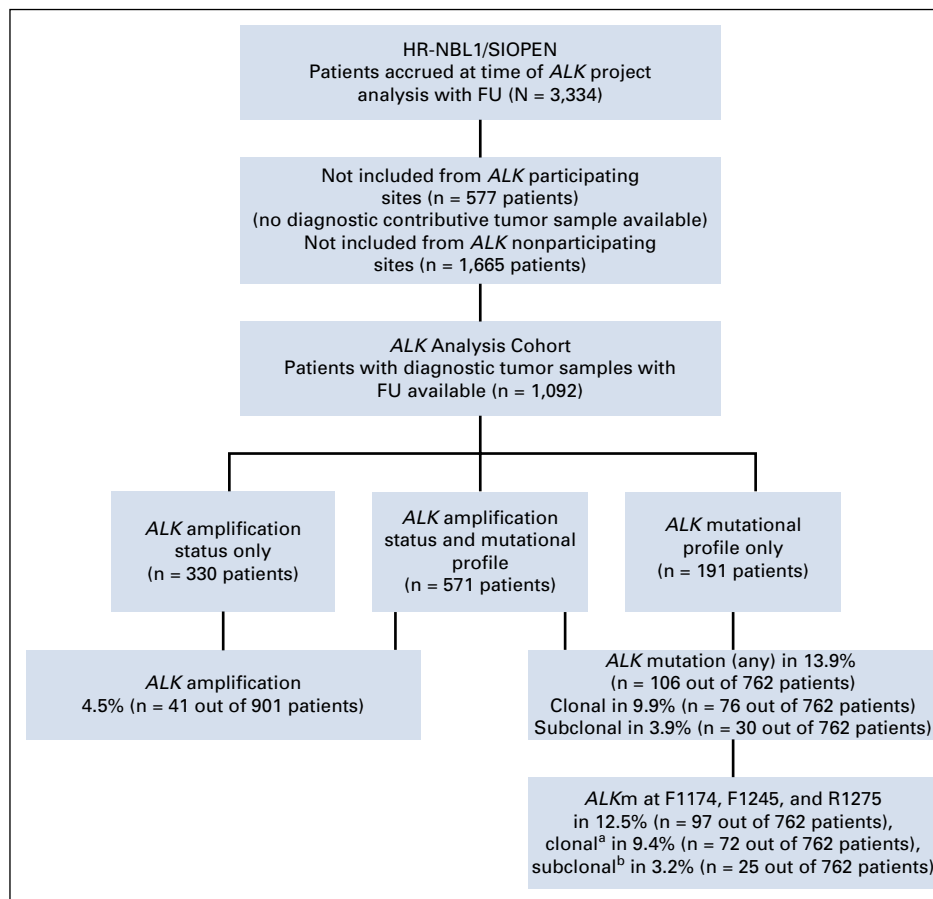


FIG 1. Flow diagram of patient inclusion. A total of 3,334 patients with HR-NB were enrolled in the HR-NBL1 trial from 188 centers. Among these, 2,350 patients were not included in this study, either because no contributive tumor material was available, or because there was no FU data, or both. Thus, 1,092 patients from 132 centers were included in this study. ^aClonal level: > 20% MAF. ^bSubclonal level: MAF 0.1%-20%. FU, follow-up; HR-NB, high-risk neuroblastoma; MAF, mutated allele fraction.

The *ALK* amplification (*ALK*_a) status was evaluated using either fluorescence in situ hybridization and/or multiplex ligation polymerase chain reaction–dependent amplification, array comparative genomic hybridization (aCGH), and/or array single-nucleotide polymorphism according to established guidelines.^{10,33,34,37} *ALK* gene amplification was defined as more than fourfold increase of *ALK* signals in relation to numbers of chromosome 2 by fluorescence in situ hybridization, or as more than 10 copies of the gene estimated by multiplex ligation–dependent amplification, aCGH, or array single-nucleotide polymorphism.

The *ALK* mutational (*ALK*_m) status was determined by Sanger sequencing, next-generation sequencing (NGS) techniques (coverage > 80×), targeted deep sequencing (TDS), or a combination of the latter techniques, covering the *ALK* regions of interest (exon 23: chr2:29443647-29443776; exon 24: chr2:29436830-29436935; exon 25: chr2:29432603-29432704; UCSC Genome Browser Home,³⁸ hg19) containing the *ALK* mutational hotspots

F1174 (exon 23), F1245 (exon 24), and R1275 (exon 25).^{20,22}

MAF ≥ 20% were defined as clonal events and MAF < 20% as subclonal events, as reported previously.^{20,22} No correction for tumor cell content was undertaken when reporting MAF. Mutations identified by Sanger sequencing were considered clonal. All detected mutations were validated by a second independent experiment: for clonal events, TDS data were validated by Sanger sequencing, and for subclonal events, NGS or TDS was validated in an independent second experiment.

Standard bioinformatics were used to detect mutations in NGS experiments as previously reported. Mutations in TDS experiments were determined as described previously.^{20,22} In brief, to highlight mutations, in each NB sample, the frequencies of each base at each position of the analyzed regions were compared with those observed in all other samples and controls. This approach enabled the identification of mutations with a statistically significant increase in percentage of a variant base, compared with background noise.

Statistical Analysis

Event-free survival (EFS) was calculated from diagnosis to the first relapse, progressive disease, secondary malignancy, or death from any cause, or until last patient contact. Overall survival (OS) was calculated from diagnosis to death from any cause, or until the last patient contact. EFS and OS were estimated using the Kaplan-Meier method and

compared using the logrank test, and if indicated with pseudo-value regression for 5-year OS.³⁹⁻⁴¹ EFS and OS are presented as 5-year point estimates together with 95% CIs using log-log transformation.⁴¹ To adjust for established risk-factors (age at diagnosis, stage, number of metastatic compartments, and *MYCN* amplification), a Cox proportional hazards regression model was used.

TABLE 1. Characteristics of Patients According to the *ALK* Amplification or *ALK* Mutation Status

Clinical Parameters	Known <i>ALK</i> Amplification Status (N = 901)					Known <i>ALK</i> Mutation Status (N = 762)						
	No		Yes		P	No Mutation		Clonal Mutation		Subclonal Mutation		P
	n	%	n	%		n	%	n	%	n	%	
Total	860	100	41	100		656	100	76	100	30	100	
Sex												
Female	376	44	16	39	.553	278	42	38	50	11	37	.348
Male	484	56	25	61		378	58	38	50	19	63	
Age, years												
< 1	51	6	7	17	.005	38	6	5	7	0	0	.348
1-1.5	101	12	9	22		79	12	15	20	3	10	
1.5-5	572	67	20	49		428	65	47	62	21	70	
> 5	136	16	5	12		111	17	9	12	6	20	
Stage												
Loc, MNA+	83	10	13	32	< .001	63	10	9	12	4	13	.890
Stage 4	768	89	26	63		586	89	66	87	26	87	
Stage 4s, MNA+	9	1	2	5		7	1	1	1	0	0	
MYCN status												
MNA-	466	54	1	2	< .001	365	56	26	34	9	30	< .001
MNA+	394	46	40	98		291	44	50	66	21	70	
Primary tumor site												
Unknown	20		1		.362	21		1		1		.278
Abdominal adrenal±	606	72	25	63		452	71	47	63	22	76	
Abdominal other±	169	20	10	25		124	20	22	29	6	21	
Other only	65	8	5	13		59	9	6	8	1	3	
Stage 4: MYCN status												
MNA-	466	61	1	4	< .001	365	62	26	39	9	35	< .001
MNA+	302	39	25	96		221	38	40	61	17	65	
Stage 4: MC												
1 MC	91	12	1	4	.091	70	13	11	17	4	17	.788
2 MC	231	32	12	52		177	32	19	29	9	38	
> 2 MC	411	56	10	43		302	55	35	54	11	46	
Overall response: end of induction												
Evaluable	804		39			607		72		28		
CR or VGPR or PR	628	78	31	79	.839	472	78	53	74	24	86	.421
MR or SD or PD	176	22	8	21		135	22	19	26	4	14	

NOTE. Patients studied for *ALK* amplifications (n = 901) and *ALK* mutations (n = 762).

Abbreviations: CR, complete response; MC, metastatic compartments; MNA, *MYCN* amplification; MR, minor response; PD, progressive disease; PR, partial response; SD, stable disease; VGPR, very good partial response.

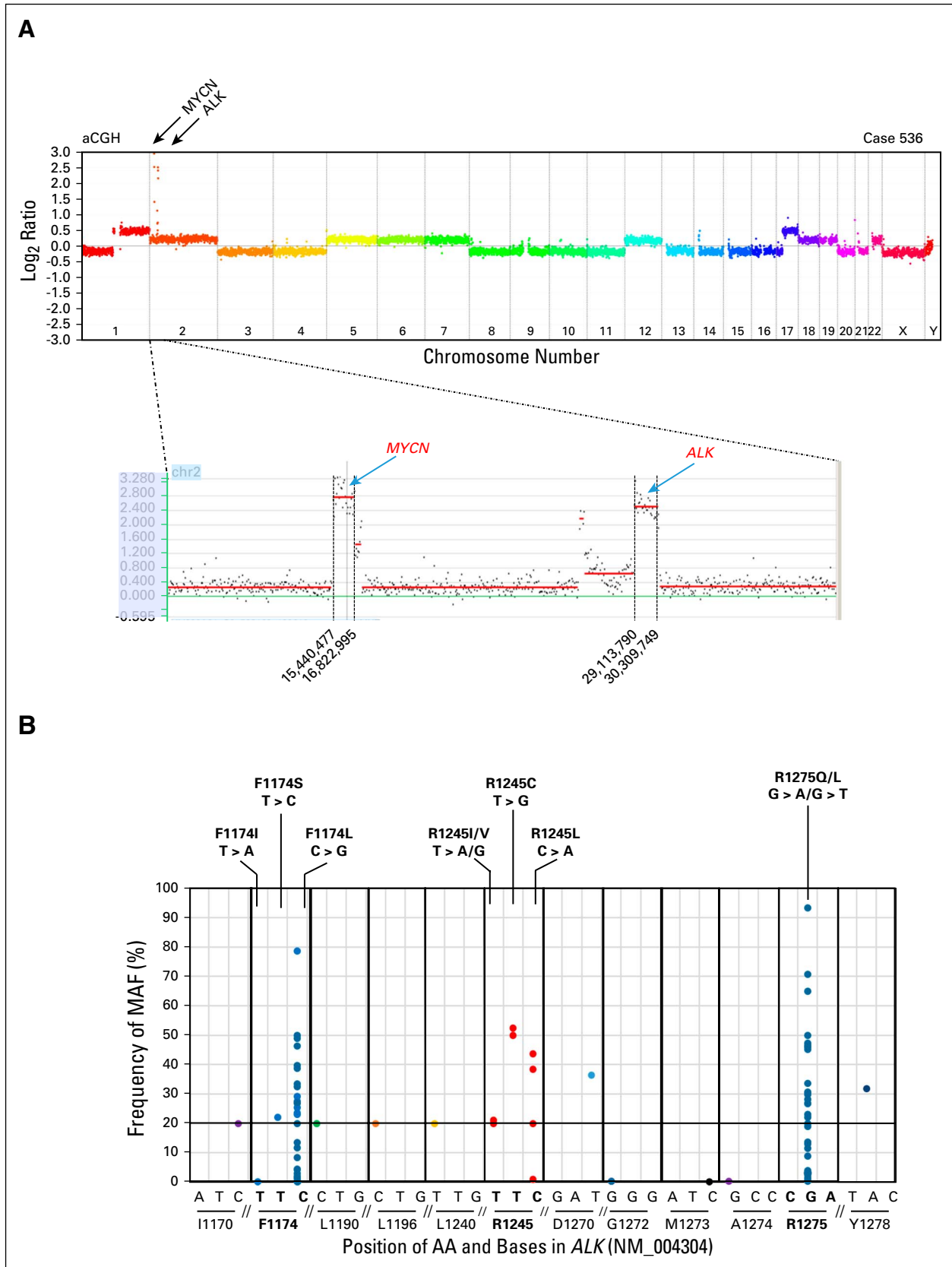


FIG 2. Genetic alterations of *ALK* in patients with HR-NB. (A) Copy-number profile of case 536. Genomic coamplification of *MYCN* and *ALK* is observed on chromosome 2, encompassing the regions between position 15,440,477 and 16,822,999 and between 29,113,790 and 30,309,749 bp (human genome assembly hg19; UCSC Genome Browser Home³⁸). (B) Frequency distribution (continued on following page)

FIG 2. (Continued). of mutated *ALK* alleles at the studied chromosome regions, encompassing the AA positions F1174, L1190, L1196, R1245, D1270, G1272, M1273, A1274, R1275, and Y1278 detected, in 762 samples. *ALK* mutations involved the common mutational hotspots (F1174, F1245, and R1275) in 12.5% (97 out of 772) of cases, at a clonal level (MAF 20%-93%) in 72 cases, and at a subclonal level (MAF < 20%) in 25 cases. At the F1174 hotspot (chr2: 29,443,695-29,443,697), alterations were observed in 44 cases: 42 cases harbored a mutation leading to the AA change F1174L, one case with F1174I, and one case with F1174S, with MAFs ranging from 0.12% to 78%. At the R1275 hotspot (chr2: 29,432,849-29,430,139), mutations were detected in 43 cases: 38 cases harbored a mutation leading to the AA change R1275Q and five cases with R1275L, with the MAFs ranging from 0.2% to 93%. Ten cases showed *ALK* mutations at the F1245 hotspot (chr2: 29,436,858-29,436,860) within exon 24. Three samples showed the F1245L mutation, three cases carried the F1245C mutation, three showed the F1245I mutation, and one showed mutation F1245V mutation (Fig 1 and Appendix Table A1). Other *ALK* mutations were detected at residues I1170, L1190 (two cases), L1196, D1270, G1272, M1273, A1274, and Y1278 within the explored regions, leading to a nonsynonymous AA change with a predicted functional impact. All these mutations were clonal (MAF > 20%) except for M1273I (MAF 0.2%) and I1170 (MAF 2.8%). AA, amino acid; aCGH, array comparative genomic hybridization; bp, base pair; HR-NB, high-risk neuroblastoma; MAF, mutated allele fraction; UCSC, University of California, Santa Cruz.

Correlations between patient and disease characteristics and *ALK* genetic alterations were explored using chi-square tests.

To allow for sufficient follow-up time, only patients enrolled until December 31, 2019, were considered. The data cutoff for the final analysis was October 3, 2020. We calculated median follow-up using the inverse Kaplan-Meier estimate. Statistical analysis was performed using SAS (version 9.4).

RESULTS

Of 3,334 patients enrolled on the HR-NBL1/SIOPEN trial between November 24, 2002, and December 31, 2019, 1,092 patients were included in the *ALK* analysis cohort (Fig 1; Appendix Table A1, online only). Patients were accrued from 132 SIOPEN member institutions or hospitals in 19 countries (Appendix Table A2, online only). Among these 1,092 patients, 81% (889 out of 1,092) were > 18 months of age at diagnosis, 47% (521 out of 1,092) showed MNA, and 88% (966 out of 1,092) had stage 4 disease, with no statistically significant difference in EFS or OS between the *ALK* analysis cohort and the overall HR-NBL1 cohort (Appendix Fig A2, online only).⁴² The median follow-up period was 6.8 years (0.1-17.4 years).

ALK Alterations

Within the *ALK* cohort, the *ALK*m status was analyzed in 762 patients, the *ALK*a status in 901 cases, with both *ALK*m and *ALK*a studied in 571 patients (Fig 1, Table 1).

ALK alterations were detected in 146 out of 1,092 patients with *ALK*a occurring in 4.5% (41 out of 901 cases) and *ALK*m in 13.9% (106 out of 762 cases). Only one case showed *ALK*a and a concomitant *ALK* R1275Q mutation with an MAF of 93%, suggesting that the mutated allele is contained in the amplicon (Appendix Fig A3, online only).

ALK Amplification and Correlation With Risk Factors

High-level genomic amplification of the *ALK* gene was found in 4.5% (41 out of 901) of cases (Fig 2A, Table 1). All but one also had MNA. *ALK*a significantly correlated with MNA ($P < .001$), non-stage 4 disease ($P < .001$), and age at diagnosis < 18 months ($P = .005$). No correlation between the presence of *ALK*a and response at the end of induction treatment was observed.

A statistically significant poorer 5-year OS was observed in patients whose tumors harbored *ALK*a (5-year OS: *ALK*a 28% [95% CI, 15 to 42] v non-*ALK*a 51% [95% CI, 47 to 54]; $P < .0001$; Fig 3A, Table 2) with a stronger prognostic effect in patients with stage 4 or 4S MNA.

ALK Mutation and Correlation With Risk Factors

ALK mutational status was studied in 762 cases by Sanger sequencing ($n = 163$), by NGS techniques ($n = 13$), or by TDS ($n = 650$, including 64 by TDS and Sanger). The biologic data for 52 cases have been reported previously.²²

Among these, 13.9% (106 out of 762) showed at least one *ALK*m within the explored *ALK* regions of interest, with 10% (76 out of 762) harboring mutations at a clonal level (MAF > 20%) and 3.9% (30 out of 762) at a subclonal level (MAF ≤ 20%): nine cases—MAF 0.1% to < 1%, 10 cases MAF 1% to < 5%, two cases MAF 5% to < 10%, and nine cases MAF 10% to < 20% (Figs 1 and 2B; Table 1).

Concordance between results analyzed by two different techniques was observed in 64 cases with clonal *ALK*m (TDS and Sanger). Subclonal *ALK*m were validated by a second independent TDS experiment, with an excellent correlation of MAF between the two experiments ($R^2 = 0.9924$; $P < .0001$) (Appendix Fig A4, online only).

*ALK*m involved the common mutational hotspots (F1174, F1245, and R1275) in 12.5% (97 out of 762) of cases, comprising 91% (97 out of 106) of all detected *ALK*m (Fig 2B).

Interestingly, three cases harbored two or more distinct mutations. In the first case, both F1174L and F1245L mutations were observed (MAF 2% and 0.8%, respectively). The second case showed three subclonal mutations F1174L, R1275Q, and R1275L (MAF 2.9%, 8.9%, and 2.9%, respectively). A third case harbored a mutation at the F1174 and R1275 hotspots (MAF 27% and 1.3%, respectively).

There were no statistically significant correlations between *ALK*m and stage, age at diagnosis, or localization of the primary tumor (adrenal, abdominal, or other) (Table 1). However, a significant correlation was observed between the presence of an *ALK*m and MNA ($P < .001$), with an enrichment of *ALK*m F1174 in MNA tumors ($P = .0005$).

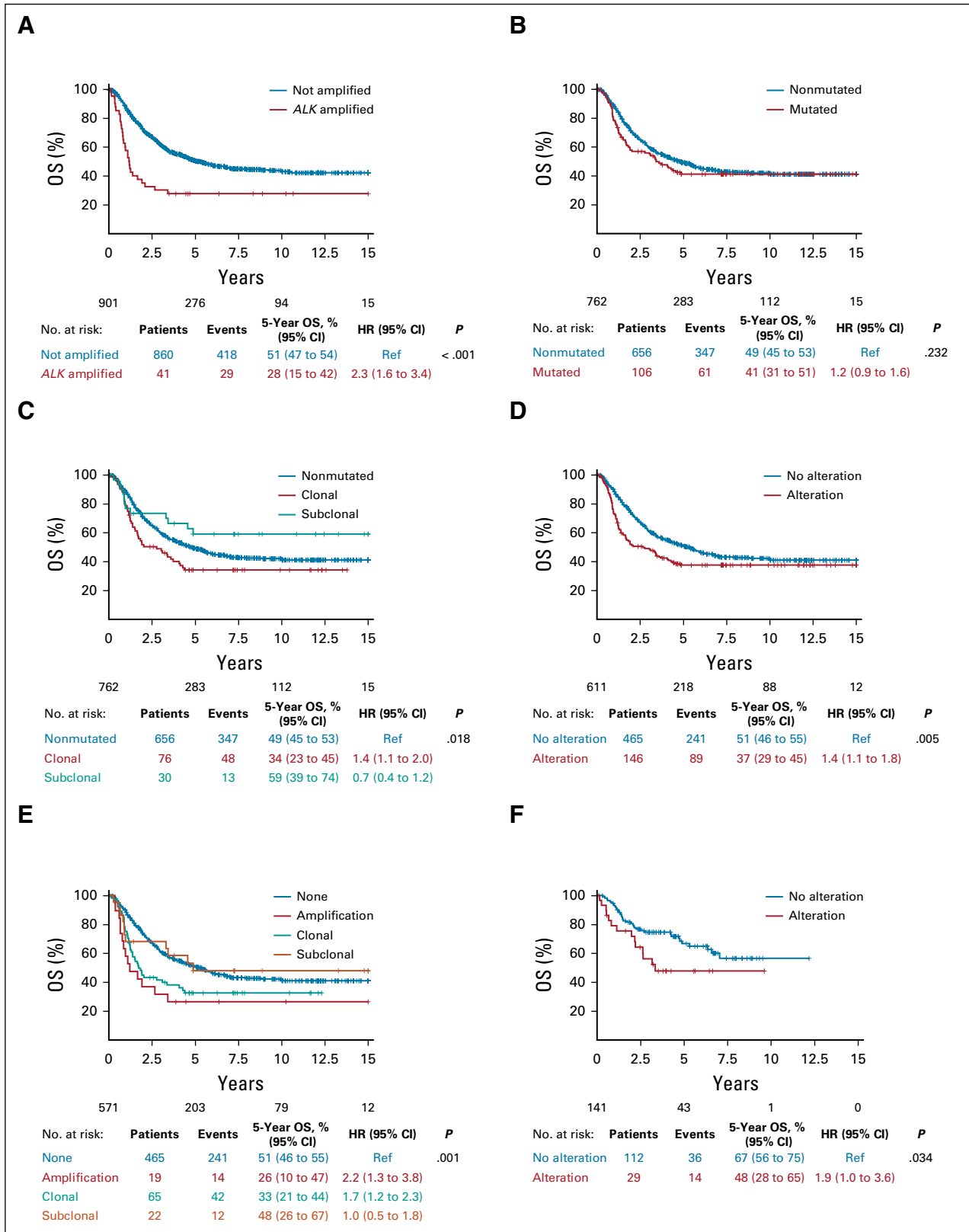


FIG 3. Survival in the ALK analysis cohort. (A) OS according to ALK amplification status in 901 patients: presence of ALK amplification (n = 41), 5-year OS 28% (95% CI, 15 to 42) versus absence of ALK amplification (n = 860), 5-year OS 51% (95% CI, 47 to 54); P < .0001. (B) OS according to ALK mutation status in 762 patients: presence of an ALK mutation (n = 106), 5-year OS 41% (95% CI, 31 to 51) versus absence of an ALK mutation (n = 656), 5-year OS 49% (95% CI, 45 to 53); P = NS. (C) OS according to ALK clonal or subclonal (continued on following page)

FIG 3. (Continued). mutation status in 762 patients: no mutation (n = 656), 5-year OS 49% (95% CI, 45 to 53); clonal mutations (n = 76), 5-year OS 34% (95% CI, 23 to 45); and subclonal mutations (n = 30), 5-year OS 59% (95% CI, 39 to 74), respectively; $P = .018$. (D) OS according to the presence of any *ALK* alterations in 611 patients with known *ALK* amplification and *ALK* mutation status: presence of an *ALK* alteration (n = 146), 5-year OS 37% (95% CI, 29 to 45); versus absence of *ALK* alterations (n = 465), 5-year OS 51% (95% CI, 46 to 55); $P = .005$. (E) OS according to the type of *ALK* alteration in the cohort of 571 patients with known *ALK* amplification and *ALK* mutation status: no alteration (n = 465), 5-year OS 51% (95% CI, 46 to 55); clonal mutations (n = 65), 5-year OS 33% (95% CI, 21 to 44); subclonal mutations (n = 12), 5-year OS 48% (95% CI, 26 to 67); and *ALK* amplification (n = 19), 5-year OS 26% (95% CI, 10 to 47), respectively; $P = .001$. (F) OS according to *ALK* alterations (*ALK* amplification or clonal *ALK* mutation) in patients who received immunotherapy (n = 141): To evaluate the impact of *ALK* alterations (*ALK* amplification or clonal *ALK* mutation) in patients who received dinutuximab beta, OS was calculated from the start of dinutuximab beta treatment and evaluated using the same approaches as described in the Materials and Methods section. *ALK* alteration (*ALK* amplification or clonal *ALK* mutation, n = 29, 5-year OS 48% [95% CI, 28 to 65]) versus no *ALK* alteration (n = 112) 67% (95% CI, 56 to 75); $P = .034$. Patient details: Appendix Table A3. HR, hazard ratio; NS, not significant; OS, overall survival; ref, reference.

This was also observed when analyzing only stage 4 tumors. No correlation between *ALKm* and response at the end of induction treatment was observed.

No statistically significant difference in outcome was observed between patients harboring any *ALKm* versus none (Fig 3B, Table 2). However, when distinguishing clonal and subclonal mutations, a poorer OS was observed only in patients with clonal *ALKm*, as opposed to subclonal or no mutations (5-year OS, clonal *ALKm* 34% [95% CI, 23 to 45], subclonal *ALKm* 59% [95% CI, 39 to 74], and no *ALKm* 49% [95% CI, 45 to 53]; $P = .018$) (Fig 3C, Table 2).

Patients with metastatic disease (stage 4 or 4S MNA) and a clonal *ALKm* showed a trend toward poorer OS. However, in patients with localized disease, the presence of *ALKm* did not confer poorer survival (Table 2).

Overall Prognostic Impact of *ALK* Genetic Alterations

To determine the overall prognostic impact of *ALK* genetic alterations, we focused on the subgroup of 571 patients with both known *ALKa* and *ALKm* status. In this subgroup of patients, a statistically significant poorer OS was observed in patients whose tumors harbored any *ALK* alteration (5-year OS, any alteration 37% [95% CI, 29 to 45] v no alteration 51% [95% CI, 46 to 55]; $P = .005$; Fig 3D). *ALKa* or clonal *ALKm* were associated with a poorer outcome (5-year OS, *ALKa* 26% [95% CI, 10 to 47], clonal *ALKm* 33% [95% CI, 21 to 44], subclonal *ALKm* 48% [95% CI, 26 to 67], and no *ALK* alteration 51% [95% CI, 46 to 55]; $P = .001$; Fig 3E, Table 2).

Among the subgroup of patients with known *ALK* status, we sought to determine the prognostic impact of *ALK* alterations according to the different treatment arms of HR-NBL1. Indeed, in the HR-NBL01/SIOPEN trial, the introduction of busulfan and melphalan as standard for HDC, and anti-GD₂ maintenance therapy as a new standard since 2010, has led to significantly improved survival (Appendix Fig A5F, online only).³⁻⁵ Importantly, when considering patients treated according to the SIOPEN standard with busulfan and melphalan HDC and maintenance immunotherapy, the presence of an *ALK* alteration (*ALKa* or clonal *ALKm*) remained associated with a poorer 5-year OS of 48% (95% CI, 28 to 65), versus no *ALK* alteration 67% (95% CI, 56 to 75); $P = .03$ (Fig 3F, Appendix Table A3,

online only), with a trend also observed when taking into account all *ALKm* (clonal and subclonal, $P = .059$).

Based on univariate risk factor exploration of the whole *ALK* analysis cohort (Appendix Fig A5), we developed a Cox model for multivariate analysis including clinical and biologic parameters previously shown to be of prognostic impact (n = 571 patients). Involvement of two or more metastatic compartments (OS: hazard ratio [HR], 2.87 [95% CI, 1.73 to 4.78]; $P = .001$) and the presence of *ALKa* (OS: HR, 2.38 [95% CI, 1.32 to 4.27]; $P = .004$) and clonal *ALKm* (OS: HR, 1.77 [95% CI, 1.25 to 2.49]; $P = .001$) were of independent prognostic significance, whereas *MNA* and age were not (Table 3).

DISCUSSION

In HR-NB, the identification of prognostic biomarkers is crucial for the development of new treatment approaches. Recent studies have shown that *MNA* is not associated with poorer outcome among the overall cohort of patients with HR-NB, but the presence of genomic amplifications other than *MYCN* might constitute a poor outcome biomarker.⁴³ We now show in this large *ALK* analysis cohort that the presence of *ALKa* or clonal *ALKm* resulted in significantly worse outcome.

Given the oncogenic driver role of *ALK* activation, and the prognostic impact of *ALKa* or clonal *ALKm*, the introduction of frontline *ALK*-targeted treatment is now strongly supported by the current study. Although early phase clinical trials of first- and second-generation *ALK* inhibitors showed modest efficacy of the first-generation inhibitor crizotinib in NB with F1174 hotspot mutations being resistant,⁴⁴ third-generation *ALK* inhibitors such as lorlatinib exhibit improved efficacy alone and when combined with chemotherapy.^{28,44-46} Crizotinib is currently being administered with chemotherapy in a phase III upfront trial for patients with HR-NB with *ALK* alterations (ClinicalTrials.gov: NCT03126916).

Improvements in HR-NB patient survival have been achieved with intensification of HDC and immunotherapy with dinutuximab (ch14.18/Sp02 and ch14.18/CHO),^{3-5,7} and our results highlight the potential of *ALK* inhibition as an attractive upfront precision-medicine strategy in patients with *ALK* alterations to further improve survival. Importantly, in patients reaching the maintenance treatment phase

TABLE 2. EFS and OS According to ALK Alterations

Parameters	OS					EFS				
	Patients, No.	Events, No.	5-Year OS, % (95% CI)	HR (95% CI)	P	Patients, No.	Events, No.	5-Year EFS, % (95% CI)	HR (95% CI)	P
Total										
<i>ALKa</i>										
No	860	418	51 (47 to 54)	Ref	< .001	860	492	40 (36 to 43)	Ref	< .001
Yes	41	29	28 (15 to 42)	2.3 (1.6 to 3.4)		41	31	24 (13 to 38)	2.0 (1.4 to 2.9)	
<i>ALKm</i>										
Nonmutated	656	347	49 (45 to 53)	Ref	.018	656	395	38 (35 to 42)	Ref	.081
<i>ALKm</i> clonal	76	48	34 (23 to 45)	1.4 (1.1 to 2.0)		76	51	31 (21 to 42)	1.3 (1.0 to 1.7)	
<i>ALKm</i> subclonal	30	13	59 (39 to 74)	0.7 (0.4 to 1.2)		30	16	49 (30 to 65)	0.8 (0.5 to 1.3)	
Known <i>ALK</i> alteration status										
Nonmutated	465	241	51 (46 to 55)	Ref	.001	465	280	38 (33 to 43)	Ref	.057
<i>ALKa</i>	19	14	26 (10 to 47)	2.2 (1.3 to 3.8)		19	14	26 (10 to 47)	1.7 (1.0 to 2.9)	
<i>ALKm</i> clonal	65	42	33 (21 to 44)	1.7 (1.2 to 2.3)		65	43	33 (22 to 44)	1.4 (1.0 to 1.9)	
<i>ALKm</i> subclonal	22	12	48 (26 to 67)	1.0 (0.5 to 1.8)		22	14	39 (19 to 59)	1.0 (0.6 to 1.8)	
Stage 4, 4s										
<i>ALKa</i>										
No	777	394	48 (44 to 52)	Ref	< .001	777	467	37 (33 to 40)	Ref	< .001
Yes	28	22	19 (7 to 35)	2.9 (1.8 to 4.6)		28	23	18 (7 to 34)	2.9 (1.8 to 4.6)	
<i>ALKm</i>										
Nonmutated	593	328	47 (43 to 51)	Ref	.068	593	375	35 (31 to 39)	Ref	.216
<i>ALKm</i> clonal	67	43	33 (22 to 45)	1.4 (1.0 to 1.9)		67	46	30 (19 to 41)	1.4 (1.0 to 1.9)	
<i>ALKm</i> subclonal	26	13	52 (31 to 70)	0.8 (0.4 to 1.4)		26	16	41 (22 to 59)	0.8 (0.4 to 1.4)	
Known <i>ALK</i> alteration status										
Nonmutated	419	228	48 (43 to 53)	Ref	.000	419	266	35 (30 to 39)	Ref	.042
<i>ALKa</i>	15	12	20 (5 to 42)	2.6 (1.3 to 4.7)		15	12	20 (5 to 42)	1.8 (1.0 to 3.4)	
<i>ALKm</i> clonal	57	38	30 (18 to 43)	1.7 (1.2 to 2.4)		57	39	30 (19 to 42)	1.4 (1.0 to 1.9)	
<i>ALKm</i> subclonal	21	12	45 (23 to 65)	1.0 (0.5 to 1.8)		21	14	36 (16 to 56)	1.0 (0.6 to 1.8)	
Stage 4, MNA–										
<i>ALKa</i>										
No	466	236	49 (44 to 54)	NA		466	292	33 (28 to 38)	NA	
Yes	1	1	NA	NA		1	1	NA	NA	
<i>ALKm</i>										
Nonmutated	365	202	49 (43 to 54)	Ref	.202	365	238	33 (28 to 38)	Ref	.245
<i>ALKm</i> clonal	26	18	28 (13 to 46)	1.5 (0.9 to 2.5)		26	20	23 (9 to 40)	1.5 (0.9 to 2.3)	
<i>ALKm</i> subclonal	9	4	53 (18 to 80)	0.9 (0.3 to 2.3)		9	5	42 (11 to 71)	0.9 (0.4 to 2.3)	
Known <i>ALK</i> alteration status										
Nonmutated	269	146	50 (43 to 56)	Ref	.010	269	174	32 (27 to 38)	Ref	.029
<i>ALKa</i>	1	1	NA	NA		1	1	NA	NA	
<i>ALKm</i> clonal	20	15	22 (7 to 42)	2.1 (1.3 to 3.6)		20	16	20 (6 to 39)	1.8 (1.1 to 2.9)	
<i>ALKm</i> subclonal	6	3	44 (7 to 78)	1.2 (0.4 to 3.7)		6	4	25 (1 to 65)	1.4 (0.5 to 3.9)	

(continued on following page)

TABLE 2. EFS and OS According to *ALK* Alterations (continued)

Parameters	OS					EFS				
	Patients, No.	Events, No.	5-Year OS, % (95% CI)	HR (95% CI)	<i>P</i>	Patients, No.	Events, No.	5-Year EFS, % (95% CI)	HR (95% CI)	<i>P</i>
Stage 4, 4s MNA+										
<i>ALKa</i>										
No	311	158	48 (42 to 54)	Ref	< .001	311	175	43 (37 to 48)	Ref	< .001
Yes	27	21	19 (7 to 36)	2.3 (1.4 to 3.7)		27	22	19 (7 to 35)	2.0 (1.3 to 3.3)	
<i>ALKm</i>										
Nonmutated	228	126	44 (38 to 51)	Ref	.453	228	137	40 (33 to 46)	Ref	.666
<i>ALKm</i> clonal	41	25	37 (22 to 51)	1.2 (0.8 to 1.8)		41	26	34 (20 to 49)	1.2 (0.8 to 1.8)	
<i>ALKm</i> subclonal	17	9	52 (27 to 73)	0.8 (0.4 to 1.5)		17	11	41 (19 to 63)	0.9 (0.5 to 1.7)	
Known <i>ALK</i> alteration status										
Nonmutated	150	82	46 (37 to 54)	Ref	.085	150	92	39 (31 to 47)	Ref	.372
<i>ALKa</i>	14	11	21 (5 to 45)	1.9 (1.0 to 3.7)		14	11	21 (5 to 45)	1.6 (0.8 to 3.0)	
<i>ALKm</i> clonal	37	23	35 (20 to 51)	1.3 (0.8 to 2.1)		37	23	36 (20 to 51)	1.2 (0.7 to 1.9)	
<i>ALKm</i> subclonal	15	9	46 (20 to 68)	0.9 (0.4 to 1.8)		15	10	40 (16 to 63)	1.0 (0.5 to 1.9)	
Localized, MNA+										
<i>ALKa</i>										
No	83	24	71 (59 to 80)	Ref	.059	83	25	69 (57 to 78)	Ref	.015
Yes	13	7	46 (19 to 70)	2.2 (0.9 to 5.1)		13	8	38 (14 to 63)	2.6 (1.2 to 5.8)	
<i>ALKm</i>										
Nonmutated	63	19	70 (57 to 80)	Ref	.114	63	20	67 (54 to 77)	Ref	.098
<i>ALKm</i> clonal	9	5	42 (11 to 71)	2.2 (0.8 to 5.8)		9	5	42 (11 to 71)	2.2 (0.8 to 5.9)	
<i>ALKm</i> subclonal	4	0	NA	NA		4	0	NA		
Known <i>ALK</i> alteration status										
Nonmutated	46	13	73 (57 to 83)	Ref	.440	46	14	68 (52 to 80)	Ref	.410
<i>ALKa</i>	4	2	50 (6 to 84)	2.0 (0.4 to 8.7)		4	2	50 (6 to 84)	1.8 (0.4 to 7.9)	
<i>ALKm</i> clonal	8	4	50 (15 to 77)	2.1 (0.7 to 6.5)		8	4	50 (15 to 77)	2.2 (0.7 to 6.8)	
<i>ALKm</i> subclonal	1	0	NA	NA		1	0	NA		

NOTE. EFS and OS in the *ALK* analysis cohort, according to different clinical parameters: complete summary of all risk-factor–based 5-year EFS and OS rates in patients according to the *ALK* amplification status (*ALKa*, *n* = 901 patients), *ALK* mutational status (*ALKm*, *n* = 762 patients), or in patients for whom both the *ALKa* status and *ALKm* status are known (known *ALK* alteration status, *n* = 571).

Abbreviations: EFS, event-free survival; MNA, MYCN-amplified; NA, not available; OS, overall survival; ref, reference.

with dinutuximab beta in the HR-NBL1/SIOPEN trial, the presence of an *ALK* alteration was still associated with poorer survival, thus strongly suggesting that integration of *ALK*-targeted therapy is warranted throughout all treatment phases of modern-era HR-NB therapy.

ALKa was observed in 4% of NB cases, accounting for approximately 1 out of 3 of *ALK*-activated NB cases. To date, co-occurrence of *ALK* hotspot mutations and genomic amplification has rarely been reported in NB.¹⁷ In this extensive cohort of patients, one case harboring both *ALKa* and an R1275 *ALKm* was identified. This indicates that these alterations are not fully mutually exclusive, although co-occurrence is extremely rare.

ALKm were found in 13.9% of cases at the studied exonic regions harboring known *ALK* mutational hotspots.^{17,24} This is higher than previously reported frequencies of *ALKm* in HR-NB of approximately 10%, most likely as previous reports using Sanger sequencing or standard-resolution NGS approaches.^{24,26} Sanger sensitivity is limited to the detection of MAF > 15%-20%, but in NB, *ALK* mutations with lower MAFs have been reported.^{14,19-21}

Ultradeep sequencing used in this analysis has a sensitivity limit of MAF of 0.1%.^{19,20} This approaches the theoretical limit of detection based on the genomic DNA input of 50 ng for one experiment, equivalent to 5,000 diploid genomes.

TABLE 3. Multivariate Analysis in 571 Patients With a Known *ALK* Amplification and *ALK* Mutation Status

Clinical Parameters	OS			EFS		
	P	HR	95% CI	P	HR	95% CI
Age, years						
< 1		1.00			1.00	
1-1.5	.269	0.72	0.40 to 1.30	.636	0.87	0.49 to 1.56
1.5-5	.265	0.75	0.45 to 1.24	.830	0.95	0.57 to 1.56
> 5	.662	0.88	0.50 to 1.55	.935	1.02	0.59 to 1.78
Metastatic compartments						
Localized-none		1.00			1.00	
1 MC	.122	1.60	0.88 to 2.90	.096	1.63	0.92 to 2.88
2 MC	.001	2.41	1.44 to 4.04	.001	2.38	1.44 to 3.94
> 2 MC	< .0001	2.87	1.73 to 4.78	< .0001	2.88	1.76 to 4.72
MYCN amplification						
MNA+	.135	1.23	0.94 to 1.62	.797	1.03	0.80 to 1.34
<i>ALK</i> alteration						
No alteration		1.00				1.00
ALKa	.004	2.38	1.32 to 4.27	.026	1.94	1.08 to 3.47
<i>ALKm</i> clonal	.001	1.77	1.25 to 2.49	.017	1.50	1.08 to 2.10
<i>ALKm</i> subclonal	.696	0.88	0.46 to 1.68	.934	1.02	0.58 to 1.81

Abbreviations: EFS, event-free survival; MC, metastatic compartments; MNA, MYCN-amplified; OS, overall survival.

This study demonstrates that use of higher-resolution techniques enables a higher detection rate of *ALKm*. The MAF distribution indicated a majority of clonal events (76 out of 106 cases). Importantly, clonal *ALKm* were associated with poorer outcome and were of independent prognostic significance, but subclonal events were not. Subclonal events, defined in this study by MAF < 20%, comprised 28% (30 out of 106) of all *ALKm*, with a very low MAF (< 5%) observed in 19 cases.

However, when considering *ALKm*, the OS remains poor in all patient subgroups (5-year OS < 62%). Furthermore, although of different prognostic impact in this study, the biomarker (*ALK* mutation) might not be of distinct predictive impact, and even in patients with subclonal *ALK* mutations, *ALK* inhibitor treatment might be effective in the targeted cell population. Thus, future upfront trials should consider *ALK*-targeted treatment based on clinically applicable reliable detection limits (for instance MAF 5% for NGS techniques) rather than the MAF defining prognostic subgroups.

As tumor samples harbored at least 20% tumor cells by pathologic examination, with additional confirmation provided by a dynamic aCGH or SNP α profile in the majority of cases, the observed low MAF is likely to correspond to intratumoral heterogeneity. In NB, intratumor heterogeneity has been reported for MNA and segmental chromosome alterations.⁴⁷⁻⁴⁹ The coexistence of *ALK* nonmutated and mutated cells within a single tumor suggests that these different subclones might coexist in an advantageous

equilibrium, which might crucially affect the dynamics of cancer progression.^{50,51} Correlation with pathologic findings, single-cell RNA or DNA experiments, and in situ approaches will elucidate how *ALK*-mutated cells are distributed throughout an NB. A higher frequency of *ALKm* at NB relapse has been demonstrated, suggesting clonal evolution of a minor *ALK*-mutated subclone to a dominant *ALK* mutated clone at relapse, but these cases might not represent clinically unfavorable cases initially.^{23,52,53} Further studies focusing on serial blood samples for ctDNA studies will further elucidate clonal evolution, also under targeted therapy.⁵⁴

In HR-NB, mutations in the p53 or RAS-MAPK pathways, including *ALK*, together with telomere maintenance caused by induction of telomerase or ALT (alternative lengthening of telomere) are thought to increase tumor aggressiveness, resulting in even poorer survival among patients with HR-NB.^{55,56} As *MYCN* leads to upregulation of *TERT* expression, MNA associated with any *ALK* alteration might lead to inferior outcome. Cases with *ALKa* show both *ALK* pathway activation and activation of telomere maintenance through MNA, with a suggested additive effect of these genetic events. The very poor survival of *ALKa* patients is concordant with this observation. However, survival of patients whose tumors harbored *ALKm* and MNA was not different from those without MNA, suggesting that *ALKm* cases constitute a more heterogeneous group with regards to the mechanistic tumor classification.⁵⁵

ALKa and *ALK* clonal mutation were both independent predictors of poor outcome in our multivariate Cox model. Notably, the end-of-induction response rate was not associated with *ALK* genetic alterations, suggesting that *ALK*-altered tumor cells are unlikely to be primarily chemotherapy resistant.

In summary, our data contribute to the rationale for future clinical trials introducing *ALK*-targeted treatment in the frontline setting together with chemotherapy and immunotherapy, and the distinct prognostic impact of different *ALK* alterations (*ALKa* and *ALKm*) needs to be considered.

AFFILIATIONS

¹Equipe SiRIC RTOP Recherche Translationnelle en Oncologie Pédiatrique, Institut Curie, Paris, France

²INSERM U830, Laboratoire de Génétique et Biologie des Cancers, Institut Curie, Paris, France

³SIREDO: Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer, Institut Curie, Paris, France

⁴Department for Studies and Statistics and Integrated Research, Vienna, Austria

⁵St Anna Children's Cancer Research Institute, Vienna, Austria

⁶Institut Curie Genomics of Excellence (ICGex) Platform, Research Center, Institut Curie, Paris, France

⁷Unité de Génétique Somatique, Service de Génétique, Hospital Group, Institut Curie, Paris, France

⁸Service de Génétique des tumeurs; Institut Gustave Roussy, Villejuif, France

⁹Department of Pathology, Oslo University Hospital, and Medical Faculty, University of Oslo, Oslo, Norway

¹⁰Department of Clinical Genetics, Children's Health Ireland at Crumlin, Dublin, Ireland

¹¹Northern Genetics Service, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom

¹²Ghent University, Ghent, Belgium

¹³Translational Research Laboratory, Centre Léon Bérard, Lyon, France

¹⁴Instituto de Investigación Sanitaria La Fe, Valencia, Spain

¹⁵Division of Clinical Studies, The Institute of Cancer Research, London, United Kingdom

¹⁶Schneider Children's Medical Center of Israel, Tel Aviv University, Tel Aviv, Israel

¹⁷Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon, Portugal

¹⁸Sahlgrenska University Hospital, Göteborg, Sweden

¹⁹Department of Pathology, IRCCS Istituto Giannina Gaslini, Genova, Italy

²⁰Laboratory of Molecular Biology, IRCCS Istituto Giannina Gaslini, Genova, Italy

²¹Pediatric Hematology-Oncology Research Laboratory, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

²²Department of Pathology, Medical School, University of Valencia-Incliva Health Research Institute/CIBERONC, Madrid, Spain

²³Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

²⁴Department of Paediatric Haematology and Oncology, Second Faculty of Medicine, Charles University and University Hospital Motol, Prague, Czech Republic

²⁵Paediatric Tumour Biology, Division of Clinical Studies, The Institute of Cancer Research, Sutton, United Kingdom

²⁶Department of Pediatric Oncology and Hematology, Institute of Pediatrics, Jagiellonian University Medical College, Krakow, Poland

²⁷Clinical and Translational Oncology Research Group, Health Research Institute La Fe, Valencia, Spain

²⁸Leeds Children's Hospital, Leeds General Infirmary, Leeds, United Kingdom

²⁹Karolinska University Hospital, Stockholm, Sweden

³⁰Department of Paediatric Haematology and Oncology, Princess Elisabeth Children's Hospital, Ghent University Hospital, Ghent, Belgium

³¹Paediatric Oncology, Fondazione IRCCS, Istituto Nazionale dei Tumori, Milan, Italy

³²Pediatric Hematology-Oncology Unit, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

³³Ruth Rappaport Children's Hospital, Rambam Health Care Campus, Haifa, Israel

³⁴Département d'Oncologie Pédiatrique, Gustave Roussy, Villejuif, France

³⁵Wolfson Childhood Cancer Research Centre, Newcastle Centre for Cancer, Translational & Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

³⁶Department for Studies and Statistics and Integrated Research, St Anna Children's Hospital, St Anna Children's Cancer Research Institute, Vienna, Austria

³⁷Department of Paediatrics, Medical University of Vienna, Vienna, Austria

CORRESPONDING AUTHOR

Gudrun Schleiermacher, MD, PhD, SIREDO Pediatric Oncology Center, Laboratory of Translational Research in Pediatric Oncology—INSERMU830 Institut Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France; e-mail: gudrun.schleiermacher@curie.fr.

EQUAL CONTRIBUTION

A.B. and U.P. contributed equally to this work and are to be considered as joint first authors. D.A.T., R.L., and G.S. contributed equally to this work and are to be considered joint senior authors.

SUPPORT

Supported by the Annenberg Foundation and the Association Hubert Gouin Enfance et Cancer, France. This study was also funded by the Fédération Enfants Cancers Santé, Les Bagouz à Manon, Les amis de Claire. Funding was also obtained from SiRIC/INCa (Grant INCa-DGOS-4654) and PHRC IC2007-09 grant.

High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche (Investissements d'Avenir program), by the Canceropole Ile-de-France, and by the SiRIC-Curie program - SiRIC Grant INCa-DGOS-4654.

In the United Kingdom, this work was supported by Neuroblastoma UK, Cancer Research UK, Department of Health, Families against Neuroblastoma, Solving Kids' Cancer, and Action Medical Research/ Great Ormond Street Hospital Charity. The CCLG Tissue Bank is funded by Cancer Research UK and CCLG.

The funding of the European Commission 5th Frame Work Grant (SIOPEN-R-NET EC Grant No. QLRI-CT-2002-01768, www.siopen-net.org) supporting the HR-NB11/SIOPEN trial is disclosed as funding source in the author statement. Pierre Fabre Médicament providing Busilvex (Paris, France), APEIRON (Vienna, Austria) providing dinutuximab beta (ch14.18/CHO) and the St Anna Kinderkrebsforschung GmbH (Vienna, Austria). The St Anna Kinderkrebsforschung was the academic sponsor of the HR-NBL1/SIOPEN trial providing resources for the remote trial data base and central trial management.

Recloning and production of the ch14.18 monoclonal antibody was done at Polymun, Vienna, Austria, and was enabled by a SIOPEN fundraising effort in 2001. APEIRON provided additional product at a later stage. The authors express their gratitude and appreciation to SIOPEN investigators, treating physicians, clinical research and care teams, and most

importantly to patients and families facing high-risk neuroblastoma for their committed participation in the trial. The European Commission, Pierre Fabre Médicament, and Apeiron had no involvement in the conduct of the research and preparation of the article.

In addition, this work was supported as follows: Belgium: vzw Kinderkankerfonds and Kom op tegen Kanker. Czech Republic: MH CZ—DRO, University Hospital Motol, Prague, Czech Republic; Israel: Hayim Association—for Children with Cancer in Israel, Ramat Gan. Italy: Fondazione Italiana per la Lotta al Neuroblastoma O.N.L.U.S. c/o Istituto G. Gaslini, Genova, Associazione Bianca Garavaglia O.N.L.U.S., Busto Arsizio. Spain: Grant FIS EC10/303, Asociación Pablo Ugarte, Cáncer Xavia, Sumemos Muchas Manos, Heath Institute Carlos III (ISCIII) and FEDER (European Regional Development Fund): Grants PI17/01558 and CIBERONC-CB16/12/00484. NEN Association (Nico contra el cáncer infantil 2017-PVR00157). Switzerland: Oncosuisse, Bern; Swiss Cancer League, Bern; Fond'action contre le Cancer, Lausanne; FORCE Fondation Recherche sur le Cancer de l'Enfant, Ecublens.

CLINICAL TRIAL INFORMATION

NCT01704716 (HR-NBL1/SIOPEN)

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI <https://doi.org/10.1200/JCO.21.00086>.

AUTHOR CONTRIBUTIONS

Conception and design: Angela Bellini, Ulrike Pötschger, Tommy Martinsson, Louis Chesler, Dominique Valteau-Couanet, Deborah A. Tweddle, Ruth Ladenstein, Gudrun Schleiermacher

Financial support: Deborah A. Tweddle, Gudrun Schleiermacher

Administrative support: Louis Chesler, Olivier Delattre, Gudrun Schleiermacher

Provision of study materials or patients: Angela Bellini, Peter F. Ambros, Nathalie Auger, Klaus Beiske, David R. Betts, Katleen de Preter, Nathalie

Clément, Valérie Combaret, Jaime Font de Mora, Irene Jiménez, Marta Jeison, Tommy Martinsson, Katia Mazzocco, Martina Morini, Annick Mühlethaler-Mottet, Rosa Noguera, Gaëlle Pierron, Sabine Taschner-Mandl, Nadine Van Roy, Louis Chesler, Victoria Castel, Martin Elliott, Per Kogner, Geneviève Laureys, Josef Malis, Maja Popovic-Beck, Shifra Ash, Olivier Delattre, Dominique Valteau-Couanet, Deborah A. Tweddle, Ruth Ladenstein, Gudrun Schleiermacher

Collection and assembly of data: Angela Bellini, Ulrike Pötschger, Eve Lapouble, Sylvain Baulande, Nathalie Auger, Klaus Beiske, Marie Bernkopf, Jaydutt Bhalshankar, Nick Bown, Katleen de Preter, Nathalie Clément, Valérie Combaret, Jaime Font de Mora, Sally L. George, Irene Jiménez, Marta Jeison, Tommy Martinsson, Katia Mazzocco, Martina Morini, Annick Mühlethaler-Mottet, Rosa Noguera, Gaëlle Pierron, Maria Rossing, Sabine Taschner-Mandl, Nadine Van Roy, Ales Vicha, Louis Chesler, Walentyna Balwierz, Victoria Castel, Martin Elliott, Per Kogner, Geneviève Laureys, Josef Malis, Maja Popovic-Beck, Shifra Ash, Olivier Delattre, Dominique Valteau-Couanet, Deborah A. Tweddle, Ruth Ladenstein, Gudrun Schleiermacher

Data analysis and interpretation: Angela Bellini, Ulrike Pötschger, Virginie Bernard, Peter F. Ambros, Nathalie Auger, Marie Bernkopf, David R. Betts, Jaime Font de Mora, Barbara Marques, Tommy Martinsson, Sabine Taschner-Mandl, Nadine Van Roy, Per Kogner, Geneviève Laureys, Roberto Luksch, Deborah A. Tweddle, Ruth Ladenstein, Gudrun Schleiermacher

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

ACKNOWLEDGMENT

The authors would like to thank the following Biobanks for providing samples: In Italy, the BIT-Gaslini Biobank, IRCCS Istituto Giannina Gaslini, Via G. Gaslini 5, Genova. In Spain, the Clinic Hospital INCLIVA-Valencia NB Tissue Bank (ISCIII, Reference: B0000339). They also thank the Children's Cancer & Leukemia Group (CCLG) Tissue Bank for access to DNA samples (CCLG 2015 BS 04), and contributing CCLG Centers, including members of the Experimental Cancer Medicine Centers Pediatric network.

REFERENCES

- Matthay KK, Maris JM, Schleiermacher G, et al: Neuroblastoma. *Nat Rev Dis Primers* 2:16078, 2016
- Holmes K, Pötschger U, Pearson ADJ, et al: Influence of surgical excision on the survival of patients with stage 4 high-risk neuroblastoma: A report from the HR-NBL1/SIOPEN study. *J Clin Oncol* 38:2902-2915, 2020
- Ladenstein R, Pötschger U, Pearson ADJ, et al: Busulfan and melphalan versus carboplatin, etoposide, and melphalan as high-dose chemotherapy for high-risk neuroblastoma (HR-NBL1/SIOPEN): An international, randomised, multi-arm, open-label, phase 3 trial. *Lancet Oncol* 18:500-514, 2017
- Ladenstein R, Pötschger U, Valteau-Couanet D, et al: Investigation of the role of dinutuximab beta-based immunotherapy in the SIOPEN high-risk neuroblastoma 1 trial (HR-NBL1). *Cancers (Basel)* 12:309, 2020
- Ladenstein R, Pötschger U, Valteau-Couanet D, et al: Interleukin 2 with anti-GD2 antibody ch14.18/CHO (dinutuximab beta) in patients with high-risk neuroblastoma (HR-NBL1/SIOPEN): A multicentre, randomised, phase 3 trial. *Lancet Oncol* 19:1617-1629, 2018
- Ozkaynak MF, Gilman AL, London WB, et al: A comprehensive safety trial of chimeric antibody 14.18 with GM-CSF, IL-2, and isotretinoin in high-risk neuroblastoma patients following myeloablative therapy: Children's Oncology Group study ANBL0931. *Front Immunol* 9:1355, 2018
- Park JR, Kreissman SG, London WB, et al: Effect of tandem autologous stem cell transplant vs single transplant on event-free survival in patients with high-risk neuroblastoma: A randomized clinical trial. *JAMA* 322:746-755, 2019
- Pinto N, Naranjo A, Hibbits E, et al: Predictors of differential response to induction therapy in high-risk neuroblastoma: A report from the Children's Oncology Group (COG). *Eur J Cancer* 112:66-79, 2019
- Seeger RC, Brodeur GM, Sather H, et al: Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N Engl J Med* 313:1111-1116, 1985
- Janoueix-Lerosey I, Schleiermacher G, Michels E, et al: Overall genomic pattern is a predictor of outcome in neuroblastoma. *J Clin Oncol* 27:1026-1033, 2009
- Peifer M, Hertwig F, Roels F, et al: Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 526:700-704, 2015
- Molenaar JJ, Koster J, Zwijnenburg DA, et al: Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 483:589-593, 2012
- Pugh TJ, Morozova O, Attiyeh EF, et al: The genetic landscape of high-risk neuroblastoma. *Nat Genet* 45:279-284, 2013
- Sausen M, Leary RJ, Jones S, et al: Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat Genet* 45:12-17, 2012
- Chen Y, Takita J, Choi YL, et al: Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* 455:971-974, 2008
- George RE, Sanda T, Hanna M, et al: Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* 455:975-978, 2008

17. Janoueix-Lerosey I, Lequin D, Brugieres L, et al: Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature* 455:967-970, 2008
18. Mosse YP, Laudenslager M, Longo L, et al: Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 455:930-935, 2008
19. Bellini A, Bessoltane-Bentahar N, Bhalshankar J, et al: Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma. *Int J Cancer* 145:2781-2791, 2019
20. Javanmardi N, Fransson S, Djos A, et al: Low frequency ALK hotspots mutations in neuroblastoma tumours detected by ultra-deep sequencing: Implications for ALK inhibitor treatment. *Sci Rep* 9:2199, 2019
21. Combaret V, Iacono I, Bellini A, et al: Detection of tumor ALK status in neuroblastoma patients using peripheral blood. *Cancer Med* 4:540-550, 2015
22. Bellini A, Bernard V, Leroy Q, et al: Deep sequencing reveals occurrence of subclonal ALK mutations in neuroblastoma at diagnosis. *Clin Cancer Res* 21:4913-4921, 2015
23. Eleveld TF, Oldridge DA, Bernard V, et al: Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* 47:864-871, 2015
24. Bresler SC, Weiser DA, Huwe PJ, et al: ALK mutations confer differential oncogenic activation and sensitivity to ALK inhibition therapy in neuroblastoma. *Cancer Cell* 26:682-694, 2014
25. Fransson S, Hansson M, Ruuth K, et al: Intragenic anaplastic lymphoma kinase (ALK) rearrangements: Translocations as a novel mechanism of ALK activation in neuroblastoma tumors. *Genes Chromosomes Cancer* 54:99-109, 2015
26. De Brouwer S, De Preter K, Kumps C, et al: Meta-analysis of neuroblastomas reveals a skewed ALK mutation spectrum in tumors with MYCN amplification. *Clin Cancer Res* 16:4353-4362, 2010
27. Friboulet L, Li N, Katayama R, et al: The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. *Cancer Discov* 4:662-673, 2014
28. Guan J, Tucker ER, Wan H, et al: The ALK inhibitor PF-06463922 is effective as a single agent in neuroblastoma driven by expression of ALK and MYCN. *Dis Model Mech* 9:941-952, 2016
29. Solomon BJ, Besse B, Bauer TM, et al: Lorlatinib in patients with ALK-positive non-small-cell lung cancer: Results from a global phase 2 study. *Lancet Oncol* 19:1654-1667, 2018
30. Ladenstein R, Potschger U, Siabalis D, et al: Dose finding study for the use of subcutaneous recombinant interleukin-2 to augment natural killer cell numbers in an outpatient setting for stage 4 neuroblastoma after megatherapy and autologous stem-cell reinfusion. *J Clin Oncol* 29:441-448, 2010
31. Ladenstein R, Valteau-Couanet D, Brock P, et al: Randomized trial of prophylactic granulocyte colony-stimulating factor during rapid COJEC induction in pediatric patients with high-risk neuroblastoma: The European HR-NBL1/SIOPEN study. *J Clin Oncol* 28:3516-3524, 2010
32. Garaventa A, Poetschger U, Valteau-Couanet D, et al: Randomized trial of two induction therapy regimens for high-risk neuroblastoma: HR-NBL1.5 International Society of Pediatric Oncology European Neuroblastoma Group study. *J Clin Oncol* 39:2552-2563, 2021
33. Ambros PF, Ambros IM, Brodeur GM, et al: International consensus for neuroblastoma molecular diagnostics: Report from the International Neuroblastoma Risk Group (INRG) Biology Committee. *Br J Cancer* 100:1471-1482, 2009
34. Ambros IM, Brunner B, Aigner G, et al: A multilocus technique for risk evaluation of patients with neuroblastoma. *Clin Cancer Res* 17:792-804, 2011
35. Schleiermacher G, Michon J, Ribeiro A, et al: Segmental chromosomal alterations lead to a higher risk of relapse in infants with MYCN-non-amplified localised unresectable/disseminated neuroblastoma (a SIOPEN collaborative study). *Br J Cancer* 105:1940-1948, 2011
36. Schleiermacher G, Mosseri V, London WB, et al: Segmental chromosomal alterations have prognostic impact in neuroblastoma: A report from the INRG project. *Br J Cancer* 107:1418-1422, 2012
37. Ambros IM, Brunner C, Abbasi R, et al: Ultra-high density SNParray in neuroblastoma molecular diagnostics. *Front Oncol* 4:202, 2014
38. UCSC Genome Browser Home: <https://genome.ucsc.edu/index.html>
39. Kaplan E, Meier P: Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457-481, 1958
40. Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *Br J Cancer* 35:1-39, 1977
41. Andersen PK, Perme MP: Pseudo-observations in survival analysis. *Stat Methods Med Res* 19:71-99, 2011
42. Morgenstern DA, Potschger U, Moreno L, et al: Risk stratification of high-risk metastatic neuroblastoma: A report from the HR-NBL-1/SIOPEN study. *Pediatr Blood Cancer* 65:e27363, 2018
43. Depuydt P, Boeva V, Hocking TD, et al: Genomic amplifications and distal 6q loss: Novel markers for poor survival in high-risk neuroblastoma patients. *J Natl Cancer Inst* 110:1084-1093, 2018
44. Mosse YP, Lim MS, Voss SD, et al: Safety and activity of crizotinib for paediatric patients with refractory solid tumours or anaplastic large-cell lymphoma: A Children's Oncology Group phase 1 consortium study. *Lancet Oncol* 14:472-480, 2013
45. Guan J, Fransson S, Siaw JT, et al: Clinical response of the novel activating ALK-11171T mutation in neuroblastoma to the ALK inhibitor ceritinib. *Cold Spring Harb Mol Case Stud* 4:a002550, 2018
46. Krytska K, Ryles HT, Sano R, et al: Crizotinib synergizes with chemotherapy in preclinical models of neuroblastoma. *Clin Cancer Res* 22:948-960, 2016
47. Bogen D, Brunner C, Walder D, et al: The genetic tumor background is an important determinant for heterogeneous MYCN-amplified neuroblastoma. *Int J Cancer* 139:153-163, 2016
48. Marrano P, Irwin MS, Thorne PS: Heterogeneity of MYCN amplification in neuroblastoma at diagnosis, treatment, relapse, and metastasis. *Genes Chromosomes Cancer* 56:28-41, 2017
49. Berbegall AP, Bogen D, Potschger U, et al: Heterogeneous MYCN amplification in neuroblastoma: A SIOP Europe Neuroblastoma study. *Br J Cancer* 118:1502-1512, 2018
50. Turajlic S, Sottoriva A, Graham T, et al: Resolving genetic heterogeneity in cancer. *Nat Rev Genet* 20:404-416, 2019
51. Williams JB, Li S, Higgs EF, et al: Tumor heterogeneity and clonal cooperation influence the immune selection of IFN-gamma-signaling mutant cancer cells. *Nat Commun* 11:602, 2020
52. Padovan-Merhar OM, Raman P, Ostrovskaya I, et al: Enrichment of targetable mutations in the relapsed neuroblastoma genome. *PLoS Genet* 12:e1006501, 2017
53. Schleiermacher G, Javanmardi N, Bernard V, et al: Emergence of new ALK mutations at relapse of neuroblastoma. *J Clin Oncol* 32:2727-2734, 2014
54. Chicard M, Colmet-Daage L, Clement N, et al: Whole-exome sequencing of cell-free DNA reveals temporo-spatial heterogeneity and identifies treatment-resistant clones in neuroblastoma. *Clin Cancer Res* 24:939-949, 2018
55. Ackermann S, Cartolano M, Hero B, et al: A mechanistic classification of clinical phenotypes in neuroblastoma. *Science* 362:1165-1170, 2018
56. Koneru B, Lopez G, Farooqi A, et al: Telomere maintenance mechanisms define clinical outcome in high-risk neuroblastoma. *Cancer Res* 80:2663-2675, 2020



AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Frequency and Prognostic Impact of ALK Amplifications and Mutations in the European Neuroblastoma Study Group (SIOPEX) High-Risk Neuroblastoma Trial (HR-NBL1)

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/authors/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](#)).

Walentyna Balwierz

Honoraria: Shire, Gilead Sciences, Novartis, Amgen

Consulting or Advisory Role: Amgen, Novartis, Roche, Takeda

Travel, Accommodations, Expenses: Jazz Pharmaceuticals, Shire, Roche, Servier

Martin Elliott

Consulting or Advisory Role: Bayer

Dominique Valteau-Couanet

Consulting or Advisory Role: EUSA Pharma

Research Funding: Orphelia Pharma

Patents, Royalties, Other Intellectual Property: Royalties from Apeiron to SIOPEX

Travel, Accommodations, Expenses: EUSA Pharma, Jazz Pharmaceuticals

Deborah A. Tweddle

Honoraria: EUSA Pharma

Travel, Accommodations, Expenses: EUSA Pharma

Ruth Ladenstein

Honoraria: Apeiron Biologics, Boehringer Ingelheim, EUSA Pharma

Consulting or Advisory Role: Apeiron Biologics, Boehringer Ingelheim, EUSA Pharma

Research Funding: Apeiron Biologics, EUSA Pharma

Patents, Royalties, Other Intellectual Property: Apeiron Biologics, EUSA Pharma

Expert Testimony: Apeiron Biologics, EUSA Pharma

Travel, Accommodations, Expenses: Apeiron Biologics, EUSA Pharma

Gudrun Schleiermacher

Honoraria: BMS

Research Funding: Bristol Myers Squibb, Pfizer, MSDAvenir, Roche

Travel, Accommodations, Expenses: Roche

No other potential conflicts of interest were reported.

APPENDIX

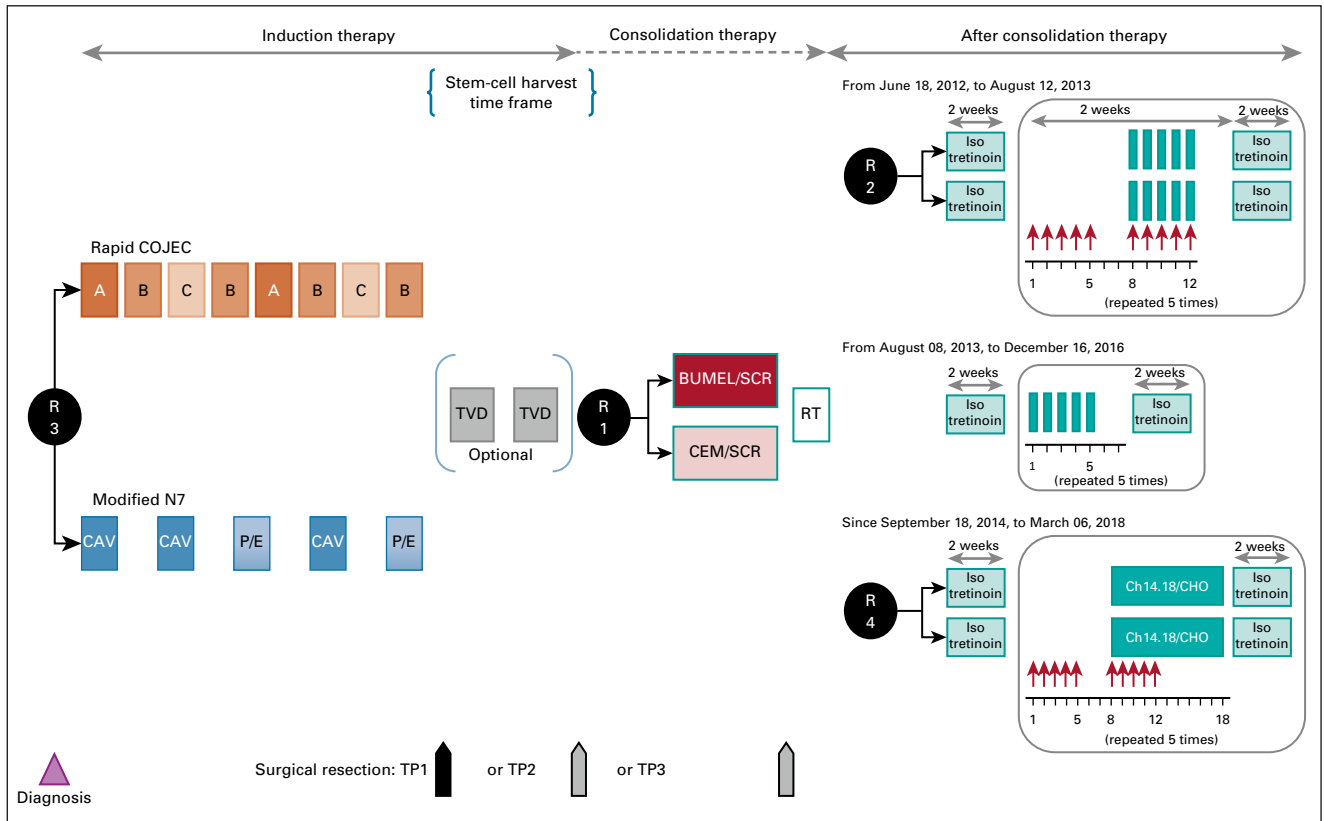


FIG A1. Treatment flowchart of the HR-NBL1 Protocol (ClinicalTrials.gov: [NCT01704716](https://clinicaltrials.gov/ct2/show/study/NCT01704716), EudraCT: 2006-001489-17) over the whole period. ^aInfants and children with a body weight below 12 kg will be dosed at 0.67 mg/kg/d. In infants weighing ≤ 5 kg, a further 1/3 dose reduction is advised. AUC, area under the curve; BUMEL, busulfan and melphalan; CAV, cyclophosphamide plus doxorubicin or vincristine; CEM, carboplatin, etoposide, and melphalan; CH14.18/CHO, human-mouse chimeric monoclonal anti-disialoganglioside GD2 antibody ch14.18 produced in Chinese hamster ovary (CHO) cells; COJEC, chemotherapy schedule COJEC defined below; GFR, glomerular filtration rate; IL-2, interleukin-2; IV, intravenous; P or E, cisplatin or etoposide; R1, randomization 1; R2, randomization 2; R3, randomization 3; R4, randomization 4; RT, radiotherapy; SCR, stringent complete response; TP, time period; TVD, topotecan-vincristine-doxorubicin. (continued on following page)




A	Course A	Vincristine Carboplatin Etoposide	1.5 mg/m ² (maximum dose 2 mg) x 1 day 750 mg/m ² x 1 day 175 mg/m ² x 2 days
B	Course B	Vincristine Cisplatin	1.5 mg/m ² (maximum dose 2 mg) x 1 day 80 mg/m ² /ctn over 24 hours x 1 day
C	Course C	Vincristine Etoposide Cyclophosphamide	Vincristine 1.5 mg/m ² (maximum dose 2 mg) x 1 day 175 mg/m ² x 2 days 1,050 mg/m ² x 2 days
CAV	Course CAV	Cyclophosphamide Doxorubicin Vincristine	70 mg/kg x 2 days 25 mg/m ² x 3 days 0.022 mg/kg x 3 days
P/E	Course P/E	Cisplatin Etoposide	50 mg/m ² x 4 days 200 mg/m ² on 3 days
TVD	Course TVD (optional)	Topotecan Vincristine Doxorubicin	100 ml/m ² x 5 days 1 mg/m ² x 2 days 22 mg/m ² x 2 days
BUMEL/SCR	BUMEL	Busivex Melphalan	< 9 kg: 1 mg/kg; 9 kg to < 16 kg: 1.2 mg/kg; 16 kg to 23 kg: 1.1 mg/kg; > 23 kg to 34 kg: 0.95 mg/kg; > 34 kg: 0.8 mg/kg x 5 days 140 mg/m ² x 1 day Autologous stem-cell reinfusion
CEM/SCR	CEM	Carboplatin Etoposide Melphalan	AUC 4.1 mg/ml min/d x 4 days (based on the GFR) ≤ 12 kg: 11.3 mg/kg/d; > 12 kg: 338 mg/m ² /d x 4 days ≤ 12 kg: 2.3 mg/kg/d; > 12 kg: 70 mg/m ² /d x 3 days
RT	Radiotherapy		Fractionated radiotherapy (21 Gy) given in 14 fractions of 1.5 Gy over not more than 21 days
		Aldesleukin (IL-2)	6 MIU/m ² /d subcutaneously on 5 consecutive days over 2 weeks
	Immunotherapy	Ch14.18/CHO	20 mg/m ² /d ^a over 5 days every 4 weeks for five courses
Ch14.18/CHO		Ch14.18/CHO	10 mg/m ² /d continuous IV infusion over 10 days
	Possible TPs for surgical resection		

FIG A1. (Continued).

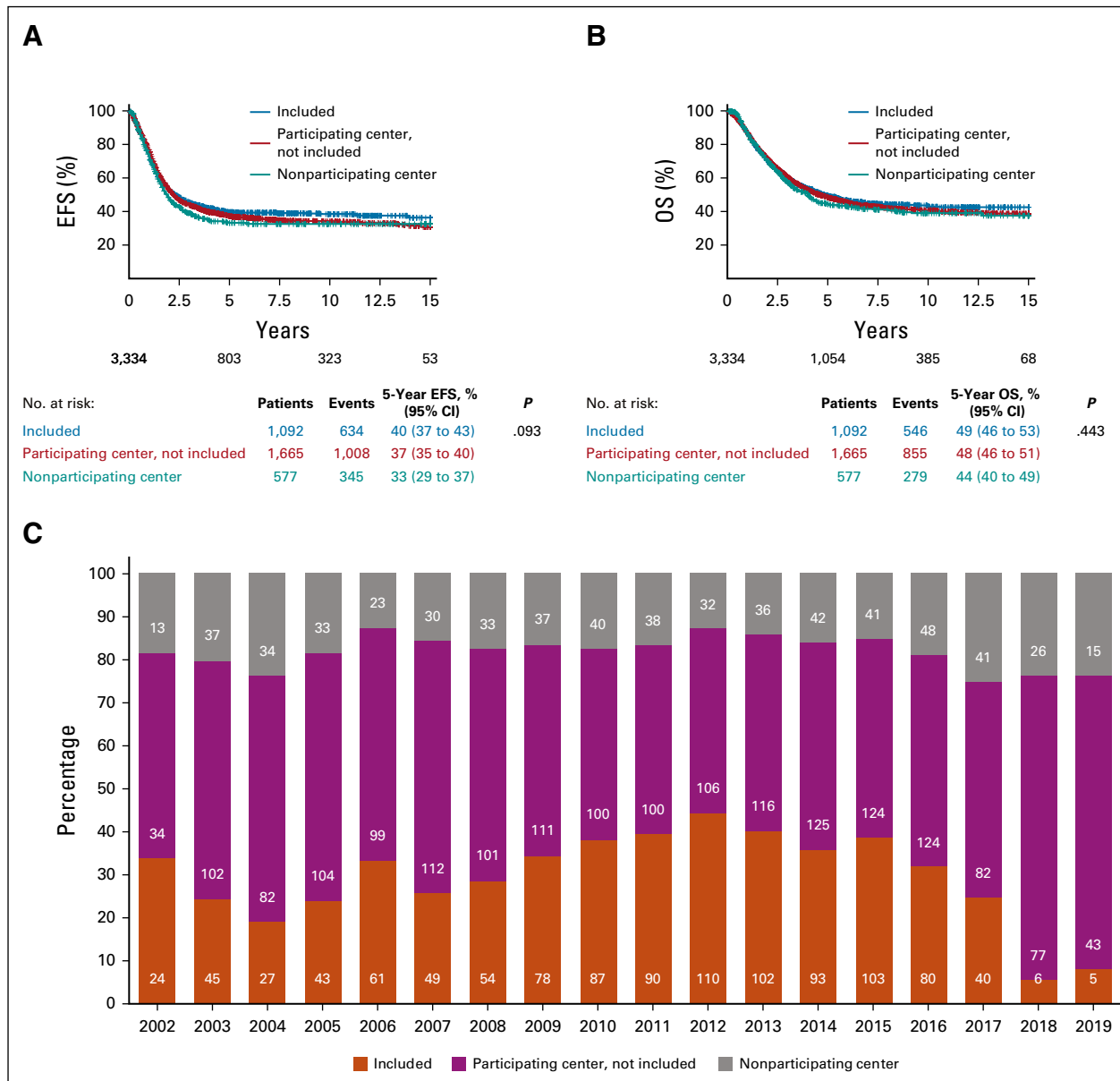


FIG A2. Comparison of patients in the ALK analysis cohort and patients not in the ALK analysis cohort. (A and B) EFS and OS of the ALK analysis cohort and patients not in the ALK cohort. (A) No statistically significant difference in EFS and (B) OS was observed between patients included in the ALK analysis cohort (n = 1,092, from 132 centers; red line), patients not included in this study from the same centers (n = 1,665, blue line) and patients not included in this study from centers not participating in this study (n = 577, green line) (5-year EFS: 40% [95% CI, 37 to 43] v 37% [95% CI, 35 to 40] v 33% [95% CI, 29 to 37]; 5-year OS: 49% [95% CI, 46 to 53] v 48% [95% CI, 46 to 51] v 44% [95% CI, 40 to 49]; P = NS). (C) Recruitment, by year (x-axis), in the ALK analysis cohort (% of patients: y-axis; absolute numbers: in the blue bars). The % and number of patients not included in the ALK analysis cohort from centers participating, and from nonparticipating centers, are indicated in orange and gray, respectively. EFS, event-free survival; NS, not significant; OS, overall survival.

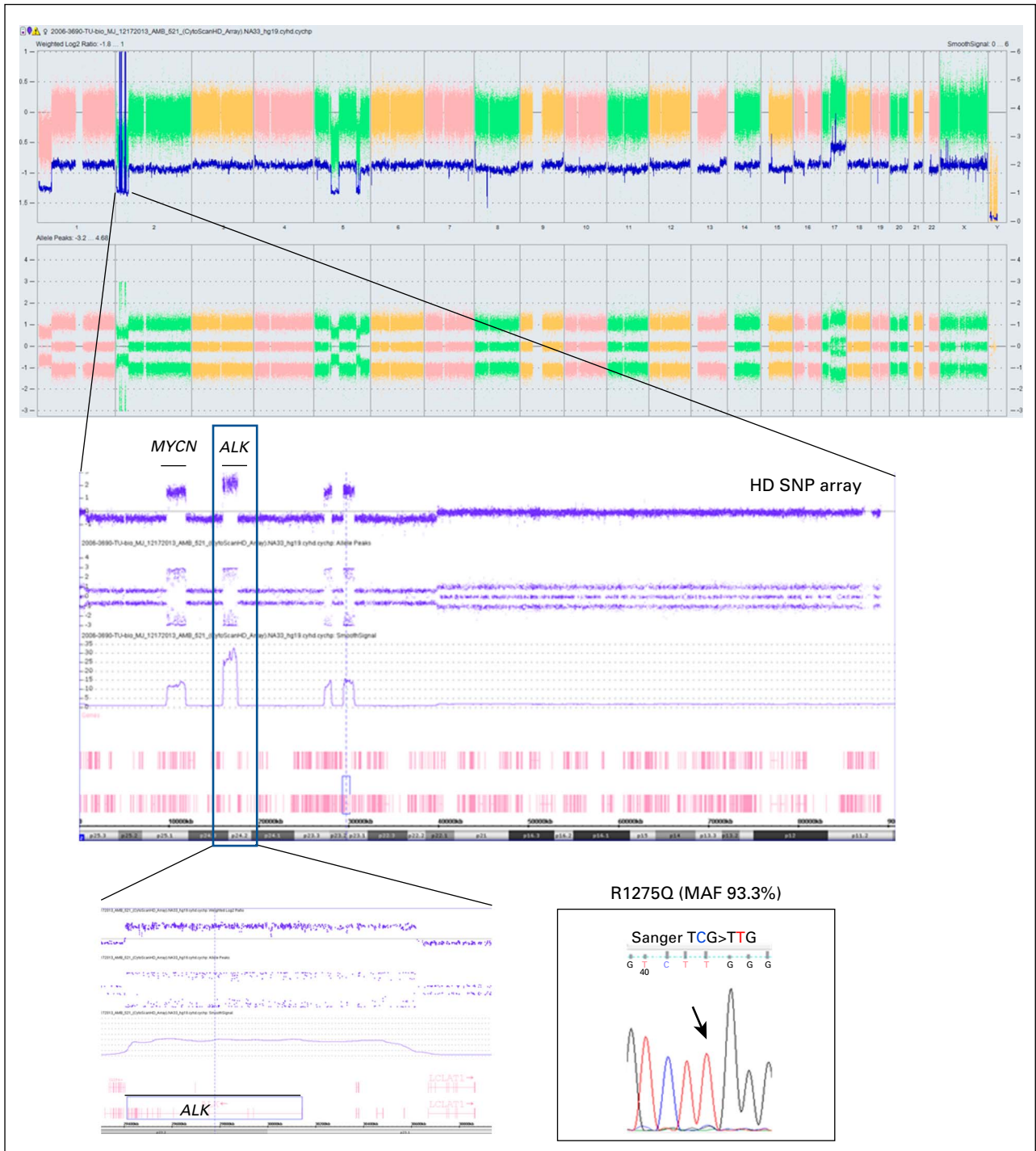


FIG A3. Double event of *ALK* amplification and *ALK* mutation detected in one case (case 15). The SNP array shows an amplified region in chromosome 2 encompassing the *ALK* gene. Sanger sequencing profile shows R1275Q mutation (MAF = 93.3%) in the same case. HD, high definition; MAF, mutated allele fraction; SNP, single-nucleotide polymorphism.

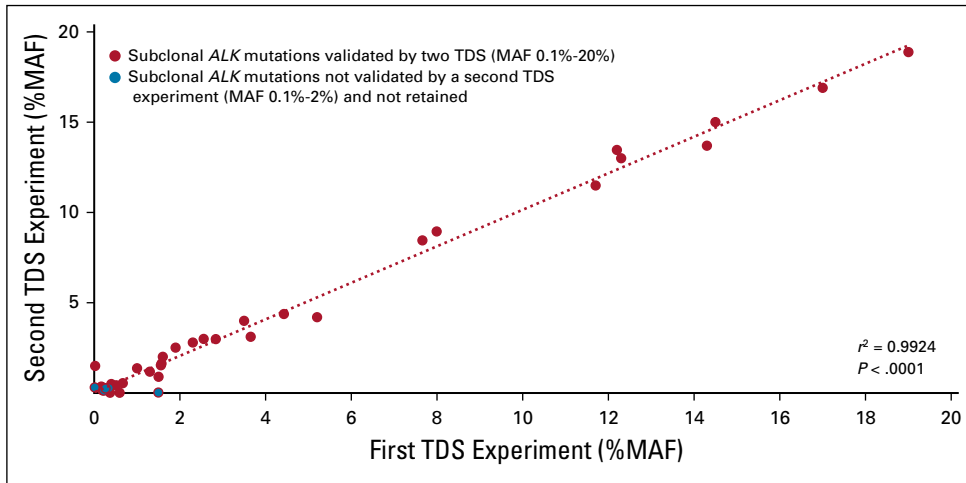


FIG A4. MAF of subclonal *ALK* mutations detected by TDS and confirmed by a second independent TDS experiment. Red spots representing the MAF for each *ALK* mutation are plotted on the *x*-axis (first TDS experiment) and *y*-axis (second TDS experiment), with a strong correlation between the two independent experiments ($r^2 = 0.9924$, $P < .0001$). Blue spots represent subclonal *ALK* mutations with a very low MAF ($< 0.1\%$) not confirmed in an independent experiment and not retained in the analysis ($n = 6$). MAF, mutated allele fraction; TDS, targeted deep sequencing.

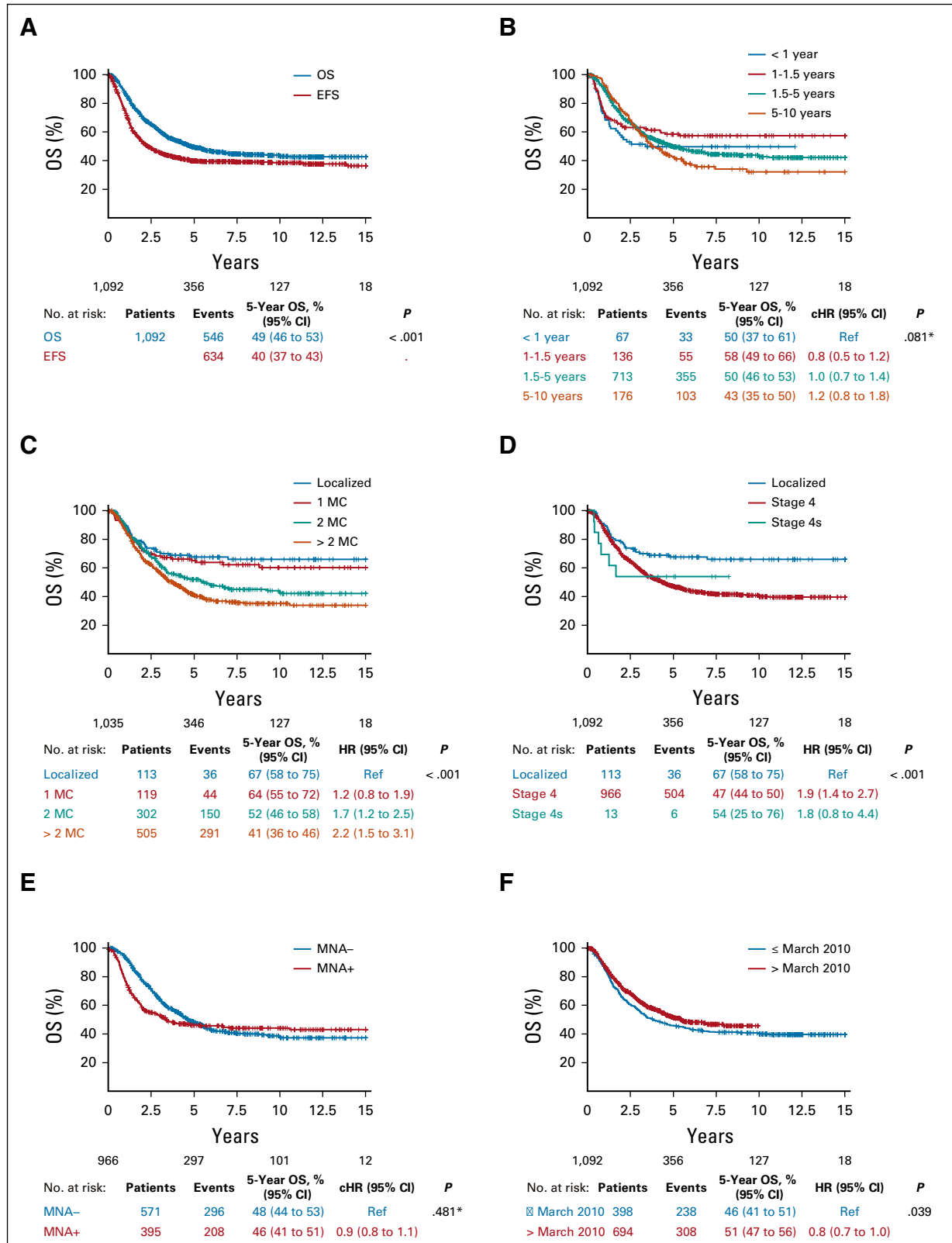


FIG A5. Survival in the ALK analysis cohort (n = 1,092 patients) according to known prognostic factors. (A) EFS and OS in the ALK analysis cohort population (n = 1,092 patients). Five-year EFS (blue line) 40% (95% CI, 37 to 43); 5-year OS (red line) 49% (95% CI, 46 to 53). (B) OS according to age. Five-year OS in patients < 1 year of age at diagnosis (red line) 50% (95% CI, 37 to 61); in patients 1-1.5 years of age at diagnosis (blue line) 58% (95% CI, 49 to 66); in patients 1.5-5 years of age at diagnosis (green line) 50% (95% CI, 46 to 53); and in patients > 5 years of age at diagnosis (purple line) 43% (95% CI, 35 to 50); *P* = NS (pseudo-value regression). (continued on following page)

FIG A5. (Continued). (C) OS according to number of involved MCs. Five-year OS in patients with localized disease (red line) 67% (95% CI, 58 to 75), in patients with involvement of one MC (blue line) 65% (95% CI, 55 to 73), two MCs (green line) 52% (95% CI, 46 to 58), or over two MCs (purple line) 41% (95% CI, 36 to 46); $P < .001$. (D) OS according to stage. Five-year OS in patients with localized disease (red line) 67% (95% CI, 58 to 75), in patients with stage 4 disease (blue line) 47% (95% CI, 44 to 50), or stage 4s disease (green line) 54% (95% CI, 25 to 76); $P < .001$. (E) OS according to *MYCN* amplification in stage 4 disease. Five-year OS in patients with MNA (blue line) 46% (95% CI, 41 to 51), in patients without MNA (red line) 48% (95% CI, 44 to 53), NS (pseudo-value regression). (F) OS according to treatment period, before (< March 2010) or after (> March 2010) the definition of HDC by BUMEL and immunotherapy maintenance as standard treatment. A significant improvement survival because of BUMEL and GD2 standard therapy is observed. Five-year OS in patients having been treated before March 2010 (red line) 46% (95% CI, 41 to 51) versus after March 2010 (blue line) 51% (95% CI, 47 to 56); $P = .039$.³⁻⁵ BUMEL, busulfan and melphalan; CHR, crude hazard ratio; EFS, event-free survival; HDC, high-dose chemotherapy; HR, hazard ratio; MC, metastatic compartment; MNA, MYCN-amplified; NS, not significant; OS, overall survival; ref, reference.

TABLE A1. Clinical Characteristics of 1,092 Patients Included in the *ALK* Analysis Cohort

	Stage 4				Stage 4s MNA+	Total
	Localized MNA+	Total	MNA-	MNA+		
Total	113	966	571	395	13	1,092
Sex, No. (%)						
Female	45 (40)	423 (44)	258 (45)	165 (42)	5 (38)	473 (43)
Male	68 (60)	543 (56)	313 (55)	230 (58)	8 (62)	619 (57)
Age at diagnosis, years						
< 1, No. (%)	5 (4)	50 (5)	0 (0)	49 (12)	13 (100)	67 (6)
1-1.5, No. (%)	22 (19)	113 (12)	39 (7)	75 (19)	0 (0)	136 (12)
1.5-5, No. (%)	79 (70)	634 (66)	392 (69)	242 (61)	0 (0)	713 (65)
5-10, No. (%)	7 (6)	169 (17)	140 (25)	29 (7)	0 (0)	176 (16)
Median (min-max)	2.1 (0.6-8.3)	2.9 (0.12-20)	3.5 (1-20)	2 (0.12-12)	0.23 (0-0.65)	2.8 (0-20)
Primary tumor, No. (%)						
No data	1	31	21	10	—	32
Cervical	5 (4)	54 (6)	37 (7)	17 (4)	0 (0)	59 (6)
Thoracic	4 (4)	157 (17)	108 (20)	49 (13)	0 (0)	161 (15)
Abdominal adrenal	85 (76)	655 (70)	341 (62)	314 (82)	13 (100)	753 (71)
Abdominal other	41 (37)	329 (35)	203 (37)	126 (33)	3 (23)	373 (35)
Pelvic	4 (4)	59 (6)	30 (5)	29 (8)	0 (0)	63 (6)
Metastatic sites, No. (%)						
None	113	—	—	—	13	113
Not specified		55	26	29	2	
1 MC		111 (12)	51 (9)	60 (16)	4 (36)	
2 MC		299 (33)	180 (33)	119 (33)	3 (27)	
> 2 MC		501 (55)	314 (58)	187 (51)	4 (36)	
ALK alteration, No. (%)						
Yes	25 (22)	118 (12)	36 (6)	82 (21)	3 (23)	146 (13)
No	88 (78)	848 (88)	535 (94)	313 (79)	10 (77)	946 (87)
ALK amplification, No. (%)						
Yes	13 (12)	26 (3)	1 (0)	25 (6)	2 (15)	41 (4)
No	83 (73)	768 (80)	466 (82)	302 (76)	9 (69)	860 (79)
Missing data	17 (15)	172 (18)	104 (18)	68 (17)	2 (15)	191 (17)
ALK mutations, No. (%)						
<i>ALK</i> m clonal	9 (8)	66 (7)	26 (5)	40 (10)	1 (8)	76 (7)
<i>ALK</i> m subclonal	4 (4)	26 (3)	9 (2)	17 (4)	0 (0)	30 (3)
No	63 (56)	586 (61)	365 (64)	221 (56)	7 (54)	656 (60)
Missing data	37 (33)	288 (30)	171 (30)	117 (30)	5 (38)	330 (30)

Abbreviations: MC, metastatic compartments; MNA, MYCN-amplified.

TABLE A2. Number of Patients Included in the ALK Analysis Cohort by Country and Center

Country	Center	Patients, No.
Total		1,092
FR	Total	344
	Institut Curie	65
	Center Léon Berard	34
	Hopitaux de Marseille La Timone	30
	Center Oscar Lambret de Lille	26
	CHR de Nantes	23
	Hopital Hautepierre-CHU Strasbourg	20
	Hôpital Trousseau Paris	18
	Institut Gustave Roussy	17
	Hôpital D'Enfants de Toulouse	14
	CHU de Grenoble	13
	CHU de Nancy Brabois	11
	CHU Montpellier Hôpital Arnaud Villeneuve	11
	CHU Rouen	10
	Hopital Jean Bernard La Miletrie Poitiers	8
	CHR de Caen	8
	CHU-Saint Etienne	6
	Hôpital de L'Archet Nice	5
	CHR Hôpital Sud de rennes	5
	Center Hospitalier Angers	5
	CHU Morvan de Brest	4
	Hotel Dieu de Clermont-ferrand	4
CHRU Nord d'Amiens	4	
Hopital d'Enfants Dijon	2	
Hopital Americain de Reims	1	
UK	Total	292
	Great Ormond Street Hospital	40
	Royal Marsden Hospital Surrey	34
	Newcastle: Royal Victoria Infirmary	29
	Dublin: OLHSC	13
	Oxford: John Radcliffe Hospital	20
	Bristol Royal Hospital for Children	19
	Glasgow Royal Hospital for Sick Children	19
	Manchester: Royal Manchester Children's Hospital	18
	Southampton General Hospital	16
	Cambridge: Addenbrooke's NHS Trust	14
	Liverpool: Alder Hey Children's Hospital	14
	Birmingham Children's Hospital	11
	Leeds: St James's University Hospital	11
	Belfast: Royal Belfast Hospital for Sick Children	9
	Sheffield Children's Hospital	7
	Cardiff: Llandough Hospital	5
	Aberdeen: Royal Aberdeen Children's Hospital	4
	Edinburgh Royal Hospital for Sick Children	4
	Leicester Royal Infirmary	3
UCLH University College London Hospital	2	

(continued on following page)

TABLE A2. Number of Patients Included in the *ALK* Analysis Cohort by Country and Center (continued)

Country	Center	Patients, No.
ES	Total	152
	H Nino Jesus	15
	Hospital Infantil La Fe	13
	Carlos Haya	11
	H Central de Asturias	10
	Hospital Infantil La Paz	10
	H. Virgen de la Arrixaca	8
	Hospital de Cruces	7
	Hospital materno infantil Virgen de las Nieves	7
	Hospital Vall d'Hebron	6
	H. Miguel Servet	6
	Hospital Clinico	5
	H. Virgen del Camino	4
	H. Son Dureta	5
	H. General de Galicia	4
	Hospital Gregorio Maranon	4
	Hospital 12 de Octubre	4
	H. de Donostia Ntra. Sra. de Aranzazu.	4
	Materno Infantil de Badajoz	3
	H. General de Alicante	3
	Virgen del Rocio	3
	Hospital Germans Triasi Pujol	2
	H Sant Pau	2
	Hospital Universitario de Canarias	2
	H. Torrecardenas	2
	Hospital Reina Sofia	2
	H. C. U. de Salamanca	2
	H. Virgen de la Salud	1
	H. Materno-Infantil Teresa Herrera	1
	H. SanT Joan de Deu	1
H. Montepincipe	1	
Complejo Hospitalario de Jaen	1	
H. Virgen de la Macarena	1	
Hospital Universitario Nuestra Sra de la Candelaria	1	
Hospital Xeral-Cies	1	
AT	Total	57
	St Anna Kinderspital	23
	Landes-Kinderklinik Linz	12
	Univ.Klinik f. Kinder-u. Jugendheilkunde Innsbruck	10
	Univ.-Klinik für Kinder- und Jugendheilkunde Graz	6
	St Johanns Spital LKH Salzburg	6

(continued on following page)

TABLE A2. Number of Patients Included in the ALK Analysis Cohort by Country and Center (continued)

Country	Center	Patients, No.
SE	Total	44
	Stockholm	14
	Lund	11
	Uppsala	8
	Children's Hospital Linköping	5
	Queen Silvia's Children's Hospital (Gothenburg)	5
	Reykjavik	1
CZ	Total	38
	University Hospital Motol, Prague 5	38
IT	All	29
	Ospedale S. Orsola	7
	Clinica di Oncoematologia Pediatrica Padova	5
	Istituto per l'Infanzia Burlo Garofolo	3
	Ospedale Bambino Gesù	3
	Policlinico Universitario	2
	Istituto Giannina Gaslini	2
	Istituto Nazionale Tumori di Milano	2
	Policlinico San Matteo	1
	Ospedali Riuniti	1
	Ospedale dei bambini, Palermo	1
	Azienda Ospedaliera Universitaria di Parma-Oncoematologia Pediatrica	1
	Policlinico Borgo Roma	1
CH	Total	25
	CHUV	11
	University Children's Hospital (Geneva)	5
	Inselspital Bern	3
	Kantonspital Aarau	3
	Ostschweizer Kinderspital	2
	Luzerner Kantonspital - Kinderspital Luzern	1
PL	Total	23
	University Children's Hospital Krakow	14
	Wroclaw Medical University	3
	Children's Hospital in Chorzów	2
	University of Medical Sciences Poznan	2
	Medical University of Bydgoszcz	1
	Medical University in Gdansk	1
BE	Total	21
	University Hospital Gent	9
	UZ Gasthuisberg	8
	Clinique de l'Espérance,	2
	Cliniques universitaires St-Luc	1
	CHR Citadelle	1
IL	Total	18
	Schneider Children's Medical Center of Israel	17
	Dana Children's Hosp., Suraski Tel-Aviv Med. Cent.	1

(continued on following page)

TABLE A2. Number of Patients Included in the *ALK* Analysis Cohort by Country and Center (continued)

Country	Center	Patients, No.
PT	Total	14
	IPOFG-CRL	14
HK	Total	10
	University of Hong Kong	10
NO	Total	10
	Rikshospitalet	5
	Haukeland University Hospital	4
	St Olavs Hospital Trondheim	1
IE	Total	7
	Dublin: OLHSC	7
FI	Total	4
	University of Tampere	4
DK	Total	2
	Aarhus Universitetshospital	1
	University Hospital of Odense	1
GR	Total	1
	Aghia Sophia Children's Hospital, Athens	1
SI	Total	1
	University Children's Hospital Ljubljana	1

Abbreviations: AT, Austria; BE, Belgium; CH, Switzerland; CZ, Czech Republic; DK, Denmark; ES, Spain; FI, Finland; FR, France; GR, Greece; HK, Hong Kong; IE, Ireland; IL, Israel; IT, Italy; NO, Norway; PL, Poland; PT, Portugal; SE, Sweden; SI, Slovenia; UK, United Kingdom.

TABLE A3. Clinical Characteristics of 35 Patients Treated by Immunotherapy Whose Tumors Harbored ALK Genetic Alterations

Patient No.	Sex	Age at Diagnosis, years	INSS Stage	Induction Treatment	Status Post Induction	HDC	Relapse	Last Status	MYCN Status	ALK Amplification Status	ALK Mutations	Type Mut
1	M	2.0	4	Rapid COJEC	CR	CEM	No	Alive	MN-NA	ALK-NA	Yes	R12
2	M	2.2	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-A	No	N
3	F	4.9	Loc	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-A	No	N
4	M	1.9	Loc	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-A	No	N
5	F	3.5	4	Rapid COJEC	VGPR	BUMEL	Yes	Dead	MNA	ALK-A	No	N
6	M	2.3	4	Rapid COJEC	MR	BUMEL	Yes	Dead	MN-NA	ALK-NA	Yes	R12
7	M	2.5	Loc	Rapid COJEC	SD	BUMEL	No	Alive	MNA	ALK-NA	Yes	F11
8	F	1.5	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MN-NA	ALK-NA	Yes	F12
9	F	2.0	4	Rapid COJEC	VGPR	CEM	Yes	Dead	MNA	ALK-NA	Yes	R12
10	M	2.6	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	F11
11	M	2.3	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-A	No	N
12	F	1.2	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-A	No	N
13	M	2.6	4	Rapid COJEC	VGPR	BUMEL	No	Alive	MNA	ALK-NA	Yes	R12
14	M	4.8	4	MOD. N7	PR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	I111
15	F	1.3	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F11
16	F	2.0	4	MOD. N7	PR	BUMEL	No	Alive	MN-NA	ALK-NA	Yes	R12
17	M	4.0	4	Rapid COJEC	VGPR	BUMEL	Yes	Alive	MNA	ALK-NA	Yes	A12 G1 G1
18	M	1.3	4	MOD. N7	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F11
19	F	4.3	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	F11
20	M	1.1	4	Rapid COJEC	PR	BUMEL	No	Alive	MN-NA	ALK-NA	Yes	F11
21	M	9.7	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	F11
22	M	2.0	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F11
23	F	1.6	4	Rapid COJEC	VGPR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	R12
24	F	6.8	4	Rapid COJEC	PR	BUMEL	No	Alive	MN-NA	NA	Yes	I111
25	F	2.1	4	Rapid COJEC	PR	CEM	Yes	Dead	MNA	ALK-A	No	N
26	M	2.7	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MN-NA	ALK-NA	Yes	F11
27	M	1.7	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-A	No	N
28	M	1.7	4	Rapid COJEC	VGPR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	F12
29	F	3.9	4	Rapid COJEC	VGPR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F12
30	M	2.8	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F11

(continued on following page)

TABLE A3. Clinical Characteristics of 35 Patients Treated by Immunotherapy Whose Tumors Harbored ALK Genetic Alterations (continued)

Patient No.	Sex	Age at Diagnosis, years	INSS Stage	Induction Treatment	Status Post Induction	HDC	Relapse	Last Status	MYCN Status	ALK Amplification Status	ALK Mutations	Type Mut
31	M	2.1	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	L12
32	F	2.2	4	Rapid COJEC	VGPR	BUMEL	Yes	Alive	MN-NA	ALK-NA	Yes	R12
33	F	2.2	4	Rapid COJEC	PR	BUMEL	Yes	Dead	MNA	ALK-NA	Yes	F12
34	M	1.9	Loc	Rapid COJEC	VGPR	BUMEL	No	Alive	MNA	ALK-NA	Yes	F12
35	F	2.0	4	Rapid COJEC	PR	BUMEL	No	Alive	MNA	ALK-NA	Yes	L11

NOTE: Among these patients, ALK amplifications were detected in eight cases, and clonal ALK mutations were detected in 21 cases. In addition, six of eight cases harbored MYCN amplification. Abbreviations: ALK-A, ALK-amplified; ALK-NA, ALK not amplified; BUMEL, busulfan and melphalan; CEM, carboplatin, etoposide, and melphalan; COJEC, CR, complete remission; F, female; HDC, high-dose chemotherapy; INSS, International Neuroblastoma Staging System; M, male; MAF, mutated allele fraction; MR, minor response; NA, not applicable; NGS, next-generation sequencing; PR, partial remission; SD, stable disease; TDS, targeted deep sequencing.

4 DISCUSSION AND PERSPECTIVES

In this research, we undertook a comprehensive analysis of neuroblastoma ITH using a multimodal data approach, encompassing bulk whole exome sequencing (WES), single-cell whole genome sequencing (scWGS), replication timing (RT), and single-cell RNA sequencing (scRNA-seq). Our bioinformatics strategy required an integrative approach, leading to the development of a robust pipeline with various modules due to the absence of a singular tool for such complex analysis.

Module 1 focused on bulk and pseudobulk data analysis to identify single nucleotide polymorphisms (SNPs) and somatic single nucleotide variants (SNVs), along with allele-specific copy number analysis. This preliminary analysis was crucial for subsequent mutational timing, where advanced tools like Mobster and MutationTimeR were employed. Mobster combines machine learning with theoretical population genetics for model-based tumor subclonal reconstruction, while MutationTimeR offers mutational timing analysis. Pseudobulk whole genome sequencing with >40X coverage proved ideal for these analyses. **Module 2** handled single-cell whole genome sequencing data, integrating tools like CHISEL for allele- and haplotype-specific copy number inference, and CopyKit for resolving clonal substructure and reconstructing genetic lineages in tumors. While CHISEL's strength lay in allele- and haplotype-specific analysis, CopyKit provided a comprehensive method for single-cell absolute copy number data analysis. This module also involved rigorous quality control to categorize cells into G1/G2 and S-phase populations. **Module 3** was dedicated to replication timing analysis using S-phase cell data, integrating the Kronos.scRT tool. **Module 4** focused on preprocessing and quality control of single-cell RNA sequencing data, utilizing the Seurat tool to ensure the selection of high-quality cells for downstream analysis.

Module 5 explored the genotype-to-phenotype relationship using the Numbat tool (T. Gao et al. 2022), which utilized phased SNPs from scWGS for genotyping scRNA-seq data and identifying clones driven by genotypic expression. Finally, **Module 6** involved the analysis of bulk WES data from treatment studies, incorporating tools like PyClone-VI and CloneEvol for analyzing clonal clusters pre- and post-treatment.

This integrative bioinformatics approach allowed us to dissect the complex clonal dynamics in neuroblastoma, providing insights into the genetic underpinnings of treatment response and resistance. The combination of diverse datasets and analytical methods underscores the intricacies of tumor evolution and the necessity of advanced computational tools to unravel them. The complete workflow is depicted in Figure 3.1.1.6.22

The definition of clones and subclones was largely based on the calling of allele and haplotype specific CNA. However, Tools such as CHISEL present a number of limitations in Allele- and Haplotype-specific analysis using tool like CHISEL. First, Germline sequencing data, either by WES or WGS gives more power to identify heterozygous SNPs. In the absence of germline, identification of heterozygous SNPs in the LoH or CN-LoH region in a given tumor sample is challenging. In the clonal LoH region(s), the identification of heterozygous SNPs becomes increasingly difficult, limiting the effectiveness of tools like CHISEL in performing allele and haplotype-specific analysis in these genomic regions. Furthermore, typical analyses often exclude chromosome X, which can lead to the oversight of subclonal alterations specific to chromosome X. As a result, potentially significant genetic changes on chromosome X might remain undetected, skewing the overall understanding of the tumor's genomic landscape. Third, a notable challenge with CHISEL analysis is its tendency to favor higher ploidy solutions, even in cases where the genome or cell is near diploid. This inclination, as also reported by (Schneider et al. 2022) can lead to inaccuracies in the determination of the clonal architecture of a tumor. Due to these limitations, our study opted for CopyKit as the primary tool for delineating clonal architecture.

The clustering of single-cell copy number data plays a pivotal role in accurately identifying the number of subclones within a tumor. To enhance the precision of subclone identification, we employed the Jaccard similarity index. This index functions as a helper tool, conducting a grid search to maximize the Jaccard Similarity across various parameter values, thereby optimizing the evaluation of the clustering process.

While this technique offers improved precision in identifying subclones, it's important to note that relying solely on data from a single time point can limit the depth of our understanding of tumor evolution. Temporal or spatial sampling, coupled with sequencing, provides a more comprehensive picture. By analyzing tumor samples at different time points or from various spatial locations within the tumor, we can gain insights into the dynamic nature of tumor evolution. This approach allows for a more stable and accurate identification of subclones, as it captures the evolutionary trajectory of the tumor, including the emergence, expansion, and potential regression of various subclonal populations over time. (Nam, Chaligne, and Landau 2021) (Black and McGranahan 2021).

Future advancements in techniques like in situ single-cell RNA sequencing (scRNA-seq) and high-throughput single-molecule fluorescence in situ hybridization (smFISH) are poised to provide experimental validation of clonal versus subclonal architecture in neuroblastoma (NB) and other cancers. smFISH, for instance, offers high spatial resolution for detecting individual RNA

molecules, allowing for precise quantification and localization of gene expression within cells and tissues (Lubeck et al. 2014). This technique, in conjunction with scRNA-seq, can greatly enhance our understanding of the heterogeneity and spatial organization of tumor cells at a single-cell level.

This study delves into the intricacies of neuroblastoma ITH, a cancer characterized by varying genomic complexities. Our analysis led to several major conclusions; By 1. distinguishing between monoclonal and polyclonal tumors, we gain vital insights into the diverse evolutionary paths. In monoclonal neuroblastoma, we observe a linear evolution pattern, where tumors maintain a consistent genetic profile over time. This clonal stasis, particularly prevalent in tumors diagnosed at a later stage, sheds light on the stability and persistence of certain genetic configurations within the tumor. Conversely, polyclonal neuroblastoma exhibits more dynamic evolutionary models, including branched and punctuated evolution. These patterns suggest early seeding or parallel evolution, where distinct subclones evolve simultaneously but independently, driven by unique molecular mechanisms. 2. Another key aspect of our study is the examination of replication timing variability among subclones within the same tumor. This approach reveals how different subclones may exhibit early or late replication timings, influencing their growth rate, genetic stability, and response to therapeutic interventions. 3. The study also explores the relationship between genetic variations and transcriptomic expression. We investigate how certain subclones, due to their unique genetic makeup and replication timing, preferentially express certain genetic alterations. This differential expression can lead to varied sensitivities to treatments among genetically distinct clones within the same tumor. 4. By integrating these findings, our study contributes significantly to the understanding of neuroblastoma's heterogeneity and its implications for treatment. Identifying subclones with distinct replication timings and transcriptomic profiles can guide the development of more targeted and effective therapeutic strategies, potentially improving patient outcomes in this complex and challenging disease.

5 CONCLUSION

Our comprehensive study on neuroblastoma intratumoral heterogeneity (ITH) has shed new light on the intricate dynamics of tumor evolution and the impact of therapy at the subclonal level. Utilizing a variety of patient-derived xenografts (PDX) models, four patient tumors at diagnosis, and an array of single-cell sequencing techniques, we have unearthed a complex narrative of tumor behavior and response.

We identified three distinct patterns of clonal evolution within neuroblastoma: linear, branched, and punctuated evolution. Linear evolution, often observed in monoclonal tumors, tended to display a phase of genomic stability, especially in cases diagnosed at a later stage. This finding points towards a pivotal phase in tumor development where genomic alterations reach a plateau, potentially offering a strategic moment for therapeutic intervention.

In contrast, branched evolution, which was prominent in cases that exhibited early seeding for progression and relapse, showcased the adaptive capabilities of tumor cells. This pattern emphasizes the urgency of initiating early and comprehensive treatment strategies. The punctuated evolution model, marked by abrupt genomic shifts, posed a significant challenge, especially in interpreting relapse cases lacking paired diagnostic samples. This evolutionary model accentuates the necessity for continuous and adaptive treatment approaches, catering to the tumor's ability to rapidly evolve and develop new resistance mechanisms.

A key revelation of our study was the influential role of subclonal populations. Even when present in minority, these subclones demonstrated notable transcriptomic activity and growth potential, significantly influencing the tumor's behavior and response to treatment. Instances where subclones with early replication timing showed enhanced proliferative and transcriptional activity, as compared to their late-replicating counterparts, possibly rendering them more susceptible to certain therapeutic interventions were particularly illustrative.

Our approach to analyzing this complex dataset involved an intricate bioinformatics workflow, integrating over ten advanced tools. This multifaceted analysis brought to light the interplay between genetic alterations, evolutionary models, replication timing, and transcriptional outcomes, thereby providing profound insights into the molecular mechanisms driving tumor evolution and response to treatment. This emphasizes the need for personalized and dynamic treatment strategies in neuroblastoma.

Furthermore, our findings indicate that CRG/EMG variations with potential functional impacts are present in a notable percentage of neuroblastoma patients. The involvement of genes such as SMARCA4 in neuroblastoma development is particularly noteworthy. While genetic alterations in ALK, including clonal mutations and amplifications, emerged as independent predictors of poorer survival in high-risk neuroblastoma, the identification of a substantial number of subclonal ALK mutations also highlights a risk of these mutations evolving during tumor progression.

In summary, our extensive research enriches our understanding of the genetic heterogeneity and evolutionary landscape of neuroblastoma. It underscores the importance of integrating multimodal data analysis in developing more targeted and effective therapeutic strategies, tailored to the complex nature of this disease.

REFERENCES

Ackermann, Sandra, Maria Cartolano, Barbara Hero, Anne Welte, Yvonne Kahlert, Andrea Roderwieser, Christoph Bartenhagen, et al. 2018. 'A Mechanistic Classification of Clinical Phenotypes in Neuroblastoma'. *Science* 362 (6419): 1165–70.

<https://doi.org/10.1126/science.aat6768>.

Adey, Andrew, Hilary G. Morrison, null Asan, Xu Xun, Jacob O. Kitzman, Emily H. Turner, Bethany Stackhouse, et al. 2010. 'Rapid, Low-Input, Low-Bias Construction of Shotgun Fragment Libraries by High-Density in Vitro Transposition'. *Genome Biology* 11 (12): R119.

<https://doi.org/10.1186/gb-2010-11-12-r119>.

Akerman, Ildem, Bahar Kasaai, Alina Bazarova, Pau Biak Sang, Isabelle Peiffer, Marie Artufel, Romain Derelle, et al. 2020. 'A Predictable Conserved DNA Base Composition Signature Defines Human Core DNA Replication Origins'. *Nature Communications* 11 (1): 4826.

<https://doi.org/10.1038/s41467-020-18527-0>.

Al Bakir, Maise, Ariana Huebner, Carlos Martínez-Ruiz, Kristiana Grigoriadis, Thomas B. K. Watkins, Oriol Pich, David A. Moore, et al. 2023. 'The Evolution of Non-Small Cell Lung Cancer Metastases in TRACERx'. *Nature*, April. <https://doi.org/10.1038/s41586-023-05729-x>.

Alexandrov, Ludmil B., Jaegil Kim, Nicholas J. Haradhvala, Mi Ni Huang, Alvin Wei Tian Ng, Yang Wu, Arnoud Boot, et al. 2020. 'The Repertoire of Mutational Signatures in Human Cancer'. *Nature* 578 (7793): 94–101. <https://doi.org/10.1038/s41586-020-1943-3>.

Alfieri, Fabio, Giulio Caravagna, and Martin H. Schaefer. 2023. 'Cancer Genomes Tolerate Deleterious Coding Mutations through Somatic Copy Number Amplifications of Wild-Type Regions'. *Nature Communications* 14 (1): 3594. <https://doi.org/10.1038/s41467-023-39313-8>.

Amiel, Jeanne, Béatrice Laudier, Tania Attié-Bitach, Ha Trang, Loïc De Pontual, Blanca Gener, Delphine Trochet, et al. 2003. 'Polyalanine Expansion and Frameshift Mutations of the Paired-like Homeobox Gene PHOX2B in Congenital Central Hypoventilation Syndrome'. *Nature Genetics* 33 (4): 459–61. <https://doi.org/10.1038/ng1130>.

Andersson, Natalie, Bjorn Bakker, Jenny Karlsson, Anders Valind, Linda Holmquist Mengelbier, Diana C.J. Spierings, Floris Fojer, and David Gisselsson. 2020. 'Extensive Clonal Branching Shapes the Evolutionary History of High-Risk Pediatric Cancers'. *Cancer Research* 80 (7): 1512–23. <https://doi.org/10.1158/0008-5472.CAN-19-3468>.

- Andor, Noemi, Billy T Lau, Claudia Catalanotti, Anuja Sathe, Matthew Kubit, Jiamin Chen, Cristina Blaj, et al. 2020. 'Joint Single Cell DNA-Seq and RNA-Seq of Gastric Cancer Cell Lines Reveals Rules of in Vitro Evolution'. *NAR Genomics and Bioinformatics* 2 (2): lqaa016. <https://doi.org/10.1093/nargab/lqaa016>.
- Baca, Sylvan C., Davide Prandi, Michael S. Lawrence, Juan Miguel Mosquera, Alessandro Romanel, Yotam Drier, Kyung Park, et al. 2013. 'Punctuated Evolution of Prostate Cancer Genomes'. *Cell* 153 (3): 666–77. <https://doi.org/10.1016/j.cell.2013.03.021>.
- Baslan, Timour, Jude Kendall, Linda Rodgers, Hilary Cox, Mike Riggs, Asya Stepansky, Jennifer Troge, et al. 2016. 'Erratum: Corrigendum: Genome-Wide Copy Number Analysis of Single Cells'. *Nature Protocols* 11 (3): 616–616. <https://doi.org/10.1038/nprot0316.616b>.
- Beerenwinkel, Niko, Roland F. Schwarz, Moritz Gerstung, and Florian Markowetz. 2015. 'Cancer Evolution: Mathematical Models and Computational Inference'. *Systematic Biology* 64 (1): e1–25. <https://doi.org/10.1093/sysbio/syu081>.
- Bellini, Angela, Ulrike Pötschger, Virginie Bernard, Eve Lapouble, Sylvain Baulande, Peter F. Ambros, Nathalie Auger, et al. 2021. 'Frequency and Prognostic Impact of *ALK* Amplifications and Mutations in the European Neuroblastoma Study Group (SIOPEN) High-Risk Neuroblastoma Trial (HR-NBL1)'. *Journal of Clinical Oncology* 39 (30): 3377–90. <https://doi.org/10.1200/JCO.21.00086>.
- Ben-David, Uri, Rameen Beroukhim, and Todd R. Golub. 2019. 'Genomic Evolution of Cancer Models: Perils and Opportunities'. *Nature Reviews Cancer* 19 (2): 97–109. <https://doi.org/10.1038/s41568-018-0095-3>.
- Berbegall, A P, E Villamón, M Piqueras, I Tadeo, A Djios, P F Ambros, T Martinsson, et al. 2016. 'Comparative Genetic Study of Intratumoral Heterogenous MYCN Amplified Neuroblastoma versus Aggressive Genetic Profile Neuroblastic Tumors'. *Oncogene* 35 (11): 1423–32. <https://doi.org/10.1038/onc.2015.200>.
- Biegel, J. A., P. S. White, H. N. Marshall, M. Fujimori, E. H. Zackai, C. D. Scher, G. M. Brodeur, and B. S. Emanuel. 1993. 'Constitutional 1p36 Deletion in a Child with Neuroblastoma'. *American Journal of Human Genetics* 52 (1): 176–82.
- Black, James R. M., and Nicholas McGranahan. 2021. 'Genetic and Non-Genetic Clonal Diversity in Cancer Evolution'. *Nature Reviews Cancer* 21 (6): 379–92. <https://doi.org/10.1038/s41568-021-00336-2>.

- Boeva, Valentina, Caroline Louis-Brennetot, Agathe Peltier, Simon Durand, Cécile Pierre-Eugène, Virginie Raynal, Heather C Etchevers, et al. 2017. 'Heterogeneity of Neuroblastoma Cell Identity Defined by Transcriptional Circuitries'. *Nature Genetics* 49 (9): 1408–13. <https://doi.org/10.1038/ng.3921>.
- Bourdeaut, Franck, Delphine Trochet, Isabelle Janoueix-Lerosey, Agnès Ribeiro, Anne Deville, Carole Coz, Jean-François Michiels, Stanislas Lyonnet, Jeanne Amiel, and Olivier Delattre. 2005. 'Germline Mutations of the Paired-like Homeobox 2B (PHOX2B) Gene in Neuroblastoma'. *Cancer Letters* 228 (1–2): 51–58. <https://doi.org/10.1016/j.canlet.2005.01.055>.
- Brady, Samuel W., Yanling Liu, Xiaotu Ma, Alexander M. Gout, Kohei Hagiwara, Xin Zhou, Jian Wang, et al. 2020. 'Pan-Neuroblastoma Analysis Reveals Age- and Signature-Associated Driver Alterations'. *Nature Communications* 11 (1): 5183. <https://doi.org/10.1038/s41467-020-18987-4>.
- Braekeveldt, Noémie, Kristoffer Von Stedingk, Susanne Fransson, Angela Martinez-Monleon, David Lindgren, Håkan Axelsson, Fredrik Levander, et al. 2018. 'Patient-Derived Xenograft Models Reveal Intratumor Heterogeneity and Temporal Stability in Neuroblastoma'. *Cancer Research* 78 (20): 5958–69. <https://doi.org/10.1158/0008-5472.CAN-18-0527>.
- Brennan, Cameron W., Roel G.W. Verhaak, Aaron McKenna, Benito Campos, Houtan Noshmehr, Sofie R. Salama, Siyuan Zheng, et al. 2013. 'The Somatic Genomic Landscape of Glioblastoma'. *Cell* 155 (2): 462–77. <https://doi.org/10.1016/j.cell.2013.09.034>.
- Briu, Lina-Marie, Chrystelle Maric, and Jean-Charles Cadoret. 2021. 'Replication Stress, Genomic Instability, and Replication Timing: A Complex Relationship'. *International Journal of Molecular Sciences* 22 (9): 4764. <https://doi.org/10.3390/ijms22094764>.
- Brodeur, Garrett M., Robert C. Seeger, Manfred Schwab, Harold E. Varmus, and J. Michael Bishop. 1984. 'Amplification of N- *Myc* in Untreated Human Neuroblastomas Correlates with Advanced Disease Stage'. *Science* 224 (4653): 1121–24. <https://doi.org/10.1126/science.6719137>.
- Campbell, Kieran R., Adi Steif, Emma Laks, Hans Zahn, Daniel Lai, Andrew McPherson, Hossein Farahani, et al. 2019. 'Clonealign: Statistical Integration of Independent Single-Cell RNA and DNA Sequencing Data from Human Cancers'. *Genome Biology* 20 (1): 54. <https://doi.org/10.1186/s13059-019-1645-z>.
- Caravagna, Giulio, Timon Heide, Marc J. Williams, Luis Zapata, Daniel Nichol, Ketevan Chkhaidze, William Cross, et al. 2020. 'Subclonal Reconstruction of Tumors by Using Machine

Learning and Population Genetics'. *Nature Genetics* 52 (9): 898–907.

<https://doi.org/10.1038/s41588-020-0675-5>.

Caravagna, Giulio, Guido Sanguinetti, Trevor A. Graham, and Andrea Sottoriva. 2020. 'The MOBSTER R Package for Tumour Subclonal Deconvolution from Bulk DNA Whole-Genome Sequencing Data'. *BMC Bioinformatics* 21 (1): 531. <https://doi.org/10.1186/s12859-020-03863-1>.

Cheloni, Stefano, Roman Hillje, Lucilla Luzi, Pier Giuseppe Pelicci, and Elena Gatti. 2021. 'XenoCell: Classification of Cellular Barcodes in Single Cell Experiments from Xenograft Samples'. *BMC Medical Genomics* 14 (1): 34. <https://doi.org/10.1186/s12920-021-00872-8>.

Cheung, Nai-Kong V., and Michael A. Dyer. 2013. 'Neuroblastoma: Developmental Biology, Cancer Genomics and Immunotherapy'. *Nature Reviews Cancer* 13 (6): 397–411. <https://doi.org/10.1038/nrc3526>.

Chicard, Mathieu, Leo Colmet-Daage, Nathalie Clement, Adrien Danzon, Mylène Bohec, Virginie Bernard, Sylvain Baulande, et al. 2018. 'Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-Spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma'. *Clinical Cancer Research* 24 (4): 939–49. <https://doi.org/10.1158/1078-0432.CCR-17-1586>.

Cingolani, Pablo, Adrian Platts, Le Lily Wang, Melissa Coon, Tung Nguyen, Luan Wang, Susan J. Land, Xiangyi Lu, and Douglas M. Ruden. 2012. 'A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of *Drosophila Melanogaster* Strain w¹¹¹⁸; Iso-2; Iso-3'. *Fly* 6 (2): 80–92. <https://doi.org/10.4161/fly.19695>.

Coorens, Tim H.H., Sarah J. Farndon, Thomas J. Mitchell, Neha Jain, Sangjin Lee, Michael Hubank, Neil Sebire, John Anderson, and Sam Behjati. 2020. 'Lineage-Independent Tumors in Bilateral Neuroblastoma'. *New England Journal of Medicine* 383 (19): 1860–65. <https://doi.org/10.1056/NEJMoa2000962>.

Cross, William, Michal Kovac, Ville Mustonen, Daniel Temko, Hayley Davis, Ann-Marie Baker, Sujata Biswas, et al. 2018. 'The Evolutionary Landscape of Colorectal Tumorigenesis'. *Nature Ecology & Evolution* 2 (10): 1661–72. <https://doi.org/10.1038/s41559-018-0642-z>.

Davis, Alexander, Ruli Gao, and Nicholas Navin. 2017. 'Tumor Evolution: Linear, Branching, Neutral or Punctuated?' *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1867 (2): 151–61. <https://doi.org/10.1016/j.bbcan.2017.01.003>.

Dentro, Stefan C., Ignaty Leshchiner, Kerstin Haase, Maxime Tarabichi, Jeff Wintersinger, Amit G. Deshwar, Kaixian Yu, et al. 2021. ‘Characterizing Genetic Intra-Tumor Heterogeneity across 2,658 Human Cancer Genomes’. *Cell* 184 (8): 2239-2254.e39.

<https://doi.org/10.1016/j.cell.2021.03.009>.

DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. ‘A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data’. *Nature Genetics* 43 (5): 491–98.

<https://doi.org/10.1038/ng.806>.

Dey, Siddharth S, Lennart Kester, Bastiaan Spanjaard, Magda Bienko, and Alexander Van Oudenaarden. 2015. ‘Integrated Genome and Transcriptome Sequencing of the Same Cell’. *Nature Biotechnology* 33 (3): 285–89. <https://doi.org/10.1038/nbt.3129>.

Dileep, Vishnu, and David M. Gilbert. 2018. ‘Single-Cell Replication Profiling to Measure Stochastic Variation in Mammalian Replication Timing’. *Nature Communications* 9 (1): 427.

<https://doi.org/10.1038/s41467-017-02800-w>.

Egolf, Laura E., Zalman Vaksman, Gonzalo Lopez, Jo Lynne Rokita, Apexa Modi, Patricia V. Basta, Hakon Hakonarson, Andrew F. Olshan, and Sharon J. Diskin. 2019. ‘Germline 16p11.2 Microdeletion Predisposes to Neuroblastoma’. *American Journal of Human Genetics* 105 (3): 658–68. <https://doi.org/10.1016/j.ajhg.2019.07.020>.

Eleveld, Thomas F, Derek A Oldridge, Virginie Bernard, Jan Koster, Leo Colmet Daage, Sharon J Diskin, Linda Schild, et al. 2015. ‘Relapsed Neuroblastomas Show Frequent RAS-MAPK Pathway Mutations’. *Nature Genetics* 47 (8): 864–71. <https://doi.org/10.1038/ng.3333>.

Fan, Jean, Hae-Ock Lee, Soohyun Lee, Da-eun Ryu, Semin Lee, Catherine Xue, Seok Jin Kim, et al. 2018. ‘Linking Transcriptional and Genetic Tumor Heterogeneity through Allele Analysis of Single-Cell RNA-Seq Data’. *Genome Research* 28 (8): 1217–27.

<https://doi.org/10.1101/gr.228080.117>.

Favero, F., T. Joshi, A.M. Marquard, N.J. Birkbak, M. Krzystanek, Q. Li, Z. Szallasi, and A.C. Eklund. 2015. ‘Sequenza: Allele-Specific Copy Number and Mutation Profiles from Tumor Sequencing Data’. *Annals of Oncology* 26 (1): 64–70. <https://doi.org/10.1093/annonc/mdu479>.

Fernando, Tharu M., Robert Piskol, Russell Bainer, Ethan S. Sokol, Sally E. Trabucco, Qing Zhang, Huong Trinh, et al. 2020. ‘Functional Characterization of SMARCA4 Variants Identified

by Targeted Exome-Sequencing of 131,668 Cancer Patients'. *Nature Communications* 11 (1): 5551. <https://doi.org/10.1038/s41467-020-19402-8>.

Field, Matthew G., Michael A. Durante, Hima Anbunathan, Louis Z. Cai, Christina L. Decatur, Anne M. Bowcock, Stefan Kurtenbach, and J. William Harbour. 2018. 'Punctuated Evolution of Canonical Genomic Aberrations in Uveal Melanoma'. *Nature Communications* 9 (1): 116. <https://doi.org/10.1038/s41467-017-02428-w>.

Fittall, Matthew W., and Peter Van Loo. 2019. 'Translating Insights into Tumor Evolution to Clinical Practice: Promises and Challenges'. *Genome Medicine* 11 (1): 20. <https://doi.org/10.1186/s13073-019-0632-z>.

Frankell, Alexander M., Michelle Dietzen, Maise Al Bakir, Emilia L. Lim, Takahiro Karasaki, Sophia Ward, Selvaraju Veeriah, et al. 2023. 'The Evolution of Lung Cancer and Impact of Subclonal Selection in TRACERx'. *Nature*, April. <https://doi.org/10.1038/s41586-023-05783-5>.

Funnell, Tyler, Ciara H. O'Flanagan, Marc J. Williams, Andrew McPherson, Steven McKinney, Farhia Kabeer, Hakwoo Lee, et al. 2022. 'Single-Cell Genomic Variation Induced by Mutational Processes in Cancer'. *Nature*, October. <https://doi.org/10.1038/s41586-022-05249-0>.

Gao, Ruli, Shanshan Bai, Ying C. Henderson, Yiyun Lin, Aislyn Schalck, Yun Yan, Tapsi Kumar, et al. 2021. 'Delineating Copy Number and Clonal Substructure in Human Tumors from Single-Cell Transcriptomes'. *Nature Biotechnology* 39 (5): 599–608. <https://doi.org/10.1038/s41587-020-00795-2>.

Gao, Ruli, Alexander Davis, Thomas O McDonald, Emi Sei, Xiuqing Shi, Yong Wang, Pei-Ching Tsai, et al. 2016. 'Punctuated Copy Number Evolution and Clonal Stasis in Triple-Negative Breast Cancer'. *Nature Genetics* 48 (10): 1119–30. <https://doi.org/10.1038/ng.3641>.

Gao, Teng, Ruslan Soldatov, HIRAK SARKAR, Adam Kurkiewicz, Evan Biederstedt, Po-Ru Loh, and Peter V. Kharchenko. 2022. 'Haplotype-Aware Analysis of Somatic Copy Number Variations from Single-Cell Transcriptomes'. *Nature Biotechnology*, September. <https://doi.org/10.1038/s41587-022-01468-y>.

Gatenby, Robert A., and Thomas L. Vincent. 2003. 'An Evolutionary Model of Carcinogenesis'. *Cancer Research* 63 (19): 6212–20.

Gemble, Simon, René Wardenaar, Kristina Keuper, Nishit Srivastava, Maddalena Nano, Anne-Sophie Macé, Andréa E. Tijhuis, et al. 2022. 'Genetic Instability from a Single S Phase after

Whole-Genome Duplication'. *Nature* 604 (7904): 146–51. <https://doi.org/10.1038/s41586-022-04578-4>.

George, Rani E., Wendy B. London, Susan L. Cohn, John M. Maris, Cynthia Kretschmar, Lisa Diller, Garrett M. Brodeur, Robert P. Castleberry, and A. Thomas Look. 2005. 'Hyperdiploidy Plus Nonamplified *MYCN* Confers a Favorable Prognosis in Children 12 to 18 Months Old With Disseminated Neuroblastoma: A Pediatric Oncology Group Study'. *Journal of Clinical Oncology* 23 (27): 6466–73. <https://doi.org/10.1200/JCO.2005.05.582>.

Gerlinger, Marco, Nicholas McGranahan, Sally M. Dewhurst, Rebecca A. Burrell, Ian Tomlinson, and Charles Swanton. 2014. 'Cancer: Evolution Within a Lifetime'. *Annual Review of Genetics* 48 (1): 215–36. <https://doi.org/10.1146/annurev-genet-120213-092314>.

Gerlinger, Marco, Andrew J. Rowan, Stuart Horswell, James Larkin, David Endesfelder, Eva Gronroos, Pierre Martinez, et al. 2012. 'Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing'. *New England Journal of Medicine* 366 (10): 883–92. <https://doi.org/10.1056/NEJMoa1113205>.

Gerstung, Moritz, Clemency Jolly, Ignaty Leshchiner, Stefan C. Dentre, Santiago Gonzalez, Daniel Rosebrock, Thomas J. Mitchell, et al. 2020. 'The Evolutionary History of 2,658 Cancers'. *Nature* 578 (7793): 122–28. <https://doi.org/10.1038/s41586-019-1907-7>.

Gillis, Sierra, and Andrew Roth. 2020. 'PyClone-VI: Scalable Inference of Clonal Population Structures Using Whole Genome Data'. *BMC Bioinformatics* 21 (1): 571. <https://doi.org/10.1186/s12859-020-03919-2>.

Gnan, Stefano, Joseph M. Josephides, Xia Wu, Manuela Spagnuolo, Dalila Saulebekova, Mylène Bohec, Marie Dumont, et al. 2022. 'Kronos scRT: A Uniform Framework for Single-Cell Replication Timing Analysis'. *Nature Communications* 13 (1): 2329. <https://doi.org/10.1038/s41467-022-30043-x>.

Gogolin, Sina, Volker Ehemann, Gabriele Becker, Lena M. Brueckner, Daniel Dreidax, Steffen Bannert, Ingo Nolte, Larissa Savelyeva, Emma Bell, and Frank Westermann. 2013. 'CDK4 Inhibition Restores G₁-S Arrest in *MYCN* - Amplified Neuroblastoma Cells in the Context of Doxorubicin-Induced DNA Damage'. *Cell Cycle* 12 (7): 1091–1104. <https://doi.org/10.4161/cc.24091>.

Gröbner, Susanne N., Barbara C. Worst, Joachim Weischenfeldt, Ivo Buchhalter, Kortine Kleinheinz, Vasilisa A. Rudneva, Pascal D. Johann, et al. 2018. 'The Landscape of Genomic

Alterations across Childhood Cancers'. *Nature* 555 (7696): 321–27.

<https://doi.org/10.1038/nature25480>.

Gundem, Gunes, Max F. Levine, Stephen S. Roberts, Irene Y. Cheung, Juan S. Medina-Martínez, Yi Feng, Juan E. Arango-Ossa, et al. 2023. 'Clonal Evolution during Metastatic Spread in High-Risk Neuroblastoma'. *Nature Genetics* 55 (6): 1022–33. <https://doi.org/10.1038/s41588-023-01395-x>.

Hanahan, Douglas. 2022. 'Hallmarks of Cancer: New Dimensions'. *Cancer Discovery* 12 (1): 31–46. <https://doi.org/10.1158/2159-8290.CD-21-1059>.

Hao, Yuhan, Stephanie Hao, Erica Andersen-Nissen, William M. Mauck, Shiwei Zheng, Andrew Butler, Maddie J. Lee, et al. 2021. 'Integrated Analysis of Multimodal Single-Cell Data'. *Cell* 184 (13): 3573-3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.

Hennig, C. 2020. 'Hennig, C. Fpc: Flexible Procedures for Clustering. R Package Version 2.2-9. (2020).'

Hiley, Crispin, Elza C de Bruin, Nicholas McGranahan, and Charles Swanton. 2014.

'Deciphering Intratumor Heterogeneity and Temporal Acquisition of Driver Events to Refine Precision Medicine'. *Genome Biology* 15 (8): 453. <https://doi.org/10.1186/s13059-014-0453-8>.

Hiratani, Ichiro, and Saori Takahashi. 2019. 'DNA Replication Timing Enters the Single-Cell Era'. *Genes* 10 (3): 221. <https://doi.org/10.3390/genes10030221>.

Hua, Xing, Wei Zhao, Angela C. Pesatori, Dario Consonni, Neil E. Caporaso, Tongwu Zhang, Bin Zhu, et al. 2020. 'Genetic and Epigenetic Intratumor Heterogeneity Impacts Prognosis of Lung Adenocarcinoma'. *Nature Communications* 11 (1): 2459. <https://doi.org/10.1038/s41467-020-16295-5>.

Huang, Xianjie, and Yuanhua Huang. 2021. 'Cellsnp-Lite: An Efficient Tool for Genotyping Single Cells'. Edited by Can Alkan. *Bioinformatics* 37 (23): 4569–71. <https://doi.org/10.1093/bioinformatics/btab358>.

Hwang, Michael S., Brian J. Mog, Jacqueline Douglass, Alexander H. Pearlman, Emily Han-Chung Hsiue, Suman Paul, Sarah R. DiNapoli, et al. 2021. 'Targeting Loss of Heterozygosity for Cancer-Specific Immunotherapy'. *Proceedings of the National Academy of Sciences* 118 (12): e2022410118. <https://doi.org/10.1073/pnas.2022410118>.

Jamal-Hanjani, Mariam, Gareth A. Wilson, Nicholas McGranahan, Nicolai J. Birkbak, Thomas B.K. Watkins, Selvaraju Veeriah, Seema Shafi, et al. 2017. 'Tracking the Evolution of Non–

Small-Cell Lung Cancer’. *New England Journal of Medicine* 376 (22): 2109–21.

<https://doi.org/10.1056/NEJMoa1616288>.

Janoueix-Lerosey, Isabelle, Philippe Hupé, Zofia Maciorowski, Philippe La Rosa, Gudrun Schleiermacher, Gaëlle Pierron, Stéphane Liva, Emmanuel Barillot, and Olivier Delattre. 2005.

‘Preferential Occurrence of Chromosome Breakpoints within Early Replicating Regions in Neuroblastoma’. *Cell Cycle* 4 (12): 1842–46. <https://doi.org/10.4161/cc.4.12.2257>.

Janoueix-Lerosey, Isabelle, Gudrun Schleiermacher, Evi Michels, Véronique Mosseri, Agnès Ribeiro, Delphine Lequin, Joëlle Vermeulen, et al. 2009. ‘Overall Genomic Pattern Is a Predictor of Outcome in Neuroblastoma’. *Journal of Clinical Oncology* 27 (7): 1026–33.

<https://doi.org/10.1200/JCO.2008.16.0630>.

Jansky, Selina, Ashwini Kumar Sharma, Verena Körber, Andrés Quintero, Umut H. Toprak, Elisa M. Wecht, Moritz Gartlgruber, et al. 2021. ‘Single-Cell Transcriptomic Analyses Provide Insights into the Developmental Origins of Neuroblastoma’. *Nature Genetics* 53 (5): 683–93.

<https://doi.org/10.1038/s41588-021-00806-1>.

Jiménez, Carlos, Roberta Antonelli, Mariona Nadal-Ribelles, Laura Devis-Jauregui, Pablo Latorre, Carme Solé, Marc Masanas, et al. 2022. ‘Structural Disruption of BAF Chromatin Remodeller Impairs Neuroblastoma Metastasis by Reverting an Invasiveness Epigenomic Program’. *Molecular Cancer* 21 (1): 175. <https://doi.org/10.1186/s12943-022-01643-4>.

Karlsson, Jenny, Anders Valind, Linda Holmquist Mengelbier, Sofia Bredin, Louise Cornmark, Caroline Jansson, Amina Wali, et al. 2018. ‘Four Evolutionary Trajectories Underlie Genetic Intratumoral Variation in Childhood Cancer’. *Nature Genetics* 50 (7): 944–50.

<https://doi.org/10.1038/s41588-018-0131-y>.

Kaufmann, Tom L., Marina Petkovic, Thomas B. K. Watkins, Emma C. Colliver, Sofya Laskina, Nisha Thapa, Darlan C. Minussi, et al. 2022. ‘MEDICC2: Whole-Genome Doubling Aware Copy-Number Phylogenies for Cancer Evolution’. *Genome Biology* 23 (1): 241.

<https://doi.org/10.1186/s13059-022-02794-9>.

Kim, Charissa, Ruli Gao, Emi Sei, Rachel Brandt, Johan Hartman, Thomas Hatschek, Nicola Crosetto, Theodoros Foukakis, and Nicholas E. Navin. 2018. ‘Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing’. *Cell* 173 (4): 879–893.e13.

<https://doi.org/10.1016/j.cell.2018.03.041>.

- Körber, Verena, Sabine A. Stainczyk, Roma Kurilov, Kai-Oliver Henrich, Barbara Hero, Benedikt Brors, Frank Westermann, and Thomas Höfer. 2023. 'Neuroblastoma Arises in Early Fetal Development and Its Evolutionary Duration Predicts Outcome'. *Nature Genetics*, March. <https://doi.org/10.1038/s41588-023-01332-y>.
- Krupina, Ksenia, Alexander Goginashvili, and Don W. Cleveland. 2021. 'Causes and Consequences of Micronuclei'. *Current Opinion in Cell Biology* 70 (June): 91–99. <https://doi.org/10.1016/j.ceb.2021.01.004>.
- La Manno, Gioele, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, et al. 2018. 'RNA Velocity of Single Cells'. *Nature* 560 (7719): 494–98. <https://doi.org/10.1038/s41586-018-0414-6>.
- Laks, Emma, Andrew McPherson, Hans Zahn, Daniel Lai, Adi Steif, Jazmine Brimhall, Justina Biele, et al. 2019. 'Clonal Decomposition and DNA Replication States Defined by Scaled Single-Cell Genome Sequencing'. *Cell* 179 (5): 1207-1221.e22. <https://doi.org/10.1016/j.cell.2019.10.026>.
- Leighton, Jake, Min Hu, Emi Sei, Funda Meric-Bernstam, and Nicholas E. Navin. 2021. 'Reconstructing Mutational Lineages in Breast Cancer by Multi-Patient-Targeted Single Cell DNA Sequencing'. Preprint. *Cancer Biology*. <https://doi.org/10.1101/2021.11.16.468877>.
- Li, H., and R. Durbin. 2009. 'Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform'. *Bioinformatics* 25 (14): 1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
- Ling, Shaoping, Zheng Hu, Zuyu Yang, Fang Yang, Yawei Li, Pei Lin, Ke Chen, et al. 2015. 'Extremely High Genetic Diversity in a Single Tumor Points to Prevalence of Non-Darwinian Cell Evolution'. *Proceedings of the National Academy of Sciences* 112 (47). <https://doi.org/10.1073/pnas.1519556112>.
- Liu, Yu, John Easton, Ying Shao, Jamie Maciaszek, Zhaoming Wang, Mark R. Wilkinson, Kelly McCastlain, et al. 2017. 'The Genomic Landscape of Pediatric and Young Adult T-Lineage Acute Lymphoblastic Leukemia'. *Nature Genetics* 49 (8): 1211–18. <https://doi.org/10.1038/ng.3909>.
- Loh, Po-Ru, Petr Danecek, Pier Francesco Palamara, Christian Fuchsberger, Yakir A Reshef, Hilary K Finucane, Sebastian Schoenherr, et al. 2016. 'Reference-Based Phasing Using the Haplotype Reference Consortium Panel'. *Nature Genetics* 48 (11): 1443–48. <https://doi.org/10.1038/ng.3679>.

López-Carrasco, Amparo, Ana P. Berbegall, Susana Martín-Vañó, Maite Blanquer-Maceiras, Victoria Castel, Samuel Navarro, and Rosa Noguera. 2021. ‘Intra-Tumour Genetic Heterogeneity and Prognosis in High-Risk Neuroblastoma’. *Cancers* 13 (20): 5173.

<https://doi.org/10.3390/cancers13205173>.

Lubeck, Eric, Ahmet F Coskun, Timur Zhiyentayev, Mubhij Ahmad, and Long Cai. 2014. ‘Single-Cell in Situ RNA Profiling by Sequential Hybridization’. *Nature Methods* 11 (4): 360–61.

<https://doi.org/10.1038/nmeth.2892>.

Ludmil B. Alexandrov, Serena Nik-Zainal, David C. Wedge, Samuel A. J. R. Aparicio, Sam Behjati, Andrew V. Biankin, Graham R. Bignell, et al. 2013. ‘Signatures of Mutational Processes in Human Cancer’. *Nature* 500 (7463): 415–21. <https://doi.org/10.1038/nature12477>.

Macaulay, Iain C, Wilfried Haerty, Parveen Kumar, Yang I Li, Tim Xiaoming Hu, Mabel J Teng, Mubeen Goolam, et al. 2015. ‘G&T-Seq: Parallel Sequencing of Single-Cell Genomes and Transcriptomes’. *Nature Methods* 12 (6): 519–22. <https://doi.org/10.1038/nmeth.3370>.

Mallory, Xian F., Mohammadamin Edrisi, Nicholas Navin, and Luay Nakhleh. 2020. ‘Methods for Copy Number Aberration Detection from Single-Cell DNA-Sequencing Data’. *Genome Biology* 21 (1): 208. <https://doi.org/10.1186/s13059-020-02119-8>.

Maris, John M, Michael D Hogarty, Rochelle Bagatell, and Susan L Cohn. 2007. ‘Neuroblastoma’. *The Lancet* 369 (9579): 2106–20. [https://doi.org/10.1016/S0140-6736\(07\)60983-0](https://doi.org/10.1016/S0140-6736(07)60983-0).

Marques Da Costa, Maria Eugénia, Sakina Zaidi, Jean-Yves Scoazec, Robin Droit, Wan Ching Lim, Antonin Marchais, Jerome Salmon, et al. 2023. ‘A Biobank of Pediatric Patient-Derived Xenograft Models in Cancer Precision Medicine Trial MAPPYACTS for Relapsed and Refractory Tumors’. *Communications Biology* 6 (1): 949. <https://doi.org/10.1038/s42003-023-05320-0>.

Martínez-Ruiz, Carlos, James R. M. Black, Clare Puttick, Mark S. Hill, Jonas Demeulemeester, Elizabeth Larose Cadieux, Kerstin Thol, et al. 2023. ‘Genomic–Transcriptomic Evolution in Lung Cancer and Metastasis’. *Nature*, April. <https://doi.org/10.1038/s41586-023-05706-4>.

Matthay, Katherine K., John M. Maris, Gudrun Schleiermacher, Akira Nakagawara, Crystal L. Mackall, Lisa Diller, and William A. Weiss. 2016. ‘Neuroblastoma’. *Nature Reviews Disease Primers* 2 (1): 16078. <https://doi.org/10.1038/nrdp.2016.78>.

- McCarthy, Davis J., Raghd Rostom, Yuanhua Huang, Daniel J. Kunz, Petr Danecek, Marc Jan Bonder, Tzachi Hagai, et al. 2020. 'Cardelino: Computational Integration of Somatic Clonal Substructure and Single-Cell Transcriptomes'. *Nature Methods* 17 (4): 414–21. <https://doi.org/10.1038/s41592-020-0766-3>.
- McGranahan, Nicholas, Francesco Favero, Elza C. De Bruin, Nicolai Juul Birkbak, Zoltan Szallasi, and Charles Swanton. 2015. 'Clonal Status of Actionable Driver Events and the Timing of Mutational Processes in Cancer Evolution'. *Science Translational Medicine* 7 (283). <https://doi.org/10.1126/scitranslmed.aaa1408>.
- McGranahan, Nicholas, and Charles Swanton. 2017. 'Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future'. *Cell* 168 (4): 613–28. <https://doi.org/10.1016/j.cell.2017.01.018>.
- McInnes, Leland, John Healy, and Steve Astels. 2017. 'Hdbscan: Hierarchical Density Based Clustering'. *The Journal of Open Source Software* 2 (11): 205. <https://doi.org/10.21105/joss.00205>.
- McInnes, Leland, John Healy, Nathaniel Saul, and Lukas Großberger. 2018. 'UMAP: Uniform Manifold Approximation and Projection'. *Journal of Open Source Software* 3 (29): 861. <https://doi.org/10.21105/joss.00861>.
- Minussi, Darlan C., Michael D. Nicholson, Hanghui Ye, Alexander Davis, Kaile Wang, Toby Baker, Maxime Tarabichi, et al. 2021. 'Breast Tumours Maintain a Reservoir of Subclonal Diversity during Expansion'. *Nature* 592 (7853): 302–8. <https://doi.org/10.1038/s41586-021-03357-x>.
- Minussi, Darlan Conterno, Emi Sei, Junke Wang, Aislyn Schalck, Yun Yan, Alexander Davis, Hua-Jun Wu, et al. 2022. 'Resolving Clonal Substructure from Single Cell Genomic Data Using CopyKit'. Preprint. Genomics. <https://doi.org/10.1101/2022.03.09.483497>.
- Molenaar, Jan J., Jan Koster, Marli E. Ebus, Peter van Sluis, Ellen M. Westerhout, Katleen de Preter, David Gisselsson, et al. 2012. 'Copy Number Defects of G1-Cell Cycle Genes in Neuroblastoma Are Frequent and Correlate with High Expression of E2F Target Genes and a Poor Prognosis'. *Genes, Chromosomes and Cancer* 51 (1): 10–19. <https://doi.org/10.1002/gcc.20926>.
- Monclair, Tom, Garrett M. Brodeur, Peter F. Ambros, Hervé J. Brisse, Giovanni Cecchetto, Keith Holmes, Michio Kaneko, et al. 2009. 'The International Neuroblastoma Risk Group (INRG)

Staging System: An INRG Task Force Report'. *Journal of Clinical Oncology* 27 (2): 298–303. <https://doi.org/10.1200/JCO.2008.16.6876>.

Mossé, Yaël P., Marci Laudenslager, Luca Longo, Kristina A. Cole, Andrew Wood, Edward F. Attiyeh, Michael J. Laquaglia, et al. 2008. 'Identification of ALK as a Major Familial Neuroblastoma Predisposition Gene'. *Nature* 455 (7215): 930–35. <https://doi.org/10.1038/nature07261>.

Nam, Anna S., Ronan Chaligne, and Dan A. Landau. 2021. 'Integrating Genetic and Non-Genetic Determinants of Cancer Evolution by Single-Cell Multi-Omics'. *Nature Reviews Genetics* 22 (1): 3–18. <https://doi.org/10.1038/s41576-020-0265-5>.

Navin, Nicholas, Jude Kendall, Jennifer Troge, Peter Andrews, Linda Rodgers, Jeanne McIndoo, Kerry Cook, et al. 2011. 'Tumour Evolution Inferred by Single-Cell Sequencing'. *Nature* 472 (7341): 90–94. <https://doi.org/10.1038/nature09807>.

Navin, Nicholas, Alexander Krasnitz, Linda Rodgers, Kerry Cook, Jennifer Meth, Jude Kendall, Michael Riggs, et al. 2010. 'Inferring Tumor Progression from Genomic Heterogeneity'. *Genome Research* 20 (1): 68–80. <https://doi.org/10.1101/gr.099622.109>.

Némati, Fariba, Xavier Sastre-Garau, Cécile Laurent, Jérôme Couturier, Pascale Mariani, Laurence Desjardins, Sophie Piperno-Neumann, et al. 2010. 'Establishment and Characterization of a Panel of Human Uveal Melanoma Xenografts Derived from Primary and/or Metastatic Tumors'. *Clinical Cancer Research* 16 (8): 2352–62. <https://doi.org/10.1158/1078-0432.CCR-09-3066>.

Nik-Zainal, Serena, Peter Van Loo, David C. Wedge, Ludmil B. Alexandrov, Christopher D. Greenman, King Wai Lau, Keiran Raine, et al. 2012. 'The Life History of 21 Breast Cancers'. *Cell* 149 (5): 994–1007. <https://doi.org/10.1016/j.cell.2012.04.023>.

Nowell, Peter C. 1976. 'The Clonal Evolution of Tumor Cell Populations: Acquired Genetic Lability Permits Stepwise Selection of Variant Sublines and Underlies Tumor Progression.' *Science* 194 (4260): 23–28. <https://doi.org/10.1126/science.959840>.

Patel, Anoop P., Itay Tirosh, John J. Trombetta, Alex K. Shalek, Shawn M. Gillespie, Hiroaki Wakimoto, Daniel P. Cahill, et al. 2014. 'Single-Cell RNA-Seq Highlights Intratumoral Heterogeneity in Primary Glioblastoma'. *Science* 344 (6190): 1396–1401. <https://doi.org/10.1126/science.1254257>.

- Peneder, Peter, Adrian M. Stütz, Didier Surdez, Manuela Krumbholz, Sabine Semper, Mathieu Chicard, Nathan C. Sheffield, et al. 2021. 'Multimodal Analysis of Cell-Free DNA Whole-Genome Sequencing for Pediatric Cancers with Low Mutational Burden'. *Nature Communications* 12 (1): 3230. <https://doi.org/10.1038/s41467-021-23445-w>.
- Pugh, Trevor J., Olena Morozova, Edward F. Attiyeh, Shahab Asgharzadeh, Jun S. Wei, Daniel Auclair, Scott L. Carter, et al. 2013. 'The Genetic Landscape of High-Risk Neuroblastoma'. *Nature Genetics* 45 (3): 279–84. <https://doi.org/10.1038/ng.2529>.
- Rambow, Florian, Aljosja Rogiers, Oskar Marin-Bejar, Sara Aibar, Julia Femel, Michael Dewaele, Panagiotis Karras, et al. 2018. 'Toward Minimal Residual Disease-Directed Therapy in Melanoma'. *Cell* 174 (4): 843-855.e19. <https://doi.org/10.1016/j.cell.2018.06.025>.
- Rosenthal, Rachel, Nicholas McGranahan, Javier Herrero, Barry S. Taylor, and Charles Swanton. 2016. 'deconstructSigs: Delineating Mutational Processes in Single Tumors Distinguishes DNA Repair Deficiencies and Patterns of Carcinoma Evolution'. *Genome Biology* 17 (1): 31. <https://doi.org/10.1186/s13059-016-0893-4>.
- Sanders, Mathijs A., Harald Vöhringer, Victoria J. Forster, Luiza Moore, Brittany B. Campbell, Yvette Hooks, Melissa Edwards, et al. 2021. 'Life without Mismatch Repair'. Preprint. Genomics. <https://doi.org/10.1101/2021.04.14.437578>.
- Sanz-Gómez, Natalia, María González-Álvarez, Javier De Las Rivas, and Guillermo De Cárcer. 2023. 'Whole-Genome Doubling as a Source of Cancer: How, When, Where, and Why?' *Frontiers in Cell and Developmental Biology* 11 (June): 1209136. <https://doi.org/10.3389/fcell.2023.1209136>.
- Sausen, Mark, Rebecca J. Leary, Siân Jones, Jian Wu, C. Patrick Reynolds, Xueyuan Liu, Amanda Blackford, et al. 2013. 'Integrated Genomic Analyses Identify ARID1A and ARID1B Alterations in the Childhood Cancer Neuroblastoma'. *Nature Genetics* 45 (1): 12–17. <https://doi.org/10.1038/ng.2493>.
- Schleiermacher, Gudrun, Isabelle Janoueix-Lerosey, Valérie Combaret, Josette Derré, Jérôme Couturier, Alain Aurias, and Olivier Delattre. 2003. 'Combined 24-Color Karyotyping and Comparative Genomic Hybridization Analysis Indicates Predominant Rearrangements of Early Replicating Chromosome Regions in Neuroblastoma'. *Cancer Genetics and Cytogenetics* 141 (1): 32–42. [https://doi.org/10.1016/S0165-4608\(02\)00644-1](https://doi.org/10.1016/S0165-4608(02)00644-1).

Schleiermacher, Gudrun, Niloufar Javanmardi, Virginie Bernard, Quentin Leroy, Julie Cappel, Thomas Rio Frio, Gaele Pierron, et al. 2014. 'Emergence of New *ALK* Mutations at Relapse of Neuroblastoma'. *Journal of Clinical Oncology* 32 (25): 2727–34.

<https://doi.org/10.1200/JCO.2013.54.0674>.

Schmelz, Karin, Joern Toedling, Matt Huska, Maja C. Cwikla, Louisa-Marie Kruetzfeldt, Jutta Proba, Peter F. Ambros, et al. 2021. 'Spatial and Temporal Intratumour Heterogeneity Has Potential Consequences for Single Biopsy-Based Neuroblastoma Treatment Decisions'. *Nature Communications* 12 (1): 6804. <https://doi.org/10.1038/s41467-021-26870-z>.

Schneider, Michael P., Amy Cullen, Justina Pangonyte, Jason Skelton, Harvey Major, Elke Van Oudenhove, Maria J. Garcia, et al. 2022. 'scAbsolute: Measuring Single-Cell Ploidy and Replication Status'. Preprint. Genomics. <https://doi.org/10.1101/2022.11.14.516440>.

Schuh, Lea, Michael Saint-Antoine, Eric M. Sanford, Benjamin L. Emert, Abhyudai Singh, Carsten Marr, Arjun Raj, and Yogesh Goyal. 2020. 'Gene Networks with Transcriptional Bursting Recapitulate Rare Transient Coordinated High Expression States in Cancer'. *Cell Systems* 10 (4): 363-378.e12. <https://doi.org/10.1016/j.cels.2020.03.004>.

Schwarz, Roland F., Charlotte K. Y. Ng, Susanna L. Cooke, Scott Newman, Jillian Temple, Anna M. Piskorz, Davina Gale, et al. 2015. 'Spatial and Temporal Heterogeneity in High-Grade Serous Ovarian Cancer: A Phylogenetic Analysis'. Edited by Christopher Kemp. *PLOS Medicine* 12 (2): e1001789. <https://doi.org/10.1371/journal.pmed.1001789>.

Seeger, Robert C., Garrett M. Brodeur, Harland Sather, Andree Dalton, Stuart E. Siegel, Kwan Y. Wong, and Denman Hammond. 1985. 'Association of Multiple Copies of the N- *Myc* Oncogene with Rapid Progression of Neuroblastomas'. *New England Journal of Medicine* 313 (18): 1111–16. <https://doi.org/10.1056/NEJM198510313131802>.

Serin Harmanci, Akdes, Arif O. Harmanci, and Xiaobo Zhou. 2020. 'CaSpER Identifies and Visualizes CNV Events by Integrative Analysis of Single-Cell or Bulk RNA-Sequencing Data'. *Nature Communications* 11 (1): 89. <https://doi.org/10.1038/s41467-019-13779-x>.

Shaffer, Sydney M., Margaret C. Dunagin, Stefan R. Torborg, Eduardo A. Torre, Benjamin Emert, Clemens Krepler, Marilda Beqiri, et al. 2017. 'Rare Cell Variability and Drug-Induced Reprogramming as a Mode of Cancer Drug Resistance'. *Nature* 546 (7658): 431–35. <https://doi.org/10.1038/nature22794>.

Sherr, Charles J., and Jiri Bartek. 2017. ‘Cell Cycle–Targeted Cancer Therapies’. *Annual Review of Cancer Biology* 1 (1): 41–57. <https://doi.org/10.1146/annurev-cancerbio-040716-075628>.

Shi, Hui, Ting Tao, Brian J. Abraham, Adam D. Durbin, Mark W. Zimmerman, Cigall Kadoch, and A. Thomas Look. 2020. ‘ARID1A Loss in Neuroblastoma Promotes the Adrenergic-to-Mesenchymal Transition by Regulating Enhancer-Mediated Gene Expression’. *Science Advances* 6 (29): eaaz3440. <https://doi.org/10.1126/sciadv.aaz3440>.

Sima, Jiao, and David M. Gilbert. 2014. ‘Complex Correlations: Replication Timing and Mutational Landscapes during Cancer and Genome Evolution’. *Current Opinion in Genetics & Development* 25 (April): 93–100. <https://doi.org/10.1016/j.gde.2013.11.022>.

Sottoriva, Andrea, Haeyoun Kang, Zhicheng Ma, Trevor A Graham, Matthew P Salomon, Junsong Zhao, Paul Marjoram, et al. 2015. ‘A Big Bang Model of Human Colorectal Tumor Growth’. *Nature Genetics* 47 (3): 209–16. <https://doi.org/10.1038/ng.3214>.

Steele, Christopher D., Ammal Abbasi, S. M. Ashiqul Islam, Amy L. Bowes, Azhar Khandekar, Kerstin Haase, Shadi Hames-Fathi, et al. 2022. ‘Signatures of Copy Number Alterations in Human Cancer’. *Nature* 606 (7916): 984–91. <https://doi.org/10.1038/s41586-022-04738-6>.

Su, Xianbin, Linan Zhao, Yi Shi, Rui Zhang, Qi Long, Shihao Bai, Qing Luo, et al. 2021. ‘Clonal Evolution in Liver Cancer at Single-Cell and Single-Variant Resolution’. *Journal of Hematology & Oncology* 14 (1): 22. <https://doi.org/10.1186/s13045-021-01036-y>.

Taliun, Daniel, Daniel N. Harris, Michael D. Kessler, Jedidiah Carlson, Zachary A. Szpiech, Raul Torres, Sarah A. Gagliano Taliun, et al. 2021. ‘Sequencing of 53,831 Diverse Genomes from the NHLBI TOPMed Program’. *Nature* 590 (7845): 290–99. <https://doi.org/10.1038/s41586-021-03205-y>.

Tarabichi, Maxime, Iñigo Martincorena, Moritz Gerstung, Armand M. Leroi, Florian Markowetz, Paul T. Spellman, Quaid D. Morris, Ole Christian Lingjærde, David C. Wedge, and Peter Van Loo. 2018. ‘Neutral Tumor Evolution?’ *Nature Genetics* 50 (12): 1630–33. <https://doi.org/10.1038/s41588-018-0258-x>.

Thatikonda, Venu, S. M. Ashiqul Islam, Robert J. Autry, Barbara C. Jones, Susanne N. Gröbner, Gregor Warsow, Barbara Hutter, et al. 2023. ‘Comprehensive Analysis of Mutational Signatures Reveals Distinct Patterns and Molecular Processes across 27 Pediatric Cancers’. *Nature Cancer* 4 (2): 276–89. <https://doi.org/10.1038/s43018-022-00509-4>.

Theissen, Jessica, Marc Boensch, Ruediger Spitz, David Betts, Sabine Stegmaier, Holger Christiansen, Felix Niggli, et al. 2009. 'Heterogeneity of the *MYCN* Oncogene in Neuroblastoma'. *Clinical Cancer Research* 15 (6): 2085–90. <https://doi.org/10.1158/1078-0432.CCR-08-1648>.

Thirant, Cécile, Agathe Peltier, Simon Durand, Amira Kramdi, Caroline Louis-Brennetot, Cécile Pierre-Eugène, Margot Gautier, et al. 2023. 'Reversible Transitions between Noradrenergic and Mesenchymal Tumor Identities Define Cell Plasticity in Neuroblastoma'. *Nature Communications* 14 (1): 2575. <https://doi.org/10.1038/s41467-023-38239-5>.

Tirosh, Itay, Benjamin Izar, Sanjay M. Prakadan, Marc H. Wadsworth, Daniel Treacy, John J. Trombetta, Asaf Rotem, et al. 2016. 'Dissecting the Multicellular Ecosystem of Metastatic Melanoma by Single-Cell RNA-Seq'. *Science* 352 (6282): 189–96. <https://doi.org/10.1126/science.aad0501>.

TRACERx Consortium, Saioa López, Emilia L. Lim, Stuart Horswell, Kerstin Haase, Ariana Huebner, Michelle Dietzen, et al. 2020. 'Interplay between Whole-Genome Doubling and the Accumulation of Deleterious Alterations in Cancer Evolution'. *Nature Genetics* 52 (3): 283–93. <https://doi.org/10.1038/s41588-020-0584-7>.

Turajlic, Samra, Hang Xu, Kevin Litchfield, Andrew Rowan, Stuart Horswell, Tim Chambers, Tim O'Brien, et al. 2018. 'Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal'. *Cell* 173 (3): 595-610.e11. <https://doi.org/10.1016/j.cell.2018.03.043>.

Ullah, Ikram, Govindasamy-Muralidharan Karthik, Amjad Alkodsai, Una Kjällquist, Gustav Stålhammar, John Lövrot, Nelson-Fuentes Martinez, et al. 2018. 'Evolutionary History of Metastatic Breast Cancer Reveals Minimal Seeding from Axillary Lymph Nodes'. *Journal of Clinical Investigation* 128 (4): 1355–70. <https://doi.org/10.1172/JCI96149>.

Velazquez-Villarreal, Enrique I., Shamoni Maheshwari, Jon Sorenson, Ian T. Fiddes, Vijay Kumar, Yifeng Yin, Michelle G. Webb, et al. 2020. 'Single-Cell Sequencing of Genomic DNA Resolves Sub-Clonal Heterogeneity in a Melanoma Cell Line'. *Communications Biology* 3 (1): 318. <https://doi.org/10.1038/s42003-020-1044-8>.

Vendramin, Roberto, Kevin Litchfield, and Charles Swanton. 2021. 'Cancer Evolution: Darwin and Beyond'. *The EMBO Journal* 40 (18). <https://doi.org/10.15252/emboj.2021108389>.

Vitak, Sarah A., Kristof A. Torkenczy, Jimi L. Rosenkrantz, Andrew J. Fields, Lena Christiansen, Melissa H. Wong, Lucia Carbone, Frank J. Steemers, and Andrew Adey. 2017. ‘Sequencing Thousands of Single-Cell Genomes with Combinatorial Indexing’. *Nature Methods* 14 (3): 302–8. <https://doi.org/10.1038/nmeth.4154>.

Wang, Qingbo, Emma Pierce-Hoffman, Beryl B. Cummings, Jessica Alföldi, Laurent C. Francioli, Laura D. Gauthier, Andrew J. Hill, Anne H. O’Donnell-Luria, Konrad J. Karczewski, and Daniel G. MacArthur. 2020. ‘Landscape of Multi-Nucleotide Variants in 125,748 Human Exomes and 15,708 Genomes’. *Nature Communications* 11 (1): 2539. <https://doi.org/10.1038/s41467-019-12438-5>.

Wang, Rujin, Dan-Yu Lin, and Yuchao Jiang. 2020. ‘SCOPE: A Normalization and Copy-Number Estimation Method for Single-Cell DNA Sequencing’. *Cell Systems* 10 (5): 445-452.e6. <https://doi.org/10.1016/j.cels.2020.03.005>.

Wang, Yong, Jill Waters, Marco L. Leung, Anna Unruh, Whijae Roh, Xiuqing Shi, Ken Chen, et al. 2014. ‘Clonal Evolution in Breast Cancer Revealed by Single Nucleus Genome Sequencing’. *Nature* 512 (7513): 155–60. <https://doi.org/10.1038/nature13600>.

Weiner, Adam C., Marc J. Williams, Hongyu Shi, Ignacio Vázquez-García, Sohrab Salehi, Nicole Rusk, Samuel Aparicio, Sohrab P. Shah, and Andrew McPherson. 2023. ‘Single-Cell DNA Replication Dynamics in Genomically Unstable Cancers’. Preprint. *Cancer Biology*. <https://doi.org/10.1101/2023.04.10.536250>.

Williams, Marc J., Andrea Sottoriva, and Trevor A. Graham. 2019. ‘Measuring Clonal Evolution in Cancer with Genomics’. *Annual Review of Genomics and Human Genetics* 20 (1): 309–29. <https://doi.org/10.1146/annurev-genom-083117-021712>.

Williams, Marc J, Benjamin Werner, Chris P Barnes, Trevor A Graham, and Andrea Sottoriva. 2016. ‘Identification of Neutral Tumor Evolution across Cancer Types’. *Nature Genetics* 48 (3): 238–44. <https://doi.org/10.1038/ng.3489>.

Wolock, Samuel L., Romain Lopez, and Allon M. Klein. 2019. ‘Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data’. *Cell Systems* 8 (4): 281-291.e9. <https://doi.org/10.1016/j.cels.2018.11.005>.

Worst, Barbara C., Cornelis M. Van Tilburg, Gnana Prakash Balasubramanian, Petra Fiesel, Ruth Witt, Angelika Freitag, Miream Boudalil, et al. 2016. ‘Next-Generation Personalised Medicine

for High-Risk Paediatric Cancer Patients – The INFORM Pilot Study’. *European Journal of Cancer* 65 (September): 91–101. <https://doi.org/10.1016/j.ejca.2016.06.009>.

Wu, Chi-Yun, Billy T. Lau, Heon Seok Kim, Anuja Sathe, Susan M. Grimes, Hanlee P. Ji, and Nancy R. Zhang. 2021. ‘Integrative Single-Cell Analysis of Allele-Specific Copy Number Alterations and Chromatin Accessibility in Cancer’. *Nature Biotechnology* 39 (10): 1259–69. <https://doi.org/10.1038/s41587-021-00911-w>.

Yu, Guangchuang. 2020. ‘Using Ggtree to Visualize Data on Tree-Like Structures’. *Current Protocols in Bioinformatics* 69 (1): e96. <https://doi.org/10.1002/cpbi.96>.

Zaccaria, Simone, and Benjamin J. Raphael. 2021. ‘Characterizing Allele- and Haplotype-Specific Copy Numbers in Single Cells with CHISEL’. *Nature Biotechnology* 39 (2): 207–14. <https://doi.org/10.1038/s41587-020-0661-6>.

Zack, Travis I, Steven E Schumacher, Scott L Carter, Andrew D Cherniack, Gordon Saksena, Barbara Tabak, Michael S Lawrence, et al. 2013. ‘Pan-Cancer Patterns of Somatic Copy Number Alteration’. *Nature Genetics* 45 (10): 1134–40. <https://doi.org/10.1038/ng.2760>.

Zahn, Hans, Adi Steif, Emma Laks, Peter Eirew, Michael VanInsberghe, Sohrab P. Shah, Samuel Aparicio, and Carl L. Hansen. 2017. ‘Scalable Whole-Genome Single-Cell Library Preparation without Preamplification’. *Nature Methods* 14 (2): 167–73. <https://doi.org/10.1038/nmeth.4140>.

Zhao, Hao, Zhifu Sun, Jing Wang, Haojie Huang, Jean-Pierre Kocher, and Liguang Wang. 2014. ‘CrossMap: A Versatile Tool for Coordinate Conversion between Genome Assemblies’. *Bioinformatics* 30 (7): 1006–7. <https://doi.org/10.1093/bioinformatics/btt730>.

Zimmerman, Mark W., Yu Liu, Shuning He, Adam D. Durbin, Brian J. Abraham, John Easton, Ying Shao, et al. 2018. ‘MYC Drives a Subset of High-Risk Pediatric Neuroblastomas and Is Activated through Mechanisms Including Enhancer Hijacking and Focal Enhancer Amplification’. *Cancer Discovery* 8 (3): 320–35. <https://doi.org/10.1158/2159-8290.CD-17-0993>.

APPENDICES

Appendix 1: Contribution towards other publication during PhD

Tucker ER, Jiménez I, Chen L, Bellini A, Gorrini C, Calton E, Gao Q, Che H, Poon E, Jamin Y, Martins Da Costa B, Barker K, Shrestha S, Hutchinson JC, Dhariwal S, Goodman A, Del Nery E, Gestraud P, **Bhalshankar J**, Iddir Y, Saberi-Ansari E, Saint-Charles A, Geoerger B, Marques Da Costa ME, Pierre-Eugène C, Janoueix-Lerosey I, Decaudin D, Nemati F, Carcaboso AM, Surdez D, Delattre O, George SL, Chesler L, Tweddle DA, Schleiermacher G. Combination Therapies Targeting ALK-aberrant Neuroblastoma in Preclinical Models. *Clin Cancer Res.* 2023 Apr 3;29(7):1317-1331. doi: 10.1158/1078-0432.CCR-22-2274. PMID: 36602782; PMCID: PMC10068437.

Thirant C, Peltier A, Durand S, Kramdi A, Louis-Brennetot C, Pierre-Eugène C, Gautier M, Costa A, Grelier A, Zaïdi S, Gruel N, Jimenez I, Lapouble E, Pierron G, Sitbon D, Brisse HJ, Gauthier A, Fréneaux P, Grossetête S, Baudrin LG, Raynal V, Baulande S, Bellini A, **Bhalshankar J**, Carcaboso AM, Geoerger B, Rohrer H, Surdez D, Boeva V, Schleiermacher G, Delattre O, Janoueix-Lerosey I. Reversible transitions between noradrenergic and mesenchymal tumor identities define cell plasticity in neuroblastoma. *Nat Commun.* 2023 May 4;14(1):2575. doi: 10.1038/s41467-023-38239-5. PMID: 37142597; PMCID: PMC10160107.

Gamble LD, Purgato S, Henderson MJ, Di Giacomo S, Russell AJ, Pignini P, Murray J, Valli E, Milazzo G, Giorgi FM, Cowley M, Ashton LJ, **Bhalshankar J**, Schleiermacher G, Rihani A, Van Maerken T, Vandesompele J, Speleman F, Versteeg R, Koster J, Eggert A, Noguera R, Stallings RL, Tonini GP, Fong K, Vaksman Z, Diskin SJ, Maris JM, London WB, Marshall GM, Ziegler DS, Hogarty MD, Perini G, Norris MD, Haber M. A G316A Polymorphism in the Ornithine Decarboxylase Gene Promoter Modulates MYCN-Driven Childhood Neuroblastoma. *Cancers (Basel).* 2021 Apr 9;13(8):1807. doi: 10.3390/cancers13081807. PMID: 33918978; PMCID: PMC8069650.



Institut Curie 26 rue d'Ulm 75248 Paris cedex 05 France

RÉSUMÉ

Dans cette étude approfondie, nous nous penchons sur l'hétérogénéité génétique du neuroblastome (NB), une tumeur solide pédiatrique difficile à traiter, en tirant parti de la puissance des technologies de séquençage de cellules uniques et de masse.

Notre étude comporte deux volets : premièrement, nous explorons l'hétérogénéité génétique intratumorale (ITH) au niveau d'une seule cellule. Nous cherchons à comprendre comment les altérations du nombre de copies d'allèles spécifiques d'une cellule à l'autre contribuent aux variations génomiques et phénotypiques, influençant ainsi la trajectoire évolutive du NB. Il s'agit notamment d'examiner la chronologie des aberrations génomiques spécifiques, la chronologie de la réplication des cellules individuelles et la coévolution du génome et du transcriptome. En outre, nous examinons minutieusement la dynamique des sous-clones et leurs réponses ou résistances spécifiques aux thérapies ciblées.

Pour ce faire, nous avons combiné le séquençage de l'exome entier (WES) de la lignée germinale et de la tumeur avec le séquençage de l'ADN du génome entier d'une cellule unique à très faible profondeur (scDNAseq) en utilisant le kit CNV à cellule unique de 10x Genomics Chromium. Au total, 9 410 cellules tumorales ont été caractérisées à partir de 14 modèles de xénogreffes dérivées de patients (PDX) de neuroblastome (NB) et de 4 biopsies tumorales de patients NB. En outre, 6 modèles PDX ont été soumis à différentes combinaisons de traitement et des WES en vrac ont été réalisés avant et après le traitement.

Notre analyse scDNAseq a permis d'identifier des génomes de neuroblastomes polyclonaux (11 cas) et monoclonaux (7 cas) par le biais d'altérations du nombre de copies spécifiques aux allèles et aux haplotypes. Nous avons observé de 2 à 11 clones distincts par tumeur polyclonale. Les événements de duplication du génome entier étaient évidents dans les génomes polyclonaux et monoclonaux. Les premiers événements clonaux comprenaient des altérations connues du nombre de copies (pertes des chromosomes 1p et 11q, gains du chromosome 17q), des amplifications de *MYCN/ALK*, ainsi que des mutations somatiques dans des gènes tels que *ALK*, *ATRX*, *TP53* et *NF1*. L'évolution parallèle du nombre de copies de deux clones distincts, le sous-clone s1/s2, dans un modèle PDX, avec des réponses spécifiques au clone à la thérapie ciblée (Lorlatinib), est particulièrement remarquable. Les profils de synchronisation de la réplication ont révélé une réplication précoce et tardive mutuellement exclusive dans les sous-clones s1 et s2, respectivement. L'analyse génotype à phénotype a indiqué une expression préférentielle du génotype du sous-clone s1.

Deuxièmement, notre étude s'étend à l'analyse du séquençage en masse, en se concentrant particulièrement sur la fréquence et l'impact des altérations *ALK* (lymphome anaplasique kinase), à la fois les amplifications et les variants nucléotidiques simples (SNV), dans une grande cohorte de 1092 patients atteints de NB à haut risque (HR-NB) traités dans le cadre du protocole européen HR-NBL1. Les mutations *ALK* clonales observées dans 33 % des cas ont été associées à des résultats plus défavorables, tandis que les mutations subclonales, qui représentent 28 % de l'ensemble des mutations *ALK*, n'ont pas eu un impact similaire sur le pronostic.

Nous étudions également la prévalence et la signification clinique des variations des gènes de remodelage de la chromatine (CRG) et d'autres gènes modificateurs épigénétiques (EMG) dans le NB. L'ADN normal et l'ADN tumoral diagnostiqué appariés de 55 patients ont été séquencés, soit par séquençage du génome entier (16 cas), soit par séquençage de l'exome entier (39 cas), complété par un séquençage ciblé ultra-profond sur 283 patients atteints de cancer du sein à l'aide de TSCA®. Le séquençage ciblé ultra-profond a mis en évidence des variations de nucléotides simples, de petites insertions/délétions, dans les gènes de remodelage de la chromatine et les gènes modificateurs épigénétiques dans 20 % des cas étudiés. Notre recherche identifie des variations CRG et EMG chez 8,4 % des patients atteints de NB, les gènes les plus fréquemment altérés étant *ATRX*, *SMARCA4*, *MLL3* et *ARID1B*. Les variations dans *SMARCA4* et *ATRX* étaient plus fréquentes dans les cas de NB que dans la population générale, ce qui suggère leur rôle crucial dans l'oncogenèse du NB.

Dans l'ensemble, ce travail fournit des informations essentielles sur les trajectoires évolutives des tumeurs du neuroblastome (NB), en identifiant pour la première fois une synchronisation distincte de la réplication dans des clones cellulaires distincts et le potentiel de ciblage des altérations clonales précoces. Notre travail fournit également des informations supplémentaires sur l'hétérogénéité intratumorale en ce qui concerne l'*ALK*, une altération ciblée, ainsi que le CRG.

MOTS CLÉS

ADN unicellulaire, ARN unicellulaire, Clones, Évolution, Temps de réplication, *ALK*, *SMARCA4*, Survie, Traitement ciblé

ABSTRACT

In this comprehensive study, we delve into the genetic heterogeneity of neuroblastoma (NB), a challenging pediatric solid tumor, by leveraging the power of single-cell and bulk sequencing technologies.

Our investigation is two-fold: firstly, we explore the genetic intratumoral heterogeneity (ITH) at a single-cell level. We aim to understand how cell-to-cell allele-specific copy number alterations contribute to genomic and phenotypic variations, influencing the evolutionary trajectory of NB. This includes examining the timing of specific genomic aberrations, the replication timing of individual cells, and the co-evolution of the genome and transcriptome. Additionally, we scrutinize subclonal dynamics and their specific responses or resistances to targeted therapies.

For this we combined germline and tumor bulk whole exome sequencing (WES) with ultra-low depth single-cell whole-genome DNA sequencing (scDNAseq) using the 10x Genomics Chromium single-cell CNV kit. A total of 9,410 tumor cells were characterized from 14 patient-derived xenografts (PDX) models of neuroblastoma (NB) and 4 tumor biopsies from NB patients. Additionally, 6 PDX models were subjected to different treatment combinations and bulk WES was conducted pre- and post-treatment.

Our scDNAseq analysis identified both polyclonal (11 cases) and monoclonal (7 cases) neuroblastoma genomes through allele and haplotype-specific copy number alterations. We observed 2 to 11 distinct clones per polyclonal tumor. Whole genome duplication events were evident in both polyclonal and monoclonal genomes. Early clonal events included known copy number alterations (losses of chromosomes 1p and 11q, gains in chromosome 17q), *MYCN/ALK* amplifications, as well as somatic mutations in genes such as *ALK*, *ATRX*, *TP53*, and *NFI*. Particularly notable was the parallel copy number evolution of two distinct clones, subclone s1/s2, in a PDX model, with clone-specific responses to targeted therapy (Lorlatinib). Replication timing profiles revealed mutually exclusive early and late replication in subclones s1 and s2, respectively. Genotype-to-phenotype analysis indicated preferential expression of the of subclone s1 genotype.

Secondly, our study extends to bulk sequencing analysis, particularly focusing on the frequency and impact of *ALK* (anaplastic lymphoma kinase) alterations, both amplifications and single nucleotide variants (SNVs), in a large cohort of 1092 high-risk NB (HR-NB) patients treated within the European HR-NBL1 protocol. Clonal *ALK* mutations observed in 33 % were associated with poorer outcomes, whereas subclonal mutations, accounting for 28% of all *ALK* mutations, did not demonstrate a similar impact on prognosis.

We also investigate the prevalence and clinical significance of variations in chromatin remodeling genes (CRGs) and other epigenetic modifier genes (EMGs) in NB. Paired diagnostic tumor and normal DNA from 55 patients were sequenced, either by whole-genome sequencing (16 cases) or whole-exome sequencing (39 cases), complemented by ultra-deep targeted sequencing on 283 NB patients using TSCA®. Ultra-deep targeted sequencing enhanced single nucleotide variations, small insertions/deletions, in chromatin remodeling genes and epigenetic modifier genes in 20% of the cases studied. Our research identifies CRG and EMG variations in 8.4% of NB patients, with the most frequently altered genes including *ATRX*, *SMARCA4*, *MLL3*, and *ARID1B*. Variations in *SMARCA4* and *ATRX* were more frequent in NB cases compared to the general population, suggesting their crucial role in NB oncogenesis.

Altogether, this work provides pivotal insights into the evolutionary trajectories of neuroblastoma (NB) tumors, identifying for the first time distinct replication timing in distinct cellular clones, and the potential for targeting early clonal alterations. Our work also provides further insight into intratumor heterogeneity with regards to *ALK*, a targetable alteration, as well as CRG.

KEYWORDS

Single cell DNA, Single cell RNA, Clones, Subclones, Evolution, Replication timing, *ALK*, *SMARCA4*, Survival, Targeted-Treatment