



Retinoblastoma heterogeneity : a comprehensive multi-omic and single-cell transcriptomic analysis

Jing Liu

► To cite this version:

Jing Liu. Retinoblastoma heterogeneity : a comprehensive multi-omic and single-cell transcriptomic analysis. Cancer. Université Paris-Saclay, 2022. English. NNT : 2022UPASL025 . tel-04053049

HAL Id: tel-04053049

<https://theses.hal.science/tel-04053049>

Submitted on 31 Mar 2023

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.

Retinoblastoma heterogeneity: a comprehensive multi-omic and single-cell transcriptomic analysis

*Hétérogénéité du rétinoblastome:
une analyse multi-omique et
transcriptomique de cellules uniques*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 582 : cancérologie : biologie – médecine -santé (CBMS)
Spécialité de doctorat : Aspects moléculaires et cellulaires de la biologie
Graduate School: Science de la vie et Santé
Référent : Faculté de médecine

Thèse préparée dans l'UMR 144 - CNRS - Institut Curie,
sous la direction de **François RADVANYI**, DRCE,
et la co-direction de **Aurélien DE REYNIES**, PUPH

Thèse soutenue à Paris, le 30 Mars 2022, par

Jing LIU

Composition du Jury

Muriel PERRON DR1, CNRS - Institut NeuroPSI -Université Paris Saclay	Présidente
Fatima AL-SHAHROUR Head of bioinformatics Unit, Centro Nacional de Investigaciones Oncológicas (Espagne)	Rapporteure
Francis MUNIER Professeur ordinaire, Hospital ophtalmique Jules-Gonin (Suisse)	Rapporteur
Gordon ROBERTSON Staff scientist, British Columbia Genome Sciences Center (Canada)	Examineur
Nicolas STRANSKY Head of Data Science and Computational Biology, Celsius Therapeutics (USA)	Examineur
François RADVANYI DRCE, UMR144 – CNRS - Institut Curie	Directeur de thèse
Aurélien DE REYNIES PUPH, INSERM - Centre de Recherche des Cordeliers - Université de Paris	Co-Directeur de thèse

Synthèse

Titre : Hétérogénéité du rétinoblastome : une analyse multi-omique et transcriptomique de cellules uniques

Mots clés : Rétinoblastome, hétérogénéité, microenvironnement, bioinformatique, multi-omique, transcriptome de cellules uniques.

Le rétinoblastome est un cancer pédiatrique dérivé de la rétine. Bien que rare, c'est la tumeur maligne intraoculaire la plus fréquente chez l'enfant. Le rétinoblastome est traité par une thérapie locale aux premiers stades, mais aux stades ultérieurs, il nécessite une chirurgie d'énucléation et une chimiothérapie systémique. Sans un diagnostic en temps utile et un traitement approprié, des métastases peuvent se développer et entraîner le décès de l'enfant.

Le rétinoblastome est une maladie hétérogène. Les cellules tumorales de différents patients ou différentes cellules d'un même patient peuvent présenter des caractéristiques moléculaires et phénotypiques distinctes. D'un point de vue histopathologique, les cellules tumorales peuvent présenter différents degrés de différenciation, ainsi qu'une croissance exophytique, endophytique ou mixte. En ce qui concerne la génomique, le rétinoblastome survient principalement après une inactivation bi-allélique de *RB1* et, dans de rares cas, une amplification de *MYCN* peut également déclencher la maladie. En termes de transcriptomique, quelques études ont été réalisées et ont révélé que le rétinoblastome peut présenter des degrés variés dans les signatures de différenciation des photorécepteurs.

Dans notre travail, nous avons identifié deux sous-types moléculaires basés sur l'analyse de 102 rétinoblastomes en utilisant le séquençage de l'exome entier, les SNP array, les puces pour mesurer l'expression des gènes et la méthylation de l'ADN. Nous avons réalisé le clustering en utilisant une stratégie combinant le clustering hiérarchique consensuel et le clustering basé sur les centroïdes. Nous avons démontré que les tumeurs de sous-type 1 présentaient une signature plus mature de différenciation de photorécepteurs cône et se manifestaient généralement chez les enfants de moins de 18 mois, tandis qu'un niveau plus faible de différenciation de photorécepteurs cône est corrélé à un niveau élevé de caractéristiques ganglionnaires/neurales, de signatures de type souche et d'instabilité génomique

dans les tumeurs de sous-type 2, que l'on retrouve davantage chez les enfants de plus de 18 mois. L'analyse d'une série indépendante de 112 rétinoblastomes présentant des facteurs pathologiques à haut risque a révélé que les tumeurs qui métastasent expriment toutes la protéine TFF1, correspondant au gène le plus surexprimé dans les tumeurs de sous-type 2 par rapport aux tumeurs de sous-type 1. Nous avons analysé 14 rétinoblastomes en utilisant le séquençage de l'ARN de cellules uniques (scRNA-seq), et avons confirmé cette hétérogénéité inter-tumorale. L'une des tumeurs analysées présentait une hétérogénéité intra-tumorale au niveau phénotypique et génomique : certaines cellules tumorales présentaient un degré plus élevé de différenciation des cônes et la perte de différenciation des cônes s'accompagnait d'une accumulation d'altérations génomiques, tandis que d'autres cellules présentaient des altérations génomiques et des phénotypes totalement différents. Nous avons décrit le paysage immunitaire du rétinoblastome grâce au scRNA-seq, et avons découvert que différents types de cellules immunitaires sont présents dans le microenvironnement tumoral, notamment de multiples populations de cellules de la lignée monocyttaire telles que les macrophages M2 protumoraux et les cellules présentatrices d'antigènes, de multiples populations de cellules T dont les cellules T régulatrices CD4⁺ et les cellules T cytotoxiques CD8⁺, ainsi que les cellules NK. L'infiltration des macrophages M2 a été validée par immunohistochimie et a été associée à un niveau élevé d'expression du MIF par une analyse sur puces à cytokines et d'une prédiction *in silico* de ligand-récepteur.

En conclusion, notre analyse multi-omique et transcriptomique de cellules uniques a permis de caractériser de manière détaillée l'hétérogénéité inter- et intra-tumorale et de décrire le paysage immunitaire du rétinoblastome.

Abstract

Title: Retinoblastoma heterogeneity: a comprehensive multi-omic and single-cell transcriptomic analysis

Keywords: Retinoblastoma, heterogeneity, microenvironment, bioinformatics, multi-omics, single-cell transcriptome

Retinoblastoma is a pediatric cancer derived from the retina. The global annual incidence is around 8000 new cases. Although rare, it is the most common intraocular malignancy in children. Retinoblastoma is treated with local therapy at early stages, but at later stages, it requires enucleation surgery and systemic chemotherapy. Without timely diagnosis and proper treatment, metastasis can develop and lead to lethal diseases.

Retinoblastoma is a heterogeneous disease. Tumor cells in different patients or different cells in one patient can exhibit distinct molecular and phenotypic characteristics. From the histopathological perspective, tumor cells can present differentiation at diverse levels, and exhibit exophytic, endophytic or mixed growth. With regard to genomics, retinoblastoma arises predominantly after bi-allelic RB1 inactivation and in rare cases, MYCN amplification may also initiate the disease. Other mutations and genomic alterations may also contribute to the progression of the retinoblastoma such as BCOR mutations, chromosome 1q gain, 6p gain, or 16q loss. In terms of transcriptomics, a few studies were performed and revealed that retinoblastoma can display varied degrees in photoreceptor differentiation signatures. In our work, we identified two molecular subtypes based on the analyses of a series of 102 retinoblastoma using whole-exome sequencing, SNP array, gene expression microarray and DNA methylation array. We made the clustering using a method combining consensus hierarchical clustering and centroid based clustering. We demonstrated that subtype 1 tumors exhibited a more matured cone differentiation signature and were usually manifested in children less than 18 months, while subtype 2 tumors were found more frequent in children more than 18 months, exhibited lower level of cone differentiation, elevated level of ganglion/neuronal features, stemness signatures, and genomic instability. Analysis of an independent series of 112

retinoblastomas with high risk pathological factors uncovered that metastatic tumors all expressed TFF1 protein at their primary sites, the gene that is most upregulated in subtype 2 tumors as compared to subtype 1 tumors. We confirmed this inter-tumoral heterogeneity in 14 additional retinoblastomas using single-cell RNA sequencing. One of the tumors analyzed displayed intra-tumoral heterogeneity at both phenotypic and genomic level, in that some tumor cells exhibited higher grade of cone differentiation, while other cells presented entirely different genomic alterations and phenotypes. We characterized the immune landscape of retinoblastoma through single-cell transcriptomics, and reported that various immune cells types are present in the tumor microenvironment, including multiple populations of monocytic lineage cells such as the protumoral alternative M2 macrophages and antigen presenting cells, multiple populations of T cells such as CD4⁺ regulatory T cells and CD8⁺ cytotoxic T cells, as well as NK cells. The infiltration of M2 macrophages was validated by immunohistochemistry, and was associated with high-level of MIF expression through cytokine array and *in silico* ligand-receptor prediction.

Taken together, our multi-omic and single-cell transcriptomic analysis comprehensively characterized the inter- and intra- tumoral heterogeneity and characterized the immune landscape in retinoblastoma.

Acknowledgements

First, I would like to express my sincere gratitude to my thesis director, Dr. François Radvanyi, for his guidance and support for me to accomplish this work. Thank you for giving me the opportunity in your team to work on exciting projects and thank you for all the interesting discussions over the years. Your passion for research has been and will always be an inspiration for me. Thank you.

I would like to thank my thesis co-director, Pr. Aurélien de Reyniès, for his guidance and support. Thank you for welcoming me in CIT team and for many fruitful discussions. Thank you for your good humor and your suggestions for my career development.

I would like to thank the members of my thesis committee, Dr. Fatima Al-Shahrour, Pr. Francis Munier, Dr. Muriel Perron, Dr. Gordon Robertson and Dr. Nicolas Stransky, for taking time to review and evaluate my work.

I would like to thank my mid-term thesis committee members, Dr. Josh Waterfall and Dr. Isabelle Janoueix-Lerosey, for their discussions and valuable feedbacks to improve my work.

I would also like to thank Mme Léa Poisot from the doctoral school for her help during the years and for the submission of the thesis.

I thank Fondation ARC and Institut Curie for funding my PhD study and ICGEX project for funding part of the single-cell project.

I thank the doctors at Institut Curie Hospital for their participation and support in the single-cell work. Thank you, Dr. François Doz, Dr. Isabelle Aerts, Dr. Alexandre Matet, Dr. Denis Malaise, Dr. Paul Freneaux, Dr. Arnaud Gauthier, Dr. Livia Lumbroso-Le Rouic, and Dr. Natalie Cassoux. I thank also all the members of the Genomics platform of Institut Curie for supporting our work.

I would like to thank our collaborators Dr. Guillermo Chantada, Dr. Rosario Aschero, Dr. Ángel Montero Carcaboso and Dr. Maria Cuadrado-Vilanova, for including me in these exciting collaborative projects.

I would also like to thank Dr. Olivier Goureau and Dr Sacha Reichman for collaborating with us and for many discussions allowing me to learn more about the retina.

I would like to thank Nanor, without you, our work on single-cell work would not have been possible. I also thank you for your accompanying and listening. Thank you Elodie for accommodating me in the team since the beginning and always being supportive. Thank you Daniela for working together on several projects and for our discussions on

the science and beyond. Thank you Céline and Meriem for setting up the foundation of the retino team. Thank you Clément for your huge support and for carrying on the work.

Thank you, all the members of the OM team. Thank you Isa and Florent for giving your valuable feedbacks on retino work. Thank you Hélène for your smile all the time. Thank you Aura and Jacqueline for the fun times. Thank you Clarice for your geek tips. Thank you Mingjun and Linda for enjoyable times in our office. Thank you Maca for staying at my place so we can be friends. Thank you Yves for the chocolates and including me in some of your interesting projects closer to clinical side. Thank you Luc, Xiangyu, Clémentine, Virginia, Laure, Jennifer, Narjesse, Trang, Yanish, Elo for your help in the work and lab life. I would also like to give my thanks to Michèle, for your help and also for all the history stories shared with me.

Last but not least, I would like to thank my loving family, I am so grateful to have your love and support me, thank you.

Table of Contents

CHAPTER 1. INTRODUCTION	5
THE RETINA AND RETINAL DEVELOPMENT	5
RETINOBLASTOMA: EPIDEMIOLOGY AND CLINICAL ASPECTS	11
RETINOBLASTOMA: GENETIC AND MOLECULAR BASIS	16
MOLECULAR CLASSIFICATIONS OF RETINOBLASTOMA	19
CELL ORIGIN OF RETINOBLASTOMA	23
OMICS APPROACHES	25
REFERENCES	28
CHAPTER 2. OBJECTIVES	30
CHAPTER 3. A HIGH-RISK RETINOBLASTOMA SUBTYPE WITH STEMNESS FEATURES, DE-DIFFERENTIATED CONE STATES AND NEURONAL/GANGLION CELL GENE EXPRESSION	33
ABSTRACT	35
INTRODUCTION	36
RESULTS	36
DISCUSSION	45
METHODS	47
REFERENCES	52
SUPPLEMENTARY INFORMATION	55
COMMENTS	73
CHAPTER 4. CHARACTERIZATION OF HETEROGENEITY IN RETINOBLASTOMA BY SINGLE-CELL RNA-SEQUENCING	75
INTRODUCTION	75
RESULTS	77
DISCUSSION	95
METHODS	97
REFERENCES	100
CHAPTER 5. CONCLUDING REMARKS	103
ANNEXE. IDENTIFICATION OF IMMUNOSUPPRESSIVE FACTORS IN RETINOBLASTOMA CELL SECRETOMES AND AQUEOUS HUMOR FROM PATIENTS	106
ABSTRACT	107
INTRODUCTION	108
MATERIALS AND METHODS	108
RESULTS	110
DISCUSSION	116
REFERENCES	117

List of Figures

CHAPTER 1. INTRODUCTION

Figure 1. Schema of the human retina.	5
Figure 2. Schematic of cell cycle.	8
Figure 3. Schema of retinal cell development in mouse.	9
Figure 4. Schema of retinal cell development in differentiating hiPSCs.	10
Figure 5. Leukocoria in a retinoblastoma patient.	12
Figure 6. Fundusoscopic examination of a retinoblastoma patient eye.	12
Figure 7. Schema of retinoblastoma protein structure.	16
Figure 8. Retinoblastoma express various retinal markers without clear subtype identified.	20
Figure 9. Identification of two main subtypes in retinoblastoma.	21
Figure 10. Quantification of differential signatures in relation to clinical and histopathological variables.	22
Figure 11. Schema of retinoblastoma formation from maturing (ARR3+) cone precursors.	24
Figure 12. Work flow of cyclic array sequencing.	25
Figure 13. Workflow of Chromium single-cell RNA-sequencing.	27

CHAPTER 3. A HIGH-RISK RETINOBLASTOMA SUBTYPE WITH STEMNESS FEATURES, DE-DIFFERENTIATED CONE STATES AND NEURONAL/GANGLION CELL GENE EXPRESSION

Figure 1. Multi-omics-based molecular subtypes of retinoblastoma and clinical characteristics.	37
Figure 2. Genomic characterization, somatic mutational landscape, and DNA methylation profiles of the two retinoblastoma subtype.	40
Figure 3. Transcriptomic differences between the two retinoblastoma subtypes.	42
Figure 4. Expression of cone and neuronal/ganglion cell markers in retinoblastoma and retinal organoids.	44
Figure 5. Intratumor heterogeneity at the single-cell level of a subtype 2 retinoblastoma (RBSC11).	46
Figure 6. Subtype 2 tumors are associated with a higher risk of metastasis.	48
Supplementary Figure 1. Multi-omics classification of a series of 72 retinoblastomas.	57
Supplementary Figure 2. Distribution of the differentially methylated CpG between the two subtypes in relation to CpG islands.	59
Supplementary Figure 3. Stemness Indices estimated by different signatures.	60
Supplementary Figure 4. Retinal cell gene expression analysis in the two subtypes of retinoblastoma.	61
Supplementary Figure 5. Data related to single-cell RNA-seq of a subtype 2 retinoblastoma (RBSC11).	67
Supplementary Figure 6. Expression of TFF1 and EBF3 in the normal developing retina and in the two subtypes of retinoblastoma.	69
Supplementary Figure 7. The two retinoblastoma subtypes were different in clinical and molecular features in two additional independent datasets.	71
Figure 7. Two visualization of photoreceptor signature expression in Kooi <i>et al.</i> 's data.	74

CHAPTER 4. CHARACTERIZATION OF HETEROGENEITY IN RETINOBLASTOMA BY SINGLE-CELL RNA-SEQUENCING

Figure 1. Single-cell transcriptome analysis of RBSC11 revealed intra-tumoral heterogeneity.	77
Figure 2. Single cells identified with new analysis pipeline.	78
Figure 3. New annotation of RBSC11.	80
Figure 4. Inferred copy number variation profiles for RBSC11.	82
Figure 5. Schema of a proposed progression model for RBSC11.	83
Figure 6. Single-cell transcriptomic analysis of retinoblastoma.	87
Figure 7. Immune cell type identification in retinoblastoma microenvironment.	90
Figure 8. Expression of selected genes in immune cells.	92
Figure 9. Inferred ligand-receptor pair between monocytes and other cells.	94

ANNEX. IDENTIFICATION OF IMMUNOSUPPRESSIVE FACTORS IN RETINOBLASTOMA CELL SECRETOMES AND AQUEOUS HUMOR FROM PATIENTS

Figure 1. Infiltrating lymphocytes in retinoblastoma eyes (tumor and conserved retinal areas) and in retinas of non-retinoblastoma (control) eyes.	110
Figure 2. Infiltrating macrophages and microglia in retinoblastoma eyes (tumor and conserved retinal areas) and in retinas of non-retinoblastoma (control) eyes.	111
Figure 3. Expression of immune checkpoint molecules in retinoblastoma samples.	112
Figure 4. Cytokine secretion and expression profile of retinoblastomas.	113
Figure 5. Single-cell sequencing of immune markers and cytokines in one retinoblastoma tumor.	114
Figure 6. Polarization and migration of PBMC-derived macrophages stimulated with retinoblastoma-conditioned medium and rMIF	115

List of Tables

CHAPTER 1. INTRODUCTION

Table 1. Summary of clinical staging systems.	15
Table 2. Summary of single-cell RNA-sequencing methods.	26

CHAPTER 3. A HIGH-RISK RETINOBLASTOMA SUBTYPE WITH STEMNESS FEATURES, DEDIFFERENTIATED CONE STATES AND NEURONAL/GANGLION CELL GENE EXPRESSION

Table 1. Clinical and histopathological characteristics of patients stratified by molecular subtype.	39
Table 2. Clinical and pathological characteristics of an additional series of 112 primary tumors presenting HRPFS.	47

CHAPTER 4. CHARACTERIZATION OF HETEROGENEITY IN RETINOBLASTOMA BY SINGLE-CELL RNA-SEQUENCING

Table 1. Patient Characteristics.	86
-----------------------------------	----

CHAPTER 1. INTRODUCTION

Retinoblastoma is a rare pediatric cancer but also the most common intraocular malignancy in children. It arises from the developing retina. This introduction aims to provide an overview of essential background knowledge of the disease and to highlight the research gaps in the field.

THE RETINA AND RETINAL DEVELOPMENT

The retina

The human retina is a 0.2mm-thick light-sensitive tissue in the posterior lining of the eye (Figure 1A). Six mayor types of neurons (rod and cone photoreceptors, horizontal cells, bipolar cells, amacrine cells and retinal ganglion cells) and glial cells are orderly layered in the retina (Figure 1B,C). They function collectively to receive light, process signals and transmit information to the brain.

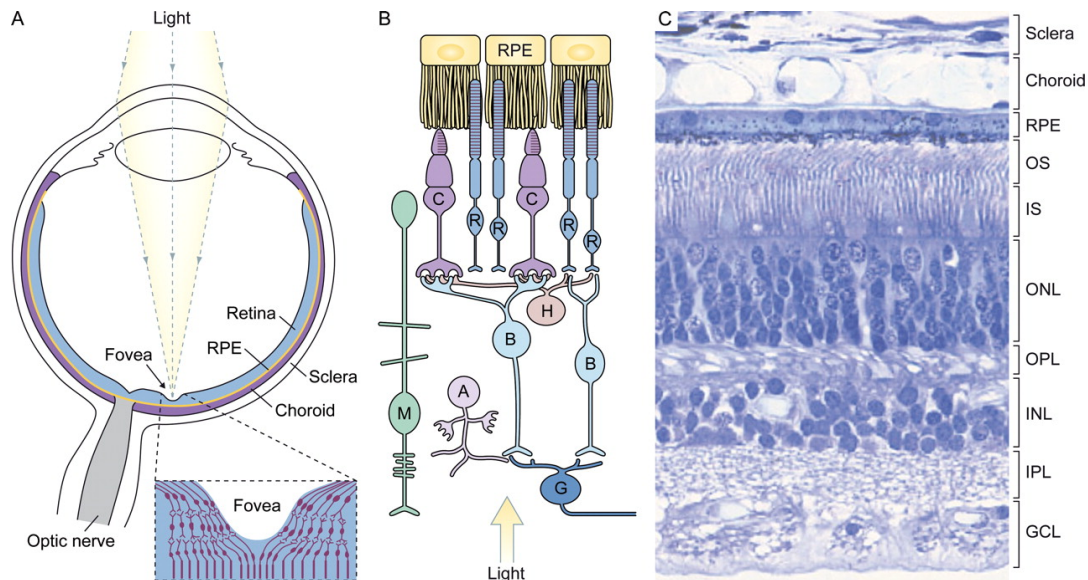


Figure 1. Schema of the human retina.

A, Diagram of the eye. An enlarged diagram of fovea region is shown. RPE: retinal pigmental epithelial.

B Diagram of the retina. Following the path of light from inner to outer layer, G: ganglion cells; A: amacrine cells; B: bipolar cells; H: horizontal cells; R: rod photoreceptors; C: cone photoreceptors; M: Müller glia cells.

C, H&E stained section of human retina. GCL: ganglion cell layer; IPL: inner plexiform layer; INL: inner nuclear layer; OPL: outer plexiform layer; ONL: outer nuclear layer; IS: inner segment; OS: outer segment. Images from (1)

Retinal cell types and their function

The retinal neurons are arranged in an orderly structure that consists of two photoreceptor segment layers (outer segment and inner segment), three nuclear layers (outer nuclear layer, inner nuclear layer, ganglion cell layer) and two synaptic layers (outer plexiform layer, inner plexiform layer) (Figure 1B,C). Light passes through transparent inner retinal layers and photons get captured by the visual pigments in inner segment and outer segment layers of the photoreceptors. Excessive light is absorbed by retinal pigmental epithelial next to the photoreceptors. After absorption by visual pigments, signal transduction cascades are initiated inside the photoreceptors, resulting in the change of membrane potential. This change regulates the releasing of neurotransmitters from the photoreceptors to the bipolar and horizontal cells in the first synaptic layer, the outer plexiform layer. Horizontal cells provide feedback and feedforward signals to both photoreceptors and bipolar cells. Bipolar cells connect to the amacrine cells and retinal ganglion cells in the second synaptic layer, the inner plexiform layer. Amacrine cells provide inhibitory feedback and feedforward signals. Finally, retinal ganglion cells integrate all the information through their dendrites and pass it on through their axons into the optic nerve and further into the brain (1,2).

The predominant cell types in the retina are the photoreceptors. Rod photoreceptors, of which there is only one type, mediate vision in dim light. Cone photoreceptors mediate vision in bright light and are responsible for color vision. There are three types of cone cells (L, M, S) and they carry different pigments (opsins) with maximal absorption for different wavelengths of light: red/long-wavelength (L), green/medium-wavelength (M), and blue/short-wavelength (S). The human retina contain about 130 million photoreceptors, about 5 million bipolar cells, and about 1 million ganglion cells. Rods outnumber cones by 20-fold, and are distributed throughout the retina except for the fovea region (1) There are also multiple subtypes of bipolar, horizontal, amacrine interneurons, and retinal ganglion cells (3,4).

In the retina there are also three major types of glial cells: Müller glia, astrocytes and microglia. The Müller glia are specialized macroglia of the retina spanning the entire thickness of neural retina (Figure 1B). They guide the light to the photoreceptors, maintain homeostasis, and are modulators of immune and inflammatory responses. Astrocytes are macroglia migrated from the brain to the retina and are located predominantly in the ganglion cell layers. They participate in maintenance of vascular stability and hemostasis in the retina. Microglia are monocytes that enter the retina and then reside in the ganglion cell layers and the two synaptic layers. They are responsible for the immune responses and neural regenerations (5).

The fovea

The fovea is 700- μ m-diameter pit near the center of the retina (Figure 1A). Among the mammals, it is only present in human and other higher primates. This region has several structural characteristics so that it has the highest visual acuity as compared to other regions of the retina. The fovea primarily consists of cone cells, the photoreceptors that mediate color vision and provide a higher resolution than rod photoreceptors, the predominant cells in non-fovea regions. Each fovea cone is connected to its exclusive bipolar and ganglion cell; unlike in other areas of the retina, multiple photoreceptors share the connecting interneurons to transmit information to the central nervous system. Light enter photoreceptors with minimal distortion in the fovea region since the cell bodies of the proximal retinal neurons have been shifted to the side (Figure 1A) (1).

Retinal development

The vertebrate retina develops from retinal progenitor cells (RPCs). They give rise to all six major types of retinal neurons and Müller glia in a conserved sequential order. Ganglion cells, cone photoreceptors, horizontal cells and amacrine cells are early-born cells, while bipolar neurons, Müller glia, and rod photoreceptors are late-born cells. During early development, RPCs are highly proliferative and divide symmetrically to grow retina into a proper size in coordination with the growth of the eye. Upregulation of the transcriptional factors (TFs) Pax6, Chx10 and Sox2 may be important for

maintaining proliferative capabilities of RPCs. Following the expansive period, RPCs tend to divide asymmetrically, producing one daughter cell that differentiates into an early-born retinal cell, and another daughter cell that is a retinal progenitor cell. At later developmental stages, both daughter cells exit cell cycle and differentiate into retinal cells (Davis and Dyer, 2010). Cell cycles are precisely coordinated during each step of retinal development to ensure a certain proportion of retinal cells to be generated in a certain order (Figure 2) (6)

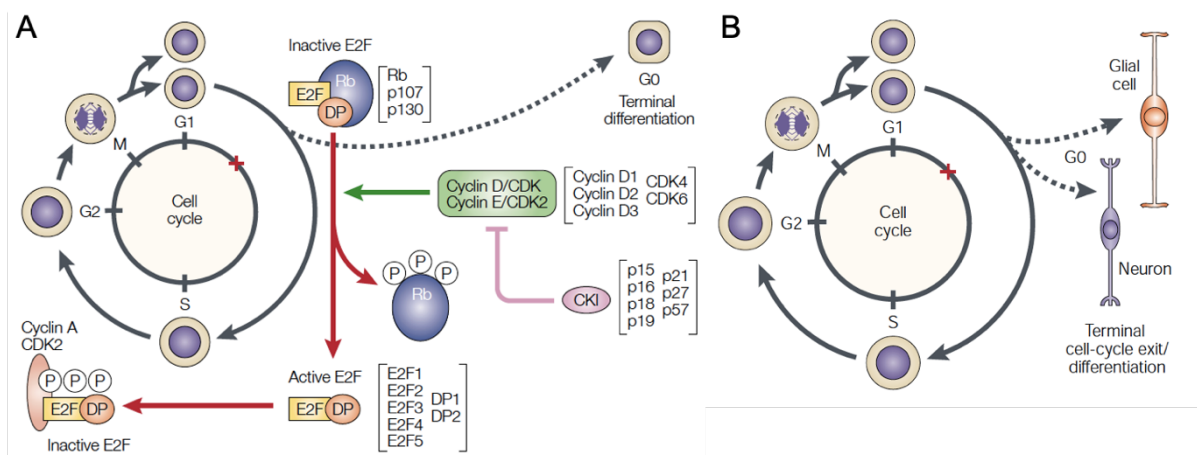


Figure 2. Schematic of cell cycle.

A, Following mitosis (M) phase, the phosphorylation state of the retinoblastoma protein (Rb), or its related family members (p107, p130) dictates whether a neuronal progenitor cell will progress through another round of cell division and pass the G1/S checkpoint, or exit the cell cycle and differentiate. It has been suggested that p130 is required to maintain cells in the differentiated state and that p107 might be important in regulating cell-cycle progression during late G1/S phase. Hypophosphorylated Rb, p107 or p130 binds to the E2F/DP heterodimer, which prevents transcription of E2F-regulated genes. Five main E2F family members have been identified and these molecules show preferential binding to different Rb family members. E2F1–3 preferentially bind to Rb, E2F5 preferentially binds to p107 and p130 and E2F4 binds to all three family members. The binding preferences and activities of the two DP genes (DP1, DP2) that have been identified in mammals have not been fully elucidated. Several different cyclin/cyclin-dependent kinase (CDK) holoenzymes have been identified that can phosphorylate Rb family members. Cyclin-D–CDK complexes are believed to act early during the G1 phase in mitotic cells. The three D-type cyclins (D1, D2 and D3) are expressed in a tissue-specific manner and associate with one of two CDKs (CDK4 and CDK6). Later during G1, the cyclin-E–CDK2 complex is believed to be the main Rb/p107/p130 kinase. Two families of cyclin-kinase inhibitors (CKIs) have been identified that can block the kinase activity of these holoenzymes. The INK family (p15, p16, p18 and p19) is believed to bind preferentially to cyclin-D–CDK complexes and leads to disruption of the holoenzyme subunits. Cip/Kip CKIs (p21, p27 and p57) can form stable complexes with either cyclin-D–CDK or cyclin-E–CDK complexes. According to one model, upregulation of an INK family member in a mitotic cell might lead

to disruption of a cyclin-D–CDK–(Cip/Kip) complex, which is followed by cyclin-E–CDK2 inactivation by the newly released Cip/Kip molecule.

B, Retinal progenitor cells exit the cell cycle during the G1 phase and differentiate into retinal neurons or Müller glia. Retinal neurons are considered in G0 phase and undergone terminal differentiations. Müller glia are considered quiescent in G0 phase but maintain the capability to re-enter cell cycle in some situations such as neural regeneration. Images and texts from (6).

Mouse retina is a model commonly used to study retinal development. In mouse, ganglion cells are generated first during embryogenesis, followed by the production of cone photoreceptors, horizontal cells, and most of the amacrine neurons. Bipolar neurons, Müller glia, the remaining amacrine neurons, and most rod photoreceptors are generated postnatally. The production of retinal cell types overlap at any given time and are regulated intrinsically by several transcriptional factors (Figure 3) (7).

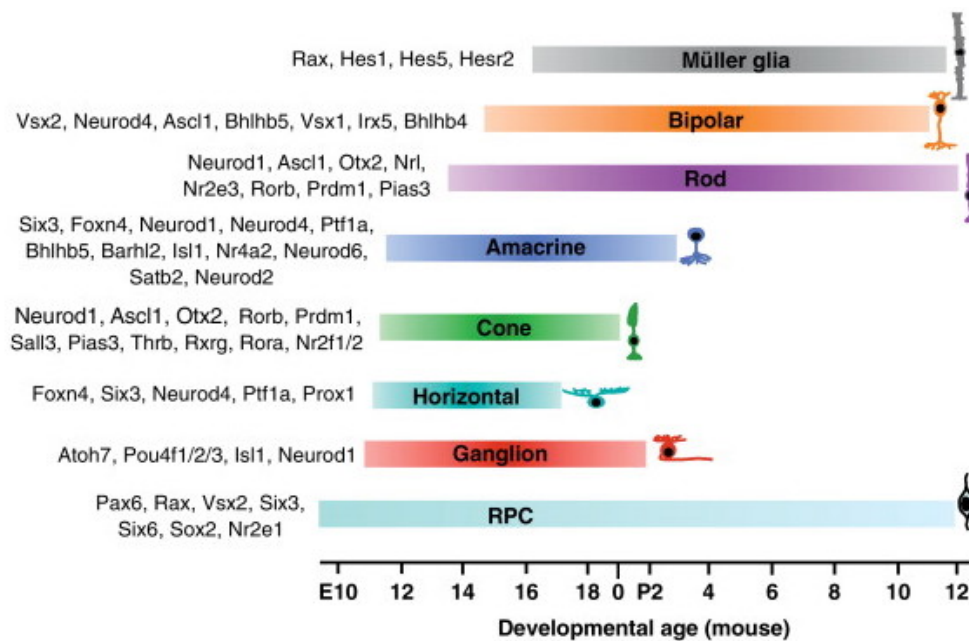


Figure 3. Schema of retinal cell development in mouse. Chronological sequence and transcriptional regulation of retinal cells and their common progenitor are shown. Developmental age is based on data from mouse. RPC: retinal progenitor cell; E: embryonic; P: postnatal. Image from (7).

Despite high levels of evolutionary conservation in vertebrate retina, there are important differences between human and mouse retinal development. A better understanding of human-specific retinal development is crucial for understanding retinoblastoma tumorigenesis and for finding new treatments for this disease.

Researchers made efforts in obtaining normal, well-preserved human fetal eyes to perform both morphological and molecular studies (8–10). Human retina also develops both prenatally and postnatally. All retinal cell types are generated before birth, and they continue to mature after birth. The generation of retinal cells in human shares a similar order as in mouse. It is also noted that human retinal development is spatially regulated, neural differentiation is initiated in the central retina near fovea and progresses toward the periphery.

Recent data have indicated that human Embryonic Stem cells (hESCs) and human induced Pluripotent Stem Cells (hiPSCs) can be used to generate retinal organoids that shares similar structure as neural retina (11,12). Different retinal cells were differentiated from the RPCs in these retinal organoids in a similar timely order to mouse (Figure 4) (12).

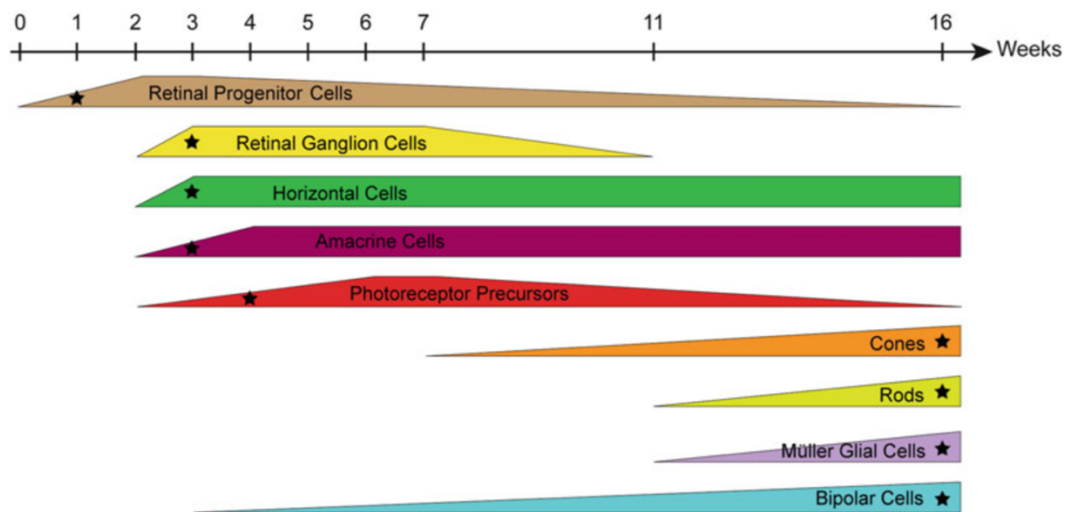


Figure 4. Schema of retinal cell development in differentiating human induced Pluripotent Stem Cells (hiPSCs). Temporal expression of retinal markers in differentiating hiPSCs-derived retinal organoids. Colored blocks represent evolution of each specific retinal cell type evaluated by RT-qPCR and immunohistochemistry. Black stars indicate the time when each specific retinal cell was undoubtedly detected by immunofluorescence. Image from (12).

RETINOBLASTOMA: EPIDEMIOLOGY AND CLINICAL ASPECTS

Incidence

Retinoblastoma is the cancer of the developing retina. It represents the most common intraocular tumor in children, and accounts for about 4% of all childhood cancers (13). Its incidence rate is about 1 in 16,000 live births per year, predicting approximately 8,000 new cases each year worldwide (14,15).

Risk factors

There are very few known risk factors for retinoblastoma. The risk of retinoblastoma is much higher in children from families with a history of retinoblastoma. Other risk factors include exposure to tobacco before pregnancy and advanced maternal age (16).

Clinical presentation

Retinoblastoma is generally presented during early childhood. More than 95% of the children diagnosed with retinoblastoma are younger than 5 years old (18). Disease can affect one (unilateral) or both eyes (bilateral), and rarely also the pineal (trilateral). Unilateral retinoblastomas count for 60% of cases and have a median age at diagnosis of two years, while 40% of cases are bilateral with a median age at diagnosis of one year (16).

The most common first sign of retinoblastoma is leukocoria, a white reflex visible through the pupil (Figure 5). Strabismus, or ocular misalignment, is the second common early sign of retinoblastoma. Early awareness of these signs are crucial, for timely treatment therefore can improve prognosis and survival rate. When diagnosed late, retinoblastoma may be presented with proptosis, buphthalmos, glaucoma and inflammation. These symptoms more often occur when the tumor spreads and are related to poor prognosis (15,17,18).

Retinoblastoma can spread through the optic nerve to the central nervous system. It can also invade choroid, sclera and orbit to enter the vessels and subsequently metastasize to bone marrow, liver and other organs (19)

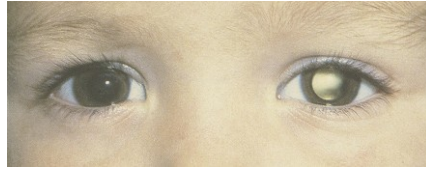


Figure 5. Leukocoria in a retinoblastoma patient. Image from (16).

Diagnosis

Dilated fundusopic examination is essential for the diagnosis of retinoblastoma as well as for the evaluation of the tumor stage. Retinoblastoma often present as one or multiple yellow-white retinal masses through funduscopy (Figure 6). There are two types of growth patterns of the retinoblastoma: endophytic and exophytic. Endophytic retinoblastoma refers to the tumor growing towards the vitreous, they often produce vitreous seeding. Exophytic tumors refer to the tumor extend beneath the retina. These tumors can cause exudative retinal detachments and may be associated with significant subretinal seeding. The appearance of seeds has been described as fine dust, spheres, or clouds, each with a progressively worse prognosis, respectively (15,18).

Ultrasound or magnetic resonance imaging can also help the diagnosis and staging. Calcification and optic nerve invasion can often be found through these examinations (15). For retinoblastoma, biopsy is not recommended, as it increases the risk of extraocular spread (15,18).

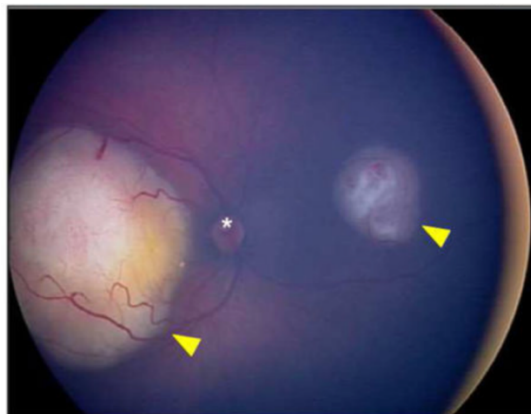


Figure 6. Funduscopy examination of a retinoblastoma patient eye. Yellow arrowheads marked two retinoblastomas. Image from (15).

Correct staging of the retinoblastoma are essential to making the treatment plan and the prediction of outcome. There are several different staging or classification systems. International Intraocular Retinoblastoma Classification (IIRC) (20) and its modified version Intraocular Classification of Retinoblastoma (ICRB) (21) are the systems generally adopted to evaluate intraocular diseases. Other classification schemes that are also used include staging of extraocular retinoblastoma (22) and staging of vitreous seeding (23). Retinoblastoma can also be staged using “Tumor Node Metastasis” (TNM) classification system (Table 1) (19).

Management

The goals of the management of retinoblastoma are to save the patient’s life, and to preserve the eye and the vision. Management plan is made based on the comprehensive evaluation of the disease by a team of experts including ophthalmologists, oncologists, radiologists, geneticists, pathologists and radiotherapists. Common primary treatments for intraocular disease include enucleation surgery, intra-venous chemotherapy with focal therapy (laser photocoagulation, thermotherapy or cryotherapy), intra-arterial chemotherapy with focal therapy, and focal therapy alone (14).

Enucleation is the surgical procedure to remove the affected eye. It is usually applied to the eye with a large tumor burden in unilateral disease. In France, it is performed in around 50% of the cases. With the development of new eye-salvage therapies, the use of enucleation will be further decreased. Cosmetic reconstruction can be performed to improve the aesthetic outcome (14).

Common used drugs in intra-venous chemotherapy include carboplatin, etoposide, vincristine and cyclosporine. In intra-arterial chemotherapy, melphalan, topotecan and carboplatin are often used (14).

When primary treatments fail, second-line treatments can be applied, which include focal therapy, repeated systemic chemotherapy, intra-arterial chemotherapy, brachytherapy, external beam radiotherapy and stereotactic conformal or proton-beam

radiotherapy. Vitreous seeding is the main reason of failure in eye preservation. Intravitreal chemotherapy may be used to improve the drug delivery in vitreous (14).

Treatments for extraocular retinoblastoma include neoadjuvant chemotherapy, enucleation, orbital radiation and adjuvant chemotherapy. In metastatic diseases, intrathecal chemotherapy (injection into the cerebrospinal fluid space), high-dose chemotherapy and stem cell rescue may be used (14).

Prognosis

Prognosis of retinoblastoma is good in high-income countries (patient survival is more than 95%). However, in low-income countries patient survival is low (30%). Poor outcome correlates with lack of retinoblastoma specific care center, late diagnosis and poor compliance of treatment due to family's poor socioeconomic status. Without timely diagnosis and appropriate treatment, lethal metastatic disease may develop (14).

Table 1. Summary of clinical staging systems. From (19).

	AJCC Clinical Staging, 8th edition (Mallipatna, et al., 2017)	IIRC Group (Murphree, 2005)	ICRB Group (Shields, 2006)	IRSS Stage (Chantada, 2006)
cT1	Intra-retinal tumour(s) with subretinal fluid ≤5 mm from base of any tumour			
cT1a	Tumors ≤3 mm and further than 1.5mm from disc and fovea	A, >3mm to fovea or B, 1.5 to 3mm	A, >3mm to fovea or B, 1.5 to 3mm	-
cT1b	Tumors >3mm or closer than 1.5mm from disc or fovea	B	B, ≤3mm or C, 3 to 5mm	-
cT2	Intraocular tumour(s) with retinal detachment, vitreous seeding, or subretinal seeding.			
cT2a	Subretinal fluid >5 mm from the base of any tumour	C, >5 mm or D, > 1 quadrant	C, or E, tumour >50% of eye volume	-
cT2b	Vitreous seeding and/or subretinal seeding	C, "local" or D, "diffuse"	C, ≤3 mm or D, > 3 mm or E, tumour >50% of eye volume	-
CT3	Advanced intraocular tumour(s)			
cT3a	Phthisis or pre-phthisis bulbi	E	E	I or II
cT3b	Tumour invasion of choroid, pars plana, ciliary body, lens, zonules, iris, or anterior chamber	E	E	I or II
cT3c	Raised intraocular pressure with neovascularization and/or buphthalmos	E	E	I or II
cT3d	Hyphaema and/or massive vitreous haemorrhage	E	E	I or II
cT3e	Aseptic orbital cellulitis	E	E	I or II
cT4	Extraocular tumour(s) involving orbit, including optic nerve			
cT4a	Radiologic evidence of retrobulbar optic nerve involvement or thickening of optic nerve or involvement of orbital tissues			I or II
cT4b	Extraocular tumour clinically evident with proptosis and/or an orbital mass			IIIa
N1	Evidence of preauricular, submandibular, and cervical lymph node involvement			IIIb
CM1	Clinical signs of distant metastasis			
cM1a	Tumour(s) involving any distant site (e.g., bone marrow, liver) on clinical or radiologic tests			IVa
cM1b	Tumour involving the CNS on radiologic imaging (not including trilateral retinoblastoma)			IVb
H	Hereditary Trait			
HX	Unknown or insufficient evidence of a constitutional RB1 gene mutation			
H1	Normal RB1 alleles in blood tested with demonstrated high-sensitivity assays			
H2	Bilateral retinoblastoma, retinoblastoma with an intracranial primitive neuroectodermal tumour (i.e., trilateral retinoblastoma), patient with family history of retinoblastoma, or molecular definition of a constitutional <i>RB1</i> gene mutation			

RETINOBLASTOMA: GENETIC AND MOLECULAR BASIS

Knudson's hypothesis

A genetic model to explain the difference of age at diagnosis and incidence rates between unilateral and bilateral retinoblastoma patients was proposed by Knudson (24). Two mutations are required to develop retinoblastoma, each of the two mutations occurs at a rather constant rate. These two mutations can both be somatic mutations, or one of the two mutations can be a germline mutation. Patients who carry one germline mutation can develop the disease earlier in their lives than the patients who need to accumulate two somatic mutations.

The *RB1* gene inactivation

The two hits underlying retinoblastoma was later recognized by the finding of chromosome 13q14 deletions in retinoblastoma patients and *RB1* gene being identified in the smallest deleted region in this deleted chromosome segment (25). Several studies further elucidated that inactivation of both alleles of *RB1* can lead to retinoblastoma formation (26). By analysis of a cohort of more than 1000 retinoblastomas, *RB1* were found inactivated in 98% of the cases (27).

The *RB1* gene is identified as the first tumor-suppressor gene. It is located on chromosome 13q14, spans 180kb, and has 27 exons. It encodes a nuclear phosphoprotein known as the retinoblastoma protein (Rb). The Rb protein has 928 amino acids and 3 domains: N-terminal domain, central pocket domain and C-terminal domain. The pocket in the central domain (Figure 7A, B) can bind to E2F, which inactivate E2F and prevent cells entering S phase (Figure 2A). When the Rb protein is phosphorylated by cyclin-dependent kinases (CDKs), the pocket domain undergoes conformation change and release E2F, the latter promotes the cell cycle G1-S transition (Figure 7C). Retinoblastoma proteins also interact with chromatin regulators (Figure 7) (28).

CHAPTER 1

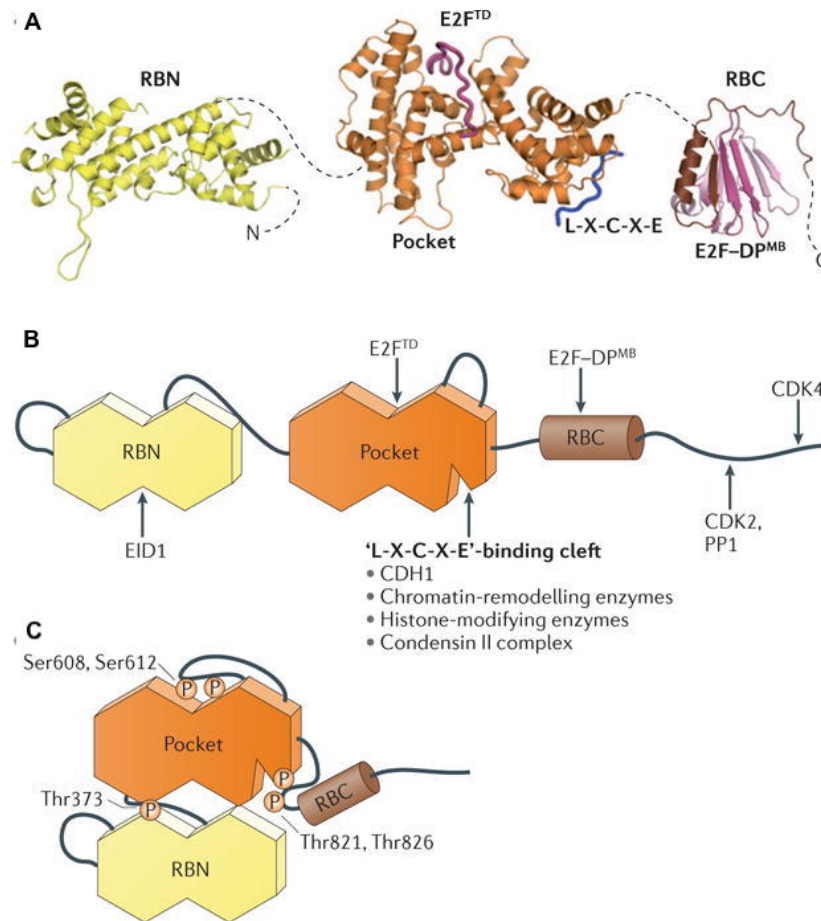


Figure 7. Schema of retinoblastoma protein structure.

A, The ribbon diagram showed the three domains of retinoblastoma protein (Rb) in active state. Central pocket domain of retinoblastoma protein binds to E2F and L-X-C-X-E peptide. RBN: Rb N-terminal domain; Pocket, Rb central pocket domain; RBC: Rb C-terminal domain; E2F^{TD}, E2F transactivation domain; DPMB, differentiation-related polypeptide marked box.

B, The schematic structure of Rb showed known binding sites for protein partners.

C, The schematic structure of Rb in its inactive state. Thr373 phosphorylation drives interdomain docking of the N-terminal domain and the pocket domain, whereas Ser608 and Ser612 and Thr821 and Thr826 phosphorylation induce binding of the pocket loop and the C-terminal domain, respectively, to the pocket domain. These different conformational changes inhibit specific RB–protein interactions with E2F and other protein partners. Images from (28).

MYCN amplification

MYCN is located in chromosome 2p and is a member of MYC oncogene family. *MYCN* amplifications are found in a variety of cancers, including neuroblastomas and retinoblastomas (14). It has been shown that *MYCN* does not likely contribute to retinoblastoma initiation but plays a role in *RB1*-mutated tumor progression (29).

However, a more recent study showed that *MYCN* amplifications were found in half of the *RB*^{+/+} retinoblastoma cases, suggesting that *MYCN* amplifications may actually be initiating events in rare cases when *RB1* gene are intact (27).

Additional genomic and epigenetic alterations

Biallelic *RB1* inactivation is necessary to initiate most retinoblastomas, but it is not sufficient, as retinoma - the benign retinal lesion - similarly involves bi-allelic *RB1* loss (30). Another recurrent mutation identified in retinoblastoma is *BCOR*, an interacting corepressor of BCL6 (31,32).

Besides mutations, recurrent DNA copy number alterations are found in retinoblastoma, include chr1q gain, chr2p gain, chr6p gain, chr13q gain, chr13q loss, chr16q loss, chr17p loss, and chr19 gain (33). These genomic alterations may map to several candidate oncogenes such as *KIF14* (chr 1), p53 regulator *MDM4* (chr 1), *E2F3* (chr 6) and *DEK* (chr 6); microRNA clusters mir-106b~25 (chr 7), mir-17~92 (chr 13); and retinoblastoma protein family member *RBL2* (chr 16) (14).

Dysregulation of chromatin regulators are also reported in retinoblastoma (34,35). This includes genes involved in DNA methylation (*DNMT1*, *DNMT3A*, *DNMT3B*, *TET1*, *TET3*), histone modification (*HDAC2*, *KDM1A*, *EZH2*), Chromatin remodeling (*HELLS*, *SMARCC1*, *SMARCD1*, *HMGA2*, *HMGB2*, *HMGB3*).

MOLECULAR CLASSIFICATIONS OF RETINOBLASTOMA

Cancer is a heterogeneous disease, therefore it is crucial to identify specific disease subtypes in order to find appropriate treatment. Clinical staging, pathological observation and genetic analysis provide hints for the right treatment. Beyond that, are there molecular subtypes that may lead to discovery of specific targeted therapeutics for the disease? Several transcriptomic analysis of retinoblastoma have been performed in the last decade, trying to answer this question.

From a transcriptomic study of 52 retinoblastomas, McEvoy *et al.* proposed that retinoblastomas express various retinal differentiation markers that are incompatible in normal development, but reported finding no distinct subtype (Figure 8) (36). Two years later, Kapatai *et al.* identified two subtypes of retinoblastoma, from a study of 21 retinoblastomas and 3 normal developing retina, with one subtype co-expressing different retinal cell markers – similar to what McEvoy *et al.* has found – and another subtype mainly expressing cone photoreceptor markers (Figure 9) (37). By applying the same clustering methods as Kapatai *et al.*, Kooi *et al.* found two clusters using transcriptomic data of a series of 76 retinoblastomas. They further proposed that the two subgroups are not dichotomous, but continuous, exhibiting gradient photoreceptor signature and ribosome/mRNA synthesis signature (Figure 10) (38).

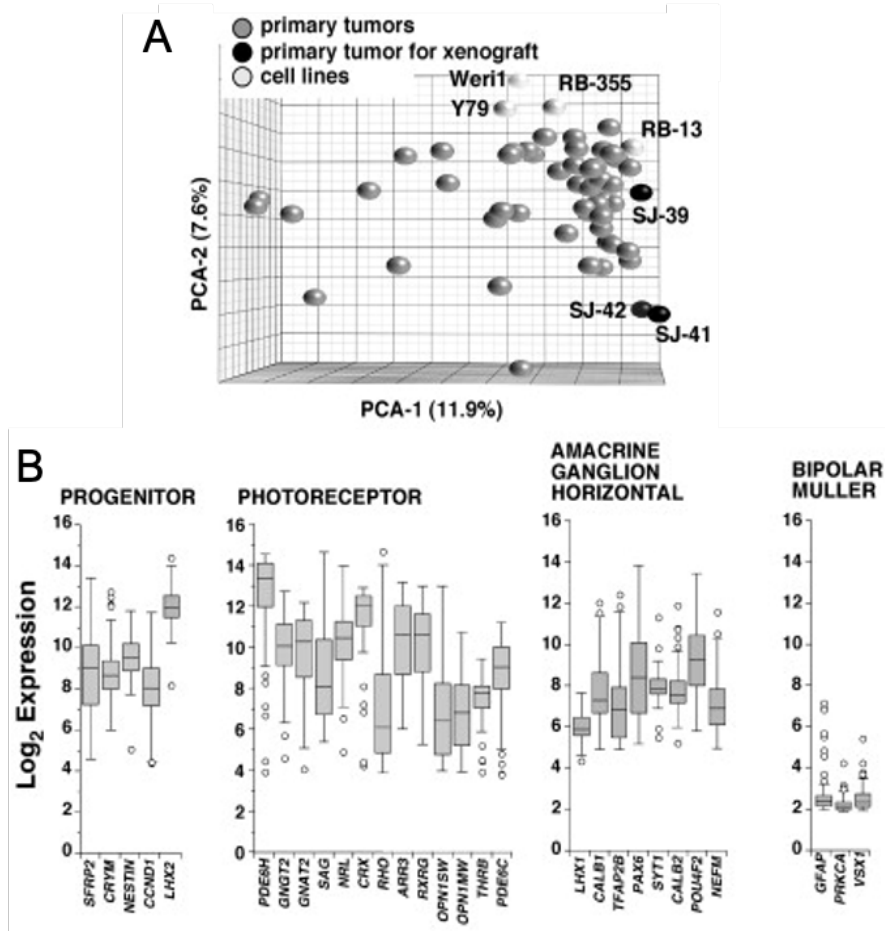


Figure 8. According to McEvoy et al., Retinoblastoma express various retinal markers without clear subtype identified.

A, 3-D PCA plot of transcriptomic data of 52 retinoblastomas and 7 cell models.

B, Box plots of the Log₂ mRNA expression for a series of retinal genes in 52 retinoblastomas. The bottom and top of the boxes are the 25th and 75th percentile and the line in the middle is the 50th percentile. The lines represent one standard deviation above and below the mean and data points outside of those lines represent outliers. Image from (36).

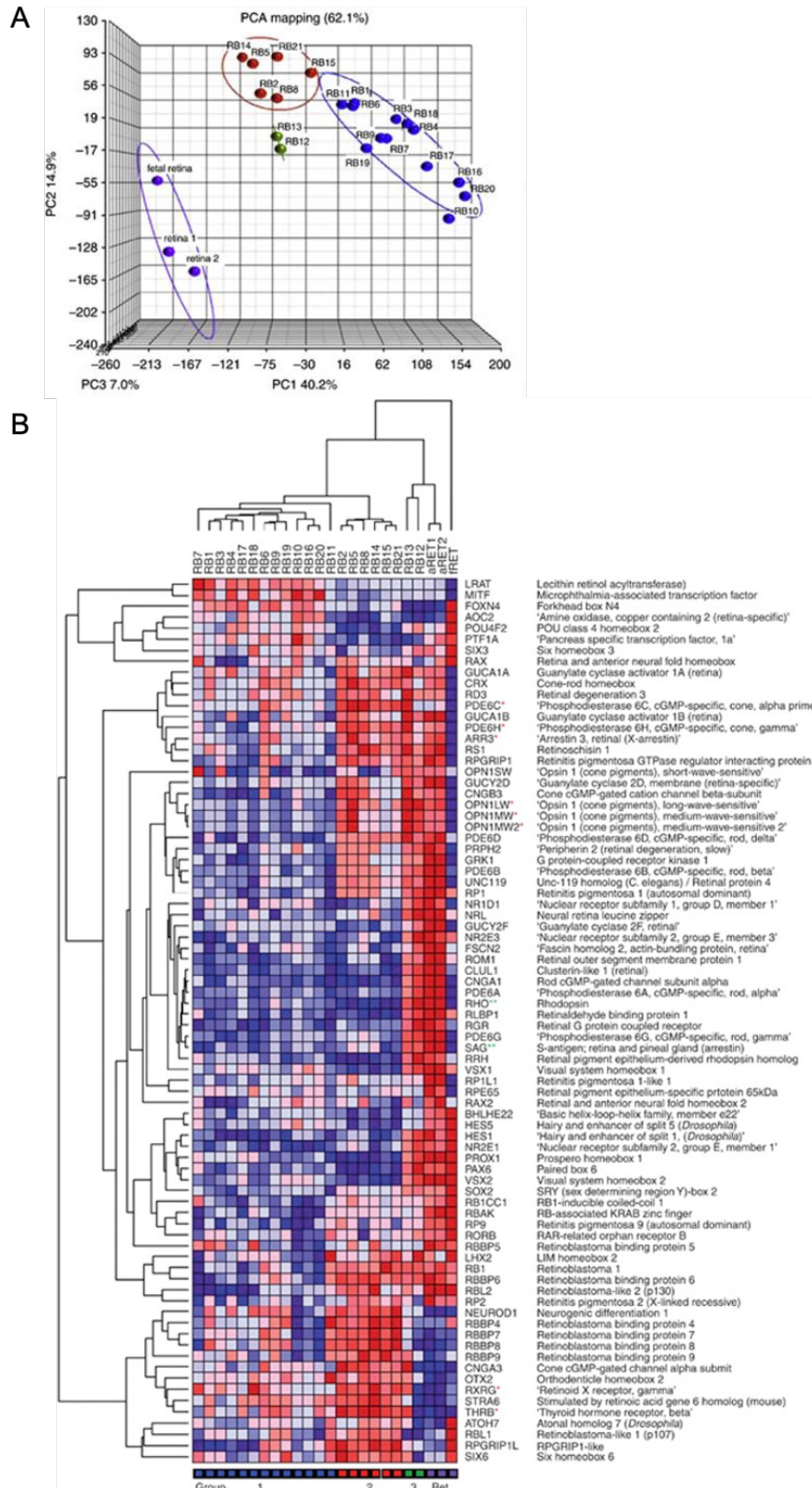


Figure 9. Identification of two main subtypes in retinoblastoma by Kapatai et al..

A, 3-D PCA plot of transcriptomic data of 21 retinoblastomas (shown as red, blue and green dots) and 3 retina samples (shown as purple dots). Color indicates different subgroup identified by clustering analysis. Two main clusters are identified in retinoblastoma samples (shown as red or blue dots). A third

CHAPTER 1

cluster of tumors (RB12 and RB13, shown as green dots) are likely contaminated with many normal cells as they expressed many genes similar to normal retina.

B, Expression of retinal markers in 21 retinoblastomas and 3 retina samples. Image from (37).

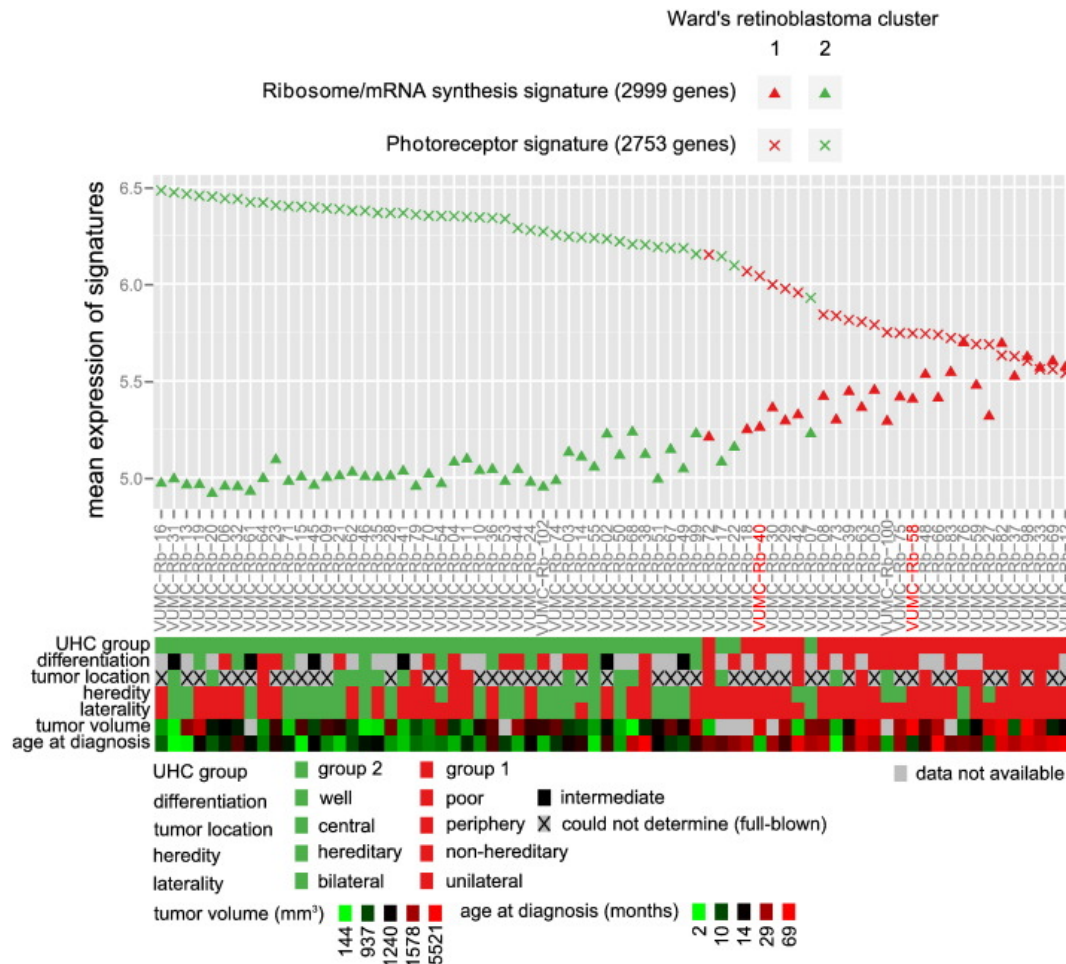


Figure 10. Quantification of differential signatures in relation to clinical and histopathological variables in the two subtypes identified by Kooi *et al.*.

Color indicates Ward's retinoblastoma cluster 1 (red) or 2 (green) in 74 retinoblastoma. The *RB1*^{+/+}*MYCN*-amplified tumors are indicated by red labels. Image from (38).

CELL-OF-ORIGIN OF RETINOBLASTOMA

RB1 loss is found in many cancers, but its predominant contribution to retinoblastoma initiation is unique. In retinoblastoma, 98% of the patients carried *RB1* mutations (27). Why is the human retina extremely sensitive to *RB1* loss? Finding out which cells are most sensitive to *RB1* loss and how they are transformed to tumor cells can shed light on the targeted treatment for retinoblastoma.

Researchers have been working on this question but have reached different conclusions: photoreceptor cone cells, progenitor cells or amacrine/horizontal cells are hypothesized as cell of origin of retinoblastoma in different studies. The finding that various retinal markers were expressed by retinoblastoma may indicate a progenitor cell of origin, but may also be due to oncogenic deregulations. Studies in mouse models suggest an amacrine/horizontal origin, as early-stage tumors expressed genes found in retinal synapse layers and exhibited similar morphological differentiations to the two retinal interneurons (39). However, retinoblastoma mouse models are different from human retinoblastoma, because tumor formation in mouse requires *Rb1*, *Rb12* or *Cdkn1b* perturbation in addition to *Rb1* loss.

Recent studies in human retinoblastoma pointed out a more probable cone lineage origin of retinoblastoma. It was first proposed by the finding that the topographic distribution of emerging retinoblastomas mimics the horizontal visual streak characteristic of long-wavelength and medium-wavelength cones (40). In accordance with this view, *RB1*^{-/-} retinoblastomas consistently express cone precursor markers but not other retinal cell-type-specific proteins. In addition, human cone precursors prominently express MDM2 and MYCN, both of which are required for retinoblastoma proliferation and survival (41). Most recent study in *RB1* depleted cultured retina led us to focus on ARR3⁺ maturing cone cells, from that retinoblastoma-like lesions are formed (Figure 11) (42).

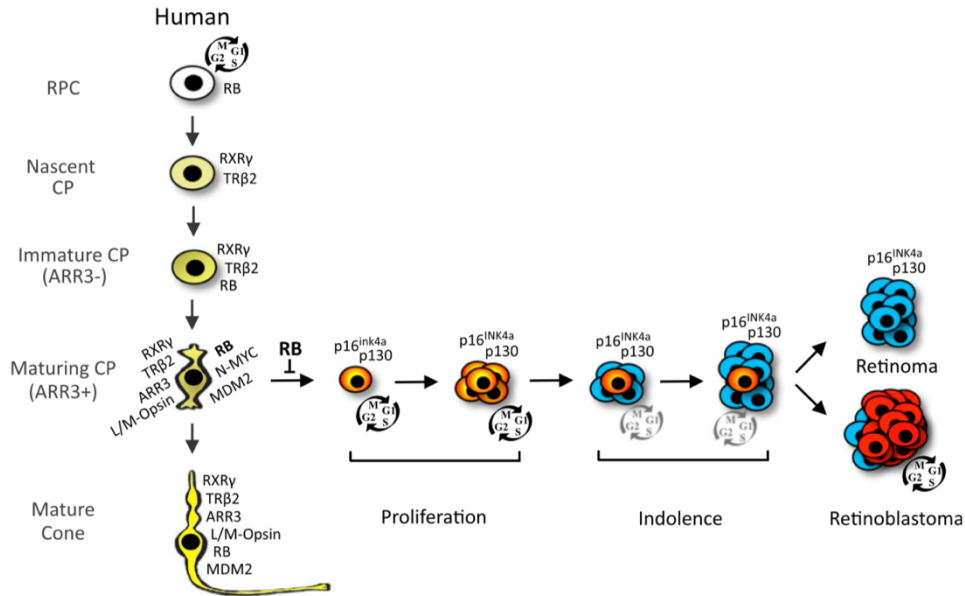


Figure 11. Schema of retinoblastoma formation from maturing (ARR3+) cone precursors.

The left part of the schema illustrated the generation of cone precursors (CPs) from retinal progenitor cells (RPCs) in human. Nascent CPs are RXR γ +, TR β 2+ and have minimal RB expression (43). RB increases during CP maturation and in association with N-MYC and MDM2 expression (Xu et al. 2009). The right part of the schema illustrated that RB loss in maturing (ARR3+) CP caused it entering cell-cycle and proliferation (orange). Cell-cycle entry after RB depletion were only observed in ARR3+ CP instead of ARR- immature CP. This cell-cycle entry maybe followed by cell-cycle withdrawal (blue) and an indolent phase that may give rise to permanently quiescent retinomas or to retinoblastoma tumors (red). Image from (42)

OMICS APPROACHES

In 1990, the ambitious Human Genome Project was launched, scientists all over the world worked together trying to decipher the entire human genome. During the next 13 years, the 3 billion base pairs of the human genome were determined by Sanger sequencing, a DNA sequencing method based on chain termination with the four color-labeled nucleotides. The huge success of the Human Genome Project was not only because it accomplished the mission to decode 99% of human euchromatin sequences, more importantly, it revealed several optimization opportunities for the sequencing methods, and provided a reference genome that enabled the potential of short-read sequencing (44).

Two years later after the completion of the Human Genome Project, in 2005, the first cyclic-array sequencing methods, that we now refer as the Next-Generation sequencing (NGS), were reported (45,46). These methods enabled parallel sequencing of millions of features at a much faster speed with a lower cost. The cyclic-array sequencing is described as the sequencing of a dense array of fragmented DNA amplicons through iterative cycles of enzymic nucleotide addition and imaging-based data acquisition (Figure 12).

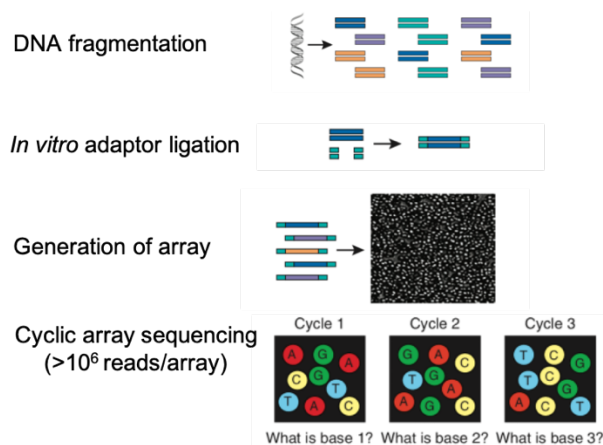


Figure 12. Work flow of cyclic array sequencing.

DNAs were fragmented and ligated *in vitro* with the common adapters. Millions of the DNA amplicons were tethered into a high-density DNA array. Reagents including 4-color-labeled dNTP were then added allowing the multiple cycles of enzymatic extension of the millions of amplicons in parallel. In each cycle, the base added to each array feature was determined by the detected color labels by imaging. successive rounds of base extension and imaging allowed sequencing for each array feature. Images from (44).

Based on the NGS technology, various high-throughput methods were developed to profile thousands to millions of DNAs, RNAs, proteins in one experiment. The advances of these omics approaches have revolutionized biomedical research in the past two decades. The International Cancer Genome Consortium (ICGC) is one of the ambitious initiatives after the Human Genome Project, it has collected clinical and molecular data in more than 24,000 tumors across 86 cancer projects. The molecular data include somatic variants, copy-number alterations, gene expression, microRNA profiles, DNA methylations and chromatin conformations. The consortium provides extensive knowledge on cancer biology and can guide the development of new therapeutics targeting the specific mutations, deregulated pathways in certain cancer subtypes.

In the recent decade, omics has entered “single-cell” age. Emerging methods were developed to profile the molecules at a single-cell scale. These methods coupled with bioinformatic algorithms allowed researchers to identify rare cell populations and new biomarkers in normal and tumor tissues, to better characterize the subclones and the microenvironment in tumors, as well as to trace tumor progression within one tumor.

Table 2 summarizes single-cell RNA sequencing methods (47). C1 and Smart-seq2 are commercialized methods that enable full length RNA sequencing in hundreds of cells (48,49). Chromium is a droplet-based 3' RNA-sequencing method that increased the throughput into profiling thousands of cells in one experiment (Figure 13) (50).

Table 2. Summary of single-cell RNA-sequencing methods. From (47)

	Transcript data	Platform	Throughput (number of cells)	Read depth (per cell)	Reaction volume	Reference
C1	Full length	Microfluidics	10^2 – 10^3	10^6	Nanoliter	Pollen et al., 2014
Smart-seq2	Full length	Plate-based	10^2 – 10^3	10^6	Microliter	Picelli et al., 2013
MATQ-seq	Full length	Plate-based	10^2 – 10^3	10^6	Microliter	Sheng et al., 2017
MARS-seq	3'-end cut	Plate-based	10^2 – 10^3	10^4 – 10^5	Microliter	Jaitin et al., 2014
CEL-seq	3'-end cut	Plate-based	10^2 – 10^3	10^4 – 10^5	Nanoliter	Hashimshony et al., 2012
Drop-seq	3'-end cut	Droplet	10^3 – 10^4	10^4 – 10^5	Nanoliter	Macosko et al., 2015
InDrop	3'-end cut	Droplet	10^3 – 10^4	10^4 – 10^5	Nanoliter	Klein et al., 2015
Chromium	3'-end cut	Droplet	10^3 – 10^4	10^4 – 10^5	Nanoliter	Zheng et al., 2017
SEQ-well	3'-end cut	Nanowell array	10^3 – 10^4	10^4 – 10^5	Nanoliter	Gierahn et al., 2017
SPLIT-seq	3'-end cut	Plate-based	10^3 – 10^5	10^4	Microliter	Rosenberg et al., 2018

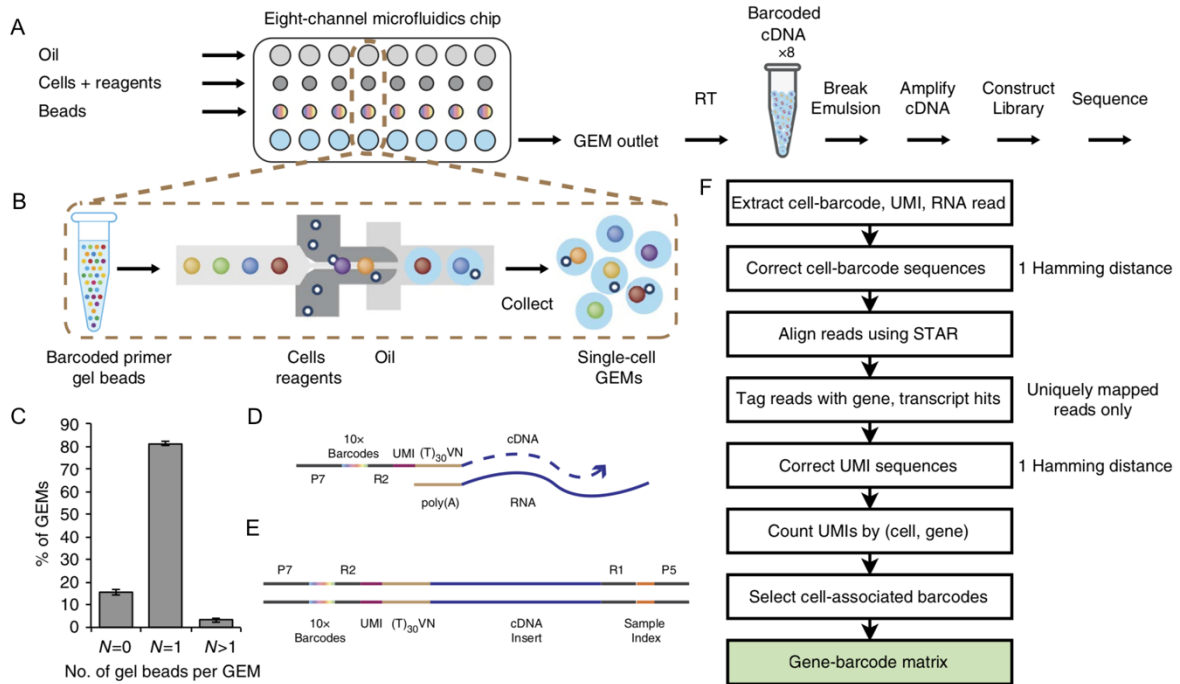


Figure 13. Workflow of Chromium single-cell RNA-sequencing.

A, Schema of the experiment workflow. Cells were combined with reagents in one channel of a microfluidic chip, and gel beads from another channel to form Gel beads in Emulsion (GEMs). Reverse transcription takes place inside each GEM, after which cDNAs are pooled for amplification and library construction in bulk.

B, Gel beads loaded with primers and barcoded oligonucleotides are first mixed with cells and reagents, and subsequently mixed with oil-surfactant solution at a microfluidic junction. Single-cell GEMs are collected in the GEM outlet.

C, Percentage of GEMs containing 0 gel bead (N=0), 1 gel bead (N=1) and more than 1 gel bead (N>1). Data include five independent runs from multiple chip and gel bead lots over >70k GEMs for each run, n=5, mean±s.e.m.

D, Gel beads contain barcoded oligonucleotides consisting of Illumina adapters, 10x barcodes, UMIs and oligo dTs, which prime RT of polyadenylated RNAs.

E, Finished library molecules consist of Illumina adapters and sample indices, allowing pooling and sequencing of multiple libraries on a next-generation short read sequencer.

F, Cell Ranger bioinformatic pipeline to obtain a gene-barcode matrix (highlighted in green).

Images from (50)

REFERENCES

1. Sung CH, Chuang JZ. The cell biology of vision. *J Cell Biol.* 2010;190(6):953–63.
2. Baden T, Euler T, Berens P. Understanding the retinal basis of vision across species. *Nat Rev Neurosci.* 2020 Jan;21(1):5-20.
3. Masland RH. Neuronal diversity in the retina. *Curr Opin Neurobiol.* 2001;11(4):431–6.
4. Sanes JR, Masland RH. The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification. *Annu Rev Neurosci.* 2015;38:221–46.
5. Reichenbach A, Bringmann A. Glia of the human retina. *Glia.* 2020;68(4):768–96.
6. Dyer MA, Cepko CL. Regulating proliferation during retinal development. *Nat Rev Neurosci.* 2001;2(5):333–42.
7. Bassett EA, Wallace VA. Cell fate determination in the vertebrate retina. *Trends Neurosci.* 2012;35(9):565–73.
8. Hendrickson A. Development of Retinal Layers in Prenatal Human Retina. *Am J Ophthalmol.* 2016 Jan;161:29-35.e1.
9. Hoshino A, Ratnapriya R, Brooks MJ, Chaitankar V, Wilken MS, Zhang C, et al. Molecular Anatomy of the Developing Human Retina. *Dev Cell.* 2017 Dec;43(6):763-779.e4.
10. Lu Y, Shiao F, Yi W, Lu S, Wu Q, Pearson JD, et al. Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. *Dev Cell.* 2020;53(4):473-491.e9.
11. Osakada F, Ikeda H, Sasai Y, Takahashi M. Stepwise differentiation of pluripotent stem cells into retinal cells. *Nat Protoc.* 2009;4(6):811–24.
12. Reichman S, Goureau O. Production of retinal cells from confluent human iPS Cells. *Methods Mol Biol.* 2016;1357:339–51.
13. Abramson DH. Retinoblastoma in the 20th century: Past success and future challenges. The Weisenfeld Lecture. *Investig Ophthalmol Vis Sci.* 2005;46(8):2684–91.
14. Dimaras H, Corson TW, Cobrinik D, White A, Zhao J, Munier FL, et al. Retinoblastoma. *Nat Rev Dis Prim.* 2015 Aug;1:15021.
15. Dimaras H, Corson TW. Retinoblastoma, the visible CNS tumor: A review. *J Neurosci Res.* 2019 Jan;97(1):29–44.
16. Aerts I, Lumbroso-Le Rouic L, Gauthier-Villars M, Brisse H, Doz F. [Retinoblastoma update]. *Arch Pediatr.* 2016 Jan;23(1):112–6.
17. Dimaras H, Kimani K, Dimba EAO, Gronsdahl P, White A, Chan HSL, et al. Retinoblastoma. *Lancet.* 2012;379(9824):1436–46.
18. Ortiz M V., Dunkel IJ. Retinoblastoma. *J Child Neurol.* 2016;31(2):227–36.
19. Mallipatna AC et al. the 8TH edition of the AJCC staging for retinoblastoma. 2017. Available from: <https://wechope.org/wp-content/uploads/2019/06/TNM-Staging-Comparison-Chart.pdf>.
20. Linn Murphree A. Intraocular retinoblastoma: the case for a new group classification. *Ophthalmol Clin North Am.* 2005 Mar;18(1):41–53, viii.
21. Shields CL, Shields JA. Basic understanding of current classification and management of retinoblastoma. *Curr Opin Ophthalmol.* 2006 Jun;17(3):228–34.
22. Chantada G, Doz F, Antoneli CBG, Grundy R, Clare Stannard FF, Dunkel IJ, et al. A proposal for an international retinoblastoma staging system. *Pediatr Blood Cancer.* 2006 Nov;47(6):801–5.
23. Munier FL. Classification and management of seeds in retinoblastoma. Ellsworth Lecture Ghent August 24th 2013. *Ophthalmic Genet.* 2014 Dec;35(4):193–207.
24. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA.* 1971;68(4):820–3.
25. Benedict WF, Murphree AL, Banerjee A, Spina CA, Sparkes MC, Sparkes RS. Patient with 13 chromosome deletion: evidence that the retinoblastoma gene is a recessive cancer gene. *Science.* 1983 Feb;219(4587):973–5.
26. Berry JL, Polski A, Cavenee WK, Dryja TP, Linn Murphree A, Gallie BL. The RB1 story: Characterization and cloning of the first tumor suppressor gene. *Genes (Basel).* 2019;10(11).
27. Rushlow DE, Mol BM, Kennett JY, Yee S, Pajovic S, Thériault BL, et al. Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *Lancet Oncol.* 2013 Apr;14(4):327-34.
28. Dick FA, Rubin SM. Molecular mechanisms underlying RB protein function. *Nat Rev Mol Cell Biol.* 2013;14(5):297–306.
29. Squire J, Goddard AD, Canton M, Becker A, Phillips RA, Gallie BL. Tumour induction by the retinoblastoma mutation is independent of N-myc expression. *Nature.* 1986 Aug;322(6079):555–7.

CHAPTER 1

30. Gallie BL, Dimaras H, Orlic M, Corson TW, Piovesan B, Prigoda NL, et al. Loss of RB1 induces non-proliferative retinoma: increasing genomic instability correlates with progression to retinoblastoma. *Hum Mol Genet.* 2008 Jan 22;17(10):1363–72.
31. Zhang J, Benavente CA, McEvoy J, Flores-Otero J, Ding L, Chen X, et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature.* 2012 Jan;481(7381):329–34.
32. Kooi IE, Mol BM, Massink MPG, Ameziane N, Meijers-Heijboer H, Dommering CJ, et al. Somatic genomic alterations in retinoblastoma beyond RB1 are rare and limited to copy number changes. *Sci Rep.* 2016 Apr;6:25264.
33. Kooi IE, Mol BM, Massink MPG, de Jong MC, de Graaf P, van der Valk P, et al. A Meta-Analysis of Retinoblastoma Copy Numbers Refines the List of Possible Driver Genes Involved in Tumor Progression. *PLoS One.* 2016;11(4):e0153323.
34. Lee C, Kim JK. Chromatin regulators in retinoblastoma: Biological roles and therapeutic applications. *J Cell Physiol.* 2021;236(4):2318–32.
35. Ganguly A, Shields CL. Differential gene expression profile of retinoblastoma compared to normal retina. *Mol Vis.* 2010;16(January):1292–303.
36. McEvoy J, Flores-Otero J, Zhang J, Nemeth K, Brennan R, Bradley C, et al. Coexpression of normally incompatible developmental pathways in retinoblastoma genesis. *Cancer Cell.* 2011 Aug;20(2):260–75.
37. Kapatai G, Brundler M-A, Jenkinson H, Kearns P, Parulekar M, Peet AC, et al. Gene expression profiling identifies different sub-types of retinoblastoma. *Br J Cancer.* 2013 Jul;109(2):512–25.
38. Kooi IE, Mol BM, Moll AC, van der Valk P, de Jong MC, de Graaf P, et al. Loss of photoreceptoriness and gain of genomic alterations in retinoblastoma reveal tumor progression. *EBioMedicine.* 2015 Jul;2(7):660–70.
39. Johnson DA, Zhang J, Frase S, Wilson M, Rodriguez-Galindo C, Dyer MA. Neuronal differentiation and synaptogenesis in retinoblastoma. *Cancer Res.* 2007;67(6):2701–11.
40. Munier FL, Balmer A, van Melle G, Gailloud C. Radial asymmetry in the topography of retinoblastoma. Clues to the cell of origin. *Ophthalmic Genet.* 1994;15(3–4):101–6.
41. Xu XL, Fang Y, Lee TC, Forrest D, Gregory-Evans C, Almeida D, et al. Retinoblastoma Has Properties of a Cone Precursor Tumor and Depends Upon Cone-Specific MDM2 Signaling. *Cell.* 2009;137(6):1018–31.
42. Singh HP, Wang S, Stachelek K, Lee S, Reid MW, Thornton ME, et al. Developmental stage-specific proliferation and retinoblastoma genesis in RB-deficient human but not mouse cone precursors. *Proc Natl Acad Sci.* 2018 Oct 2;115(40):E9391 LP-E9400.
43. Lee TC, Almeida D, Claros N, Abramson DH, Cobrinik D. Cell cycle-specific and cell type-specific expression of Rb in the developing human retina. *Investig Ophthalmol Vis Sci.* 2006;47(12):5590–8.
44. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol.* 2008;26(10):1135–45.
45. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science.* 2005 Sep;309(5741):1728–32.
46. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature.* 2005 Sep;437(7057):376–80.
47. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* 2017 Aug;9(1):75.
48. Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH, et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol.* 2014 Oct;32(10):1053–8.
49. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods.* 2013 Nov;10(11):1096–8.
50. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 2015 May 21;161(5):1202–14.

CHAPTER 2.

OBJECTIVES

Context

Retinoblastoma is a rare cancer of the developing retina. The goals of the treatment are to save a patient's life, and to preserve the eye(s) and the vision.

Institut Curie is the leading treatment and research center for retinoblastoma in France and is one of the largest in the world. The institute receives around 50 new retinoblastoma patients per year. The institute conduct clinical, translational and basic research with the aim of reducing side-effects of the treatments, improving the existing treatments, and discovering new treatments that can preserve the eyes and the vision.

The Molecular Oncology research team at Institut Curie is led by Dr. François Radvanyi and is focused on understanding molecular mechanisms of tumor progression in bladder cancer and in retinoblastoma. The team is composed of biologists, bioinformaticians and clinicians. The research strategies include identification of genetic and epigenetic events from large-scale genomic and transcriptomic data and functional validation of candidate genes using cell and animal models. The knowledge obtained from understanding cancer biology can be translated to the development of new therapeutic targets, diagnostic and prognostic markers.

The CIT research team at La Ligue Contre Le Cancer is led by Dr. Aurélien de Reyniès and aims at identifying cancer subtypes and characterizing their specific driving events. The team has extensive capabilities in bioinformatics and collaborates with multiple clinical and research teams in France to collect and analyze large-scale genomic and transcriptomic datasets in various cancer types under the CIT (Carte d'Identité des Tumeurs) program.

Rationale

Thanks to the timely diagnosis and treatments, retinoblastoma is no longer a lethal disease in high-income countries. Less-invasive local procedures and localized chemotherapies have been developed that largely decreased the side-effects of the treatments and improved quality of life for patients. However, enucleation of the affected eye(s) are still needed in around half of patients in France. No targeted therapies have been successfully developed into clinical stages.

A better understanding of the molecular mechanisms of retinoblastoma progression can give insights to a better use of the current treatment and the discovery of novel therapeutic targets. Retinoblastoma is shown to be predominantly driven by bi-allelic *RB1* inactivation. *MYCN* amplification may be another driving event as it is found in many patients with wide-type *RB1* gene. Genomic and transcriptomic studies have been performed to identify additional events that are needed to promote the retinoblastoma progression, but perhaps due to rarity of this cancer, the number of the studies as well as the characterization of the genomic or transcriptomic events are still lagging, compared to other pediatric and adult cancers.

Characterization of the molecular basis of cancer can be achieved through comparison to its normal counterpart or through comparison within its subtypes. The cell of origin of retinoblastoma has been debated, either being a cone precursor or a multipotent retinal precursor. For the existence of subtypes, three transcriptomic analyses have reached different conclusions: no subtype, two subtypes, or no subtype but with different cone differentiation states.

Hypothesis

Since 2006, Molecular Oncology team at Institut Cruie has started to work on retinoblastoma. Several datasets have been established:

- A list of mutations derived from 84 primary tumor samples studied by whole-exome sequencing.
- A somatic copy number alteration (SCNA) dataset derived from 102 samples studied by BAC array CGH, Illumina SNP 370K array or Cytoscan™ copy number array.
- A methylation array dataset derived from 66 samples studied by Illumina 450K methylation array.

- A transcriptomic dataset derived from 59 samples studied by Affymetrix Human Genome U133 plus 2.0 array.
- A matching clinical dataset.
- The list of those datasets is available from: https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM4_ESM.xlsx.

Through analyzing these data, the Molecular Oncology team has identified two molecular subtypes of retinoblastoma by analyzing gene expression, DNA methylation and somatic copy number alterations (SCNA) in 72 patient samples. They have shown that genes upregulated in subtype1 were enriched in cone photoreceptor differentiation genes (e.g. ARR3), while strongly expressed genes in subtype2 (e.g. EBF3) were involved in ganglion differentiation and axonal outgrowth. In terms of clinical features, patients with subtype1 tumors are significantly younger at diagnosis. Histologically, subtype1 tumors tended to grow into subretinal areas, whereas subtype2 tumors were more likely to grow towards the vitreous.

Based on these evidences, we hypothesized the existence of two molecular subtypes of retinoblastoma, one cone-like subtype, the other subtype with mixed cone and ganglion features. Different molecular mechanisms may be involved in these two subtypes.

Objectives

The objectives of my PhD projects are:

- 1) I sought to better depict the molecular landscapes of the two subtypes of retinoblastoma through analyzing the genomic, transcriptomic and methylomic data that have already been acquired.
- 2) In collaboration with the biologist in the team, we sought to collect new single-cell transcriptomic data to better characterize the inter-tumoral heterogeneity of the retinoblastoma, the intra-tumoral heterogeneity observed in subtype 2 retinoblastoma, and potentially the retinoblastoma tumor microenvironment.

CHAPTER 3.

A HIGH-RISK RETINOBLASTOMA SUBTYPE WITH STEMNESS FEATURES, DEDIFFERENTIATED CONE STATES AND NEURONAL/GANGLION CELL GENE EXPRESSION

In this chapter, we will address the questions: are there molecular subtypes of retinoblastoma? What are their molecular and clinical-pathological features and cells of origin?

Three classifications have been proposed based on transcriptomic studies in the past decades. The first study by McEvoy *et al.* showed that there is no clear subtype using principal component analysis and hierarchical clustering on 52 retinoblastoma samples. Both the tumors and the orthotopic retinoblastoma xenografts expressed markers of various retinal cell types, including photoreceptors, interneurons and progenitor cells. Hence, the authors suggested that retinoblastoma is relative homogeneous and retinal progenitor cells may be the cells of origin for this tumor.

Kapatai *et al.* identified two subtypes of retinoblastoma from an analysis of 21 retinoblastomas and 3 normal retina. They showed that one subtype co-expressed various retinal cell markers, are relatively less differentiated and carried more copy number alterations (1q loss, 6p gain and 16q loss). Another subtype predominantly expressed cone photoreceptor markers. The authors hypothesized that the two subtypes may originate from different retinal cells, one from retinal progenitor cells, another from retinal cone photoreceptor lineage.

Kooi *et al.* also found two clusters using transcriptomic data of a series of 76 retinoblastomas. They illustrated that the two subgroups were not dichotomous, but displayed gradients of photoreceptor and ribosome/mRNA synthesis signatures. The level of photoreceptor signature in tumors were inversely correlated with age of diagnosis and positively correlated with degree of differentiation. The authors proposed

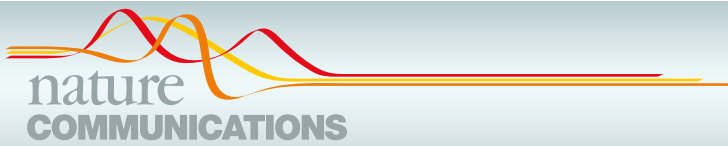
that the tumors with lower differentiation and lower degree of photoreceptor maturity may be dedifferentiated from another subtype with higher photoreceptor maturity.

In this chapter, based on a combined strategy of consensus clustering and centroid clustering of genomic, transcriptomic and methylomic data from 102 patient samples, we identified two molecular subtypes of retinoblastoma that are associated with different clinical-pathological features. We showed that all retinoblastoma expressed cone photoreceptor markers, with subtype 1 showing more matured cone signatures, and subtype 2, more aggressive, displayed less differentiated cone signatures and higher level of neuronal/ganglion and stemness features. This subtype was also associated with relative unstable genome, hypomethylated DNAs and hypermethylated CpG islands. We showed that subtype 2 tumors were more heterogeneous, in some cases both matured cone markers and neuronal/ganglion markers were expressed in one tumor. We analyzed one of such cases by single-cell RNA sequencing and revealed different genomic alterations were involved in different tumor cells with different phenotype, suggesting that different progression pathways and different cells of origin may be involved in retinoblastoma with varied levels of cone differentiation.

This article is published in *Nature Communications* (2021), available through the link: <https://doi.org/10.1038/s41467-021-25792-0>

I am the co-first author of the article with the two previous PhD students of our team, Daniela and Meriem. I participated in conceptualization, bioinformatics and statistical analyses, data visualization, writing and revision of the article.

A section of comments on the article is at the end of this chapter at page 109.



ARTICLE


<https://doi.org/10.1038/s41467-021-25792-0>

OPEN

A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression

Jing Liu^{1,2,3,34} , Daniela Ottaviani^{1,2,4,34} , Meriem Sefta^{1,2,34}, Céline Desbrousses^{1,2} , Elodie Chapeaublanc^{1,2} , Rosario Aschero⁵ , Nanor Sirab^{1,2} , Fabiana Lubieniecki⁵, Gabriela Lamas⁵, Laurie Tonon⁶, Catherine Dehainault^{7,8}, Clément Hua^{1,2} , Paul Fréneaux⁷, Sacha Reichman⁹ , Narjesse Karboul^{1,2}, Anne Biton^{1,2,10,11,30} , Liliana Mirabal-Ortega^{12,13,14} , Magalie Larcher^{12,13,14}, Céline Brulard^{1,2,31}, Sandrine Arrufat⁷, André Nicolas⁷, Nabila Elarouci³, Tatiana Popova¹⁵, Fariba Némati¹⁶ , Didier Decaudin¹⁶, David Gentien¹⁶, Sylvain Baulande¹⁷ , Odette Mariani⁷, Florent Dufour^{1,2} , Sylvain Guibert¹⁸, Céline Vallot¹⁸ , Livia Lumbroso-Le Rouic¹⁹, Alexandre Matet^{19,20} , Laurence Desjardins¹⁹, Guillem Pascual-Pasto^{21,22}, Mariona Suñol^{21,23}, Jaume Catala-Mora^{21,24} , Genoveva Correa Llano^{21,22}, Jérôme Couturier⁷, Emmanuel Barillot^{10,11} , Paula Schaiquevich^{5,25} , Marion Gauthier-Villars^{7,8,15}, Dominique Stoppa-Lyonnet^{7,8,20} , Lisa Golmard^{7,8,15}, Claude Houdayer^{7,8,15,32}, Hervé Brisse²⁶ , Isabelle Bernard-Pierrot^{1,2} , Eric Letouzé^{27,28} , Alain Viari⁶, Simon Saule^{12,13,14}, Xavier Sastre-Garau^{7,33}, François Doz^{20,29}, Angel M. Carcaboso^{21,22} , Nathalie Cassoux^{19,20}, Celio Pouponnot^{12,13,14}, Olivier Goureau⁹, Guillermo Chantada^{4,21,22,25,35} , Aurélien de Reyniès^{3,35}, Isabelle Aerts^{1,2,29,35} & François Radvanyi^{1,2,35}

Retinoblastoma is the most frequent intraocular malignancy in children, originating from a maturing cone precursor in the developing retina. Little is known on the molecular basis underlying the biological and clinical behavior of this cancer. Here, using multi-omics data, we demonstrate the existence of two retinoblastoma subtypes. Subtype 1, of earlier onset, includes most of the heritable forms. It harbors few genetic alterations other than the initiating *RB1* inactivation and corresponds to differentiated tumors expressing mature cone markers. By contrast, subtype 2 tumors harbor frequent recurrent genetic alterations including *MYCN*-amplification. They express markers of less differentiated cone together with neuronal/ganglion cell markers with marked inter- and intra-tumor heterogeneity. The cone dedifferentiation in subtype 2 is associated with stemness features including low immune and interferon response, E2F and MYC/MYCN activation and a higher propensity for metastasis. The recognition of these two subtypes, one maintaining a cone-differentiated state, and the other, more aggressive, associated with cone dedifferentiation and expression of neuronal markers, opens up important biological and clinical perspectives for retinoblastomas.

A list of author affiliations appears at the end of the paper.

Retinoblastoma is a rare childhood cancer of the developing retina with an incidence rate of about 1 in 17,000 live births^{1–3}, but is the most frequent pediatric intraocular malignancy. The main therapeutic objective for retinoblastoma is first to save the child's life through early detection, treatment of the ocular tumor, and prevention of metastatic spread. Secondary goals are eye preservation and maximization of visual potential⁴. In low-income countries, retinoblastoma is associated with low patient survival due to delayed diagnosis, poor access to multi-disciplinary retinoblastoma-specific healthcare, and socio-economic factors. In high-income countries, tumor remission is achieved in more than 95% of cases, however some patients still develop metastases⁵. Metastases can be due to dissemination through the optic nerve into the central nervous system (CNS) and through the sclera to the orbit. Retinoblastoma can also give rise to systemic metastases⁶. Several histopathological features are considered high-risk factors for tumor progression and metastasis⁷.

Retinoblastoma is usually initiated by biallelic inactivation of the *RB1* tumor suppressor gene. A minority of non-hereditary retinoblastomas (<2%) are initiated by *MYCN*-amplification without *RB1* inactivation⁸. In most cases, hereditary retinoblastomas are bilateral, whereas non-hereditary cases are always unilateral.

The retina includes six types of neurons (rod and cone photoreceptors, bipolar, amacrine, horizontal, and ganglion cells) and Müller glia, all of which are generated from multipotent retinal progenitor cells^{9,10}. Studies in human show that the cell-of-origin of retinoblastoma is a cone precursor^{11–15}.

Three studies based on gene expression profiling reached conflicting conclusions concerning the possible existence of retinoblastoma molecular subtypes and the retinal cell type-specific markers expressed in retinoblastoma^{16–18}. Beyond *RB1*, the only recurrently mutated gene in retinoblastoma (7–13% of cases) is the epigenetic modifier gene *BCOR*^{19–21}. Recurrent genomic alterations have been identified: gains and amplifications on 1q, 2p (targeting *MYCN*), and 6p, losses on 13q (targeting *RB1*) and 16q^{22–25}. Several studies have reported a positive correlation between high copy-number alterations, age at diagnosis, and other clinical and histopathological variables, including unilaterality, non-hereditary status, and low differentiation^{24,26–30}. Despite this wealth of findings, a molecular framework for understanding the biology and clinical behavior of retinoblastoma is lacking.

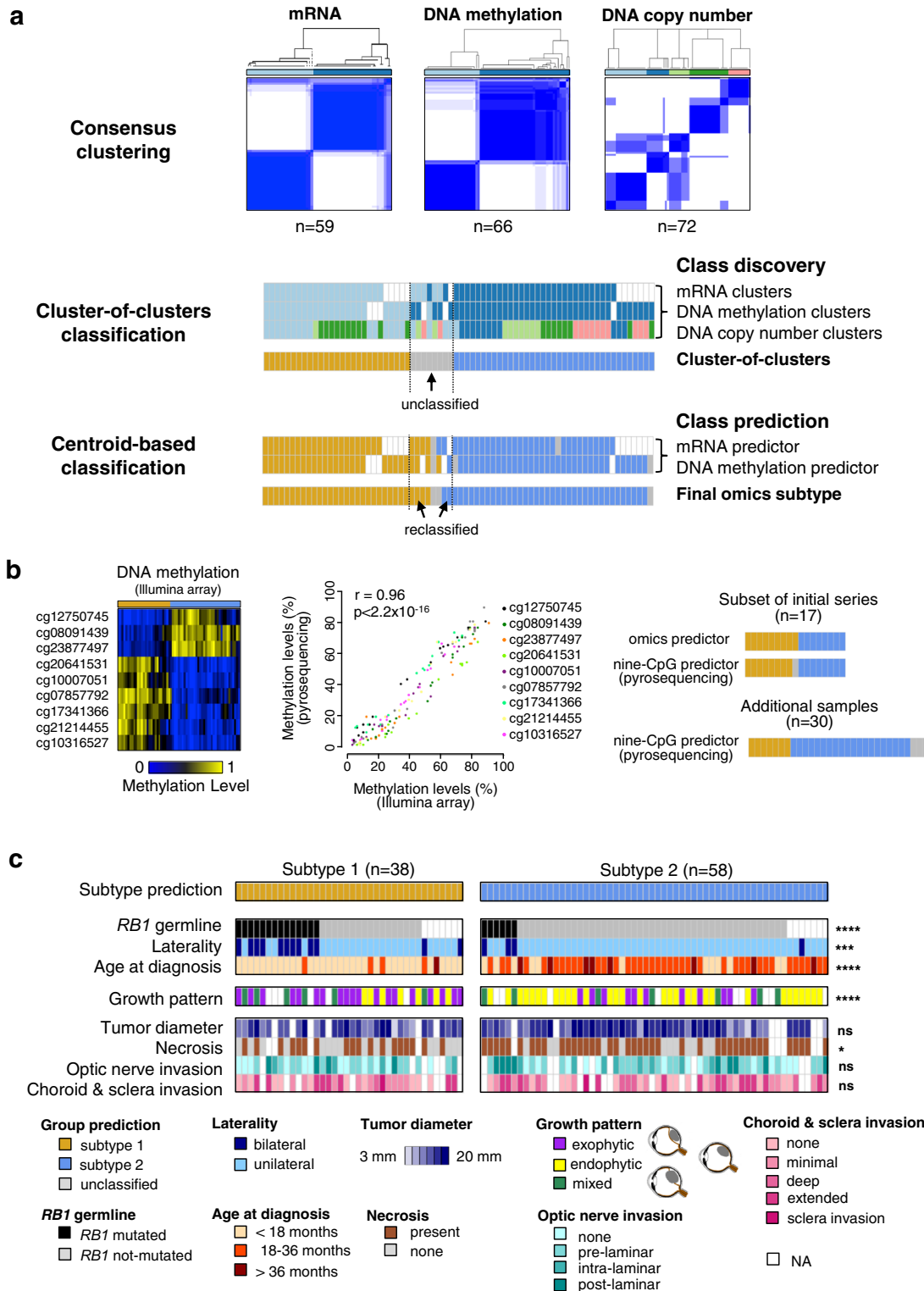
In this work, we identify two subtypes of retinoblastoma associated with different clinical and pathological features (age at diagnosis, laterality, heredity, and growth pattern) following integrative analysis of the transcriptome, methylome, and DNA copy-number alteration data from a series of 102 retinoblastomas. Further characterization provides evidence for the relevance of these two subtypes for understanding the biology of retinoblastoma, and for clinical management of this disease. Few genetic alterations other than *RB1* inactivation are associated with subtype 1 tumors. By contrast, in addition to *RB1* inactivation, almost all subtype 2 tumors harbor other recurrent genetic alterations, including *MYCN* amplifications. Consistent with a maturing cone precursor as the cell-of-origin of retinoblastoma, we find that both subtypes express cone markers. We show, by a detailed analysis of cone differentiation including the use of immunohistochemistry, retinal organoids, and single cells, that subtype 2 tumors are less differentiated than subtype 1 tumors and express neuronal/ganglion cell markers with marked inter- and intratumor heterogeneity. This lower cone differentiation in subtype 2 is associated with stemness features, including a higher propensity for metastasis, as shown by a study of an additional series of 112 retinoblastomas, including metastatic tumors.

Results

Identification of two retinoblastoma molecular subtypes with distinct clinical and pathological features. We analyzed a series of 102 enucleated retinoblastomas (Supplementary Data 1). To investigate the existence of different retinoblastoma molecular subtypes, we combined three genomic approaches, mRNA expression, DNA methylation, and somatic copy-number alterations (SCNAs) in a subset of 72 of the 102 retinoblastomas. All three datasets were available for 53 of the 72 tumors, and at least two of the three datasets were available for all 72 tumors (Supplementary Data 1). Within each of these three omics datasets, we calculated several partitions of the samples in *k* clusters (*k*-partitions), for various values of *k*, through unsupervised hierarchical clustering, using varying numbers of features and different linkages (see “Methods” section). Then, for each omics and each value of *k*, we performed a consensus clustering analysis to derive a consensus *k*-partition. Doing so the transcriptome-based and methylome-based analyses both yielded stable consensus partitions in two clusters, while the SCNA-based analysis yielded a stable consensus partition in five clusters (Fig. 1a, upper panel and Supplementary Fig. 1a). Cluster memberships from each of the three partitions were analyzed by a cluster-of-clusters approach, briefly, a sample co-classification matrix was built and was then subjected to hierarchical clustering using complete linkage. It revealed the convergence of the three partitions around two molecular subtypes gathering 89% (64/72) of the cases (Fig. 1a, middle panel and Supplementary Fig. 1b). Nearest centroid classification attributed to the same subtypes 63 of the 64 classified samples. Moreover, six of the eight unclassified samples could be attributed to a subtype, yielding a final number of 69 classified samples (69/72, 96%): 31 belonging to subtype 1 and 38 to subtype 2 (Fig. 1a lower panel, and Supplementary Fig. 1c, Supplementary Data 1).

To assign to a subtype the 30 remaining tumors of our 102 tumor series, we then established a nine-CpG-based classifier, based on the genome-wide CpG methylation array profiling (see “Methods” section) (Fig. 1b, left panel and Supplementary Data 1). We verified that there was a high concordance in quantifying the level of CpG methylation between DNA methylation arrays and pyrosequencing assays (Fig. 1b, middle panel). This nine-CpG-based classifier attributed seven of the remaining 30 samples to subtype 1, and 20 to subtype 2, while three cases remained unclassified (Fig. 1b, right panel). Altogether the majority of the tumors (96/102, 94%) could be assigned to one of the two subtypes (38 to subtype 1, 58 to subtype 2).

We then compared the clinical and pathological features of these two subtypes (Fig. 1c, Table 1, Supplementary Data 1). Patients with subtype 1 tumors were significantly younger at diagnosis (median age = 11.0 vs 23.9 months; Wilcoxon rank-sum test, $p = 8.9 \times 10^{-11}$). This subtype included 75% of the bilateral ($p = 1.51 \times 10^{-3}$) and 70% of the hereditary cases ($p = 7.68 \times 10^{-4}$). Unexpectedly, among patients with subtype 1 tumors, age at diagnosis did not differ significantly between hereditary forms (median = 10.2 months) and non-hereditary forms (median = 11.2 months) (Wilcoxon rank-sum test, $p = 0.451$). Likewise, there was also no significant difference between the age at diagnosis for hereditary and non-hereditary forms of subtype 2 tumors (median = 19.8 and 24.7 months, respectively, Wilcoxon rank-sum test, $p = 0.320$). Retinoblastomas generally display exophytic growth (into the subretinal space), endophytic growth (towards the vitreous), or, less frequently, a mixed growth pattern (both endophytic and exophytic). Subtype 1 tumors were significantly more likely to be exophytic, whereas most of the subtype 2 tumors were endophytic ($p = 7.33 \times 10^{-4}$). Necrotic areas were more frequently observed in subtype 2 tumors than in subtype 1 tumors



($p = 0.020$). Tumor diameter and histological risk features (optic nerve invasion, choroid, or sclera invasion) did not differ significantly between the two subtypes.

Subtype 2 displayed more genetic alterations than subtype 1 and included the *MYCN*-amplified tumors. We investigated the

genomic characteristics of the two tumor subtypes, by determining their SCNA profiles (Supplementary Data 2). Gains of 1q, 2p (*MYCN*), 6p, 13q, and losses/LOH of 13q (*RB1*), 16q were the most frequent alterations, consistent with reported findings for retinoblastoma^{22–25}. 6p gains and 13q losses/LOH were equally distributed between tumor subtypes, whereas 1q gains, 2p gains, and 16q losses/LOH were significantly more frequent in the

Fig. 1 Multi-omics-based molecular subtypes of retinoblastoma and clinical characteristics. **a** Consensus clustering of retinoblastomas based on transcriptomic, DNA methylation, and copy-number alteration data (top panel). Unsupervised cluster-of-clusters analysis (middle panel). Supervised centroid-based classification (bottom panel). Final omics subtype: subtype 1, $n = 31$ (gold); subtype 2, $n = 38$ (blue); unclassified, $n = 3$ (gray). **b** Heatmap showing methylation values (methylome arrays) for the nine-CpG-based classifier (left panel). Correlation between methylation values assessed by pyrosequencing and by methylome array, for 17 tumors (middle panel). A two-sided Pearson's correlation test was used. The nine-CpG-based classifier applied to a subset of 17 tumors of the initial series, led to the same classification as obtained by the -omics approach in 16 cases (one case being not classified by the nine-CpG-based classifier). Subtype assignment of 30 additional tumors based on the nine-CpG-based classifier (right panel). **c** Final molecular classification of 96 retinoblastomas and their key clinical and pathological characteristics. $p \geq 0.05$ (ns), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****). For comparisons of *RB1* germline mutation, laterity, growth pattern, tumor diameter, and necrosis between two subtypes, χ^2 tests were used. For comparisons of age at diagnosis and tumor diameter between two subtypes, two-sided Kruskal-Wallis rank tests were used. For comparisons of optic nerve invasion and choroid and sclera invasion between two subtypes, two-sided Fisher's exact tests were used. Exact p -values are provided in Table 1.

subtype 2 samples ($p = 5.5 \times 10^{-11}$, $p = 0.0037$, and $p = 1.8 \times 10^{-7}$, respectively) (Fig. 2a). *MYCN* amplifications varied from 14 to 246 copies (Supplementary Data 2) and were found only in subtype 2 tumors (10/58) ($p = 0.013$).

The overall genomic instability score, estimated as the proportion of genome with copy-number alterations, was significantly higher ($p = 3.3 \times 10^{-7}$) for subtype 2 than for subtype 1 tumors (Fig. 2b), and was also significantly higher when tumors with *MYCN* amplification were excluded from the analysis. By contrast, genomic instability scores did not differ between subtype 2 tumors with *MYCN* amplifications and subtype 1 tumors.

We then characterized the mutational landscape of the retinoblastoma subtypes. We performed whole-exome capture followed by paired-end massively parallel sequencing (WES) on genomic tumoral and matched normal DNA of 71 patients from the 102-retinoblastoma series (subtype 1, $n = 25$; subtype 2, $n = 41$; unclassified, $n = 5$). We identified 242 somatic mutations in 186 genes (Supplementary Data 2). The tumors harbored a median of two mutations. The number of somatic mutations identified by WES was significantly higher ($p = 1.2 \times 10^{-7}$) for subtype 2 than for subtype 1 tumors (Fig. 2c). Restricting subtype 2 tumors to either *MYCN*-amplified or *MYCN*-non-amplified tumors yielded the same result.

Three genes, *RB1*, *BCOR*, and *ARID1A*, were found to be recurrently mutated. We performed targeted sequencing for these three genes in 23 of the 31 samples lacking WES data. The distributions of *RB1*, *BCOR*, *ARID1A* mutations, *MYCN* amplifications, 1q gains, and 16q losses are shown by subtype in Fig. 2d. For *RB1* the germinal and somatic point mutations identified are shown, together with deletions, copy-neutral LOH, and promoter methylation. *RB1* mutations were found in most tumors, regardless of subtype, and no difference in the mutation type was observed between the two tumor subtypes. Of note, we found a tumor without *RB1* alteration, it belonged to subtype 2 and displayed a high level of *MYCN* amplification (141 copies). *BCOR* mutations ($n = 9$) were found exclusively in subtype 2 ($p = 0.02$), as were the two *ARID1A* mutations. Most of the subtype 2 tumors without *MYCN* amplification (46/48, 96%) presented gains of 1q and/or losses of 16q. By contrast, none of the *MYCN*-amplified tumors except one had a 1q gain or 16q loss ($p = 0.005$) (Fig. 2d).

Subtype 2 tumors harbored hypermethylation within CpG islands and hypomethylation outside CpG islands. We compared the methylome of subtype 1 tumors ($n = 27$) and subtype 2 tumors ($n = 36$, including 4 *MYCN*-amplified tumors). A heatmap representing the methylation levels of the 6607 CpGs significantly differentially methylated between the two subtypes (Supplementary Data 2) is shown in Fig. 2e. Subtype 2 tumors showed more frequent hypermethylation within CpG islands, and

a more frequent hypomethylation outside CpG islands, than subtype 1 tumors (Fig. 2f, g and Supplementary Fig. 2). The four *MYCN*-amplified subtype 2 tumors studied presented a hypomethylation outside CpG islands and did not present hypermethylation within CpG islands (Fig. 2g).

The two subtypes exhibited differences in the expression of cone and ganglion/neuronal markers and in stemness. We compared the transcriptome of the two subtypes. Almost one-third of the genes were found differentially expressed between the two subtypes (6207/20408, adjusted p -value < 0.05) (Supplementary Data 3).

Cone markers (such as *GUCA1C*, *GNAT2*, *ARR3*, *GUCA1A*, *GUCA1B*, *GNGT2*, *PDE6C*, *PDE6H*, *OPN1SW*) and neuronal/ganglion markers (such as *EBF3*, *DCX*, *ROBO1*, *SOX11*, *GAP43*, *PCDHB10*, *STMN2*, *NEFM*, *POU4F2*, *EBF1*) were among the most differentially expressed genes. Cone markers were overexpressed in subtype 1 tumors, whereas neuronal/ganglion markers were overexpressed in subtype 2 tumors (Fig. 3a). Among the genes known to be involved in retinoblastoma^{1,31}, several were found to be differentially expressed between the two subtypes (*KIF14*, *MDM4*, *MIR17HG*, *MYCN*, *SKP2* upregulated in subtype 2; *RBL2* downregulated in subtype 2) (Supplementary Data 3). Some of these genes were located in gained/amplified (*KIF14* and *MDM4* at 1q32.1 and *MYCN* at 2p24.3), or lost (*RBL2* at 16q12.2) chromosomal regions, whereas others were involved in the *MYC/MYCN* pathway (*MIR17HG*, *SKP2*). Hierarchical clustering of the 6207 genes identified three main gene clusters: two upregulated in subtype 1 (gene cluster 1.1 consisting of 1201 genes and gene cluster 1.2 consisting of 1788 genes) and one containing all the genes upregulated in subtype 2 (3112 genes; gene cluster 2) (Fig. 3b). We performed enrichment analysis using the gene sets from gene ontology biological processes (GOBP) and MSigDB hallmarks (HALLMARK) (Fig. 3c and Supplementary Data 3). Cluster 1.1 genes mainly upregulated in a subset of subtype 1 tumors, were associated with tumor microenvironment (immune response, inflammation, interferon response, complement, glial cells) and rod cells markers. Cluster 1.2 was enriched in genes related to fatty acid metabolism, oxidative phosphorylation, and photoreceptor/cone cells. Cluster 2 was enriched in genes associated with the cell cycle, E2F target genes, RNA processing, *MYC* pathway, and neuron morphogenesis.

The lack of an inflammation/immune signature and the enrichment in *MYC* and E2F target genes in subtype 2 was evocative of stemness features^{32,33}. Moreover, *CD24*, one of the two most overexpressed genes in subtype 2 tumors (Fig. 3a and Supplementary Data 3), has been shown to be a neuronal stem cell marker and a cancer stem cell marker for several tumor types³⁴. Stemness indices, based on transcriptomic data, allowed relative evaluation of the degree of stemness in tumor samples. We applied four different stemness signatures^{32,33,35,36} to the 59

Table 1 Clinical and histopathological characteristics of patients stratified by molecular subtype.

	Subtype 1 <i>n</i> (%)	Subtype 2 <i>n</i> (%)	<i>N</i>	<i>p</i> -value ^a
Patients	38 (40)	58 (60)	96	
Clinical Center				
Institut Curie	31 (42)	43 (58)	74	0.655 ^b
Hospital Garrahan	6 (33)	12 (66)	18	
Hospital Sant Joan de Déu	1 (25)	3 (75)	4	
Sex				
Female	17 (35)	31 (65)	48	0.403 ^c
Male	21 (44)	27 (56)	48	
<i>RB1</i> germline mutation				
Yes	14 (70)	6 (30)	20	7.681 × 10 ⁻⁴ ^c
No	17 (28)	44 (72)	61	
NA	7 (47)	8 (53)	15	
Laterality				
Bilateral	12 (75)	4 (25)	16	1.506 × 10 ⁻³ ^c
Unilateral	26 (33)	54 (66)	80	
Age at diagnosis				
<18 months	33 (73)	12 (27)	45	2.132 × 10 ⁻⁹ ^d
18–36 months	4 (10)	38 (90)	42	
>36 months	1 (11)	8 (89)	9	
Growth pattern				
Endophytic	7 (18)	31 (82)	38	7.332 × 10 ⁻⁴ ^c
Exophytic	19 (63)	11 (37)	30	
Mixed	6 (46)	7 (54)	13	
NA	6 (40)	9 (60)	15	
Tumor diameter (mm)				
(3.98–6.67]	1 (50)	1 (50)	2	0.2094 ^d
(6.67–9.33]	1 (25)	3 (75)	4	
(9.33–12]	7 (50)	7 (50)	14	
(12–14.7]	9 (64)	5 (36)	14	
(14.7–17.3]	9 (27)	24 (73)	33	
(17.3–20]	5 (31)	11 (69)	16	
NA	6 (46)	7 (54)	13	
Necrosis				
Yes	18 (31)	40 (69)	58	0.0203 ^c
None	16 (57)	12 (43)	28	
NA	4 (40)	6 (60)	10	
Optic nerve invasion				
None	12 (48)	13 (52)	25	0.7467 ^b
Prelaminar	12 (39)	19 (61)	31	
Intralaminar	4 (33)	8 (66)	12	
Post-laminar	4 (31)	9 (69)	13	
NA	6 (40)	9 (60)	15	
Choroid and sclera invasion				
None	10 (40)	15 (60)	25	0.6468 ^b
Minimal	10 (48)	11 (52)	21	
Deep	1 (14)	6 (86)	7	
Extended	8 (38)	13 (62)	21	
Sclera invasion	1 (50)	1 (50)	2	
NA	8 (40)	12 (60)	20	

NA not available, *n* number in each subtype, *N* total number.

^aSignificant *p*-value < 0.05.

^bTwo-sided Fisher's exact test.

^cChi² test.

^dTwo-sided Kruskal–Wallis rank test.

retinoblastoma samples for which transcriptomic data were available. The stemness indices assessed by these signatures were significantly higher in subtype 2 than in subtype 1 (Fig. 3d, upper panel and Supplementary Fig. 3a). In addition, the stemness indices obtained with the different signatures were highly correlated (Supplementary Fig. 3b). We searched for hallmark gene sets associated with stemness (Supplementary Data 3). The hallmarks positively correlated with stemness included E2F

targets, MYC targets V2, MYC targets V1 and G2/M checkpoint (Fig. 3d, lower panel and Supplementary Data 3). These hallmarks were the same as those identified in cluster 2 (cluster of genes overexpressed in subtype 2). The hallmarks negatively correlated with stemness included interferon-alpha response, interferon-gamma response, and complement (Fig. 3d, lower panel and Supplementary Data 3), and were the same as those identified in cluster 1.1 (cluster of genes overexpressed in subtype 1 and associated with the tumor microenvironment). We also assessed the relationship between stemness and the abundance of the various immune cells, as estimated with the Microenvironment Cell Population (MCP)—counter score³⁷. Stemness indices were negatively correlated with the MCP scores of monocytic lineage, B lineage, and cytotoxic lymphocytes (Fig. 3d, lower panel and Supplementary Data 3). Altogether, we showed that subtype 2 was associated with high stemness.

The upregulation of cone-related genes in subtype 1 and of neuronal/ganglion cell-related genes in subtype 2 (Fig. 3a) led us to analyze in detail the expression of genes associated with the different retinal cell types (rod and cone photoreceptors, ganglion, amacrine, bipolar, and horizontal cells, and Müller glia). The list of retinal cell type markers was selected from a systematic literature search and from single-cell RNA-seq (scRNA-seq) data obtained at different time points during human retinal development³⁸. From the annotated cell types defined by Lu et al.³⁸, we identified lists of candidate markers associated with each retinal cell type (Supplementary Data 3). In order to choose the most specific markers, we developed a tool for visualizing gene expression profiles in the different retinal cell types (see “Methods” section) (<https://retinoblastoma-retina-markers.curie.fr/>). Based on an analysis of the expression profiles of the candidate markers obtained from Lu et al.'s data and of markers found in the literature we proposed markers for the different retinal cell types (given in Supplementary Data 3).

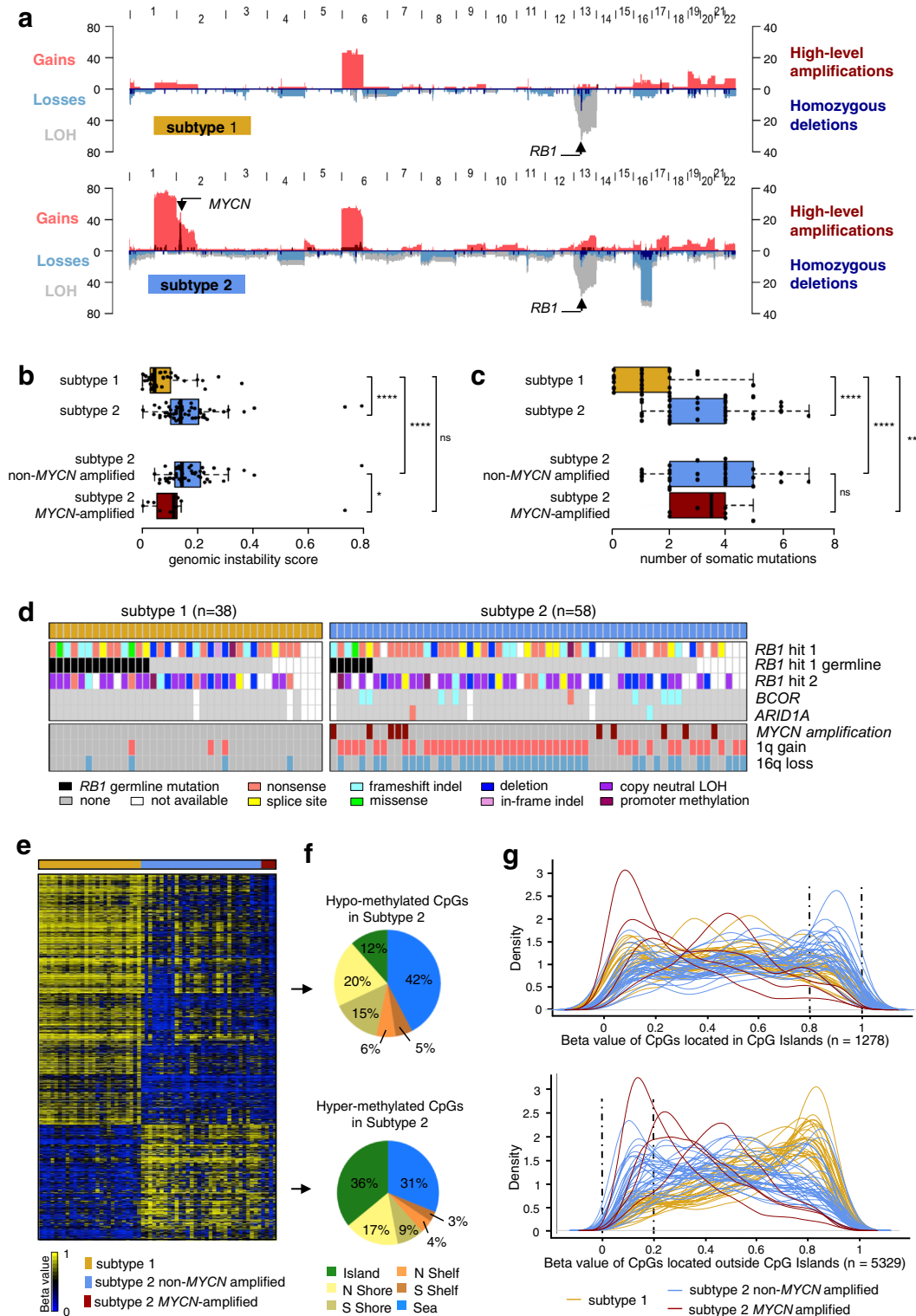
Cone markers were overall expressed in both subtype 1 and 2 retinoblastomas, with different expression levels between subtypes depending on the markers (Fig. 3e, upper panel). Among the 24 ganglion cell markers analyzed, a small subset (*EBF3*, *EBF1*, *GAP43*, *POU4F2*, *NEFM*, *ALCAM*, *NRN1*, *CNTN2*) were consistently overexpressed in subtype 2 tumors (Supplementary Fig. 4a and Fig. 3e, lower panel).

Using the lists of candidate markers associated with each retinal cell type obtained from Lu et al.'s data³⁸, we provided further evidence for an enrichment of markers associated with ganglion cells in subtype 2 tumors (Supplementary Data 3). These genes overexpressed in subtype 2 tumors can be considered both as ganglion and neuronal genes. Indeed, although specific to ganglion cells in the context of the retina (Supplementary Fig. 4b), all displayed expression in the brain and played different functions in the central nervous system^{39–47}.

Most of the markers of other retinal cell types (rods, amacrine, bipolar, horizontal, and Müller glia cells) were not expressed in retinoblastomas or were only expressed in a subset of tumors (Supplementary Fig. 4a). The expression of these markers was likely due to the presence of normal retinal cells in some retinoblastomas. Indeed non-neoplastic rods and Müller glial cells have been shown to be present in some retinoblastomas¹³.

State of cone differentiation and expression of neuronal/ganglion cell markers distinguished the two subtypes.

The expression of cone markers observed in both subtypes of retinoblastoma is consistent with the retinoblastoma cell-of-origin being a committed cone cell. Differences in cone marker expression were observed between the two subtypes, raising the question of whether these differences could correspond to



different stages of cone differentiation. Retinal organoids are three-dimensional structures derived from human induced pluripotent stem (iPS) cells that recapitulate the spatial and temporal differentiation of the retina providing powerful in vitro models of human retinal development^{48,49}. We measured the level of expression of early and late cone markers in retinal organoids at various time points (d35, d49, d56, d84, d112, d175) after the

differentiation of human iPS cells into the retina, and in subtype 1 ($n = 23$) and subtype 2 ($n = 44$) retinoblastomas, using the NanoString technology (Fig. 4a and Supplementary Data 4). As expected, in iPS cell-derived retinal organoids, the expression of early photoreceptor/cone markers (*OTX2*, *CRX*, *THRB*, *RXRG*) appeared at earlier time points than late cone markers (*PDE6H*, *GNAT2*, *ARR3*, *GUCA1C*). *GUCA1C* was the last marker to be

Fig. 2 Genomic characterization, somatic mutational landscape, and DNA methylation profiles of the two retinoblastoma subtypes. **a** Pattern of somatic copy-number alterations in subtype 1 (top, $n = 38$) and subtype 2 (bottom, $n = 58$) retinoblastomas. **b** Boxplots comparing genomic instability between subtype 1 tumors ($n = 38$) and subtype 2 tumors ($n = 58$). Among the subtype 2 tumors, non-MYC N -amplified ($n = 48$) and MYC N -amplified ($n = 10$) tumors are also shown. Significant differences were tested by two-sided Wilcoxon tests for Subtype 1 vs Subtype 2: $p = 3.3 \times 10^{-7}$; Subtype 1 vs Subtype 2 non-MYC N : $p = 1.2 \times 10^{-7}$; Subtype 1 vs Subtype 2 MYC N -amplified: $p = 0.147$; and Subtype 2 non-MYC N -amplified vs Subtype 2 MYC N -amplified: $p = 0.014$. **c** Boxplots comparing the number of somatic mutations between subtype 1 tumors ($n = 25$) and subtype 2 tumors ($n = 41$). Among the subtype 2 tumors, non-MYC N -amplified ($n = 33$) and MYC N -amplified ($n = 8$) tumors are also shown. Significance differences were tested by two-sided Wilcoxon tests for Subtype 1 vs Subtype 2: $p = 8.1 \times 10^{-7}$; Subtype 1 vs Subtype 2 non-MYC N -amplified: $p = 3.5 \times 10^{-6}$; Subtype 1 vs Subtype 2 MYC N -amplified: $p = 0.001$; and Subtype 2 non-MYC N -amplified vs Subtype 2 MYC N -amplified: $p = 0.775$. **b, c** In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. The whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Note: $p \geq 0.05$ (ns), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****). **d** Somatic mutations of the three genes recurrently altered by tumor subtype. For *RB1* are indicated the germline mutations. MYC N amplifications, 1q gains, and 16q losses are also shown. **e** Heatmap of the 6607 differentially methylated CpGs (difference of methylation level > 0.2 , adjusted $p < 0.05$, two-sided Wilcoxon test and BH correction) between subtype 1 and subtype 2. **f** Distribution, in subtype 2 as compared to subtype 1, of hypomethylated CpGs (upper panel) and hypermethylated CpGs (lower panel), by CpG content and neighborhood context. **g** Density plots showing the distribution of methylation levels of the differentially methylated CpGs located in CpG islands (upper panel) and outside CpG islands (lower panel).

expressed, consistent with previous in vitro and in vivo observations^{50,51}. Early cone markers were expressed in both tumor subtypes, at very similar levels. By contrast, late cone markers were expressed, on average, at lower levels in subtype 2 tumors, the most downregulated marker *GUCA1C* being the latest cone marker expressed. These results indicated that subtype 1 tumors corresponded to a more differentiated stage of cone development than subtype 2 tumors.

As several neuronal/ganglion cell lineage-related genes were shown to be differentially expressed between tumor subtypes (Fig. 3), we also compared their levels of expression in retinal organoids and in tumor samples of the two subtypes (Fig. 4a and Supplementary Data 4). Ganglion-cell markers were expressed at early time points of retinal differentiation (from d49), and their expression levels decreased after d84, consistent with the loss of ganglion cells in retinal organoids at late time points⁵². These ganglion markers were upregulated in subtype 2 compared to subtype 1 tumors (Fig. 4a). Two of them, *EBF3* and *GAP43*, were expressed in subtype 2 tumors with levels comparable to those observed in retinal organoids between d49 and d84.

To more precisely determine the cone development stage corresponding to subtype 1 and subtype 2 tumor cells, we calculated, for each time point after the induction of retinal differentiation, the correlation coefficient between the centroid of each tumor subtype and those of the organoids using cone marker expression (Fig. 4b). Subtype 1 tumors were closest to later cone differentiation (highest correlation observed at d173), whereas subtype 2 tumors were closest to earlier cone differentiation (highest correlation observed between d84 and d112).

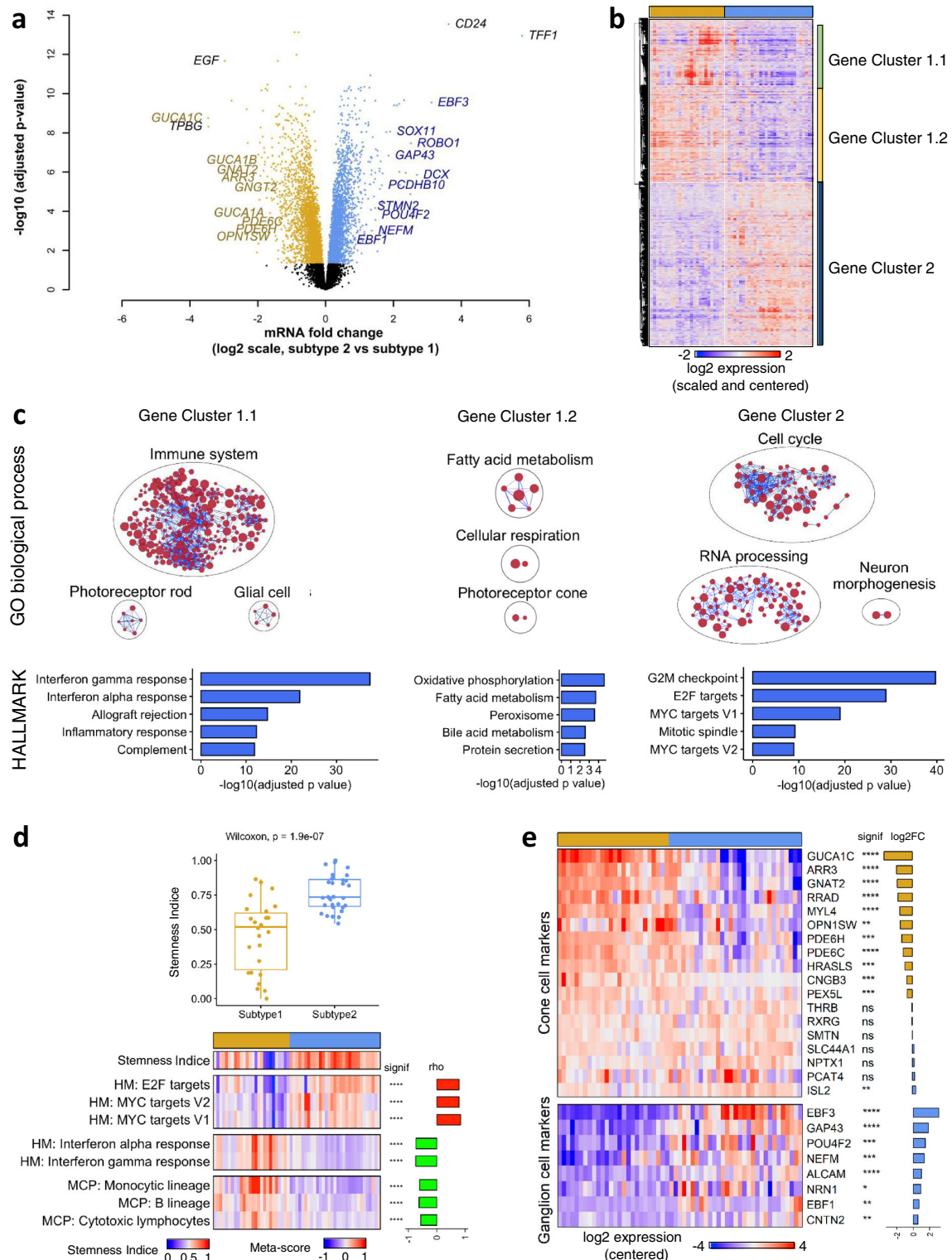
To illustrate the degree of cone differentiation achieved by individual retinoblastoma cases of each subtype, we generated a phylogenetic tree using photoreceptor/cone marker expression, incorporating retinal organoid samples at various time points after the induction of differentiation, and retinoblastoma samples (Fig. 4c). All subtype 1 tumors were close to iPS cell-derived retinal organoids at a late time point of differentiation (d173). Subtype 2 tumors were spread out from d84 to d173 retinal organoids.

To explore further the heterogeneity in terms of cone differentiation in retinoblastoma, we studied by immunohistochemistry the distribution of an early photoreceptor marker (CRX), and a later marker specific to the cone lineage (ARR3). We also assessed the expression of one ganglion cell marker (EBF3). Immunohistochemical staining was performed on paraffin-embedded samples of subtype 1 ($n = 9$) and subtype 2

($n = 25$) retinoblastomas (Supplementary Data 4). Two examples of each tumor subtype are presented in Fig. 4d. As expected, in the peritumoral normal retina, the transcription factor CRX was expressed in all cells of the outer nuclear layer (ONL), whereas ARR3 was expressed in a subset of cells in the ONL. EBF3 was expressed in ganglion cells, but also in some amacrine cells in the inner nuclear layer, as previously reported^{51,53–55}. All tumors, regardless of the subtype, expressed the photoreceptor marker CRX in agreement with retinoblastoma being derived from cone-committed cells. The ARR3⁺/EBF3[−] pattern was the only pattern observed in subtype 1 tumors (Fig. 4d, e and Supplementary Data 4). These tumors were positive for the proliferation marker Ki-67 (Fig. 4d, and Supplementary Data 4). Two types of expression patterns were observed for ARR3 and EBF3 in subtype 2 tumors (Fig. 4d). Most subtype 2 tumors (16/25, 64%) coexpressed ARR3 and EBF3 (ARR3⁺/EBF3⁺), as illustrated by tumor RB659 in Fig. 4d. Other subtype 2 tumors (8/25, 32%) displayed mutually exclusive expression of ARR3 and EBF3 (ARR3[−]/EBF3⁺ or ARR3⁺/EBF3[−] areas), as illustrated by tumor RB617 in Fig. 4d. One tumor (1/25) expressed EBF3 but not ARR3. Tumors of subtype 2 coexpressing ARR3 and EBF3 (ARR3⁺/EBF3⁺) were always positive for Ki-67. In subtype 2 tumors with a mutually exclusive expression of ARR3 and EBF3, the ARR3[−]/EBF3⁺ areas were always positive for Ki-67, whereas the ARR3⁺/EBF3[−] areas were mostly negative for Ki-67 (6 of 7 cases tested) (Fig. 4d and Supplementary Data 4). Histological examination of these Ki-67-negative ARR3⁺/EBF3[−] areas showed the presence of fleurettes (foci of photoreceptor differentiation) and an absence of mitoses in three of these six cases. The presence of these different areas within the tumor could reflect a range of tumor cell type stages, from stem, to progenitor to differentiating to terminally differentiated, with many of the latter being post-mitotic. Alternatively, the Ki-67-negative ARR3⁺/EBF3[−] areas could correspond to retinoma, a benign non-proliferative lesion observed adjacent to retinoblastoma^{56–58}.

Single-cell analysis of intratumoral heterogeneity in a subtype 2 tumor. To further explore the intratumoral heterogeneity of subtype 2 tumors, we performed droplet-based single-cell RNA sequencing on a subtype 2 tumor (RBSC11). Immunohistochemical analysis of this tumor showed a mutually exclusive expression of ARR3 and EBF3, defining two types of areas (CRX⁺/ARR3⁺/EBF3[−] and CRX⁺/ARR3[−]/EBF3⁺) (Supplementary Fig. 5a), as observed in about 30% of subtype 2 tumors.

We retained transcriptomes of 1198 cells after initial quality controls (Supplementary Fig. 5b). To identify the different cell



populations present in the tumor, we performed shared nearest neighbor (SNN) clustering and identified seven clusters (Fig. 5a).

To characterize the different clusters, we used (1) known cell type-specific markers, (2) cluster markers (the most upregulated genes in the cluster compared to all other clusters), (3) pathway analysis of cluster markers, (4) correlation to bulk mRNA expression profiles of purified cell types (Fig. 5b, c and

Supplementary Fig. 5c, d, Supplementary Data 5). Clusters 0–4, accounting for 89.2% of all cells analyzed, expressed early photoreceptor/cone markers (e.g., *OTX2*, *CRX*, *THRB*, and *RXRG*). Clusters 0 and 2 expressed neuronal/ganglion cell markers (e.g., *GAP43*, *SOX11*, *UCHL1*, *DCX*, *EBF3*), whereas clusters 1 and 4 expressed late cone markers (e.g., *ARR3* and *GUCA1C*). Clusters 2 and 4 expressed proliferation markers, such

Fig. 3 Transcriptomic differences between the two retinoblastoma subtypes. **a** Volcano plot with genes significantly upregulated in subtype 1 ($n = 26$) (gold) and subtype 2 ($n = 31$) (blue). The genes related to cone-cell and neuronal/ganglion-cell differentiation are indicated (in gold and blue, respectively), together with the most highly differentially expressed genes in each subtype. **b** Hierarchical clustering of the significantly differentially expressed genes identified three main gene clusters. **c** Upper panels: Gene sets from the GOBP collection enriched in clusters 1.1, 1.2, 2 in hypergeometric tests. Results are presented as networks of enriched gene sets (nodes) connected based on their overlapping genes (edges). Node size is proportional to the total number of genes in the gene set concerned. The names of the various GOBP terms are given in Supplementary Data 3. Bottom panels: Top 5 Gene sets from the HALLMARK collection enriched in clusters 1.1, 1.2, 2. **d** Upper panel: Boxplots of stemness indices, determined as in Malta et al.³², in the two subtypes of retinoblastoma (subtype 1 tumors: $n = 26$, subtype 2 tumors: $n = 31$). In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. Whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Significance was tested by a two-sided Wilcoxon test, $p = 1.9 \times 10^{-7}$. Bottom panel: Heatmap of stemness indices and meta-score of the most correlated and anti-correlated HALLMARK (HM) pathways and MCP-score of the most anti-correlated immune cells. Spearman's rho and p -value are shown in the figure. $p < 0.0001$ (****). **e** Heatmap representing expression pattern of cone- and ganglion-associated genes in the two subtypes of retinoblastoma. Statistical significance and log2 fold-change in expression between subtype 2 and subtype 1 are also shown. Adjusted, $p \geq 0.05$ (ns), adjusted, $p < 0.05$ (*), adjusted, $p < 0.01$ (**), adjusted, $p < 0.001$ (***), adjusted, $p < 0.0001$ (****). Limma moderated two-sided t -tests and BH correction were used. Exact p -values are provided in Supplementary Data 3.

as *MKI67*. Cluster 3 presented a hypoxic gene expression program, including expression of the pro-apoptotic gene *BNIP3*. Clusters 5 and 6, accounting for 10.8% (129/1198) of all cells analyzed expressed hematopoietic markers, probably corresponding to stromal cell populations. Cluster 5 expressed monocyte/microglia markers (e.g., *CD14* and *AIF1*), whereas cluster 6 expressed T-lymphocyte markers, including markers of T-cell activation (e.g., *CD3D* and *TRAC*). A visualization of the expression of markers of each cluster is shown in Supplementary Fig. 5e, together with the expression of these markers in the normal developing retina.

To analyze the genomic heterogeneity in this tumor, we inferred copy-number variations (CNVs) in each single cell from the single-cell transcriptome data (see “Methods” section) (Fig. 5d). This analysis revealed that clusters 0–4 corresponded to tumor cells (presence of genomic alterations), whereas clusters 5 and 6 corresponded to normal cells (absence of genomic alterations). Genomic alteration patterns subdivided malignant cells into three distinct cell populations: cells with multiple genomic alterations (gains of 1q, 2q, 9p, 13q, loss of 8q), cells with 2p and 10q gains, and cells with 10q gains only. All cells from clusters 0 and 2 (*CRX*⁺/*EBF3*⁺/*GAP43*⁺ tumor cells), and some cluster 3 cells, corresponded to the first profile (multiple alterations). Cells from clusters 1 and 4 (*CRX*⁺/*ARR3*⁺/*GUCA1C*⁺ tumor cells) corresponded to the last two profiles (10q gain \pm 2p gain). Lastly, some cluster 3 cells corresponded to the second profile (2p and 10q gains).

The phenotypic analysis and the inferred copy-number alterations from single-cell RNA-seq data led us to conclude that the malignant cells of the subtype 2 tumor analyzed consisted of two populations, one expressing early photoreceptor/cone markers and neuronal/ganglion cell markers (clusters 0 and 2), and the other expressing early photoreceptor/cone markers and late cone markers (clusters 1 and 4). These two cell populations existed in three states, G1/S (clusters 0 and 1), G2/M (clusters 2 and 4), and hypoxic (cluster 3). A schema summarizing the interpretation of the different clusters is shown in Fig. 5e (upper panel). The *CRX*⁺/*EBF3*⁺/*GAP43*⁺ tumor population (clusters 0 and 2), presenting numerous genomic alterations, appeared to be genomically homogeneous. The *CRX*⁺/*ARR3*⁺/*GUCA1C*⁺ tumor population (clusters 1 and 4) was less unstable and consisted of two genomically different subpopulations. A tumor progression tree constructed from the genomic alterations found in the different cell populations of this tumor is proposed in Fig. 5e (bottom panel). The co-expression of *CRX*/*EBF3*/*GAP43* (early photoreceptor/cone marker and neuronal/ganglion cell markers)

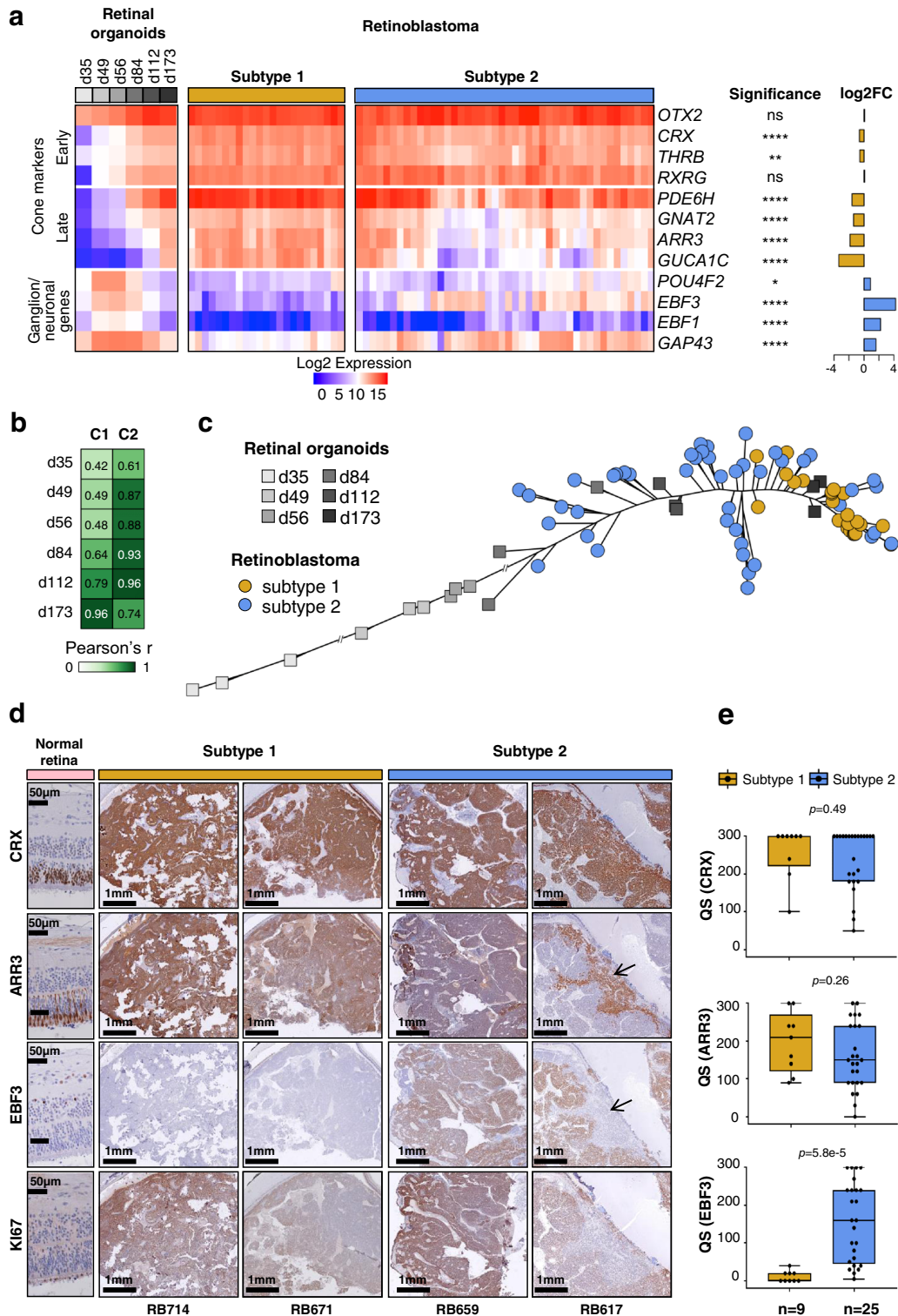
was unique to tumor cells as it was absent or very rare during normal retinal development (Supplementary Fig. 5f).

The single-cell RNA-seq analysis was performed on only one retinoblastoma. Single-cell analysis of additional tumors of both subtypes are necessary to further assess retinoblastoma heterogeneity and to investigate the relationship between retinal development and tumorigenesis using trajectory inference methods such as the ones estimating RNA velocity^{59,60}.

Subtype 2 tumors are associated with a higher risk of metastasis. We then investigated whether the retinoblastomas developing metastases belonged to a specific molecular subtype. No patients in our initial series of 102 retinoblastomas cases developed metastases. We, therefore, studied an additional series of 112 primary tumors presenting high-risk pathological features (HRPFs) at diagnosis, among which 19 tumors subsequently developed metastasis. All these patients were treated at the Garrahan Hospital (Buenos Aires, Argentina). Their clinicopathological characteristics, including HRPFs, are provided in Supplementary Data 6 and summarized in Table 2.

TFF1 belongs to a family of small secretory molecules involved in the protection and repair of the gastrointestinal tract⁶¹. *TFF1* is not expressed in the normal developing retina (Supplementary Fig. 6a). It was the top upregulated gene in subtype 2 tumors compared to subtype 1 tumors (fold-change = 55, adjusted p -value $< 10^{-12}$, Fig. 3a, Supplementary Data 3), with expression in most subtype 2 tumors but little or no expression in subtype 1 tumors (Supplementary Fig. 6b, c). These results were confirmed based on the transcriptome of two additional tumor series^{16,18} (Supplementary Figs. 6b, c and 7).

We assessed TFF1 protein expression by immunohistochemistry, in 55 of the tumors from our initial series of 102 classified retinoblastomas (18 subtype 1 and 37 subtype 2 tumors). Expression of TFF1, *CRX*, and *ARR3* are shown for representative tumors of subtypes 1 and 2 in Fig. 6a. Subtype 1 tumors displayed little or no TFF1 expression (QS ≤ 50 ; QS, quick score), whereas most subtype 2 tumors displayed high levels of expression (QS > 50 ; Fig. 6a, b, Supplementary Data 6). We then analyzed TFF1 expression in the additional series of 112 primary tumors with HRPFs including 19 metastatic cases (Garrahan series). TFF1 expression could be evaluated in 18 of the 19 primary tumors that subsequently developed metastasis. All 18 cases were positive for TFF1 (QS > 50), in contrast to the non-metastatic cases ($p = 0.00033$) (Fig. 6b and Supplementary Data 6),



suggesting that they belonged to subtype 2. Consistent with this, 15 of the 16 metastatic cases analyzed were also positive for EBF3 (QS > 270) (Supplementary Data 6), a ganglion marker specifically associated with subtype 2 (Figs. 3a, 4d, e and Supplementary Fig. 6). In seven of the 19 metastatic cases, tissues were available from both the primary tumor and the metastasis. In all but one of these cases, the metastatic sites were also positive for TFF1 (QS

range of 90–300). For EBF3, the six metastatic sites analyzed were positive (QS > 255), including the one negative for TFF1 (Fig. 6c and Supplementary Data 6). All these results suggested that subtype 2 tumors are more aggressive than subtype 1 tumors. These findings require validation by additional evidence for subtype 2 assignment, and by studies on additional series of metastatic cases.

Fig. 4 Expression of cone and neuronal/ganglion cell markers in retinoblastoma and retinal organoids. **a** Heatmap showing the expression of cone and ganglion markers in retinal organoids at different differentiation time points, and in subtype 1 and subtype 2 tumors assessed by NanoString technology. Differences in gene expression between the two subtypes were assessed by two-sided *t*-tests with BH correction. Exact *p*-values are provided in Supplementary Data 4. **b** Pearson's correlation of the expression of 8 cone markers, between the centroids of the 2 retinoblastoma subtypes and retinal organoids at different time points in differentiation. C1: centroid of subtype 1; C2: centroid of subtype 2. **c** Phylogenetic tree based on cone marker expression, for retinal organoids at different differentiation time points and for retinoblastoma samples. **d** Immunohistochemical staining of CRX, ARR3, EBF3, and Ki-67 in normal retina and retinoblastoma. For RB617, the black arrows indicate the mutually exclusive patterns for ARR3 and EBF3. Immunohistochemistry experiments were performed on 34 samples (subtype 1, *n* = 9; subtype 2, *n* = 25). Two representative images are shown for each subtype. **e** Boxplots showing the quick score (QS) for the differentiation markers used in the immunohistochemical analysis: CRX, ARR3, and EBF3. In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. The whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Two-sided Wilcoxon tests were used.

Discussion

The use of a multi-omics approach led us to the reliable identification of two main retinoblastoma molecular subtypes. The different molecular, pathological and clinical features of these two subtypes highlighted the relevance of this classification. In support of this, we could validate the transcriptomic signatures that distinguished the two subtypes in two independent series of retinoblastoma^{16,18} (Supplementary Fig. 7). The features of these two subtypes provide explanations for previous biological and clinical observations, with potential implications for retinoblastoma research and treatment.

Both subtypes expressed cone markers, consistent with the cone origin of human retinoblastoma^{11–15}. There are several possible non-exclusive explanations for the existence of two subtypes of retinoblastoma. The two subtypes may be derived from cone precursors located at different retinal positions. Several studies have reported a central-to-peripheral progression of retinoblastoma location with increasing age at diagnosis⁶². As subtype 2 tumors are diagnosed significantly later than subtype 1 tumors (median age = 23.9 vs 11 months), they are therefore likely to be more peripherally located than subtype 1 tumors. The two subtypes may be derived from different cone precursors. They may also be derived from cone precursors at different stages of maturation. Arguing against this last explanation, it has been shown that *RB1*^{−/−} retinoblastoma derived from an *ARR3*⁺ maturing cone precursor¹⁵.

We showed that subtype 1 tumors presented later markers of differentiated cones (*ARR3*⁺, *GUCA1C*⁺) and that subtype 2 tumors presented markers of earlier differentiation with an important heterogeneity between and within tumors. This is in agreement with the lower differentiation and the heterogeneity reported in older retinoblastoma patients⁵⁸. As both subtypes are likely derived from an *ARR3*⁺ maturing cone precursor, the lower differentiation and the heterogeneity of subtype 2 tumors with *RB1* inactivation probably result from a dedifferentiation process.

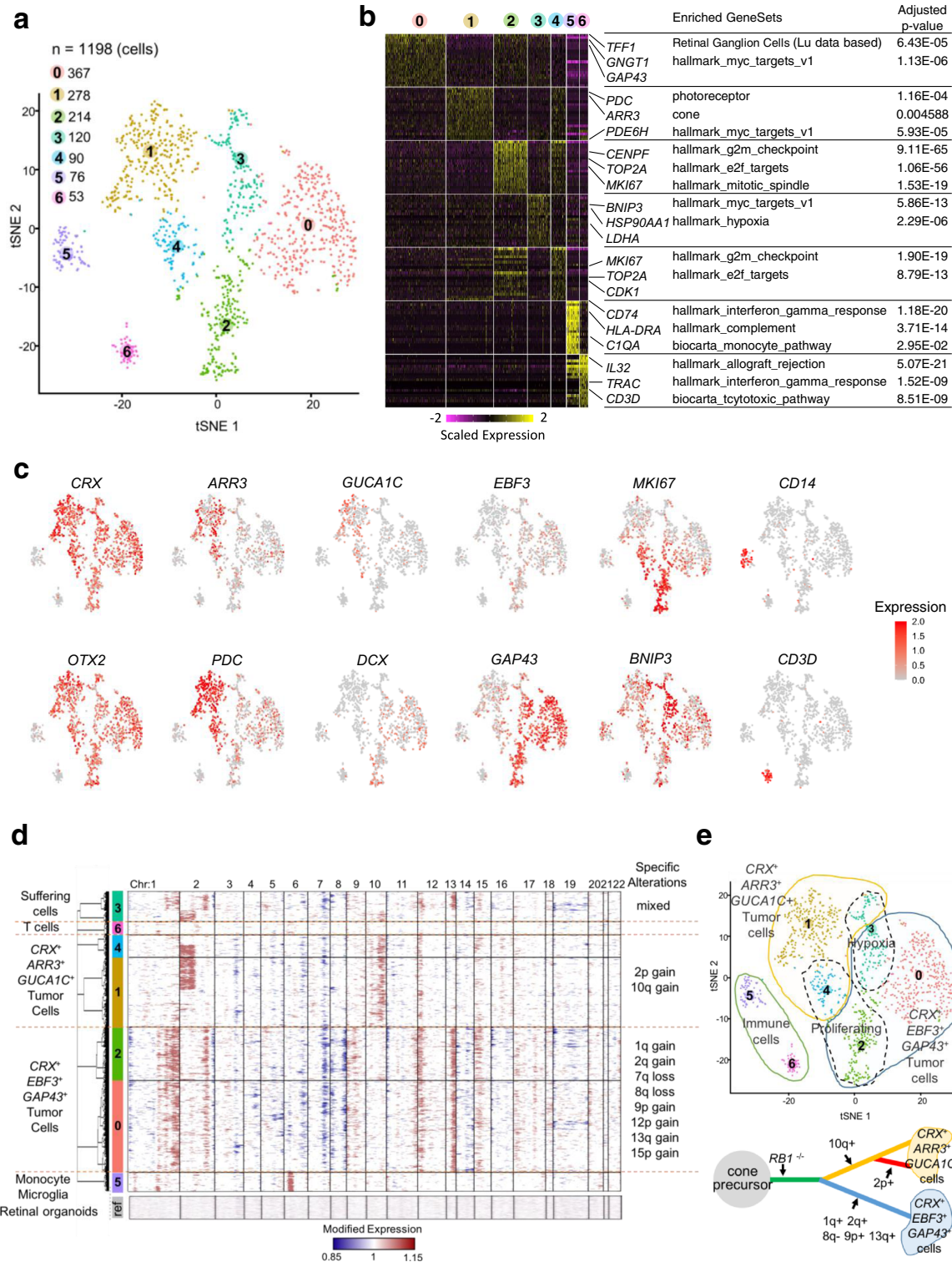
We found that subtype 2 was associated with low immune and interferon response, *E2F* and *MYC/MYCN* activation, and a higher propensity for metastasis, corresponding to stemness features recently reported^{32,33,35,36}. Genetic alterations and losses of function of *RB1* and *TP53* have also been shown to be associated with stemness in various cancers^{32,36}. *RB1* inactivation was present in most of the tumors of both retinoblastoma subtypes, but, nevertheless, a difference in stemness was observed between the two subtypes. The higher stemness in subtype 2 could be related to a decreased expression of another gene from the *RB* family, *RBL2*, located on 16q, which was lost in the majority of subtype 2 tumors. The higher stemness in subtype 2 tumors could be also related to an increased expression of *MDM4*, an inhibitor of *TP53* located on 1q which was gained/amplified in 74% of subtype 2 tumors. It has been proposed that both *MDM4* and

MDM2 abrogate *p53*-mediated tumor surveillance in retinoblastoma^{63,64}. Our results indicate that *MDM4* could be involved in subtype 2 tumors. In addition to the expression of cone markers, subtype 2 tumors overexpressed markers attributed to ganglion cell markers in the context of the retina. However, all these markers can also be viewed as neuronal markers (they correspond to genes expressed and involved in the central nervous system). Moreover, among the genes overexpressed in subtype 2 tumors, we identified neuronal genes expressed during the development of retinal ganglion cells but also of other retinal cell types (like *SOX11*, *DCX*, *STMN2*). These observations suggest that subtype 2 may be considered as a cone-neuronal subtype.

Expression of neuronal genes has now been found not only in the brain and neuroendocrine tumors, but also in some cancers of epithelial origin (breast, ovary, colon)⁶⁵. In recent years, it has become clear that tumor cells exploit neuronal and neurodevelopmental pathways to proliferate, migrate, and interact with normal cells, including endothelial cells and neurons^{65,66}. Therefore, the overexpression of neuronal genes that we found in subtype 2 tumors may contribute to the aggressiveness of these tumors.

The overexpression of *MYCN/MYC* target genes in subtype 2 tumors, and the assignment of 10 out of 11 *MYCN*-amplified tumors to subtype 2 tumors (the remaining *MYCN*-amplified tumor being unclassified) suggest that *MYCN/MYC* play an important role in this subtype. *MYC* and *MYCN* have been implicated in other pediatric tumors, including neuroblastoma and medulloblastoma, often in subsets of high-risk tumors. In neuroblastoma, *MYCN* amplification is found in approximately 20% of cases and is associated with high-risk disease and poor prognosis⁶⁷. It has recently been shown that *MYC* could also be a driver in another subset of high-risk neuroblastomas^{68,69}. Group 3 medulloblastoma are associated with *MYC* amplification (10–17%) and the worst overall survival^{70,71}. The activation of *MYC/MYCN* in subtype 2 tumors might be exploited for specific treatments of these tumors. Indeed *MYC/MYCN* can be inhibited indirectly by targeting their transcription with drugs such as *JQ1* and *OTX015*⁷², or directly, by targeting *MYC/MAX* interaction⁷³.

In the series of 102 retinoblastomas, tumors with *MYCN* amplification accounted for 17% of subtype 2 tumors. *MYCN*-amplified tumors did not cluster separately from other subtype 2 tumors on transcriptome analyses, but they nevertheless had specific features. Clinically, tumors with *MYCN* amplification were diagnosed at an earlier age than other subtype 2 tumors (median age at diagnosis: 15.9 vs 26.9 months). Molecularly, the tumors with *MYCN* amplification could be distinguished from subtype 2 tumors without *MYCN* amplification on the basis of uncommon 1q gains and 16q losses. Moreover, the tumors with *MYCN* amplification were hypomethylated outside CpG islands, as in other subtype 2 tumors, but they did not display



hypermethylation within CpG islands, by contrast to other subtype 2 tumors.

In high- and middle-income countries, the frequency of enucleation for retinoblastoma is decreasing, due to early diagnosis and the development of new conservative treatments. Techniques for analyzing tumor DNA methylation and copy-number changes in aqueous humor samples and blood from cell-free DNA have

recently been developed^{74,75}. The molecular characterization of retinoblastoma has, to date, been performed on tumor samples obtained from enucleation. The analyses of retinoblastoma through the use of liquid biopsy should provide a more comprehensive picture of the disease. Moreover, aqueous humor and blood samples could potentially be used to optimize retinoblastoma treatment through stratification by subtype.

Fig. 5 Intratumor heterogeneity at the single-cell level of a subtype 2 retinoblastoma (RBSC11). **a** 2D t-SNE plot of 1198 single retinoblastoma cells from one patient. Each dot represents one cell. **b** Heatmap of top cluster markers (top 20 most upregulated genes per cluster according to fold-change). Representative cluster markers and enriched gene sets are shown. Cluster marker *p*-values were calculated by hypergeometric tests with BH correction. **c** Expression of selected genes shown in 2D t-SNE plot (early photoreceptor markers: *CRX*, *OTX2*; late cone markers: *ARR3*, *GUCA1C*; neuronal/ganglion markers: *EBF3*, *GAP43*, *DCX*; proliferation marker: *MKI67*; pro-apoptotic marker: *BNIP3*; macrophage marker: *CD14*; T-cell marker: *CD3D*). **d** CNV profiles inferred from single-cell gene expression. Each row represents the profile of one individual cell. The genes on chromosome 6p overexpressed in the non-malignant cells monocyte/microglia correspond to HLA complex genes and should not be interpreted as CNV in cluster 5. **e** Upper panel: Diagram summarizing the interpretation of the different clusters of the 2D t-SNE plot. Lower panel: A progression model for this retinoblastoma case based on genomic alterations.

Table 2 Clinical and pathological characteristics of an additional series of 112 primary tumors presenting HRPFs.

Characteristics	Metastatic (n = 19)	Non-metastatic (n = 93)	p-value
Laterality n (%)			
Unilateral	14 (73.7%)	70 (75.3%)	0.8844 ^a
Bilateral	5 (26.3%)	23 (24.7%)	
Age at diagnosis (months)			
Median (range)	31 (10–88)	31 (1–168)	0.9166 ^b
Initial treatment n (%)			
Enucleation	15 (78.9%)	91 (97.8%)	0.007394 ^c
Pre-enucleation chemotherapy	4 (21.1%)	2 (2.2%)	
IRSS Stage I HRPF			
Isolated massive choroidal invasion (+ scleral invasion)	4 (1) (21%)	7 (6) (7.5%)	0.0312 ^c
Post-laminar optic nerve invasion (+ massive choroidal and/or scleral invasion)	9 (3) (47.4%)	83 (49) (89.3%)	
IRSS Stage II			
Tumor at the resection margin of the optic nerve	5 (26.3%)	3 (3.2%)	0.003428 ^c
IRSS not classified			
Complete necrosis	1 (5.3%)	0	
Site of metastatic relapse			
Isolated orbit	3 (15.8%)	N/A	
CNS	6 (31.6%)		
Systemic	1 (5.3%)		
Orbit and lymph node	1 (5.3%)		
Orbit and systemic relapse	3 (15.8%)		
Orbit and CNS	5 (26.3%)		

^aChi² test.

^bTwo-sided Wilcoxon rank-sum test.

^cTwo-sided Fisher's exact test.

In conclusion, the identification of two molecular subtypes—cone-like and cone/neuronal—represents a major advance in the understanding of retinoblastoma. It should redefine further studies of this pediatric cancer, including the development of models, improvement of diagnosis and prognosis, and identification of more specific treatments. The high stemness and neuronal features associated with subtype 2 tumors connect retinoblastoma with emerging fields of cancer research, and open up new opportunities for treatment.

Methods

Patient samples

Initial series of 102 retinoblastomas. We included 102 tumors from 50 male patients and 52 female patients in this study. These patients came from three different hospitals: Institut Curie in Paris, France (78 patients), the Garrahan Hospital in Buenos Aires, Argentina (19 patients), and the Sant Joan de Déu Hospital in Barcelona, Spain (5 patients). The median age at diagnosis was 19.9 months (minimum: 27 days, maximum: 9.65 years). Six patients had received chemotherapy and/or radiotherapy prior to enucleation.

Series of 112 retinoblastomas with HRPFs. We included an independent series of 112 patients with high-risk pathological features (HRPFs)⁷ from the Garrahan Hospital in this study (61 females and 51 males). The median age at diagnosis was 31 months (range: 1–168 months). Among the 112 patients, 19 subsequently developed the metastatic disease (9 females and 10 males). The median time from retinoblastoma diagnosis to metastasis was nine months (range: 4–65 months). Additional clinical characteristics are included in Table 2 and Supplementary Data 6.

Formalin-fixed paraffin-embedded tissues from the 112 tumors were analyzed. For seven metastatic patients, the metastatic sites were also available.

Additional retinoblastoma sample for single-cell RNA sequencing. One additional sample (RBSC11) was studied by single-cell RNA-seq. The sample was obtained from an enucleated patient >18 months of age with a unilateral non-hereditary form of retinoblastoma who did not receive treatment prior to enucleation.

Fetal retina. Fetal retinas were obtained from medical abortions. They were provided by the Fetal Pathology Unit of Antoine-Béclère Hospital in Paris (France). Three fetal retinas—RET215 (from a 20-week-old fetus), RET2 (23-week-old fetus), and RET1 (27-week-old fetus) were included in this study.

Ethics statement. All experiments were performed retrospectively and in accordance with the Declaration of Helsinki and the legislation of each participating country—France, Argentina, and Spain. The study was approved by the Institut Curie Review Board, the institutional review board of the Hospital de Pediatria Juan P Garrahan, and the Clinical Research Ethics Committee of Sant Joan de Déu Hospital. Written informed consent was obtained from parents or legal guardians of retinoblastoma patients, in accordance with current guidelines and legislation of each participating country.

Human fetuses (20, 23, 27 GW) were obtained from legally-induced terminations of pregnancy performed at the Antoine-Béclère Hospital in France. Fetal tissues were collected with the women's written consent, in accordance with the legal procedure agreed by the French National Agency for Biomedical Research (Agence de Biomédecine) and the approval of the local ethics committee of Antoine-Béclère Hospital.

Human iPSC maintenance and retinal organoid generation. Human-induced pluripotent stem cells (iPSCs) derived from dermal fibroblasts (hiPSC-2 clone)⁵²

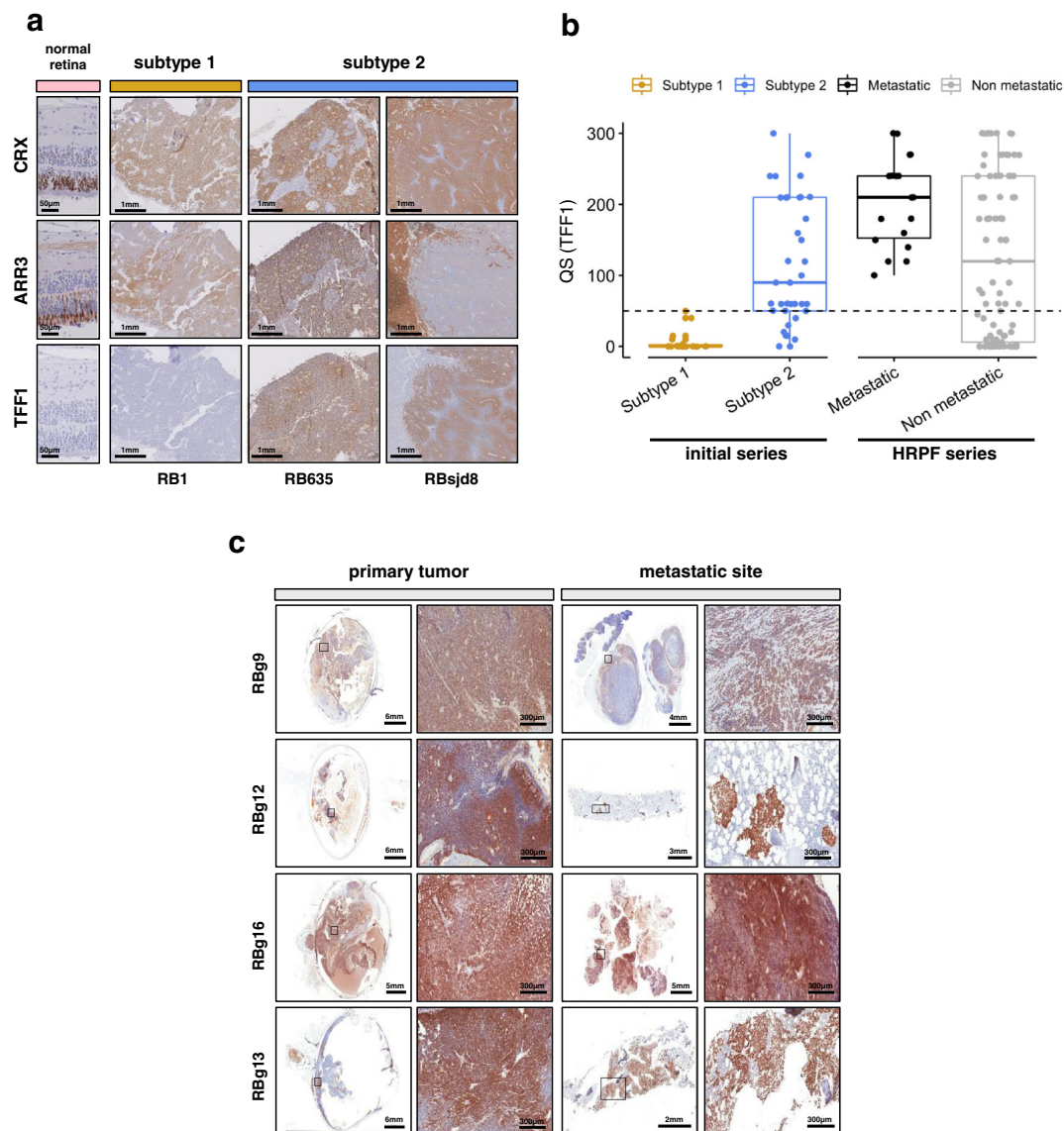


Fig. 6 Subtype 2 tumors are associated with a higher risk of metastasis. a Immunostaining of CRX, ARR3, and TFF1 in normal retina and retinoblastoma. Immunohistochemistry experiments were performed on 55 samples (subtype 1, $n = 18$; subtype 2, $n = 37$) from the initial series of 102 retinoblastomas. Representative images are shown: one subtype 1 tumor (RB1) and two subtype 2 tumors (RB635, RBsjd8). The subtype 2 tumors presented either a co-staining (RB635) or a mirror pattern (RBsjd8) for ARR3 and TFF1. **b** Boxplots showing the quick score (QS) for TFF1 in 55 tumors of the initial series (subtype 1, $n = 18$; subtype 2, $n = 37$), and in 112 tumors of the HRPF series. In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. The whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Two-sided Wilcoxon tests were used to assess the difference of the QS for Subtype 1 vs Subtype 2, $p = 1.1 \times 10^{-7}$, and metastatic vs non-metastatic, $p = 0.007$. **c** Immunostaining of TFF1 for primary tumors of metastatic retinoblastoma (left) and their metastatic sites (right), at low and high magnification. TFF1 expression could be assessed by immunohistochemistry for 6 of 7 available primary-metastasis tumor pairs. Representative images of four are shown.

were cultured on truncated recombinant human vitronectin-coated dishes in a humidified 37 °C incubator with 5% CO₂ in Essential 8™ medium (ThermoFisher Scientific) with daily medium change and weekly passage (2 ml enzyme-free Gentle cell dissociation reagent for 7 min at room temperature)⁴⁸. For retinal differentiation, adherent iPSCs were expanded to 70–80% and cultured in Essential 6™ medium (ThermoFisher Scientific) for 2 days, followed by replacing each 2–3 days Essential 6™ medium supplemented with 1% N-2 Supplement, 10 units/ml Penicillin and 10 µg/ml Streptomycin (ThermoFisher Scientific). At around day 28, retinal organoids were isolated with a needle and cultured as floating structures in ProB27 medium (DMEM:Nutrient Mixture F-12 1:1, L-glutamine, 1% MEM non-essential amino acids, 2% B27 supplement (ThermoFisher Scientific), 10 units/ml Penicillin, and 10 µg/ml Streptomycin) supplemented with recombinant human FGF2 (PeproTech) for a week and then in ProB27 medium for the next several weeks allowing retinal differentiation and maturation^{48,76}. By RT-qPCR and

immunofluorescence analysis, we previously showed that the different iPSC lines (hiPSC-2 clone⁵², AAVS1:CrxP_H2BmCherry hiPSC line⁷⁷) we used, are able to produce the whole repertoire of retinal cells, in an identical way and following the same chronological order with first the appearance of ganglion cells, then the amacrine and horizontal cells and finally the mature photoreceptors, the bipolar cells, and the Müller glial cells. The use of different markers of photoreceptor lineage (CRX, RCVRN, NRL, NR2E3, ARR3, RHO, OPSINS...) showed that the genesis of cones and rods is identical in the different iPSC lines used.

Sample collection and processing

Tumor samples. Institut Curie. Immediately after enucleation, a needle was inserted through the anterior chamber of the eye to extract a tumor sample by aspiration. The tumor specimen was placed in an RPMI medium on ice. The cells were

resuspended, counted and the suspension was split in two (for DNA and RNA preparation). The tubes were then centrifuged to remove the medium and the pellet was snap-frozen for later extraction. The remainder of the ocular globe was paraffin-embedded. For tumor DNA extraction, samples were first incubated in lysis buffer with recombinant proteinase K (Roche, Boulogne-Billancourt, France). They were next incubated with RNase A (Roche). DNA was then extracted using a standard phenol-chloroform protocol. Tumor RNA was extracted using the miRNeasy Mini Kit, according to the manufacturer's instructions (Qiagen, Courtaboeuf, France).

Garrahan Hospital and Sant Joan de Déu Hospital. Immediately after enucleation, a needle was inserted through the anterior chamber of the eye to extract a tumor sample by aspiration. The tumor specimen was either placed in guanidine thiocyanate or snap-frozen for later extraction. For tumor samples preserved in guanidine thiocyanate, alkaline phenol/chloroform/isoamyl alcohol (24:1:25) extraction was used for tumor DNA extraction. For snap-frozen tumor samples, commercial affinity columns (QIAamp DNA Mini Kit, Qiagen) or a standard phenol-chloroform protocol were used for tumor DNA extraction.

Single-cell RNA-seq sample. Tumor sample was processed immediately following needle aspiration through the anterior chamber of the eye. The sample was placed in an ice-cold CO₂-independent medium. Density gradient centrifugation by Histopaque-1077 (Sigma-Aldrich) was used to remove debris, dead cells, and erythrocytes. The isolated viable cells were mechanically dissociated, washed, and resuspended in phosphate-buffered saline supplemented with 0.04% bovine serum. Cell count and viability were determined by trypan blue exclusion on a Vi-CELL XR (Beckman Coulter Life Sciences).

Blood samples. For Curie hospital samples, normal DNA was extracted with a perchlorate/chloroform protocol or FujiFilm QuickGene technology (Kurabo Biomedical, Osaka, Japan). For Garrahan Hospital samples, normal DNA was extracted with a phenol/chloroform/isoamyl alcohol (24:1:25) protocol or with commercial affinity columns (QIAamp DNA Mini Kit, Qiagen). For Sant Joan de Déu Hospital samples, a standard isopropanol precipitation protocol was used.

Fetal retina. Fetal tissues were maintained in ice-cold Hanks balanced salt solution (HBSS) after medical abortions. For the isolation of neural retinal tissue, eyes were transferred onto a sterile Petri dish containing ice-cold PBS and maintaining a cornea side-up position with fine forceps. A small incision was made in the corneal junction using a small scalpel. The tip of the curved microscissors was inserted into the small incision. Eyes were carefully rotated of 360 degrees, and small incisions were made all the way around the eye, parallel to the corneal junction, allowing dissociation of the anterior eyecup and lens from the posterior eyecup. The posterior eyecup was passed onto a small Petri dish containing ice-cold PBS. The neural retina was carefully isolated from the underlying retinal pigmented epithelium by blunt dissection using fine forceps. RNA was extracted using the miRNeasy Mini Kit, according to the manufacturer's instructions (Qiagen, Courtaboeuf, France).

Human iPSCs. Total RNA was extracted from human iPSCs using the Nucleospin RNA II kit (Macherey-Nagel), according to the manufacturer's instructions.

Gene expression arrays. RNA of 59 samples (see Supplementary Data 1) were hybridized, in two batches, to Affymetrix Human Genome U133 plus 2.0 Array Plates (Santa Clara, CA) according to Affymetrix standard protocols. Raw CEL files were RMA⁷⁸ normalized using R package *affy* 1.60.0. Batch effects were corrected with the help of the Bioconductor package *SVA* 3.30.1. The arrays were mapped to genes with a Brainarray Custom CDF (EntrezG version 23)⁷⁹. Independent component analysis in $k = 3$ independent components (IC) was performed using R package *MineICA* 1.24.0 (JADE method)^{80,81}. The genes with high negative (< -2.5) or positive contributions (> 2.5) to IC were analyzed through pathway enrichment analysis (hypergeometric tests), seeking specifically signatures related to potential contamination by stromal cells. Genes with high positive contributions to IC #1 were found highly enriched in markers of stromal cells, and were discarded from clustering analyses.

DNA methylation arrays. Sixty-six DNA samples (Supplementary Data 1) were hybridized on Infinium HumanMethylation450 BeadChip arrays (Illumina, San Diego, CA). Four microliters of bisulfite-converted DNA were used for hybridization, following the Illumina Infinium HD Methylation protocol⁸². Data were processed using *preprocessIllumina* and *getBeta* functions in R package *Minfi* 1.28.4⁸³. Probes were annotated using the R package *IlluminaHumanMethylation450kmanifest* 0.4. Probes located in Chromosome X and Chromosome Y were discarded from subsequent analyses.

SNP arrays and BAC-CGH arrays. Ninety-five retinoblastomas were analyzed using SNP arrays or BAC-CGH arrays (Supplementary Data 1). Seventy tumor samples were analyzed on high-density SNP arrays. The B allele frequency and log-ratio signals were smoothed and analyzed using the Genome Alteration Print (GAP) algorithm (http://bioinfo-out.curie.fr/projects/snp_gap/)⁸⁴. Twenty-five

tumor samples were analyzed on BAC-CGH microarrays. These arrays consisted of 3510 or 5323 clones covering the human genome with an average resolution of 850Kb or 560Kb; they were designed by the CIT-CGH Consortium (INSERM U830, Institut Curie, Paris) and IntegrageTM⁸⁵. Hybridized slides were scanned and the scan data was pre-processed using R package *MANOR* 1.36.0⁸⁶ to correct for local spatial bias and continuous spatial gradient. Each array-CGH profile was centered on the median log₂ ratio and then analyzed to extrapolate copy-number profiles using the GLAD algorithm 2.28.1⁸⁷.

Whole-exome sequencing. Whole-exome sequencing was performed for 71 retinoblastomas and matched normal (blood) samples (Supplementary Data 2). For 32 tumor/normal sample pairs, sequence capture and exome sequencing were performed by the Sequencing Platform of Institut Curie. The Nextera exome enrichment kit (Illumina) was used for DNA library preparation. The eluted fraction was amplified by PCR and sequenced on an Illumina HiSeq 2500 sequencer as paired-end 100 × 100 bp or 150 × 150 bp reads. For the remaining 39 tumors/normal sample pairs, sequence capture and exome sequencing were performed by Integrage. The protocol followed by Integrage has been described elsewhere⁸⁸. In brief, Agilent in-solution enrichment (SureSelect Human All Exon Kit v4 + UTR) was used for DNA library preparation. The eluted fraction was amplified by PCR and sequenced on an Illumina HiSeq 2000 sequencer as paired-end 75 bp reads.

Single-cell library preparation and sequencing. Six thousand cells were loaded onto the Chromium System using the single-cell 3' reagent kits v2, in accordance with the manufacturer's protocol (10× Genomics), where single cells are partitioned in droplets. Following capture and lysis, cDNA incorporating UMI (unique molecular identifier) and cell barcode was synthesized and amplified. Amplified cDNA was fragmented and the Illumina sequencing library was constructed as per the manufacturer's protocol (Illumina). Libraries were loaded at 400pM and pair-end sequenced on Novaseq 6000 using NovaSeq 6000 S1 Reagent Kit (Illumina). Cells were sequenced at a mean depth of 100000. For quality control and quantification of cDNA and library, BioAnalyzer (Agilent BioAnalyzer High Sensitivity chip) was used.

Additional RNA quantification, DNA methylation, and mutation analyses

NanoString® codeset design and mRNA quantification. A codeset of 22 target genes was custom-designed and manufactured by NanoString® (Supplementary Data 4). One hundred nanograms of total RNA extracted from each sample was assessed on the Gen2 nCounter Analysis System from NanoString® Technologies at the Genomics Platform of the Curie Institute following the manufacturer's instructions. Samples were hybridized with multiplexed NanoString® probes containing a biotinylated capture probe and a reporter probe attached to a fluorescent barcode specific for each transcript, according to the nCounter codeset design (NanoString, Seattle, WA, USA). Hybridized samples were then purified and immobilized in a sample cartridge on the nCounter Prep Station for data collection, followed by quantification of the target mRNA in each sample using the nCounter Digital Analyzer (NanoString®). Data were normalized according to NanoString guidelines with nSolver 4.0. Briefly, the background was subtracted using the geometric mean of negative controls provided by NanoString®. The matrix was log-transformed (base 2) for further analysis.

Pyrosequencing. Forty-seven retinoblastoma samples were analyzed by performing pyrosequencing of the 9 selected CpGs (Supplementary Data 1 and Data Analysis section (Array-based methylation signature)).

Bisulfite treatment of genomic DNA (500 ng) was performed using the EZ DNA Methylation kit (Zymo Research). Primer design for each CpG target was performed using the PyroMark Assay Design software 2.0.2 (Qiagen) and pyrosequencing reaction was performed using PyroMark Q24 instrument (Qiagen). Primers used are provided in Supplementary Data 7. Pyrograms obtained were analyzed using the PyroMark Q24 software 2.0.6.20 (Qiagen) and methylation status was calculated at each CpG of interest.

Targeted sequencing. Targeted sequencing of the exonic regions of *RBI*, *BCOR*, and *ARID1A* was performed by IntegraGen SA (Evry, France) on 23 samples from the series of 102 retinoblastomas not subjected to whole-exome sequencing (Supplementary Data 2). The Fluidigm Access Array microfluidic system was used. PCR products were barcoded, pooled, and subjected to Illumina sequencing on a MiSeq instrument as paired-end 150-bp reads.

Sanger sequencing. Primer design was performed using Primer3 plus software⁸⁹. Their sequences are provided in Supplementary Data 7. PCR amplification was performed with the HotStarTaq plus DNA Polymerase (Qiagen). PCR products were purified and sequenced at the Genomics Platform of the Institut Curie, using an ABI 3730 XL (Applied Biosystems, Life Technologies). Sequence analysis was carried out using Sequencher® version 5.4.1 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI USA). One hundred nonsynonymous variants were identified by whole-exome sequencing and all variants identified by targeted sequencing were verified using Sanger dye-terminator sequencing. We validated 92

nonsynonymous mutations identified by whole-exome sequencing (of 100 variants tested, 92%) and all the mutations identified by targeted sequencing.

Immunohistochemistry. Immunohistochemical staining was performed on 3 μm -thick sections.

For the cohort of 102 retinoblastomas included in this study, automated immunostaining for CRX, ARR3, EBF3, Ki-67 (Supplementary Data 4), and TFF1 (Supplementary Data 6) was performed on the available paraffin-embedded samples with Autostainer 480 (Lab Vision) at Institut Curie. The following antibodies were used: anti-CRX (Abcam, ab140603; 1:300 for AFA/Bouin fixed tissue and 1:500 for formalin-fixed tissues), anti-ARR3 (Proteintech Group, 11100-2-AP; 1:300 for AFA/Bouin fixed tissue and 1:500 for formalin-fixed tissues), anti-EBF3 (Abnova Corporation, H00253738-M05; 1:800), anti-Ki-67 (Abcam, ab1558; 1:2500), and anti-TFF1 (Sigma-Aldrich, HPA003425; 1:1000). Additional information about the conditions used is described in Supplementary Data 4. For each slide, staining was assessed by eyeballing independently by two specialists (authors: NS and PF) blind to molecular subtype classification, taking into account the intensity (I) as null (0), mild (1), moderate (2), and strong (3), and the percentage (P) of tumor cells with stained nuclei for CRX and EBF3 and stained cytoplasm for ARR3 and TFF1. The quick score (QS) was then calculated as $I \times P$ (from 0 to 300).

For the independent series of 112 retinoblastomas with high-risk pathological features from Garrahan Hospital, immunostaining of TFF1 was performed manually in the Pathology Department of the Garrahan Hospital according to the procedure used at Institut Curie. For each slide, staining was assessed independently by three specialists (authors: R.A., F.L., and G.L.).

Bioinformatics and data analysis

GISTIC analysis. The copy-number alteration data for the 72 retinoblastomas studied by consensus clustering were first analyzed using GISTIC2.0 2.0.22⁹⁰. Twelve significant recurrent copy-number alteration regions were identified. The average copy number for each sample across these regions was then used for consensus clustering of the copy-number alteration data.

Consensus clustering. Consensus clustering was performed independently on the transcriptomic, methylomic, and GISTIC-processed copy-number alteration data of 72 retinoblastoma samples ($n = 59$ transcriptomes, $n = 66$ methylomes, $n = 72$ copy-number alteration profiles) (Supplementary Data 1). mRNA expression was assessed through Affymetrix U133plus2.0 arrays, genome methylation through Illumina Infinium Human Methylation 450 BeadChip arrays, and somatic copy-number alterations through SNP arrays or CGH-BAC arrays.

For the transcriptomic data, consensus hierarchical clustering was derived from a series of 24 dendrograms, which were obtained on all 59 retinoblastoma samples (columns) by analyzing 8 data subsets related to various numbers of genes (rows), through hierarchical clustering using 3 different linkage methods (average, complete, and Ward) and one distance metric ($1 - \text{Pearson correlation coefficient}$). To construct the 8 data subsets, various number of genes (rows) (spanning between 100 and 4709 genes) were selected based on 2 criteria: minimal robust coefficient of variation (rCV) thresholds spanning the 99.5th to the 60th percentiles, and p -value lower than 0.01 for a test of variance (we test whether the variance for a gene is higher or not than the median variance across all genes).

Having obtained these 24 dendrograms, we cut each dendrogram in k clusters, and get a series of partitions in k groups, for k ranging from 2 to 8 (NB: a partition in k groups is called a k -partition). For each value of k , we then derived a consensus k -partition from the 24 k -partitions obtained from the 24 dendrograms. To do so, we first calculated the (samples \times samples) co-classification matrix from these 24 k -partitions (NB: in the co-classification matrix, the cell (i, j) reports the number of partitions where samples i and j belong to the same group). The co-classification matrix is a similarity matrix and can be transformed into a dissimilarity matrix by replacing the value x in each cell (i, j) by $\text{MAX_VALUE} - x$ (Here $\text{MAX_VALUE} = 24$). Then this dissimilarity matrix can be used to perform the hierarchical clustering of the related samples, using the complete linkage. Finally, the obtained dendrogram is cut in k clusters to yield the consensus k -partition.

Of note, before calculating the consensus k -partitions (k from 2 to 8), we assessed the intrinsic stability of the underlying k -partitions, as compared to k -partitions obtained using the same linkage and the same set of genes, but based on “noisy” data. “Noisy” data were generated for each of the 8 data subsets (200 iterations for each) by addition of random Gaussian noise ($\mu = 0$, $\sigma = 1.5 \times \text{median variance calculated from the data set}$). The stability of each initial k -partition was then assessed using a stability score corresponding to the mean symmetric difference distance between an initial k -partition and the corresponding k -partitions derived from “noisy” data. The symmetric difference distance compares two partitions and gives the proportion of retention of the pairs of samples that are in the same group. It brings values ranging from 0 to 1: comparing two equal partitions yields a value of 1.

Consensus clustering of the methylomic data ($n = 66$ retinoblastomas) was performed in a similar manner, this time with between 2086 and 87937 CpGs selected (rCV thresholds spanning the 99.5th to the 60th percentiles and a p -value lower than 0.01 for the test of variance). Consensus clustering of the GISTIC-processed copy-number alteration data ($n = 72$ retinoblastomas) was also performed in a similar manner, this time with 3 or 4 significant copy-number

regions selected (rCV thresholds spanning the 80th to the 50th percentiles and a p -value lower than 0.01 for the test of variance). We observed both for transcriptome and methylome that the (intra-omics) consensus partition with $k = 2$ clusters was more stable than solutions with $k > 2$ clusters. We thus selected $k = 2$ clusters for these two omics. The DNA copy-number data yielded 5 clusters.

Cluster-of-clusters and centroid classification. To identify a common samples’ partition across all three genomic platforms (transcriptome, methylome, copy number), we used a cluster-of-cluster approach. Based on the three unsupervised consensus partitions previously obtained from the three omics datasets (one consensus partition per omics data set), we first built a (samples \times samples) co-classification matrix, with values ranging from 0 to 1, with 0 corresponding to a pair of samples that never co-classify in any genomic data set, and 1 corresponding to a pair of samples that always co-classify in all three genomic datasets. This matrix was then subjected to hierarchical clustering using complete linkage. Three clusters of clusters were thus identified ($n = 27$, $n = 37$, and $n = 8$). The two larger cluster-of-clusters corresponded to two core molecular subtypes, subtype 1 and subtype 2. The smallest cluster-of-clusters ($n = 8$) corresponded to ambiguous samples whose cluster assignments were not consistent across all three genomic platforms.

To classify these remaining samples according to either subtype 1 or subtype 2, we built two supervised centroid-based predictors, one transcriptomic and the other methylomic. The two core clusters of clusters defining subtype 1 and 2 were used to train these classifiers. For the transcriptomic data, the centroids of subtype 1 and subtype 2 were calculated as the intra-cluster median expression of the 800 genes most significantly differentially expressed between the two clusters (taking the 400 most upregulated genes in each subtype); similarly, for the methylomic data, the centroids of subtypes 1 and 2 were based on the median beta value of the 10,000 CpGs most significantly differentially methylated between the two clusters (5000 most methylated in each subtype). Each sample was assigned to the class whose centroid was closest to its profile, based on a Pearson’s correlation coefficient of at least 0.1 (we let unclassified samples yielding a Pearson’s correlation coefficient less than 0.1 to any of the two centroids/classes). Following this centroid-based step, we could classify 6 of the 8 samples without initial cluster-of-cluster attribution (four were assigned to cluster 1, two to cluster 2). This step also identified 3 outlier samples: two were already unclassified after the first cluster-of-clusters step, one was attributed initially to cluster 2.

Copy-number analysis. Copy-number alterations (CNAs) were analyzed using whole-exome sequencing (WES) data ($n = 63$), SNP arrays (Illumina HumanCNV370 quad, $n = 15$; Illumina Human610 quad, $n = 6$; Affymetrix Cytoscan, $n = 3$), and BAC arrays (3510 markers, $n = 12$; 5323 markers, $n = 3$). BAC arrays were analyzed using GLAD algorithm 2.28.1⁸⁷ to smooth log-ratio profiles into homogeneous segments and assign a discrete status to each segment (homozygous deletion, deletion, normal, gain, amplification). SNP arrays were analyzed using the Genome Alteration Print method⁸⁴, which takes into account both the log ratio and B allele frequency signals to determine normal cell contamination, tumor ploidy, and the absolute copy-number of each segment. The median absolute copy-number was considered to be the zero level of each sample. Segments with an absolute copy number $> \text{zero} + 0.5$ or $< \text{zero} - 0.5$ were considered to have gains and deletions, respectively. Segments with an absolute copy-number ≥ 5 or ≤ 0.5 were considered to have high-level amplifications and homozygous deletions, respectively. To identify CNAs using WES data, we calculated the log ratio of the coverage in each tumor and its matched normal sample for each bait of the exome capture kit with a coverage $\geq 30\times$ in the normal sample. Log-ratio profiles were then smoothed using the circular binary segmentation algorithm, as implemented in the Bioconductor package DNACopy 1.50.1⁹¹ (default parameters except $\text{min.width} = 4$, $\text{undo.splits} = \text{sundo}$, $\text{undo.SD} = 1.5$). The most frequent smoothed value was considered to be the zero level of each sample. Segments with a smoothed log ratio $> \text{zero} + 0.15$ or $< \text{zero} - 0.15$ were considered to have gains and deletions, respectively. High-level amplification and homozygous deletion thresholds were defined as the mean ± 5 s.d. of log ratios in regions of normal copy number. Visual inspection of the profiles allowed to validate recurrent focal amplifications and homozygous deletions.

For a given sample, the GNL (Gain = 1/Normal = 0/Loss = -1) copy-number data are aggregated by chromosome, as the proportion of features with an aberration (i.e., gain or loss). The overall genomic instability score corresponds to the mean score across all chromosomes.

Whole-exome sequencing analysis pipeline and mutation annotation. Sample reads were aligned using Burrows–Wheeler Aligner (BWA 0.7.4)⁹². Targeted regions were sequenced to an average depth of $82\times$, with 99% of the regions covered by $\geq 1\times$, 97.0% covered by $\geq 4\times$, and 87% covered by $\geq 20\times$.

For detection of somatic single-nucleotide variants (SNVs) and base insertions or deletions (indels), we used two separate variant-calling pipelines, the results of which were then merged. The first pipeline used MuTect 1.1.5⁹³ for SNV calling and the GATK SomaticIndelDetector 2.1–8 for indel calling^{94–96}. The second pipeline used VarScan 2.3.7 somatic and VarScan somatic filter for both SNV and indel calling (<http://varscan.sourceforge.net>)⁹⁷. After the variants called by both pipelines were merged, they were annotated using Annovar v2014Mar10⁹⁸. Custom

filters and manual curation using the Integrative Genomics Viewer (IGV 2.3.34)⁹⁹ were then used to maximize the number of true positive calls and to minimize the number of false positives.

Methylation analysis

Array-based methylation signature. From the methylome array data ($n = 66$), we selected the most differentially methylated CpGs between the two retinoblastoma subtypes (clusters of clusters) based on statistics of the Wilcoxon test. Out of the top 50 hypermethylated CpGs and top 30 hypomethylated CpGs of subtype 2 retinoblastoma (by p -value), top 7 hypermethylated and top 7 hypomethylated CpGs by the difference of beta value were selected for pyrosequencing. 5 of them did not perform well in pyrosequencing. This method led to the selection of 9 CpGs significantly differentially methylated that have been analyzed by pyrosequencing for sample classification. Seventeen samples from the initial series were analyzed by pyrosequencing for validation of the nine-CpG-based classifier (9 subtype 1, 8 subtype 2); from these samples, we derived subtype 1 and subtype 2 centroids based on these 9 CpGs. The nearest-centroid approach (with Pearson's metric and a minimal threshold of 0.3) correctly assigned 16 of these 17 samples to their known subtype and left unassigned the remaining sample. Additional samples analyzed by pyrosequencing for these 9 CpGs were then classified using the nearest-centroid approach (Pearson's metric) at a minimal threshold of 0.3.

Differential methylation analysis. Differential methylation analysis was performed by two-sided Wilcoxon rank-sum test and BH correction to compare methylation level of 473,864 probes between 27 subtype 1 and 36 subtype 2 retinoblastomas. 94,101 probes were found differentially methylated between the two subtypes (69,901 probes higher in subtype 1, 24,200 probes higher in subtype 2). 6607 probes had a difference of beta value of more than 0.2 (4520 higher in subtype 1, 2087 higher in subtype 2) (Supplementary Data 2).

Differential gene expression and pathway enrichment analysis. Differential gene expression analysis was performed by Limma R package 3.40.6¹⁰⁰ to compare the expression of 20,408 genes between 26 subtype 1 and 31 subtype 2 tumors. 6207 genes were found differentially expressed (adjusted p -value < 0.05) (Supplementary Data 3). Three main gene clusters were identified by hierarchical cluster analysis (mean centering of genes, $1 - \text{Pearson's correlation coefficient}$ as distance and average linkage). Visualization using heatmaps was performed with the R package ComplexHeatmap 2.1.1. Pathway enrichment analysis was performed by R clusterProfiler package 3.12.0¹⁰¹. Enriched gene sets from GOBP (Gene Ontology Biological Process) with adjusted p -value < 0.01 were selected for Cytoscape (3.7) EnrichmentMap (2.1.1) analysis¹⁰². Gene sets tested (GOBP and HALLMARK) were from the Molecular Signatures Database (MSigDB, version 6.2)¹⁰³.

Evaluation of stemness by transcriptome. Stemness indices in retinoblastoma were evaluated as described in Malta et al.³². Briefly, the weight vectors of 12,955 genes were obtained by Malta et al. as a stemness signature to identify pluripotent stem cells from progenitor cells in PCBC (Progenitor Cell Biology Consortium) transcriptomic data set. 12,364 genes were available in our data set. After mean-centering, the expression matrix, Spearman's correlation with the stemness signature vectors were calculated for each sample of retinoblastoma and then scaled to the range of 0 to 1 as the stemness indices. The other three stemness indices were estimated using three stemness gene signatures (Miranda et al., Shats et al., Smith et al. of 109, 80, and 49 genes, respectively)^{33,35,36} by `ssgsea` function of R package `gsva` 1.30.0. Boxplots were generated using R package `ggpubr` 0.2.0.

Pathway meta-score. Pathway meta-scores were calculated as the average expression of the genes involved in one selected pathway and then centered and scaled.

Analysis of two independent transcriptomic datasets. We applied the nearest-centroid approach (with Pearson's metric and a minimal threshold of 0.1) using the transcriptomic centroids calculated from our datasets to classify two publicly available transcriptomic datasets (GSE59983 and GSE29683).

In the Kooi et al.'s series¹⁸ ($n = 76$), 46 subtype 1 samples and 28 subtype 2 samples were identified, 2 samples were unable to be assigned a subtype. In the McEvoy et al.'s series¹⁶ ($n = 55$), 24 subtype 1 samples and 22 subtype 2 samples were identified among the 48 samples, 2 samples were unable to be assigned a subtype. Some samples ($n = 7$) were excluded from clustering analysis due to the high contamination of retinal pigmented epithelial (RPE) cells. We examined the average expression of an RPE gene signature (from Liao et al.¹⁰⁴, $n = 83/87$ genes present in the data) and removed the suspected outlier samples ($n = 7$) by Interquartile rule (suspected outliers are the samples when their average expression of RPE signature > Q3 + 1.5 IQR or < Q1 - 1.5 IQR).

Phylogenetic analysis of retinoblastoma and retinal organoids. Gene expression data of 8 genes related to cone-cell differentiation (*OTX2*, *CRX*, *THRB*, *RXRG*, *PDE6H*, *GNAT2*, *ARR3*, *GUCA1C*) were assessed by NanoString in 67 retinoblastomas (23 subtype 1 and 44 subtype 2) and 18 retinal organoids at 6 time

points after induction from iPSCs were used in phylogenetic analysis. Phylogenies were inferred by the minimal evolution algorithm¹⁰⁵ using `fastme.bal` function in R ape package 5.3 applied to Euclidean distance matrix based on these 8 gene expressions.

Single-cell transcriptome analysis

RBSC11 retinoblastoma. Sample demultiplexing, alignment to the reference genome (GRCh38, Ensembl 84, pre-built Cell Ranger reference version 1.2.0), quantification and initial quality control (QC) were performed using the Cell Ranger software (version 2.1.1, 10x Genomics).

Genes that were expressed in more than 3 cells and cells that expressed more than 500 genes and less than 5% of mitochondria genes were retained ($n = 1198$). The median numbers of genes and UMI counts per cell were 2911 and 7749, respectively. Normalization and clustering were performed using Seurat package version 2.3.4. UMI counts were normalized by `NormalizeData` function with `logNormalize` method, by a scaling factor of the median UMI count. UMI counts were then scaled to regress out the effect of UMI counts. Variable genes were found with `FindVariableGenes` function with `logVMR` function. Genes with an average expression more than 0.0125 and <8 and with dispersion more than 0.5 were considered as variable genes for principal component analysis (PCA). Cell clusters were identified by `FindClusters` function with shared nearest neighbor (SNN) method modularity optimization-based clustering algorithm¹⁰⁶, using the first 20 principal components. The parameter Resolution in the `FindClusters` was set between 0.4 and 1.4 and finally set to 0.6 for it provided a better biological interpretation.

Cluster markers were identified by `FindAllMarkers` function. Briefly, the expression of genes that expressed in more than 10% of cells in one cluster were compared with the expression of these genes in all other clusters, using Wilcoxon rank-sum test and corrected with BH correction. The procedure was repeated for all clusters. Genes upregulated in each cluster with more than 0.2 fold were considered as cluster markers. Pathway analysis of cluster markers was performed by R clusterProfiler package¹⁰¹. Gene sets tested were from the Molecular Signatures Database¹⁰³ (HALLMARK and BioCarta) and from Supplementary Data 3 (Cell type markers_Lu data and Selected cell type markers).

Correlation to bulk mRNA expression profiles of purified cell types was performed by R SingleR package 1.0.1¹⁰⁷. The expression profile of each cell was compared with the expression profiles of a data set that contains 713 microarray samples classified to 38 main cell types and further annotated to 169 subtypes¹⁰⁸.

Copy-number variations (CNVs) were inferred from the single-cell gene expression by `InferCNV` package 0.8.2, using normal retinal organoids derived from hiPSCs as reference.

Normal developing retina (Lu et al.³⁸, data). Normal retina scRNA-seq data from Lu et al.³⁸ were retrieved from GEO Omnibus database GSE138002. We retrieved the final filtered count matrix (GSE138002_Final_matrix.mtx.gz), gene annotations (GSE138002_genes.csv.gz), and cell annotations (GSE138002_Final_barcode.csv.gz). The latest includes, for each cell, the UMAP coordinates and the retinal cell type annotation computed by Lu et al. that was used for our analysis. Normalization of the UMI counts and identification of markers for each cell type was done with Seurat as described for the retinoblastoma sample. We also looked for pan-photoreceptor markers (markers of both cones and rods). Among the markers of Cones or Rods, genes that were found overexpressed in Cones against all other types except Rods and in Rods against all other types except Cones were assigned to pan-photoreceptor. Values indicated in Supplementary Data 3 for pan-photoreceptor markers have been computed using the `FindAllMarkers` function comparing photoreceptor cells against all other cells.

Visualization tool. A R-Shiny web-app [<https://retinoblastoma-retina-markers.curie.fr>], based on the shiny (v.1.6.0) and shinydashboard (v.0.7.1) R-packages, was developed to visualize the expression of markers of the retina cell populations, of the two subtypes of retinoblastoma and other genes of interest cited across the manuscript in the two single-cell RNA-seq datasets (from normal human developing retina³⁸ and from a subtype 2 retinoblastoma, RBSC11 (this report)). The different plots and tables are made based on the R packages `cowplot` (v.1.1.1) and the ones included in `tidyverse` (v.1.3.0).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw array data are deposited in the Gene Expression Omnibus (GEO) database under accession code [GSE58785](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58785). The raw whole-exome sequencing data are deposited in the European Genome-Phenome Archive (EGA) database under accession code [EGAS00001005248](https://www.ebi.ac.uk/ega/EGAS00001005248). The raw targeted sequencing data are deposited in the EGA database under accession code [EGAS00001005550](https://www.ebi.ac.uk/ega/EGAS00001005550). The raw single-cell RNA sequencing data are deposited in the EGA database under accession code [EGAS00001005178](https://www.ebi.ac.uk/ega/EGAS00001005178). Data in EGA is available under restricted access, access can be obtained by contacting Retinoblastoma Data Access Committee - Institut Curie (data.office@curie.fr). The public retinoblastoma

transcriptomic data used in this study are available in the GEO database under accession codes [GSE29683](#) and [GSE59983](#). The public human developing retina scRNA-seq data used in this study are available in the GEO database under accession code [GSE138002](#). The remaining data are available within the Article, Supplementary Information, or Source Data file. Additional data inquiry can be addressed to the Lead contact: francois.radvanyi@curie.fr. Source data are provided with this paper.

Code availability

Codes used to generate the analysis, figures and visualization app (<https://retinoblastoma-retina-markers.curie.fr>) are available at Github repositories (DOI: [10.5281/zenodo.5164167](https://doi.org/10.5281/zenodo.5164167), DOI: [10.5281/zenodo.5163255](https://doi.org/10.5281/zenodo.5163255))^{109,110}.

Received: 14 November 2020; Accepted: 26 August 2021;

Published online: 22 September 2021

References

- Dimaras, H. et al. Retinoblastoma. *Nat. Rev. Dis. Primers* **1**, 15021 (2015).
- Dyer, M. A. Lessons from retinoblastoma: implications for cancer, development, evolution, and regenerative medicine. *Trends Mol. Med.* **22**, 863–876 (2016).
- Munier, F. L. et al. Conservative management of retinoblastoma: challenging orthodoxy without compromising the state of metastatic grace. 'Alive, with good vision and no comorbidity'. *Prog. Retin. Eye Res.* **73**, 100764 (2019).
- AlAli, A., Kletke, S., Gallie, B. & Lam, W.-C. Retinoblastoma for pediatric ophthalmologists. *Asia-Pac. J. Ophthalmol.* **7**, 160–168 (2018).
- Lu, J. E. et al. Metastases and death rates after primary enucleation of unilateral retinoblastoma in the USA 2007–2017. *Br. J. Ophthalmol.* **103**, 1272–1277 (2019).
- Dunkel, I. J. & Chantada, G. L. in *Clinical Ophthalmic Oncology* (ed. Saunders, W. B.) 484–486 (Saunders, 2007).
- Sastre, X. et al. Proceedings of the consensus meetings from the International Retinoblastoma Staging Working Group on the pathology guidelines for the examination of enucleated eyes and evaluation of prognostic risk factors in retinoblastoma. *Arch. Pathol. Lab. Med.* **133**, 1199–1202 (2009).
- Rushlow, D. E. et al. Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *Lancet Oncol.* **14**, 327–334 (2013).
- Turner, D. L., Snyder, E. Y. & Cepko, C. L. Lineage-independent determination of cell type in the embryonic mouse retina. *Neuron* **4**, 833–845 (1990).
- Cepko, C. Intrinsically different retinal progenitor cells produce specific types of progeny. *Nat. Rev. Neurosci.* **15**, 615–627 (2014).
- Bogenmann, E., Lochrie, M. A. & Simon, M. I. Cone cell-specific genes expressed in retinoblastoma. *Science* **240**, 76–78 (1988).
- Munier, F. L., Balmer, A., van Melle, G. & Gailloud, C. Radial asymmetry in the topography of retinoblastoma. *Clues Cell Orig. Ophthalmic Genet.* **15**, 101–106 (1994).
- Xu, X. L. et al. Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell* **137**, 1018–1031 (2009).
- Xu, X. L. et al. Rb suppresses human cone-precursor-derived retinoblastoma tumours. *Nature* **514**, 385–388 (2014).
- Singh, H. P. et al. Developmental stage-specific proliferation and retinoblastoma genesis in RB-deficient human but not mouse cone precursors. *Proc. Natl Acad. Sci. USA* **115**, E9391LP–E9399400 (2018).
- McEvoy, J. et al. Coexpression of normally incompatible developmental pathways in retinoblastoma genesis. *Cancer Cell* **20**, 260–275 (2011).
- Kapatai, G. et al. Gene expression profiling identifies different sub-types of retinoblastoma. *Br. J. Cancer* **109**, 512–525 (2013).
- Kooi, I. E. et al. Loss of photoreceptor and gain of genomic alterations in retinoblastoma reveal tumor progression. *EBioMedicine* **2**, 660–670 (2015).
- Zhang, J. et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature* **481**, 329–334 (2012).
- McEvoy, J. et al. RB1 gene inactivation by chromothripsis in human retinoblastoma. *Oncotarget* **5**, 438–450 (2014).
- Kooi, I. E. et al. Somatic genomic alterations in retinoblastoma beyond RB1 are rare and limited to copy number changes. *Sci. Rep.* **6**, 25264 (2016).
- Kooi, I. E. et al. A meta-analysis of retinoblastoma copy numbers refines the list of possible driver genes involved in tumor progression. *PLoS ONE* **11**, e0153323 (2016).
- Mairal, A. et al. Detection of chromosome imbalances in retinoblastoma by parallel karyotype and CGH analyses. *Genes. Chromosomes Cancer* **28**, 370–379 (2000).
- Herzog, S. et al. Marked differences in unilateral isolated retinoblastomas from young and older children studied by comparative genomic hybridization. *Hum. Genet.* **108**, 98–104 (2001).
- Bowles, E. et al. Profiling genomic copy number changes in retinoblastoma beyond loss of RB1. *Genes. Chromosomes Cancer* **46**, 118–129 (2007).
- van der Wal, J. E. et al. Comparative genomic hybridisation divides retinoblastomas into a high and a low level chromosomal instability group. *J. Clin. Pathol.* **56**, 26–30 (2003).
- Zielinski, B. et al. Detection of chromosomal imbalances in retinoblastoma by matrix-based comparative genomic hybridization. *Genes. Chromosomes Cancer* **43**, 294–301 (2005).
- Sampieri, K. et al. Array comparative genomic hybridization in retinoma and retinoblastoma tissues. *Cancer Sci.* **100**, 465–471 (2009).
- Mol, B. M. et al. High resolution SNP array profiling identifies variability in retinoblastoma genome stability. *Genes. Chromosomes Cancer* **53**, 1–14 (2014).
- Polski, A. et al. Variability in retinoblastoma genome stability is driven by age and not heritability. *Genes. Chromosomes Cancer* **59**, 584–590 (2020).
- Priya, K., Jada, S. R., Quah, B. L., Quah, T. C. & Lai, P. S. High incidence of allelic loss at 16q12.2 region spanning RBL2/p130 gene in retinoblastoma. *Cancer Biol. Ther.* **8**, 714–717 (2009).
- Malta, T. M. et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* **173**, 338–354.e15 (2018).
- Miranda, A. et al. Cancer stemness, intratumoral heterogeneity, and immune response across cancers. *Proc. Natl Acad. Sci. USA* **116**, 9020–9029 (2019).
- Gilliam, D. T., Menon, V., Bretz, N. P. & Pruszak, J. The CD24 surface antigen in neural development and disease. *Neurobiol. Dis.* **99**, 133–144 (2017).
- Shats, I. et al. Using a stem cell-based signature to guide therapeutic selection in cancer. *Cancer Res.* **71**, 1772–1780 (2011).
- Smith, B. A. et al. A human adult stem cell signature marks aggressive variants across epithelial cancers. *Cell Rep.* **24**, 3353–3366.e5 (2018).
- Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).
- Lu, Y. et al. Single-cell analysis of human retina identifies evolutionarily conserved and species-specific mechanisms controlling development. *Dev. Cell* **53**, 473–491.e9 (2020).
- Sleven, H. et al. De novo mutations in EBF3 cause a neurodevelopmental syndrome. *Am. J. Hum. Genet.* **100**, 138–150 (2017).
- Tinteri, A. et al. Active intermixing of indirect and direct neurons builds the striatal mosaic. *Nat. Commun.* **9**, 4725 (2018).
- Gauthier-Campbell, C., Bredt, D. S., Murphy, T. H. & El-Husseini, A. E.-D. Regulation of dendritic branching and filopodia formation in hippocampal neurons by specific acylated protein motifs. *Mol. Biol. Cell* **15**, 2205–2217 (2004).
- Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
- Benzing, K., Flunkert, S., Schedl, A. & Engelkamp, D. A novel approach to selectively target neuronal subpopulations reveals genetic pathways that regulate tangential migration in the vertebrate hindbrain. *PLoS Genet.* **7**, e1002099 (2011).
- Wingo, A. P. et al. Large-scale proteomic analysis of human brain identifies proteins associated with cognitive trajectory in advanced age. *Nat. Commun.* **10**, 1619 (2019).
- Bye, C. R., Rytova, V., Alsanie, W. F., Parish, C. L. & Thompson, L. H. Axonal growth of midbrain dopamine neurons is modulated by the cell adhesion molecule ALCAM through trans-heterophilic interactions with L1cam, Chl1, and semaphorins. *J. Neurosci.* **39**, 6656–6667 (2019).
- Piras, I. S. et al. Association of AEBP1 and NRN1 RNA expression with Alzheimer's disease and neurofibrillary tangle density in middle temporal gyrus. *Brain Res.* **1719**, 217–224 (2019).
- Masuda, T. Contactin-2/TAG-1, active on the front line for three decades. *Cell Adh. Migr.* **11**, 524–531 (2017).
- Reichman, S. et al. Generation of storable retinal organoids and retinal pigmented epithelium from adherent human iPS cells in xeno-free and feeder-free conditions. *Stem Cells* **35**, 1176–1188 (2017).
- Kruczek, K. & Swaroop, A. Pluripotent stem cell-derived retinal organoids for disease modeling and development of therapies. *Stem Cells* **38**, 1206–1215 (2020).
- Kaewkhaw, R. et al. Transcriptome dynamics of developing photoreceptors in three-dimensional retina cultures recapitulates temporal sequence of human cone and rod differentiation revealing cell surface markers and gene networks. *Stem Cells* **33**, 3504–3518 (2015).
- Hoshino, A. et al. Molecular anatomy of the developing human retina. *Dev. Cell* **43**, 763–779.e4 (2017).
- Reichman, S. et al. From confluent human iPS cells to self-forming neural retina and retinal pigmented epithelium. *Proc. Natl Acad. Sci. USA* **111**, 8518–8523 (2014).
- Gill, K. P. et al. Enriched retinal ganglion cells derived from human embryonic stem cells. *Sci. Rep.* **6**, 30552 (2016).

54. Aparicio, J. G. et al. Temporal expression of CD184(CXCR4) and CD171(LICAM) identifies distinct early developmental stages of human retinal ganglion cells in embryonic stem cell derived retina. *Exp. Eye Res.* **154**, 177–189 (2017).
55. Chuang, J.-H. et al. Expression profiling of cell-intrinsic regulators in the process of differentiation of human iPSCs into retinal lineages. *Stem Cell Res. Ther.* **9**, 140 (2018).
56. Gallie, B. L. et al. Loss of RB1 induces non-proliferative retinoma: increasing genomic instability correlates with progression to retinoblastoma. *Hum. Mol. Genet.* **17**, 1363–1372 (2008).
57. Sampieri, K. et al. Genomic differences between retinoma and retinoblastoma. *Acta Oncol.* **47**, 1483–1492 (2008).
58. Eagle, R. C. J. High-risk features and tumor differentiation in retinoblastoma: a retrospective histopathologic study. *Arch. Pathol. Lab. Med.* **133**, 1203–1209 (2009).
59. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
60. Kester, L. & van Oudenaarden, A. Single-cell transcriptomics meets lineage tracing. *Cell Stem Cell* **23**, 166–179 (2018).
61. Jahan, R. et al. Odyssey of trefoil factors in cancer: diagnostic and therapeutic implications. *Biochim. Biophys. Acta Rev. cancer* **1873**, 188362 (2020).
62. King, B. A. et al. Spatiotemporal patterns of tumor occurrence in children with intraocular retinoblastoma. *PLoS ONE* **10**, e0132932 (2015).
63. Laurie, N. A. et al. Inactivation of the p53 pathway in retinoblastoma. *Nature* **444**, 61–66 (2006).
64. Qi, D.-L. & Cobrinik, D. MDM2 but not MDM4 promotes retinoblastoma cell proliferation through p53-independent regulation of MYCN translation. *Oncogene* **36**, 1760–1769 (2017).
65. Jung, E., Alfonso, J., Monyer, H., Wick, W. & Winkler, F. Neuronal signatures in cancer. *Int. J. Cancer* <https://doi.org/10.1002/ijc.33138> (2020).
66. Monje, M. Synaptic communication in brain cancer. *Cancer Res.* **80**, 2979–2982 (2020).
67. Matthay, K. K. et al. Neuroblastoma. *Nat. Rev. Dis. Primers* **2**, 16078 (2016).
68. Wang, L. L. et al. Augmented expression of MYC and/or MYCN protein defines highly aggressive MYC-driven neuroblastoma: a Children's Oncology Group study. *Br. J. Cancer* **113**, 57–63 (2015).
69. Zimmerman, M. W. et al. MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. *Cancer Discov.* **8**, 320–335 (2018).
70. Northcott, P. A. et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* **488**, 49–56 (2012).
71. Roussel, M. F. & Robinson, G. W. Role of MYC in medulloblastoma. *Cold Spring Harb. Perspect. Med.* **3**, a014308 (2013).
72. Puissant, A. et al. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
73. Han, H. et al. Small-molecule MYC inhibitors suppress tumor growth and enhance immunotherapy. *Cancer Cell* **36**, 483–497.e15 (2019).
74. Berry, J. L. et al. Potential of aqueous humor as a surrogate tumor biopsy for retinoblastoma. *JAMA Ophthalmol.* **135**, 1221–1230 (2017).
75. Jung, M. et al. Cell-free SHOX2 DNA methylation in blood as a molecular staging parameter for risk stratification in renal cell carcinoma patients: a Prospective Observational Cohort study. *Clin. Chem.* **65**, 559–568 (2019).
76. Slembrouck-Brec, A., Nanteau, C., Sahel, J.-A., Goureau, O. & Reichman, S. Defined xeno-free and feeder-free culture conditions for the generation of human iPSC-derived retinal cell models. *J. Vis. Exp.* **139**, 57795 (2018).
77. Gagliardi, G. et al. Characterization and transplantation of CD73-positive photoreceptors isolated from human iPSC-derived retinal organoids. *Stem Cell Rep.* **11**, 665–680 (2018).
78. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
79. Dai, M. et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **33**, e175 (2005).
80. Biton, A., Zinoviyev, A., Barillot, E. & Radvanyi, F. *MinICA: Independent Component Analysis of Transcriptomic Data* (2013).
81. Nordhausen, K., Gutch, H. W., Oja, H. & Theis, F. J. in *Latent Variable Analysis and Signal Separation* (eds Theis, F., Cichocki, A., Yeredor, A. & Zibulevsky, M.) 172–179 (Springer Berlin Heidelberg, 2012).
82. Bibikova, M. et al. High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288–295 (2011).
83. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
84. Popova, T. et al. Genome Alteration Print (GAP): a tool to visualize and mine complex cancer genomic profiles obtained by SNP arrays. *Genome Biol.* **10**, R128 (2009).
85. Idubai, A. et al. Two types of chromosome 1p losses with opposite significance in gliomas. *Ann. Neurol.* **58**, 483–487 (2005).
86. Neuvial, P. et al. Spatial normalization of array-CGH data. *BMC Bioinformatics* **7**, 1–20 (2006).
87. Hupé, P., Stransky, N., Thiery, J.-P., Radvanyi, F. & Barillot, E. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. *Bioinformatics* **20**, 3413–3422 (2004).
88. Guichard, C. et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44**, 694–698 (2012).
89. Untergasser, A. et al. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* **35**, W71–W74 (2007).
90. Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
91. Venkatraman, E. S. & Olshen, A. B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).
92. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
93. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
94. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013).
95. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
96. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
97. Koboldt, D. C. et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
98. Yang, H. & Wang, K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* **10**, 1556–1566 (2015).
99. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics* **14**, 178–192 (2013).
100. Smyth, G. K. *limma: Linear Models for Microarray Data*. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* 397–420 (Springer, New York, 2005).
101. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
102. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
103. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
104. Liao, J.-L. et al. Molecular signature of primary retinal pigment epithelium and stem-cell-derived RPE cells. *Hum. Mol. Genet.* **19**, 4229–4238 (2010).
105. Desper, R. & Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J. Comput. Biol.* **9**, 687–705 (2002).
106. Waltman, L. & van Eck, N. J. A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* **86**, 471 (2013).
107. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
108. Mabbott, N. A., Baillie, J. K., Brown, H., Freeman, T. C. & Hume, D. A. An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC Genomics* **14**, 632 (2013).
109. Liu, J. et al. Source codes for “A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression”. *GitHub* <https://doi.org/10.5281/zenodo.5164167> (2021).
110. Liu, J. et al. Source codes of visualisation tool for “A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression”. *GitHub* <https://doi.org/10.5281/zenodo.5163255> (2021).

Acknowledgements

We thank Cécile Reyes and Aude Vieillefont from the Genomics platform, Benoît Albaud from the NGS platform at the Institut Curie (IC), Emmanuel Martin from Integragen for sequencing, and Nadège Gruel (IC) for her help in the single-cell experiments. We thank Dr. Jelena Martinovic and Prof. Alexandra Benachi from Bécélère Hospital, Dr. M. Eugénia Riveiro from OTD Oncology, and Pr. Daniel Louvard from IC for their help with this work. Me.S. was supported by fellowships from the French MESRI and the IC, D.O. by the SFCE, the IC, and the AMCC, J.L. by the Fondation ARC, C.D. by the Barletta Foundation, Clé.H. by La Ligue Contre le Cancer. A.M.C. was funded by ISCii-FEDER (CP13/00189). This work was supported by the Retinostop Association, the Barletta Foundation, the INCA, and the INSERM in the framework of an ICGC project

(<https://icgc.org/icgc/cgp/62/355/1002881>), the IC, the ICGex Equipex program, the INCa/UNADEV, the SFCE, the FES, the Association l'Etoile de Martin, the ANR in the framework of LabEx LIFESENSES (ANR-10-LABX-65), and IHU FORESIGHT (ANR-18-IAHU-01), the XBTC sponsored by Pla Director d'Oncologia de Catalunya.

Author contributions

Resources: L.L.L.-R., A.M., L.D., J.C.-M., H.S., H.B., Fr.D., A.M.C., N.C., G.C. and I.A. (clinical data); F.L., G.L., P.F., Ma.S., X.S.G., A.M.C. and G.C. (pathological data); Ca.D., J.C., M.G.-V., D.S.-L., L.G. and Cla.H. (genetic data); R.A., F.L., G.L., O.M. and G.P.-P. (RNA, DNA preparation). Investigation: A.N., Cé.D., C.P., D.D., D.G., D.O., I.B.-P., J.C., J.G., J.L., L.M.O., F.L.D., F.N., M.L., N.K., N.S., O.G., P.S., R.A., S.A., S.B. and S.R. Methodology: E.C., Clé.H. and C.B. (development of visualization tools). Formal analysis (bioinformatics and statistical analyses): J.L., Me.S., E.C., L.T., Clé.H., A.B., C.B., T.P., S.G., C.V., E.B., E.L., A.V. and A.d.R. Data curation: E.C. and N.E. Conceptualization: J.L., D.O., Me.S., I.B.P., S.S., X.S.-G., Fr.D., N.C., C.P., O.G., G.C., A.d.R., I.A. and F.R. Visualization: J.L., D.O., Me.S., E.C., R.A., N.S., Clé.H., F.N., E.L., C.P., G.C., A.d.R. and F.R. Writing—original draft: J.L., D.O., Me.S., G.C., A.d.R. and F.R. Writing—review and editing: all authors. Funding acquisition: Me.S., Fr.D., O.G., G.C. and F.R. Supervision: A.M.C., G.C., A.d.R., I.A. and F.R. We thank the patients and their families for participating in this study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-25792-0>.

Correspondence and requests for materials should be addressed to François Radvanyi.

Peer review information *Nature Communications* thanks Quan Nguyen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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

© The Author(s) 2021

¹Institut Curie, CNRS, UMR144, Equipe Labellisée Ligue contre le Cancer, PSL Research University, 75005 Paris, France. ²Sorbonne Universités, UPMC Université Paris 06, CNRS, UMR144, 75005 Paris, France. ³Programme Cartes d'Identité des Tumeurs, Ligue Nationale Contre le Cancer, 75013 Paris, France. ⁴Precision Medicine, Hospital J.P. Garrahan, Buenos Aires, Argentina. ⁵Pathology Service, Hospital J.P. Garrahan, Buenos Aires, Argentina. ⁶Synergie Lyon Cancer, Plateforme de Bioinformatique "Gilles Thomas", Centre Léon Bérard, 69008 Lyon, France. ⁷Département de Biologie des Tumeurs, Institut Curie, 75005 Paris, France. ⁸Service de Génétique, Institut Curie, 75005 Paris, France. ⁹Institut de la Vision, Sorbonne Université, INSERM, CNRS, 75012 Paris, France. ¹⁰Institut Curie, PSL Research University, INSERM, U900, 75005 Paris, France. ¹¹Ecole des Mines ParisTech, 77305 Fontainebleau, France. ¹²Institut Curie, CNRS, UMR3347, PSL Research University, 91405 Orsay, France. ¹³Institut Curie, PSL Research University, INSERM, U1021, 91405 Orsay, France. ¹⁴Université Paris-Saclay, 91405 Orsay, France. ¹⁵Institut Curie, PSL Research University, INSERM U830, 75005 Paris, France. ¹⁶Département de Recherche Translationnelle, Institut Curie, 75005 Paris, France. ¹⁷Institut Curie, PSL Research University, NGS Platform, 75005 Paris, France. ¹⁸GeCo Genomics Consulting, Integragen, 91000 Evry, France. ¹⁹Département de Chirurgie, Service d'Ophtalmologie, Institut Curie, 75005 Paris, France. ²⁰Université de Paris, Paris, France. ²¹Institut de Recerca Sant Joan de Déu, 08950 Barcelona, Spain. ²²Pediatric Hematology and Oncology, Hospital Sant Joan de Déu, 08950 Barcelona, Spain. ²³Department of Pathology, Hospital Sant Joan de Déu, 08950 Barcelona, Spain. ²⁴Department of Ophthalmology, Hospital Sant Joan de Déu, 08950 Barcelona, Spain. ²⁵National Scientific and Technical Research Council, CONICET, Buenos Aires, Argentina. ²⁶Département d'Imagerie Médicale, Institut Curie, 75005 Paris, France. ²⁷Centre de Recherche des Cordeliers, Sorbonne Universités, INSERM, 75006 Paris, France. ²⁸Functional Genomics of Solid Tumors, équipe labellisée Ligue Contre le Cancer, Université de Paris, Université Paris 13, Paris, France. ²⁹SIREDO Center (Care, Innovation and Research in Pediatric Adolescent and Young Adult Oncology), Institut Curie, 75005 Paris, France. ³⁰Present address: Institut Pasteur – Hub Bioinformatique et Biostatistique – C3BI, USR 3756 IP CNRS, 75015 Paris, France. ³¹Present address: INSERM U930, CHU Bretonneau, 37000 Tours, France. ³²Present address: Department of Genetics, Rouen University Hospital, 76000 Rouen, France. ³³Present address: Department of Pathology, Centre Hospitalier Intercommunal de Créteil, 94000 Créteil, France. ³⁴These authors contributed equally: Jing Liu, Daniela Ottaviani, Meriem Sefta. ³⁵These authors jointly supervised this work: Guillermo Chantada, Aurélien de Reyniès, Isabelle Aerts, François Radvanyi. ✉email: francois.radvanyi@curie.fr

SUPPLEMENTARY INFORMATION

List of supplementary figures

Supplementary Figure 1. Multi-omics classification of a series of 72 retinoblastomas.

Supplementary Figure 2. Distribution of the differentially methylated CpG between the two subtypes in relation to CpG islands.

Supplementary Figure 3. Stemness Indices estimated by different signatures.

Supplementary Figure 4. Retinal cell gene expression analysis in the two subtypes of retinoblastoma.

Supplementary Figure 5. Data related to single-cell RNA-seq of a subtype 2 retinoblastoma (RBSC11).

Supplementary Figure 6. Expression of TFF1 and EBF3 in the normal developing retina and in the two subtypes of retinoblastoma.

Supplementary Figure 7. The two retinoblastoma subtypes were different in clinical and molecular features in two additional independent datasets.

Supplementary figures are available through the following link:

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM1_ESM.pdf

List of supplementary data

Supplementary Data 1. Patient information, available molecular data types, and molecular subtype assignment.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM4_ESM.xlsx

Supplementary Data 2. Genomic characterization, somatic mutational landscape and DNA methylation profile of retinoblastoma subtypes.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM5_ESM.xlsx

Supplementary Data 3. Differential gene expression analyses between retinoblastoma subtypes.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM6_ESM.xlsx

Supplementary Data 4. Late stage of cone differentiation in subtype 1 tumors and various stages of cone differentiation in subtype 2 tumors together with the expression of ganglion markers.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM7_ESM.xlsx

Supplementary Data 5. Single-cell transcriptomic analysis.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM8_ESM.xlsx

Supplementary Data 6. Clinicopathological characteristics and immunohistochemical TFF1 data of an independent series of metastatic retinoblastoma patients (n=19), and of a control group (n=93).

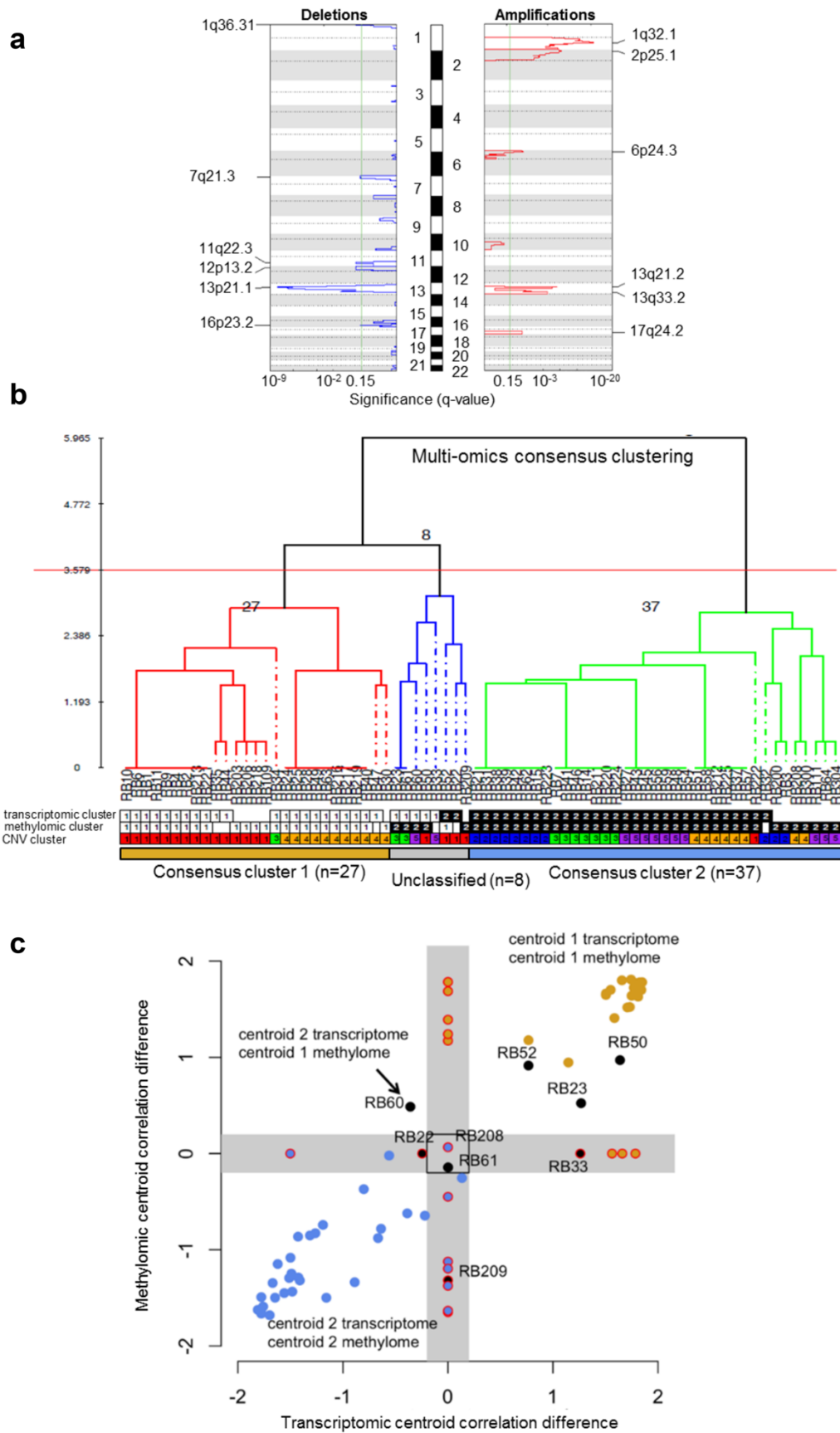
CHAPTER 3

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM9_ESM.xlsx

Supplementary Data 7. List of primers used.

https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM10_ESM.xlsx

CHAPTER 3



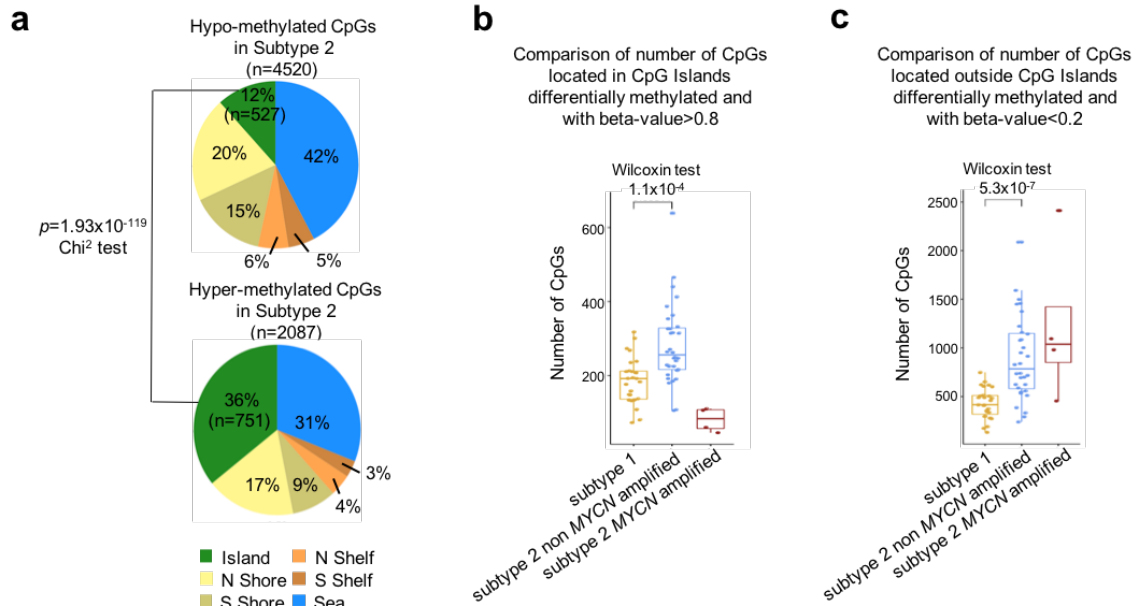
Supplementary Figure 1. Multi-omics classification of a series of 72 retinoblastomas.

CHAPTER 3

a, GISTIC plot for copy number alteration data. GISTIC plot of the 72 retinoblastoma specimens, integrating frequency and amplitude to identify significant amplifications (red) and deletions (blue) across the genome (ordered by chromosome). These significant amplifications and deletions were used for copy number alteration data clustering (Fig. 1a).

b, Cluster-of-clusters classification. Cluster-of-clusters analysis for 72 retinoblastoma cases, based on three unsupervised partitions, each partition being obtained using a different genomic platform (transcriptomic, methylomic, and copy number alteration data). A [0;1] normalized co-classification matrix was derived from the 3 initial partitions. Hierarchical clustering was then performed with inter-individual distance defined as $(1 - \text{co-classification score})$ and complete linkage. This analysis identified two major groups of 27 and 37 samples, and one ambiguous/unclassified group of 8 samples. The annotations below the dendrogram represent the clusters of samples defined independently by consensus clustering of all three genomic datasets.

c, Centroid-based classification. This is a schematic representation of the centroid classification methodology used. Each point represents a sample. Samples belonging to cluster-of-clusters 1 are in yellow ($n=27$), cluster-of-clusters 2 in blue ($n=37$), and yet unclassified samples in black ($n=8$). On the X-axis is the difference between each sample's correlation to the cluster-of-clusters 1 transcriptomic centroid and the cluster-of-clusters 2 transcriptomic centroid. The Y-axis represents the same information for the methylomic centroids. Samples with one dataset missing have this difference set to 0 (X coordinate=0 for missing transcriptome, Y coordinate=0 for missing methylome), and the corresponding data points are circled in red. Centroid correlation differences below 0.2 were considered to be outliers, and are in the gray areas. Gray areas therefore contain outliers or samples with one missing dataset. All cluster-of-clusters 1 samples re-classified correctly in methylomic centroid 1 and transcriptomic centroid 1 (or one of the two if a dataset is missing). All but one cluster-of-clusters 2 samples re-classified correctly also. The last one (RB208) was an outlier in the methylomic dataset with no transcriptomic data available, and was thus set to unclassified. For the samples with no cluster-of-clusters attribution (black points): RB52, RB23, RB50 and RB33 were assigned to the first group; RB22, and RB209 were assigned to the second; RB61 was an outlier in both datasets, and RB60 had a discrepancy in the transcriptomic centroid-based classification and the methylomic centroid-based classification, RB61 and RB60 therefore remained unclassified. Ultimately 31 samples were assigned to the first group, 38 to the second, and 3 (RB208, RB60, RB61) remained unclassified.



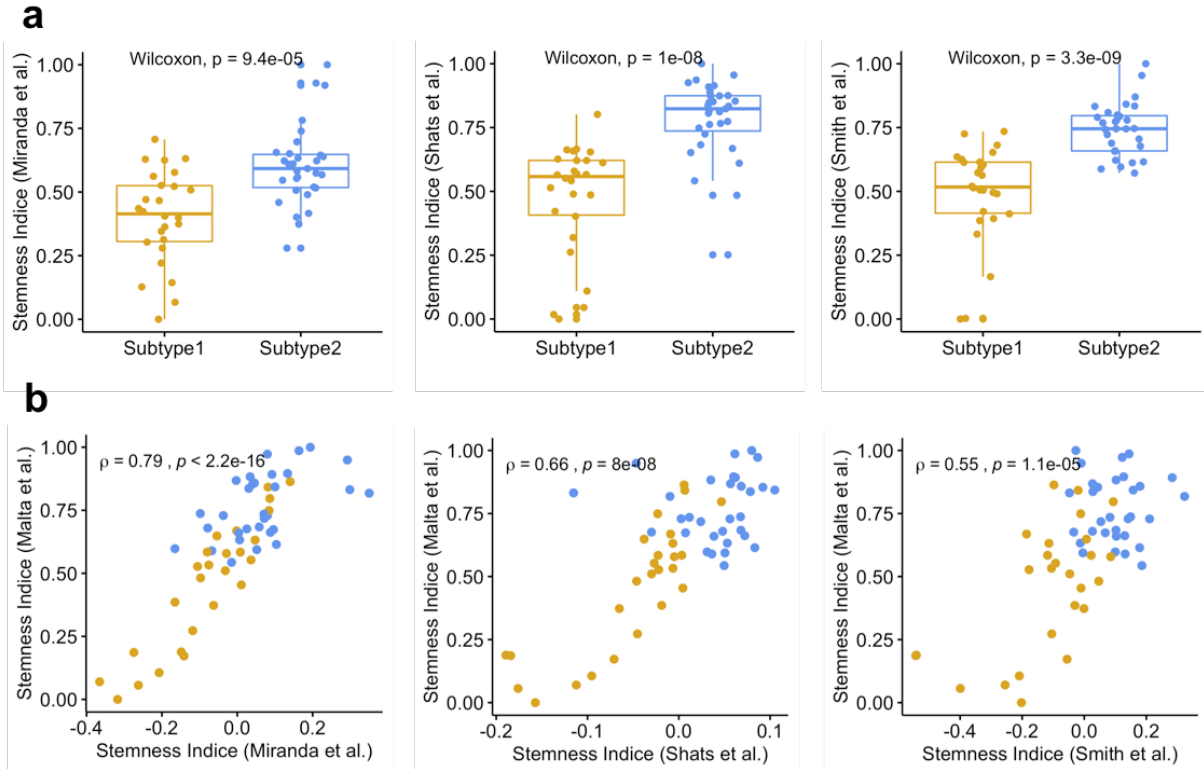
Supplementary Figure 2. Distribution of the differentially methylated CpG between the two subtypes in relation to CpG islands.

a, Distribution of the hypomethylated CpGs (upper panel) and the hypermethylated CpGs (lower panel) in subtype 2 as compared to subtype 1, by CpG content and neighborhood context. Hypermethylated CpGs were more frequently located within CpG islands than hypomethylated CpGs (751 of 2087 versus 527 of 4520, $p=1.93 \times 10^{-119}$, Chi² test) (related to Figure 2F).

b, Boxplot comparing the number of differentially methylated CpGs with high levels of methylation (beta-value within 0.8 to 1) located in CpG Islands in subtype 1 (n=27), subtype 2 non-MYCIN-amplified (n=32), and subtype 2 MYCIN-amplified (n=4) tumors (related to Figure 2g upper panel).

c, Boxplot comparing the number of differentially methylated CpGs with low levels of methylation (beta-value within 0 to 0.2) located outside CpG Islands in subtype 1 (n=27), subtype 2 non-MYCIN-amplified (n=32), and subtype 2 MYCIN-amplified (n=4) tumors (related to Figure 2g bottom panel).

b, c, In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. Whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Significance was tested by two-sided Wilcoxon test, $p=1.9 \times 10^{-7}$.



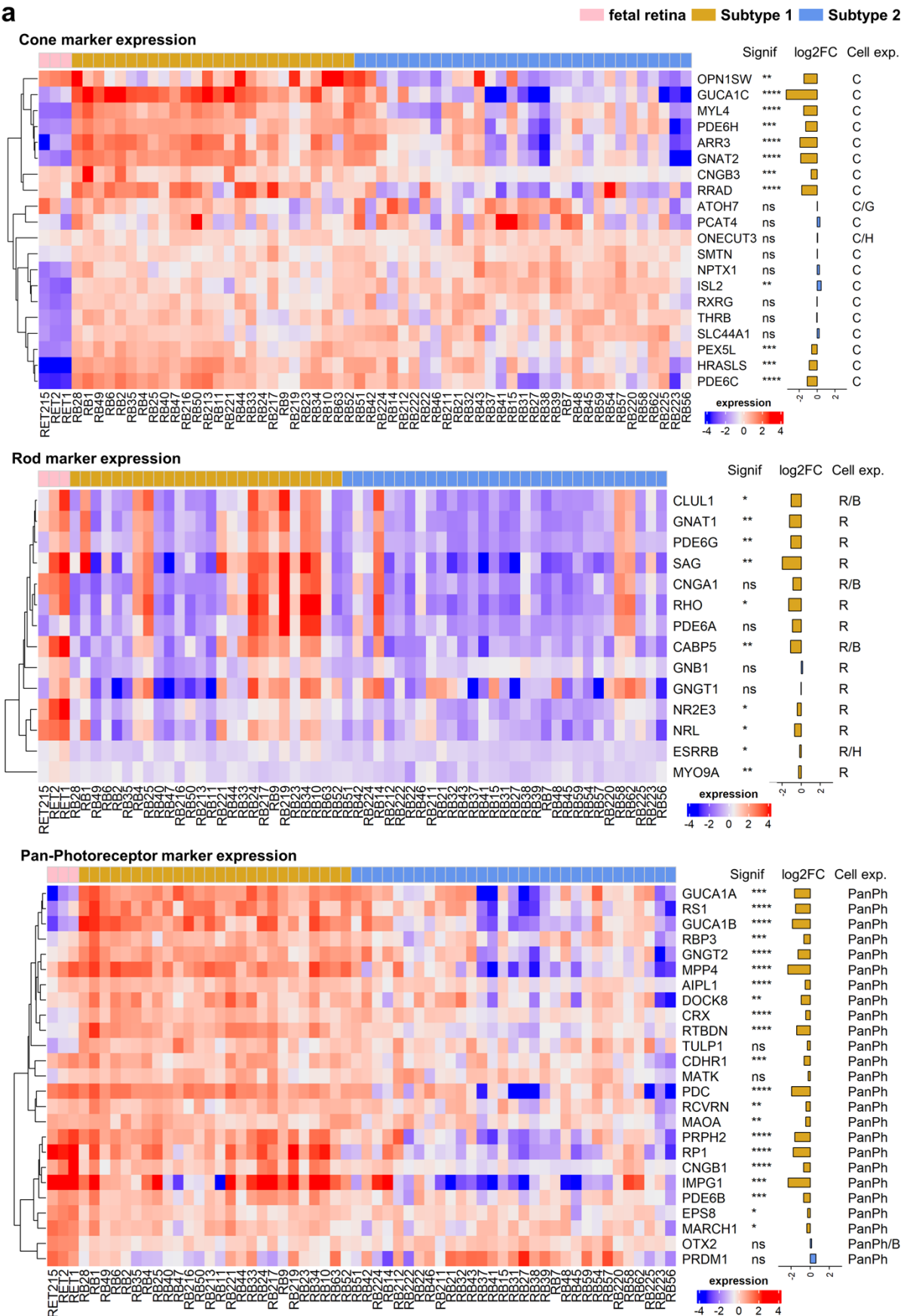
Supplementary Figure 3. Stemness Indices estimated by different signatures.

a, Boxplot of stemness indices in the two subtypes of retinoblastoma (subtype 1 tumors: $n=26$, subtype 2 tumors: $n=31$) estimated by different signatures (from left to right: Miranda et al. 201833, Shats et al. 201135, Smith et al. 201836). In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. Whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers. Statistical tests are two-sided.

b, Scatter plot showing the correlation between stemness indice estimated by Malta et al.'s method (y-axis) or by other signatures (x axis, from left to right: Miranda et al. 2018, Shats et al. 2011, Smith et al. 2018). Two-sided Spearman's correlation test was applied, rho and p value are shown.

CHAPTER 3

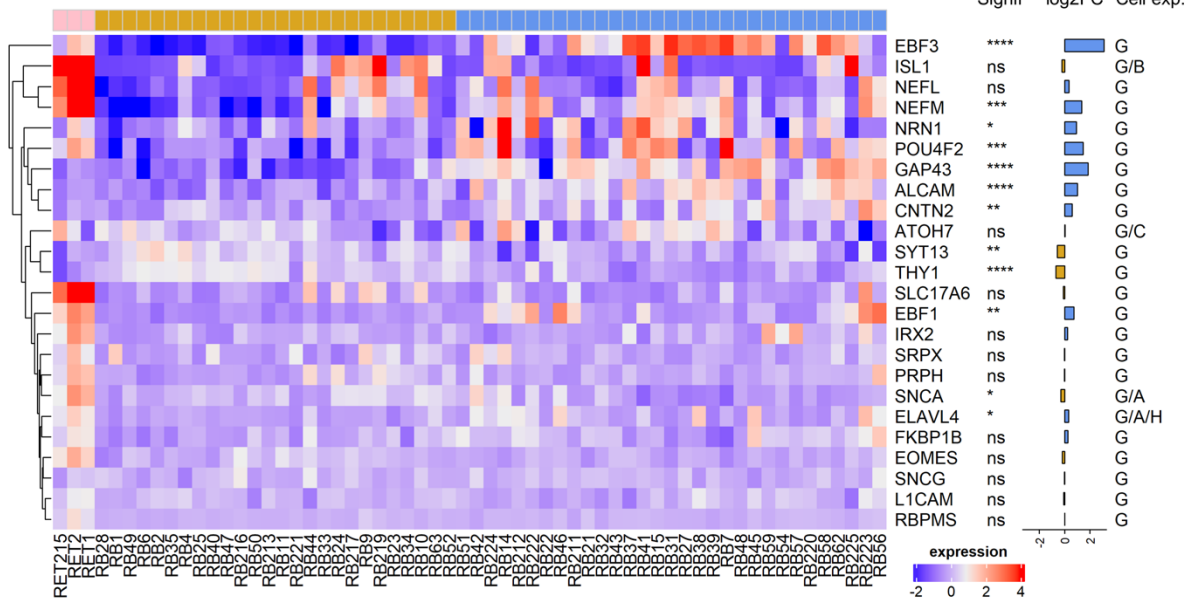
a



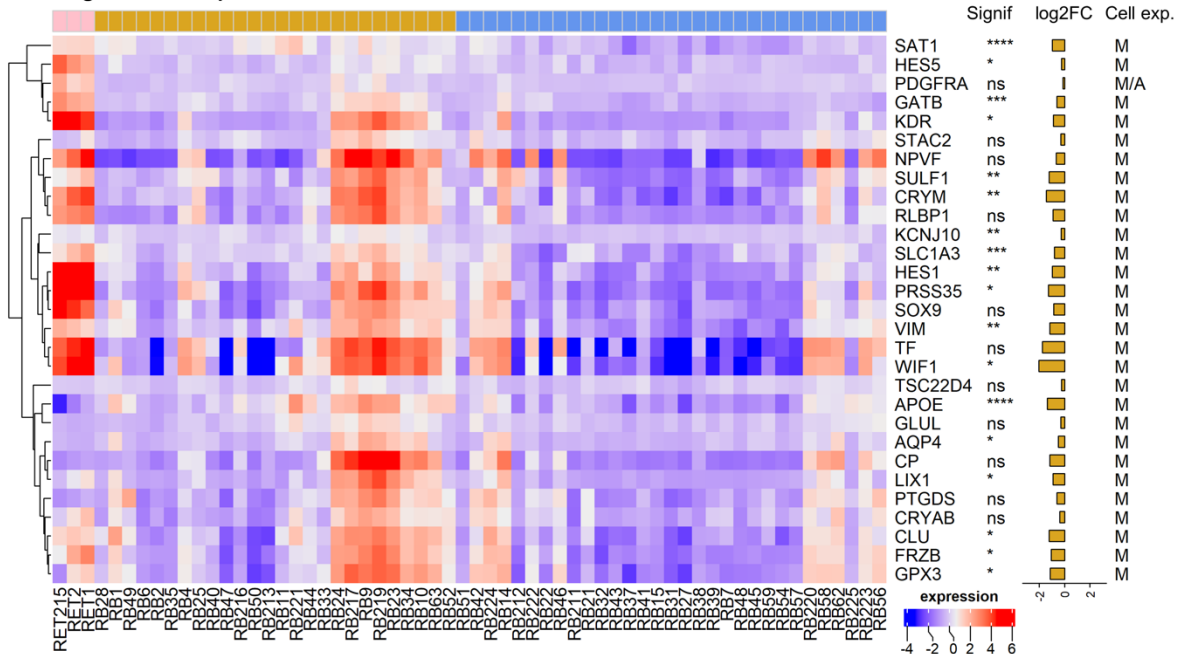
CHAPTER 3

fetal retina Subtype 1 Subtype 2

Ganglion marker expression



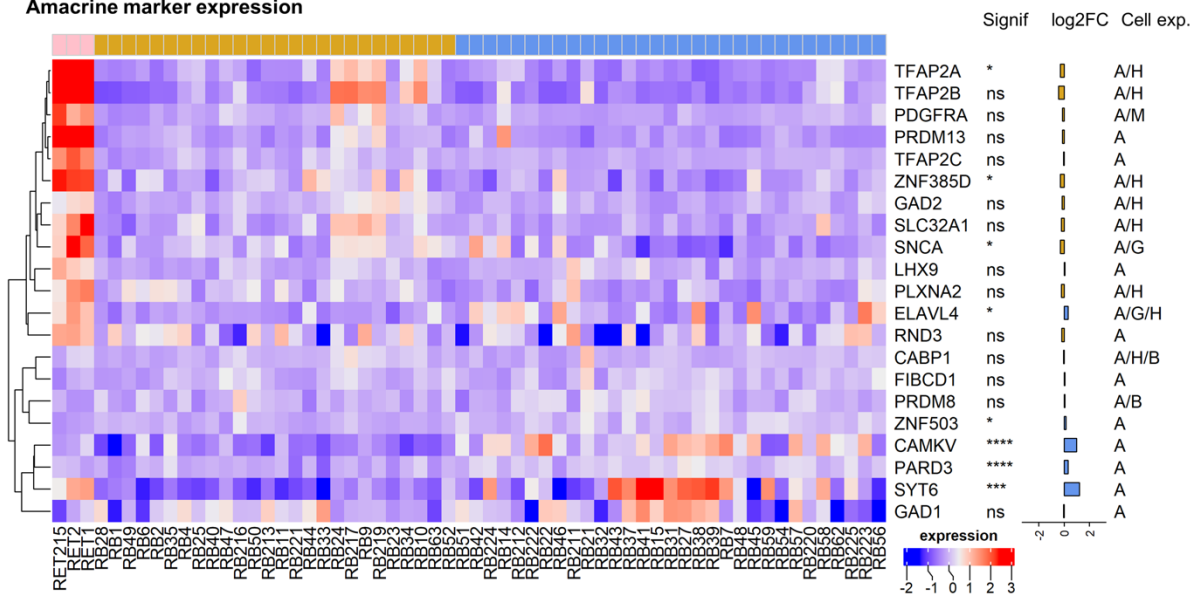
Müller glia marker expression



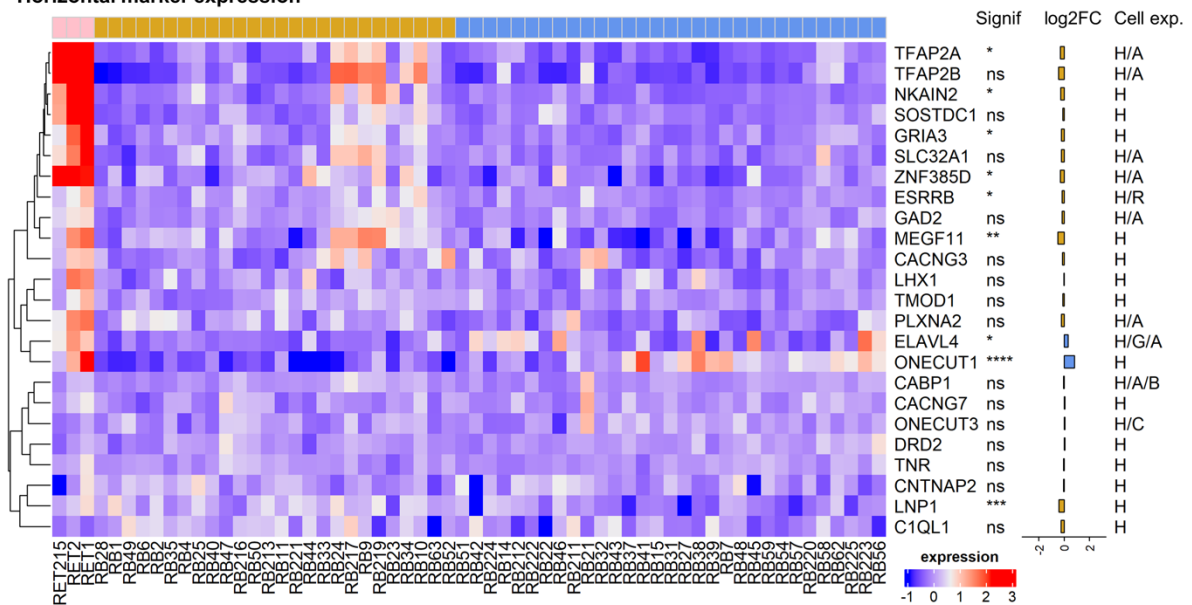
CHAPTER 3

fetal retina Subtype 1 Subtype 2

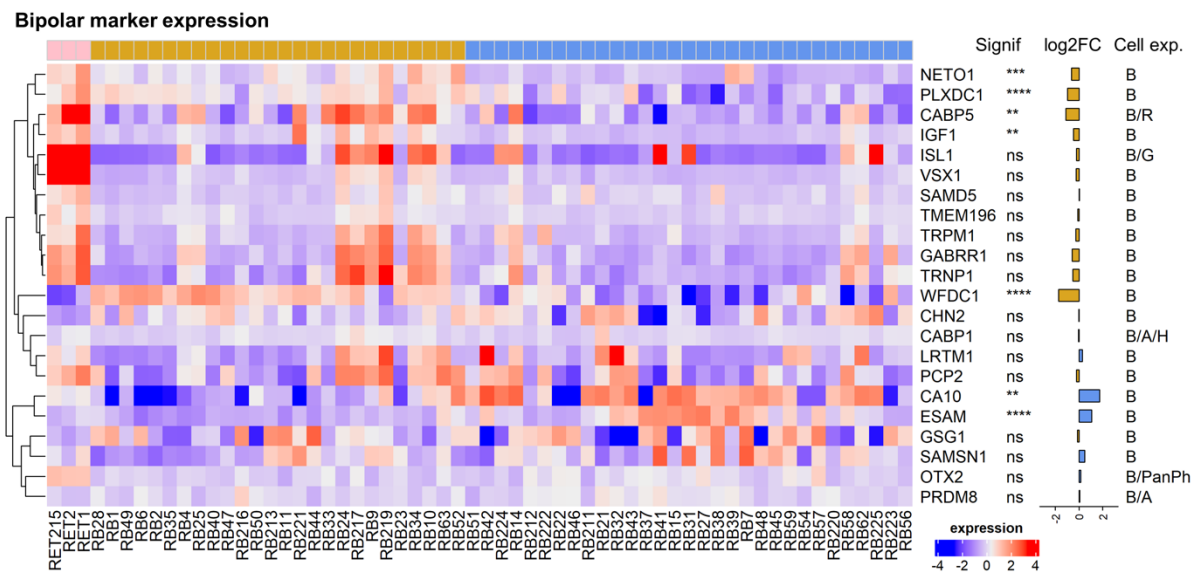
Amacrine marker expression



Horizontal marker expression



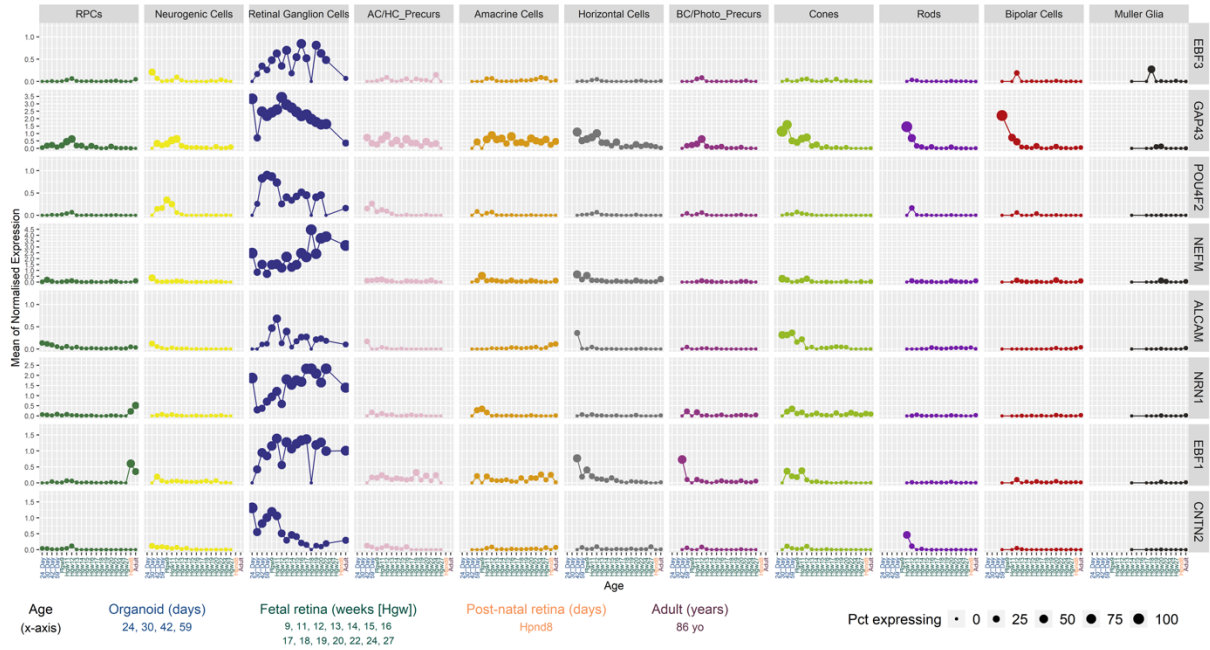
fetal retina Subtype 1 Subtype 2



C: Cone
R: Rod
PanPh: Pan-Photoreceptor
G: Ganglion
M: Müller glia
A: Amacrine
H: Horizontal
B: Bipolar

b

Expression in normal developing retina of ganglion markers upregulated in subtype 2 tumors



Supplementary Figure 4. Retinal cell gene expression analysis in the two subtypes of retinoblastoma.

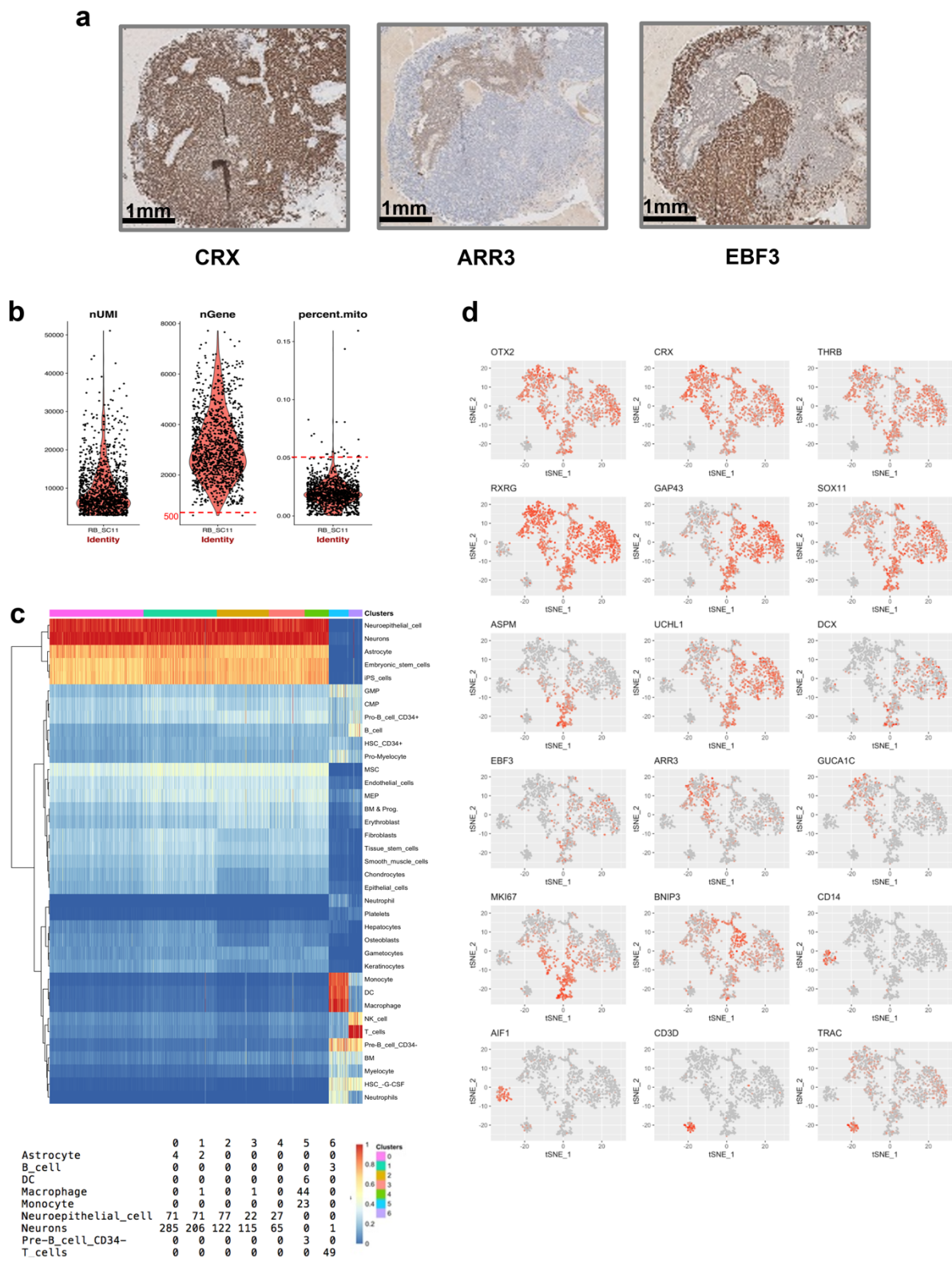
a, The expression of retinal cell-associated genes is presented as a heatmap, for three fetal retinas, 26 subtype 1 tumors and 31 subtype 2 tumors. A heatmap was constructed for each retinal cell type, indicating the level of expression, the statistical significance (Signif.) and the log2 fold change (log2FC) of expression between subtype 2 and subtype 1 tumors. It is also indicated whether the gene is expressed by more than one retinal cell type (Cell exp.). The complete list of markers is given in Supplementary Table 3. Limma moderated t-test was used for the analysis of gene expression, BH correction was applied, exact p values are provided in Supplementary Table 3.

b, Expression of ganglion markers, displayed in Figure 3e, in the normal developing retina according to cell types (single-cell RNA-seq data from Lu et al., 202038). For each gene and each cell type, a pseudo-dot plot is provided. At each age (x-axis), the dot size is proportional to the percentage of expressing cells (i.e. non-zero counts) and its y-coordinate indicates the mean expression.

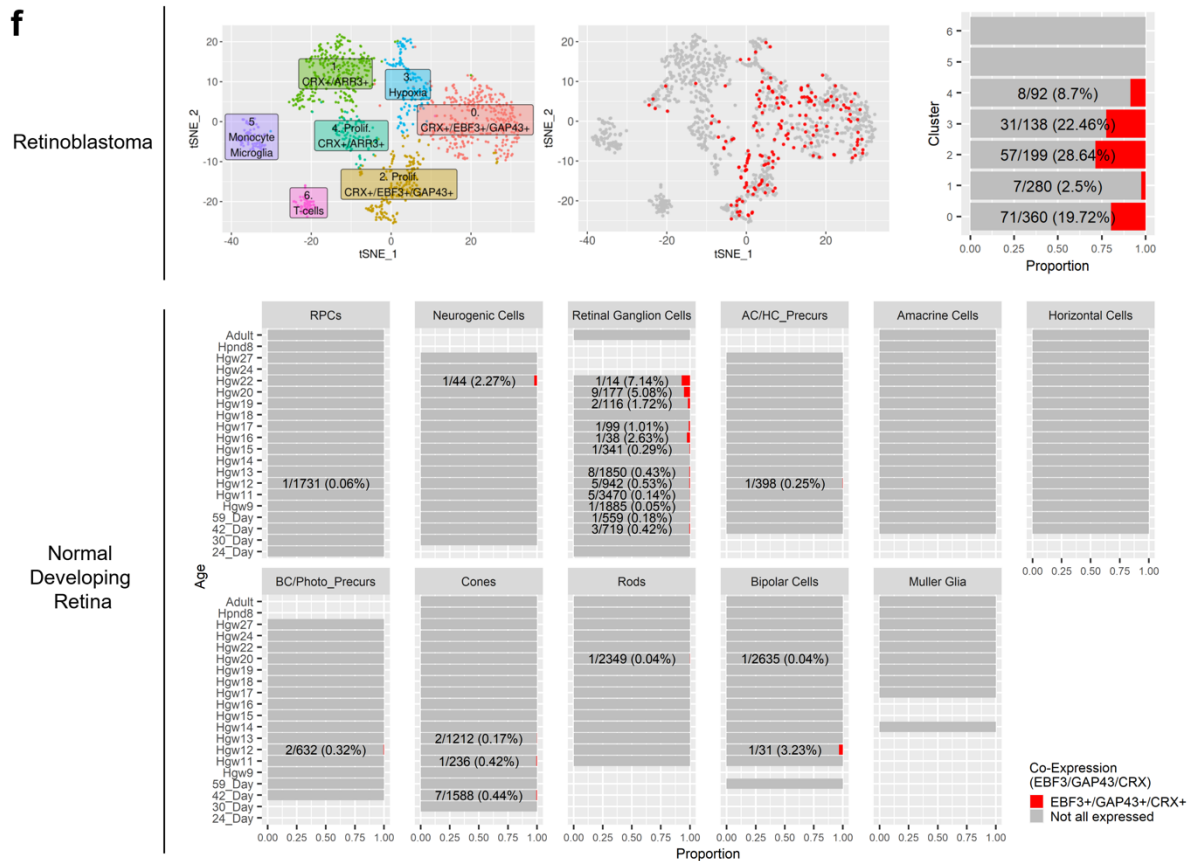
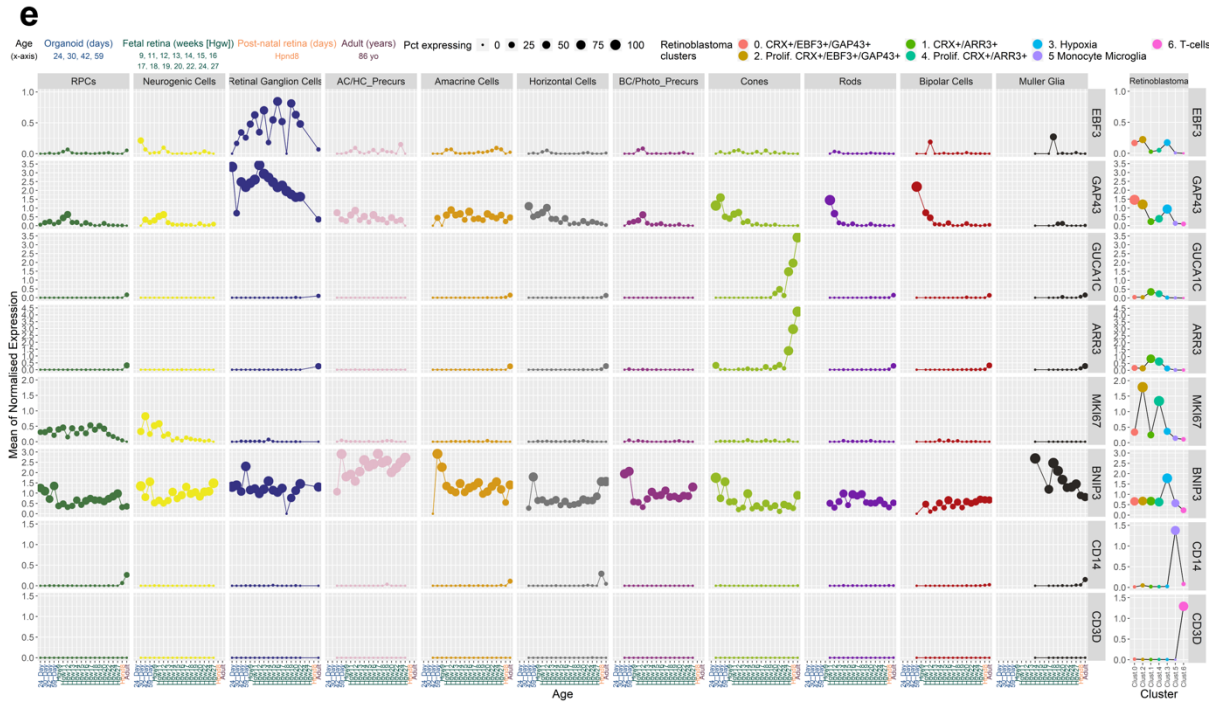
For the markers of all cell types, the online tool can be used to visualize their expression:

- cone: <https://retinoblastoma-retina-markers.curie.fr/cone>,
- rod: <https://retinoblastoma-retina-markers.curie.fr/rod>,
- pan-photoreceptor: <https://retinoblastoma-retina-markers.curie.fr/pan-photoreceptor>,
- ganglion: <https://retinoblastoma-retina-markers.curie.fr/ganglion>,
- muller glia: <https://retinoblastoma-retina-markers.curie.fr/muller>,
- amacrine: <https://retinoblastoma-retina-markers.curie.fr/amacrine>,
- horizontal: <https://retinoblastoma-retina-markers.curie.fr/horizontal>,
- bipolar: <https://retinoblastoma-retina-markers.curie.fr/bipolar>

CHAPTER 3



CHAPTER 3



Supplementary Figure 5. Data related to single-cell RNA-seq of a subtype 2 retinoblastoma (RBSC11).

a, Immunohistochemical staining of CRX (photoreceptor marker), ARR3 (late cone marker) and EBF3 (ganglion cell marker) for a new case of retinoblastoma (RBSC11).

CHAPTER 3

b, Quality control of single-cell analysis for RBSC11: UMI counts, number of genes and percentage of mitochondrial genes.

c, Heatmap of SingleR annotation scores derived by reference to the HPCA dataset with clusters superimposed for 1198 single cells in RBSC11.

d, Expression of selected retinal/neuronal genes in RBSC11 shown in 2D t-SNE plots: early photoreceptor markers (OTX2, CRX, THRB, RXRG); late cone markers (ARR3, GUCA1C); ganglion/neuronal cell markers (GAP43, SOX11, UCHL1, DCX, EBF3).

e, Expression of selected genes in the normal developing retina according to cell types (single-cell RNA-seq data from Lu et al., 202038) (left panels). For each gene and each cell type, a pseudo-dot plot is provided. At each age (x-axis), the dot size is proportional to the percentage of expressing cells (i.e. non-zero counts) and its y-coordinate indicates the mean expression. Expression of these genes in the seven cell clusters of retinoblastoma RBSC11 (right panels). For each cluster, the dot size is proportional to the percentage of expressing cells (i.e. non-zero count) and its y-coordinate indicates the mean expression.

The selected genes are representative of the different cell clusters found in tumor RBSC11.

Clusters 0 and 2, expressed early photoreceptor/cone markers (e.g. CRX) and neuronal/ganglion cell markers (e.g. EBF3, GAP43).

Clusters 1 and 4, expressed early photoreceptor/cone markers (e.g. CRX) and late cone markers (e.g. ARR3).

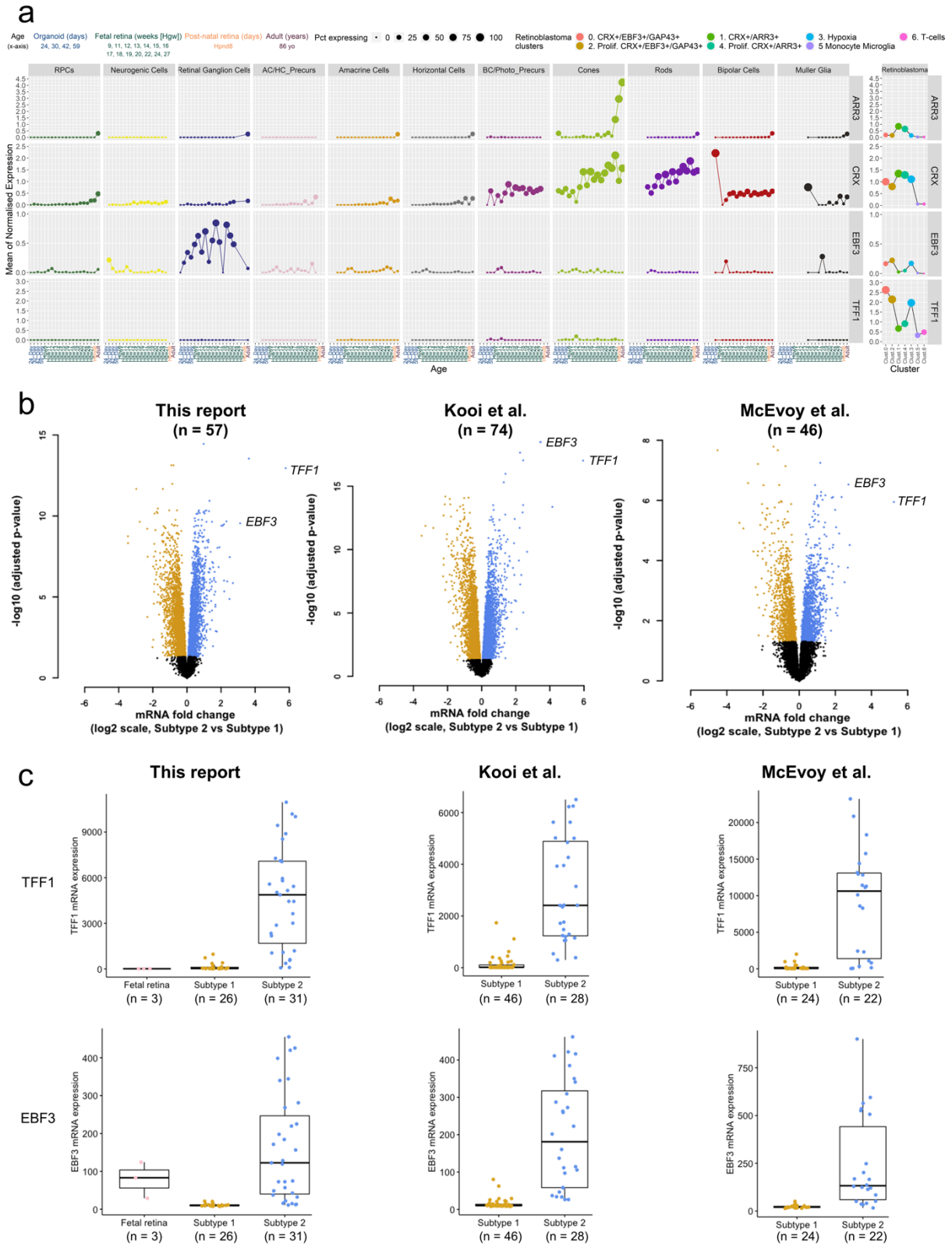
Clusters 2 and 4 correspond to G2/M cells (expressing MKI67)

Cluster 3 corresponds to hypoxic cells (of both tumor cell populations) and expressed BNIP3.

Clusters 5 and 6 correspond to normal cells, macrophage/microglia (cluster 5 expressing CD14), T-lymphocytes (cluster 6 expressing CD3D).

f, Co-expression of CRX/EBF3/GAP43 in the retinoblastoma sample (RBSC11) (upper panels) and the normal developing retina (lower panels). The bar plots represent the abundance of the co-expression pattern. When the number of cells displaying co-expression is not zero, the proportion and absolute number of co-expressing cells are displayed. For the retinoblastoma sample, cells co-expressing the three genes (CRX/EBF3/GAP43) are shown in 2D t-SNE plots. The plots can be retrieved from: <https://retinoblastoma-retina-markers.curie.fr/coexp-ExtDat>.

CHAPTER 3



Supplementary Figure 6. Expression of TFF1 and EBF3 in the normal developing retina and in the two subtypes of retinoblastoma.

a, Expression of TFF1 and EBF3 and of two cone photoreceptor genes (CRX, ARR3) in the normal developing retina according to cell types (single cell RNA-seq data from Lu et al., 202038) (left panels). For each gene and each cell type, a pseudo-dot plot is provided. At each age (x-axis), the dot size is

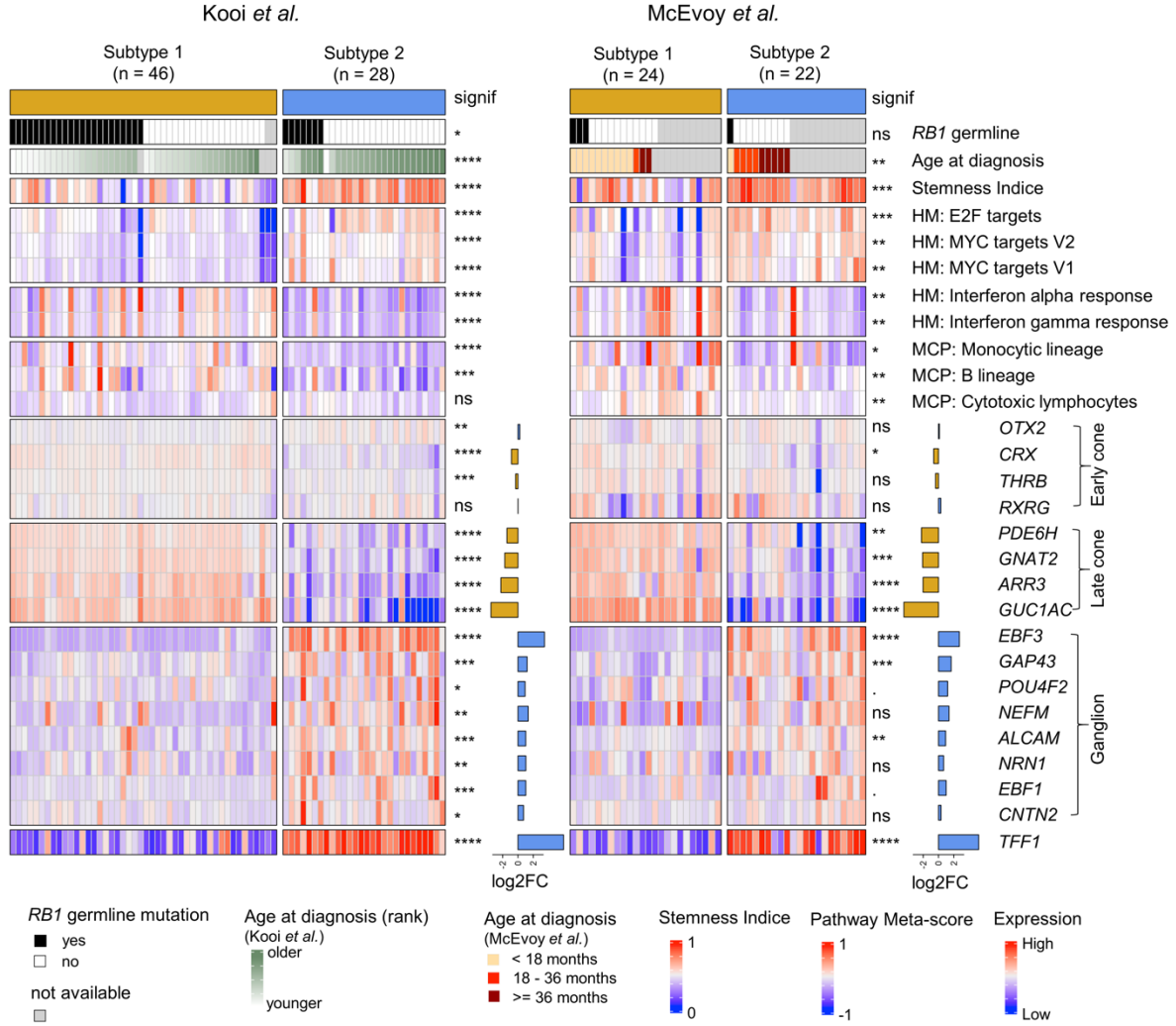
CHAPTER 3

proportional to the percentage of expressing cells (i.e. non-zero counts) and its y-coordinate indicates the mean expression. TFF1 is not expressed in the normal developing retina. Expression of these genes in the seven cell clusters of retinoblastoma RBSC11 (right panels). For each cluster, the dot size is proportional to the percentage of expressing cells (i.e. non-zero count) and its y-coordinate indicates the mean expression.

b, Volcano plots showing that TFF1 and EBF3 are among the most significantly upregulated genes in subtype 2 retinoblastoma (blue) compared to subtype 1 (gold) in our series and in two independent transcriptomic retinoblastoma datasets (Kooi et al.18, McEvoy et al.16). For the subtype assignment of the tumors of these two series, see Methods and Supplementary Fig. 7.

c, Boxplots representing the expression of TFF1 and EBF3 in the two subtypes using three independent datasets (this report, Kooi et al.18, McEvoy et al.16). In the boxplots, the central mark indicates the median and the bottom and top edges of the box the 25th and 75th percentiles. Whiskers are the smaller of 1.5 times the interquartile range or the length of the 25th percentiles to the smallest data point or the 75th percentiles to the largest data point. Data points outside the whiskers are outliers.

CHAPTER 3



		Kooi et al.			McEvoy et al.		
		Subtype 1 (n = 46)	Subtype 2 (n = 28)	p value	Subtype 1 (n = 24)	Subtype 2 (n = 22)	p value
RB1 germline mutation	yes	23	7	0.02837 ^a	3	1	0.6146 ^a
	no	21	21		11	9	
Age at diagnosis (rank)	median rank	26	56.5	2.19E-06 ^b			
Age at diagnosis	< 18 months				10	1	0.006 ^c
	18 - 36 months				1	4	
	>= 36 months				2	5	

Note: a: Fisher's exact test, b Wilcoxon test, c: Chi-square test

Supplementary Figure 7. The two retinoblastoma subtypes were different in clinical and molecular features in two additional independent datasets.

The two retinoblastoma subtypes were identified in two additional independent retinoblastoma datasets (Kooi et al. 201518, McEvoy et al. 201116) using our centroid-based transcriptomic predictor (see Methods). For each dataset, clinical (RB1 germline mutation status, age at diagnosis) and molecular (stemness, E2F targets, MYC pathways, interferon responses, estimation of abundance of various immune cells) features, and the expression of cone and ganglion markers, and of TFF1 were compared between the two subtypes. The features characteristic of each subtype identified in our initial series

CHAPTER 3

were found in these two independent series. For germline mutation status and age, the statistical tests used to evaluate the difference between the two subtypes are indicated in the Wilcoxon test was used to evaluate the differences in HALLMARK pathway meta-scores and MCP counter-estimated immune cell abundance between the two subtypes of retinoblastoma; Limma moderated t-test was used for the analysis of gene expression, significance based on adjusted p values is shown. log2 fold-changes in expression between subtype 2 and subtype 1 are also shown. $p \geq 0.1$ (ns), $p < 0.1$ (.), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****)

COMMENTS

Identification and characterization of cancer subtypes brings us knowledge on how tumors are formed and helps us to identify pertinent treatments for different disease subtypes. Previously, clustering on three different transcriptomic studies were performed and reached controversial conclusions: some researchers showed retinoblastoma is rather homogeneous while others classified patient tumors into two subgroups. A third group showed that the two clusters identified by hierarchical clustering displayed gradients of photoreceptor signature expression and tumor progression (Figure 7A). We took the data from the third group and observed a bimodal distribution for the mean expression of their photoreceptor signatures from histogram (Figure 7B).

In this chapter, we identified two molecular subtypes of retinoblastoma with different clinical-pathological features based on a multi-omic analysis of 102 retinoblastoma samples. We demonstrated that subtype 2 tumors were clinically associated with later onsets and metastasis. They exhibited higher levels of stemness and MYC pathway activation, as well as lower levels of immune and inflammation signatures. They were more heterogeneous than subtype 1 tumors; they exhibited less differentiated cone states and expressed higher levels of neuronal/ganglion signatures. They had higher genomic instability and lower levels of DNA methylation, but were associated with higher methylation levels in CpG Islands. Those molecular features give us hints on potential therapeutic targets for this subtype of disease.

Single-cell transcriptomic analysis of one retinoblastoma revealed its heterogeneity was at both genetic and phenotypic levels. These results allowed us to propose two progression paths for the two parts of tumor cells. In this tumor, less differentiated regions did not evolve from more differentiated parts as they harbored different genomic alterations. But we could not rule out the possibility of a dedifferentiation, as we are going to see in the next chapter, reanalysis of this tumor uncovered that dedifferentiation did happen in some subsets of the retinoblastoma cells with accumulation of genomic alterations.

CHAPTER 3

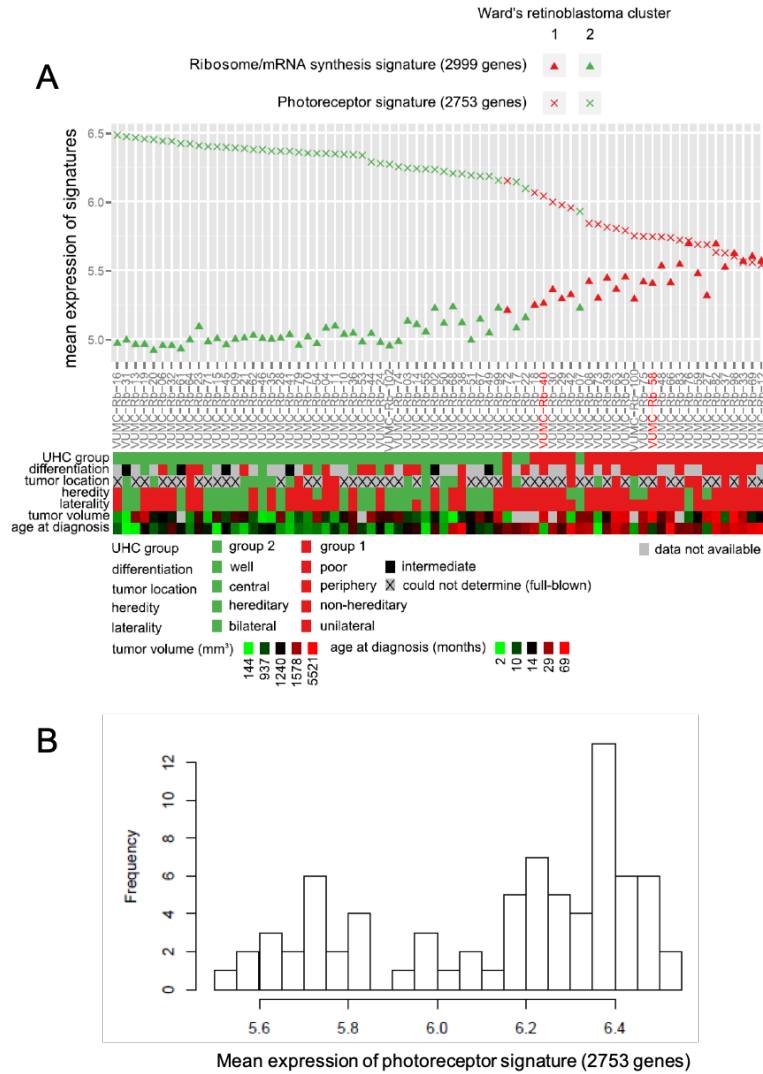


Figure 7. Two visualization of photoreceptor signature expression in Kooi *et al.*'s data.

A, Kooi *et al.* illustrated that the mean expression of photoreceptor signature (2753 genes) were continuous in two subtypes of retinoblastoma in 72 retinoblastoma samples. Figure from Kooi *et al.* 2015, EBioMedicine.

B, A histogram showing the distribution of mean expression of photoreceptor signature (the same 2753 genes) in the same 72 patient samples from Kooi *et al.*'s data.

CHAPTER 4.

CHARACTERIZATION OF HETEROGENEITY IN RETINOBLASTOMA BY SINGLE-CELL RNA-SEQUENCING

This chapter shows some unpublished single-cell data on the characterization of retinoblastoma intra- and inter- heterogeneity, and immune microenvironment .

INTRODUCTION

Cancer is a heterogeneous disease. Tumor cells in different patients or different cells in one patient can exhibit distinct molecular and phenotypic characteristics that leads to varied responses to the treatment. Heterogeneity is also reported in retinoblastoma, a cancer of the retina in children. At genetic level, the majority of the retinoblastomas develop following biallelic inactivation of *RB1* genes (1). In rare cases (<2% of non-familial diseases) when *RB1* genes are intact, *MYCN* amplifications may give rise to retinoblastoma (2). At pathological level, tumors exhibit varied growth patterns: endophytic, exophytic and mixed. At histological level, diversity exists in tumor differentiation: some tumors present with Flexner-Wintersteiner rosettes and some with fleurettes, and the degree of differentiation seemed correlate with age (3). Molecular subtypes are also reported in retinoblastoma (4–6). A multi-omics analysis on 102 tumors identified two subtypes, one with more matured cone signatures and the other high-risk subtype of retinoblastoma with higher level of genomic instability, stemness, dedifferentiated cone states and neuronal/ganglion features (7).

Single-cell transcriptomic analysis is an emerging approach that allows characterization of tumor heterogeneity at the individual cell level. Our previous single-cell transcriptomic analysis in one retinoblastoma showed remarkable intra-tumoral heterogeneity at both genomic and transcriptomic levels where some cells showed a less differentiated cone state with more and different copy number alterations from other tumor cells (Figure 1) (7). A study using single-cell RNA-seq and single-cell ATAC-seq on two retinoblastomas showed retinoblastoma cells presented two cell programs that were activated by different regulators (8). The trajectory analysis in

another study of single-cell RNA-seq with two retinoblastoma suggested that tumor cells may undergo cone dedifferentiation (9).

Tumors are heterogeneous also in a sense that they are composed of different types of cells, including tumor cells, as well as vessels, fibroblasts, and immune cells. Those cells exist in tumor microenvironment may interact with each other and lead to evasion of immune surveillance and tumor progression. Studies based on immunohistochemistry and immunoblotting on immune checkpoint proteins were carried out and revealed an immune-suppressive microenvironment in retinoblastoma (10–12). Our previous analysis in 23 patient samples with immunohistochemistry and cytokine array pointed out that secretion of cytokines such as MIF and EMMPRIN from the tumor cells may promote the infiltration of protumoral M2 macrophages in retinoblastoma (13).

In the present study, we characterized the intra-tumoral heterogeneity at a higher resolution in one retinoblastoma that we previous described. We performed single-cell RNA-sequencing in 14 patient samples (13 samples in addition to RBSC11) and validated that tumors from younger patients expressed more matured cone photoreceptor markers, while tumors from older patients expressed less differentiated cone markers and neuronal genes. We also describe, for the first time, the immune microenvironment of retinoblastoma using the single-cell transcriptome. We show that the protumoral alternative M2 macrophages were enriched in retinoblastoma microenvironment and several candidate genes and inferred several tumor-immune interactions that may lead to the polarization of M2 macrophages.

RESULTS

Intra-tumoral heterogeneity at a higher resolution: reanalyzing a retinoblastoma patient sample by single-cell RNA-sequencing

Single-cell analysis in one retinoblastoma showed remarkable intra-tumoral heterogeneity at both genomic and transcriptomic levels, in which some cells with chromosome 2p and 10q gains were more matured in terms of cone differentiation, and other cells, harboring more and different copy number alterations, showed a less-differentiated cone state (7) (Figure 1). We reanalyzed the sample aliased as RBSC11 and revealed that, in the cells expressing more matured cone signatures, a dedifferentiation in cone states was associated with gain of additional chromosome alterations.

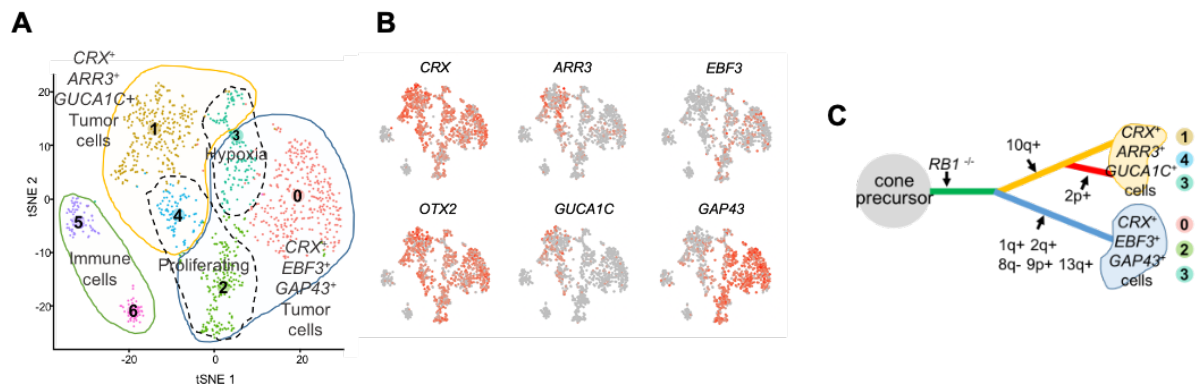


Figure 1. Single-cell transcriptome analysis of RBSC11 revealed intra-tumoral heterogeneity.

A, 2D-t-SNE plot showing 1198 retinoblastoma cells from one patient. Each dot represents one cell. Cells were clustered and annotated based on their transcriptomic profiles. Tumor cells represented 90% of all cells and were consisted of two main populations, *CRX*⁺*ARR3*⁺*GUCA1C*⁺ cells and *CRX*⁺*EBF3*⁺*GAP43*⁺ cells, each with a group of cells proliferating.

B, Expression of cone and neuronal/ganglion markers in 2D T-SNE plot. Early photoreceptor markers: *CRX* and *OTX2*; late cone markers: *ARR3* and *GUCA1C*; ganglion and neuronal markers: *EBF3* and *GAP43*.

C, A proposed tumor progression model for RBSC11 based on the inferred copy number variation profiles of the two cell populations. (From Figure 5 of (7))

More cells were identified using CellRanger V3 with EmptyDrops (14) cell calling algorithm implemented. We obtained 1274 cells after applying the same quality control

filter used in the previous analysis, which were visualized in 2D-UMAP (Uniform Manifold Approximation and Projection) (Figure 2A). Cells identified in the new analysis didn't have significant difference in the quality control metrics with previous analysis (Figure 2B).

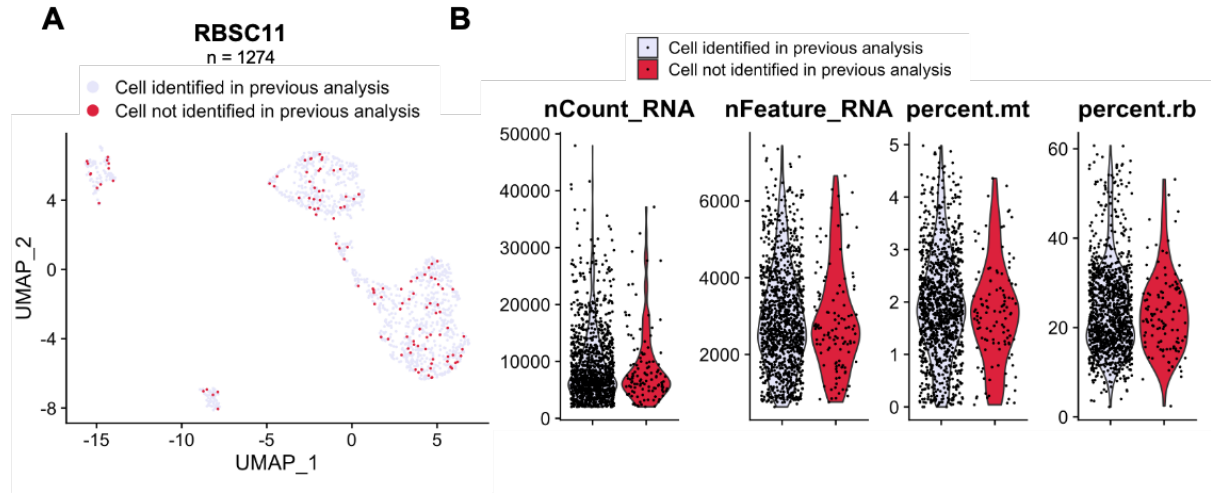


Figure 2. Single cells identified with new analysis pipeline.

A, 2D-UMAP plot showing 1274 cells that were identified with the CellRanger V3. Cells that were newly identified and not detected in previous analysis were shown in red.

B, Quality control metrics of the cells identified in both versions of analysis and only identified in the new analysis.

Expression data were normalized and scaled with the effect of number of counts and difference between cell cycle S phase and G2M score regressed out. Cells were clustered using an improved graph-based method implemented in Seurat V3. Briefly, a K-nearest neighbor graph was constructed based on the Euclidean distance in PCA space, the edge weights between any two cells were refined with Jaccard index based shared overlaps; Louvain algorithm was then applied to iteratively group cells together (15). Nine clusters were identified by the new analysis (Figure 3A). Clusters were annotated based on cluster markers, which were the genes upregulated in the cluster as compared to all other clusters. Out of the nine clusters, seven clusters (1114 cells, 87.5%) were annotated as tumor cells, one cluster (92 cells, 7.2%) was annotated as monocytic lineage (monocytes or macrophages or microglia) and one cluster (68 cells, 5.3%) was annotated as T cells. Within the tumor cells, clusters were further annotated

based on known retinoblastoma genes, retina cell type markers, cell cycle phase and apoptotic status.

All tumor cell clusters expressed early photoreceptor and early cone markers such as *OTX2*, *CRX*, *THRB* and *RXRG* (Figure 3A, B) as we previously described (7). Two clusters expressed cone markers at later stages (e.g. *ARR3*, *GUCA1C*). Interestingly, the new clustering could separate the two clusters by cone maturity, one cluster expressed *ARR3*, a maturing cone marker and the other expressed *GUCA1C*, a cone marker at a more mature stage. The other five clusters expressed retinoblastoma genes (e.g. *TFF1*) and neuronal/ganglion markers (e.g. *EBF3*, *GAP43*, *SOX11*, *UCHL1*, *DCX*) that were identified in our previous analysis. Interestingly, a new cluster of 37 cells expressing *PRL* and *CD24* was identified in the new analysis. The gene *PRL* encodes a secreted hormone prolactin, which acts primarily on the mammary gland by promoting lactation, but is also a trophic factor in the retina (17). It is expressed throughout the retina layers in mouse and monkey (16). The gene *CD24* encodes a sialoglycoprotein that plays a role in cell differentiation, and is reported as a neuronal stemness marker (18). This cluster was not identified in the previous analysis and was submerged within the apoptotic cells (Figure 3C). The proliferation marker *MKI67* was expressed by both the cells expressing cone markers and the cells expressing neuronal/ganglion markers. Its expression is high in the cluster of cells predicted as G2M phase (Figure 3A,B). The gene *BNIP3* encodes a pro-apoptotic protein of the BCL2 family, and was reported involved in hypoxia-induced apoptosis (19). Consistent with this, the cluster of cells expressing *BNIP3* also expressed genes enriched in hypoxia pathways. The expression of monocyte marker (e.g. *CD14*), macrophage marker (e.g. *CD163*), microglia marker (*AIF1*), and T cell marker (e.g. *CD3D*) are shown in the UMAP space (Figure 3B).

CHAPTER 4

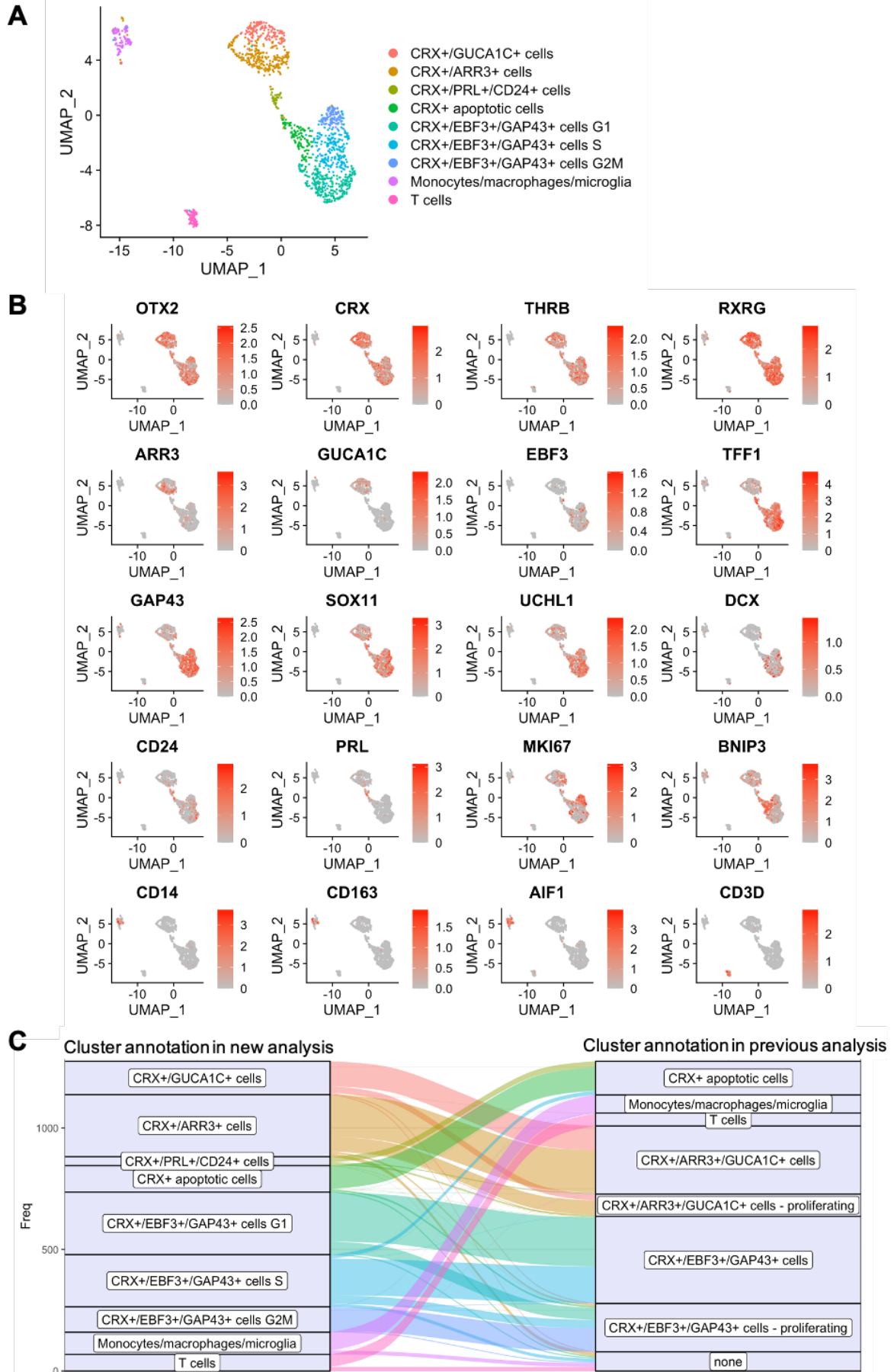


Figure 3. New annotation of RBSC11.

A, 2D-UMAP plot showing 1274 cells colored by annotated clusters.

B, Expression of selected genes are shown in a 2D-UMAP plot. Early photoreceptor markers: *CRX*, *OTX2*; early cone markers: *THRB*, *RXRG*; late cone markers: *ARR3*, *GUCA1C*; aberrant retinoblastoma gene: *TFF1*; neuronal/ganglion markers: *EBF3*, *GAP43*, *SOX11*, *UCHL1*, *DCX*; top gene in the newly identified cluster: *PRL*, *CD24*; proliferation marker: *MKI67*; pro-apoptotic marker: *BNIP3*; monocyte marker: *CD14*; macrophage marker: *CD163*; microglia marker: *AIF1*; T-cell marker: *CD3D*.

C, Alluvial plot showing the cell annotation assignment of the same cells in the current and previous analysis. Cells annotated as none in the previous analysis are the cells newly identified with the currently pipeline.

We previously proposed a tumor progression model for RBSC11 based on the inferred genomic profiles (Figure 1C). *CRX*⁺/*ARR3*⁺/*GUCA1C*⁺ cells were transformed after the two hits of *RB1*, followed by chromosome 10q gain and in some cells 2p gain. We weren't able to distinguish, in their transcriptomic profiles, the cells with 10q gain only from the cells with both 10q gain and 2p gain. In the new analysis, we were able to separate the *CRX*⁺/*ARR3*⁺/*GUCA1C*⁺ cells into two clusters, one expressing maturing cone cell marker *ARR3* and the other expressing more mature cone marker *GUCA1C*. To examine if the two clusters by cone maturity corresponded to the two groups of cells with different genomic profiles, we inferred copy number variations at single-cell level from gene expression, using inferCNV (20). Indeed, the *CRX*⁺/*GUCA1C*⁺ cells were inferred as carrying 10q gain without 2p gain; and the *CRX*⁺/*ARR3*⁺ cells were inferred as carrying both 10q gain and 2p gain (Figure 4). Interestingly, the cells carrying 10q gain and 2p gain also included the cells in the *PRL*⁺/*CD24*⁺/*TFF1*⁺ cluster (Figure 4).

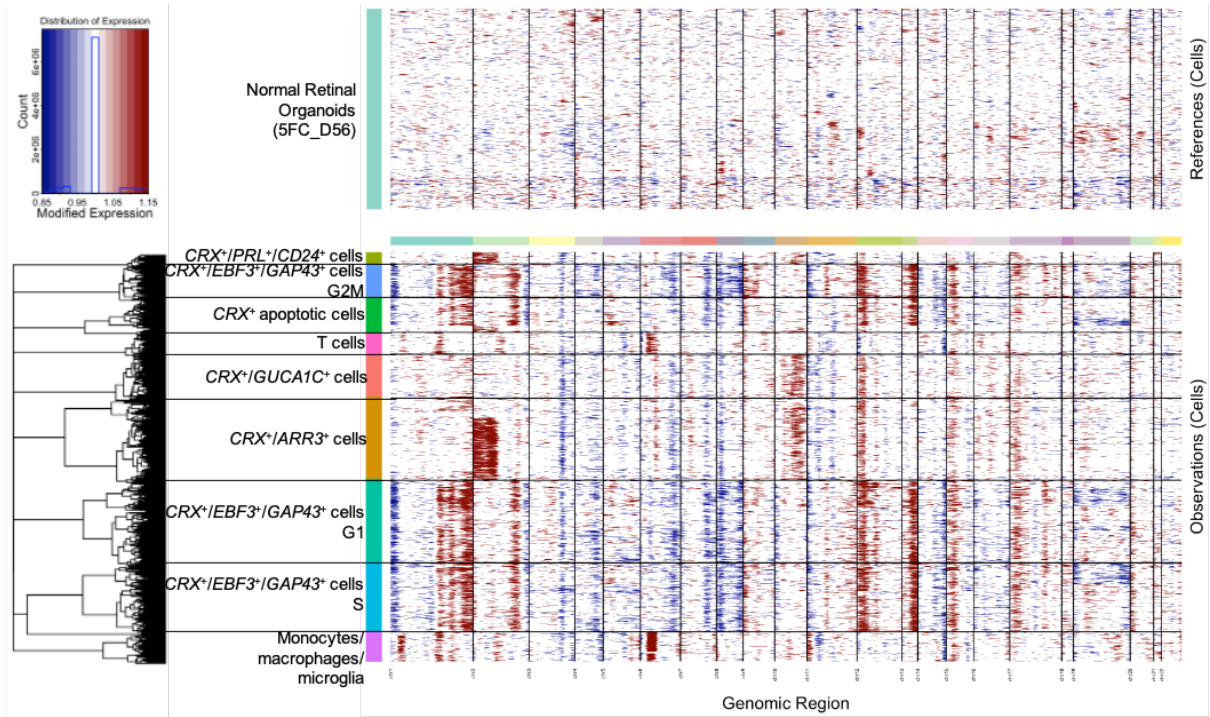


Figure 4. Inferred copy number variation profiles for RBSC11.

Combining the new clustering results and the inferred copy number profiles, we could propose a new cell progression model (Figure 5). After bi-allelic inactivation of *RB1* and the gain of 10q, *CRX*⁺*GUCA1C*⁺ cone cells were transformed into tumor cells. Addition of 2p gain induced tumor cells into less differentiated cone states that no longer expressed the mature marker *GUCA1C* but still expressed a maturing cone marker *ARR3*. Cells may further lose the *ARR3* and become *CRX*/*PRL*/*CD24*/*TFF1* positive. On the other hand, bi-allelic inactivation of *RB1* combined with numerous genomic alterations could lead to cells with neuronal features that expressing *CRX*, *EBF3*, *GAP43*, and *TFF1*. All those cells seemed to proliferate, as they expressed proliferation marker *MKI67* (Figure 3B).

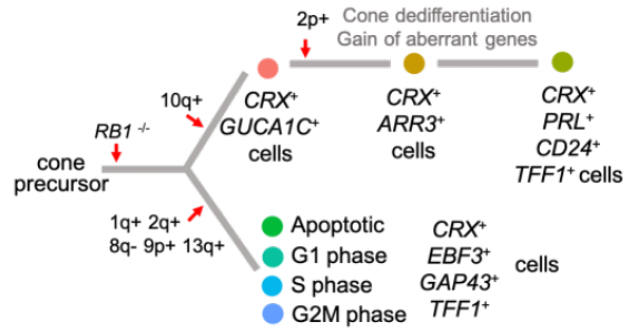


Figure 5. Schema of a proposed progression model for RBSC11.

Based on the new annotation and the inferred copy number variation profiles of the clusters identified in the new analysis, we proposed a refined progression model for RBSC11. Part of the cells came through *RB1* loss and chromosome 10q gain and became *CRX*⁺*GUCA1C*⁺ tumor cells. With gain of 2p in those tumor cells, they lost mature cone gene and expressed aberrant tumor genes, became *CRX*⁺*ARR3*⁺ and *CRX*⁺*PRL*⁺*CD24*⁺*TFF1*⁺. Other part of the cells lost *RB1* and acquired 1q, 2q, 9p, 13q gain, 8q loss and became *CRX*⁺*EBF3*⁺*GAP43*⁺*TFF1*⁺ tumor cells. Those tumor cells can be in different cell cycle states or showing apoptotic and hypoxic features.

Our results support the hypothesis that retinoblastomas originate from cone precursors, as all cells expressed early cone precursor markers. Singh *et al.* showed that cultured human *ARR3*⁺ maturing cone cells enter the cell cycle and form pre-retinoblastoma-like lesions when *RB1* is depleted by shRNAs, but they exit the cell cycle in the following months (39). Our results were in accordance that tumor cells can express maturing cone marker *ARR3* or more matured cone marker *GUCA1C*, but their transformation may require not only *RB1* mutations but also additional genomic alterations. In some of those tumor cells, dedifferentiation in cone states and gain of aberrant tumor genes may occur with the accumulation of chromosome 2p gain. At the stage of diagnosis, no tumor cells obtained absolute growth advantage, therefore the three populations coexisted. Other tumor cells expressing *CRX* and neuronal/ganglion markers *EBF3*/*GAP43* may be derived from less differentiated cone photoreceptor precursors expressing *CRX* but not *ARR3*, or they may be also derived from maturing cone precursors but underwent cone dedifferentiation, but in this case, *CRX*⁺*EBF3*⁺*GAP43*⁺*TFF1*⁺ cells outgrew the other tumor cells.

Our results was based on single-cell transcriptomic analysis of one retinoblastoma, it would be important to examine the expression of early and late cone photoreceptor markers, neuronal/ganglion markers, and aberrantly expressed genes in

retinoblastoma using immunohistochemistry and to verify the copy number variation profiles in macro-dissected tumor cells with different phenotype. Single-cell RNA-sequencing is applied to fresh tumors and viable cell enrichment may eliminate some fragile cells; therefore the samples that can be analyzed are limited and the cell population analyzed maybe biased. Spatial transcriptomics that can be applied to formalin-fixed, paraffin-embedded (FFPE) tissues could allow us to select and analyze the tumor samples with intra-tumoral heterogeneity, thus provide more interesting results.

Inter-tumoral heterogeneity confirmed at single-cell resolution

We extended the droplet-based single-cell 3' RNA-sequencing for 14 primary retinoblastomas. Retinoblastoma samples were needle-aspirated from primary enucleation in patients with unilateral and non-hereditary diseases. Age of diagnosis for patients were between 9 to 63 months and median age of diagnosis was 20 months. Patient characteristics were provided in Table 1.

We obtained 39297 cells from the 14 retinoblastomas after quality control, which were visualized in 2D-UMAP (Uniform Manifold Approximation and Projection) (Figure 6A). Cells were clustered using graph-based methods using Seurat V3 (Figure 6B). Pathway enrichment analysis of the cluster markers enabled us to annotate the cells into tumor cells (n = 31236, 79.5%), Monocytes (n = 4828, 12%), T cells (n = 2211, 6%), Rod photoreceptors (n = 579, 1.5%) and Müller glia (n = 425, 1%) (Figure 6C). Expression of the most upregulated genes in each cell type were shown in heatmap (Figure 6D).

All retinoblastoma cells expressed *CRX* and *OTX2* (Figure 6E), two early photoreceptor markers. Later cone markers *ARR3* and *GUCA1C*, that we previously reported expressed by subtype 1 retinoblastoma, were expressed by some retinoblastoma cells from patients younger than 18 months (Figure 6C,E). Ganglion/neuronal markers *EBF3*, *GAP43* and *DCX*, and *TFF1*, the genes expressed by subtype 2 retinoblastoma, were expressed by some retinoblastoma cells from patients older than 18 months (Figure 6C,E). The findings that tumor cells from younger patients expressed markers of cone precursors and more mature cone photoreceptors and that tumor cells from older patients expressed cone precursor markers and neuronal/ganglion markers were in agreement with the characteristics of the two subtypes that we previously reported. The two genes *PRL* and *CD24* were also expressed by some retinoblastoma cells from patients older than 18 months (Figure 6C,E).

CHAPTER 4

Table 1. Patient Characteristics.

sampleID	Gender (Male /Female)	Age at enucleation (month)	Laterality (Unilateral/ Bilateral)	Hereditary (Y/N)	Prior treatment	IIRC class	TNM stage	Tumor diameter (mm, by MRI)	Differentiation (Well /Intermediate /Poor)	Type of growth (Exophytic /Endophytic /Mixed)	Optic nerve invasion (None /Prelaminar /Intralaminar /Postlaminar)	Choroid and sclera invasion (None /Minimal /Deep /Extende /Extended and Sclera)
RBSC11	Female	23	Unilateral	N	no	E	cT2b	na	Intermediate	na	Prelaminar	Extended and Sclera
RBSC14	Male	28	Unilateral	N	no	E	cT2b	17	na	Endophytic	None	None
RBSC15	Male	27	Unilateral	N	no	D	cT2b	15	Intermediate	Endophytic	na	None
RBSC16	Female	22	Unilateral	N	no	E	cT3d	na	Poor	Mixed	Postlaminar	Minimal
RBSC17	Male	10	Unilateral	N	no	E	cT2a	17	Intermediate	Exophytic	Prelaminar	Deep
RBSC19	Female	63	Unilateral	N	no	E	cT2b	na	Poor	Exophytic	None	None
RBSC23	Female	17	Unilateral	N	no	E	cT2b	15	Intermediate	Exophytic	None	Minimal
RBSC24	Male	54	Unilateral	N	no	E	cT2b	19	Poor	Exophytic	Intralaminar	None
RBSC26	Female	13	Unilateral	N	no	D	cT2b	na	Intermediate	Endophytic	None	None
RBSC27	Male	20	Unilateral	N	no	E	cT2b	16	Intermediate	Mixed	Prelaminar	Minimal
RBSC30	Female	10	Unilateral	N	no	E	cT2b	17	Intermediate	Endophytic	None	None
RBSC31	Male	11	Unilateral	N	no	E	cT3c	16.5	Intermediate	Endophytic	Prelaminar	Minimal
RBSC32	Female	19	Unilateral	N	no	E	cT2b	na	Intermediate	Endophytic	None	Minimal
RBSC33	Male	9	Unilateral	N	no	E	cT2b	18	Intermediate	Exophytic	Postlaminar	Minimal

na: not available

CHAPTER 4

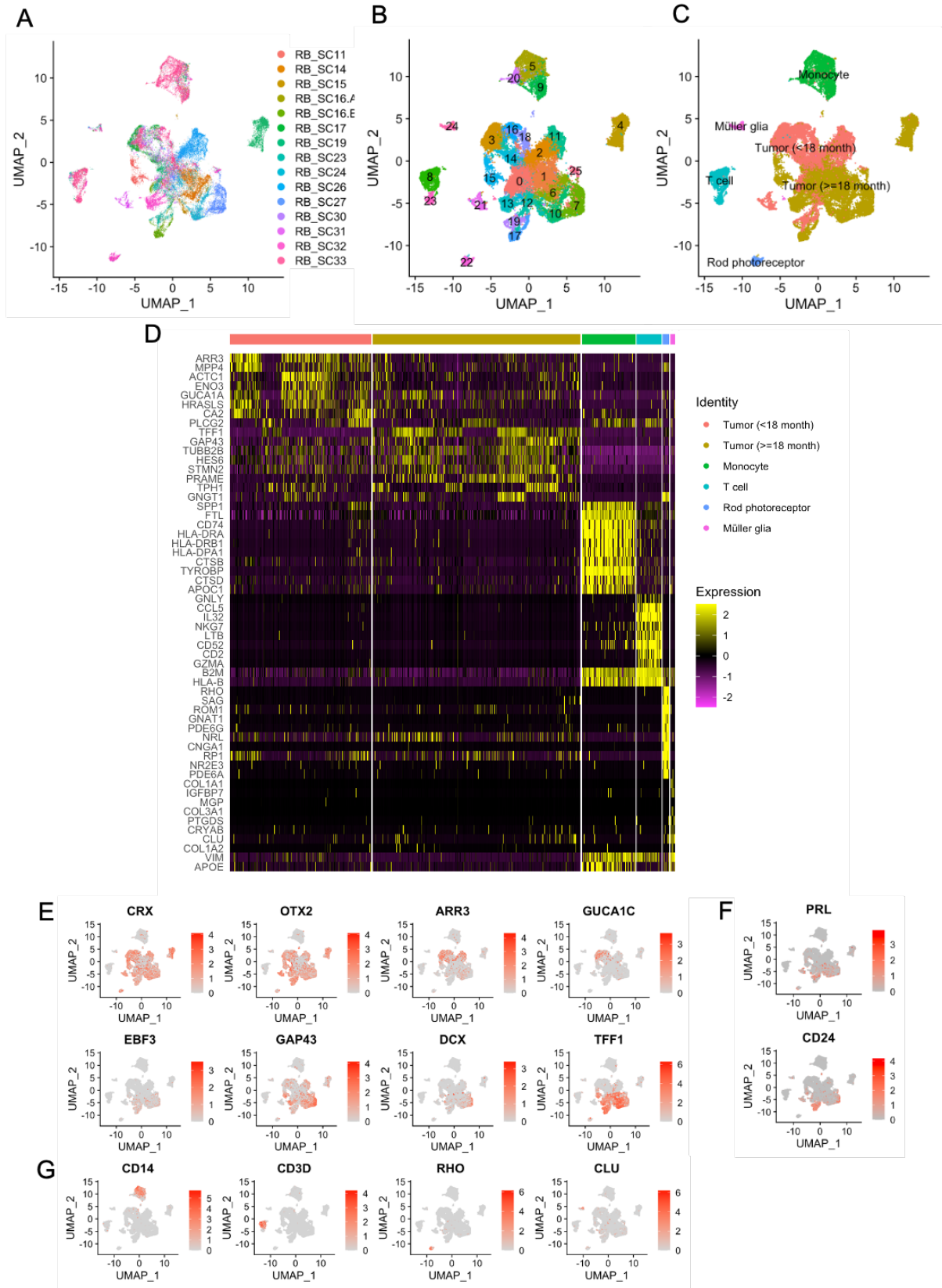


Figure 6. Single-cell transcriptomic analysis of retinoblastoma.

A, 2D-UMAP of 39297 cells from the 14 retinoblastomas colored by sample. Each dot represent a cell.

B, 2D-UMAP annotated by cluster.

C, 2D-UMAP annotated by cell type.

D, Heatmap showing the most upregulated genes in each cell type.

E, Selected genes expressed by tumors shown in 2D-UMAP. Early photoreceptor marker: *CRX*, *OTX2*; Cone photoreceptor markers: *ARR3*, *GUCA1C*; Neuronal/ganglion markers: *EBF3*, *GAP43*, *DCX*; Aberrant expressed gene in retinoblastoma: *TFF1*;

F, Expression of *PRL* and *CD24* in retinoblastoma shown in 2D-UMAP.

G, Representative genes for other cell types shown in 2D-UMAP. Monocyte marker: *CD14*; T cell marker: *CD3D*; Mature rod photoreceptor marker: *RHO*; Muller glia marker: *CLU*.

The distance between two cells in UMAP reflects the similarity of their transcriptomic profiles. Interestingly, tumor cells from younger patients were closer in 2D-UMAP space (Figure 6B) and expressing higher *ARR3* (Figure 6D,E), a maturing cone photoreceptor marker, while tumor cells from older patients were closer to each other (Figure 6B) and highly expressed *TFF1* (Figure 6D,E). We showed previously that RBSC11 expressed *ARR3* in part of the tumor cells with chr10q gain (some cells also had 2p gain), and expressed *EBF3/GAP43* - the neuronal/ganglion markers - in the rest part of the tumor with multiple copy number variations (gains of 1q, 2q, 9p, 13q, loss of 8q) (7). Intriguingly, the *ARR3*⁺ cells of RBSC11 were closer to the tumor cells from younger patients expressing *ARR3* in the 2D-UMAP, and the *EBF3*⁺/*GAP43*⁺ were closer to the tumor cells from older patients (Figure 6B,E). Among the tumor cells, RBSC19 formed a distinct cluster (Figure 6A-C), which might be due to the older age at diagnosis of the patient (63 months), which is very different from the rest of the patients.

We identified monocytes, T cells, mature rod cells and Müller glia cells in the retinoblastoma tumor microenvironment. They weren't identified as tumor cells as they did not harbor the copy number variations from inferCNV. Expression of selected markers for each cell type were shown (Figure 6G). Rod photoreceptors are the most abundant cell types in the retina that mediate vision in dim light (21). Müller glia are the major glial cell types derived from retinal progenitors that provide structure support and participate in maintenance of homeostasis in the retina (22). The rod cells and Müller glia identified in some retinoblastomas are likely the normal cells as previously shown (7,23).

Immune landscape of retinoblastoma

To investigate the immune landscape of retinoblastoma, we focused our analysis on the 7039 cells (18%) that were classified as monocytes or T cells out of 39297 cells from the 14 retinoblastomas. For a more detailed annotation, we reclassified those immune cells into 12 clusters using graph-based clustering (Figure 7A) and performed cell type prediction by label transfer from PBMC reference (Figure 7B) (24). A final annotation (Figure 7C) was made based on both cluster markers and cell type prediction. We identified various subsets of monocytic lineage cells (monocytes, macrophages, microglia, dendritic cells), T cells (CD 4 T, CD8), NK cells, B cells (Figure 7C). A heatmap illustrating the expression of top 10 genes in each cell type is shown (Figure 7D).

CHAPTER 4

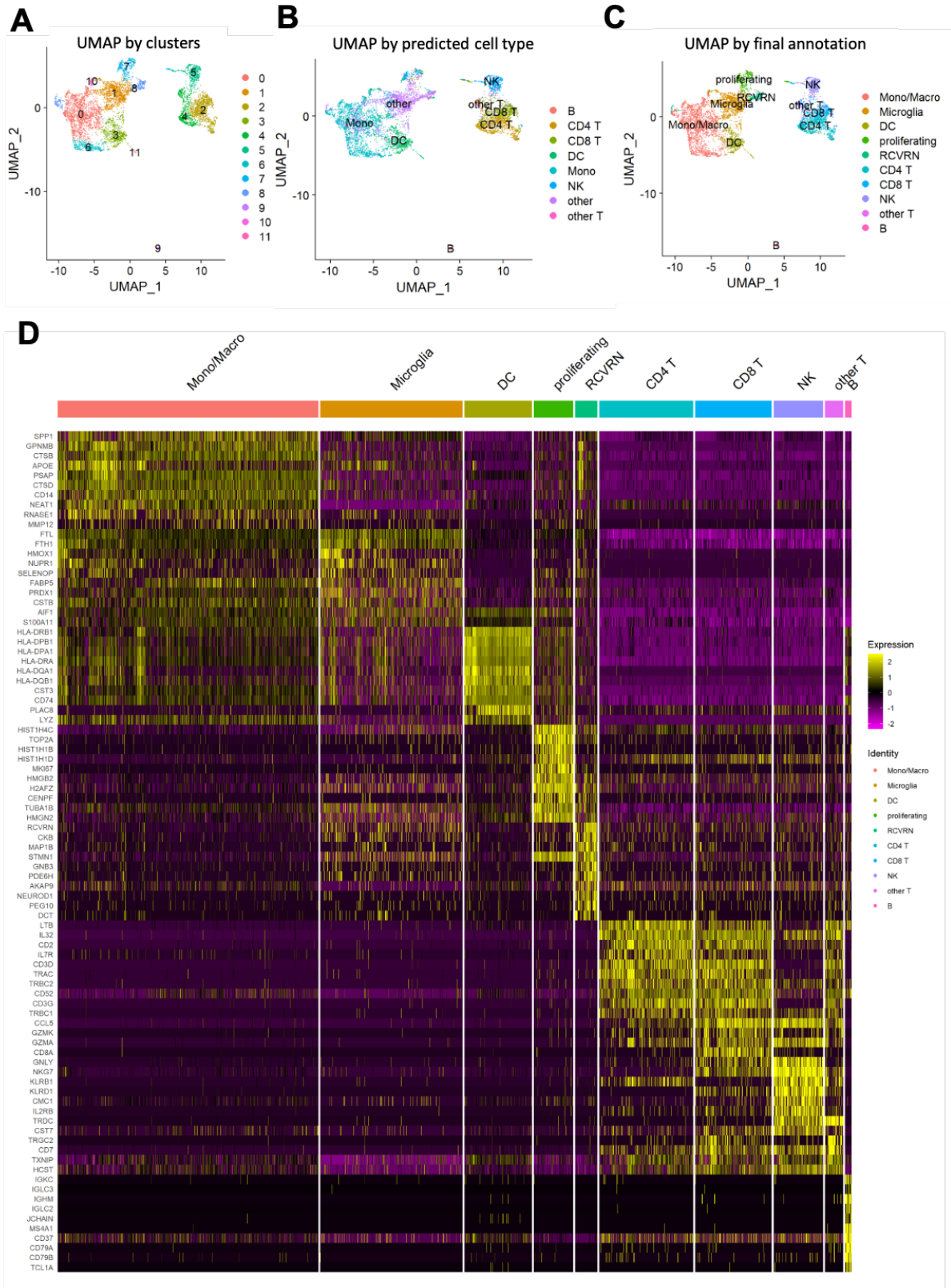


Figure 7. Immune cell type identification in retinoblastoma microenvironment.

A, UMAP colored by clusters. Each dot represent a cell.

B, UMAP colored by cell type prediction from label transfer. Mono, monocytes; DC, Dendritic cells.

C, UMAP colored by final cell type annotation that combined clusters and predicted cell types from label transfer. Mono/Macro, monocytes/macrophages.

D, A heatmap showing the top 10 most upregulated genes in each cell type as compared to other cell types.

Among monocytic lineage cell clusters, five main cell types were identified: monocytes/macrophages (n = 2363), stressed microglia (n = 1291), dendritic cells/microglia (n = 610), proliferating cells (n = 358) and a cluster of cells expressing *RCVRN* (n = 201). Markers of monocyte (e.g. *CD14*) and macrophage (e.g. *CD68*) were mainly expressed by the monocyte/macrophage cluster (Figure 8A). M2 macrophage marker such as *CD163*, *MRC1* were also expressed by the monocytes/macrophages cluster, instead of the markers of M1 microphages such as *CD80*, *NOS2* (Figure 8A). This is consistent with our previous findings that M2 macrophages were activated in the retinoblastoma tumor microenvironment (13). Marker of microglia *AIF1* was expressed by all monocytic lineage cell clusters (Figure 8B, left panel) and was relative higher in three subsets of cells that we annotated as stressed microglia, dendritic cells or proliferating cells (Figure 8B, right panel). The stressed microglia subset exhibited high mitochondria gene percentage and low total count (Figure 8C). They highly expressed genes involved in oxidative phosphorylation, reactive oxygen species pathway, iron and lipid metabolism, such as *FTL*, *FTH1*, *HMOX1*, *NURP1*, *SELENOP*, *FABP5*, *PRDX1*, *CSTB* (Figure 7D). The two genes ferritin light chain (*FTL*) and ferritin heavy chain (*FTH1*) encode the two subunits of the ferritin protein, which functions in the storage and delivery of iron. Both genes were reported upregulated in several cancers and positively correlated with tumor-associated macrophages and T regulatory cell infiltration in tumors (25,26). The gene *HMOX1* encodes Heme oxygenase-1, which participates in M2 macrophage polarization (27).

The cells annotated as dendritic cells expressed MHC class II genes (e.g. *HLA-DRB1*, *PLAC8*) (Figure 7D) and the proliferating cells expressed proliferating markers (e.g. *TOP2A*, *MKI67*) (Figure 7D). The cluster named as *RCVRN* highly expressed *RCVRN*, *CTB*, *MAP1B*, *STMN1*, *GNB3*, *PDE6H*, *AKAP9*, *NEUROD1*, *PEG10*, *DCT* (Figure 7D), many of which are photoreceptor markers (7,28). Those genes, along with the cone photoreceptor genes (e.g. *CRX*, *ARR3*, *GUCA1C*) and neuronal/ganglion genes (e.g.

GAP43, *EBF3*), and aberrantly expressed genes (e.g. *TFF1*) were expressed by the cluster RCVRN and by the stressed microglia (Figure 7D). Expression of both microglia markers and retinoblastoma-related genes may suggest the phagocytosis in the tumor associated microglia cells.

Among the lymphoid cells, we identified CD4 T cells (n = 851), CD8 T cells (n = 694), NK cells (n = 448), other T cells (n = 164) and B cells (n = 59) (Figure 7C). CD4 T cells expressed *CD3D* and either *CD4* or *CD8* (Figure 8E). A subset of the CD4 T cells expressed *FOXP3*; these are likely to be regulatory T cells, and account for around 52 cells (3% out of all T cells) (Figure 8E). CD8 T cells and NK cells expressed cytotoxic genes such as *GNLY* and *PRF1* (Figure 8E).

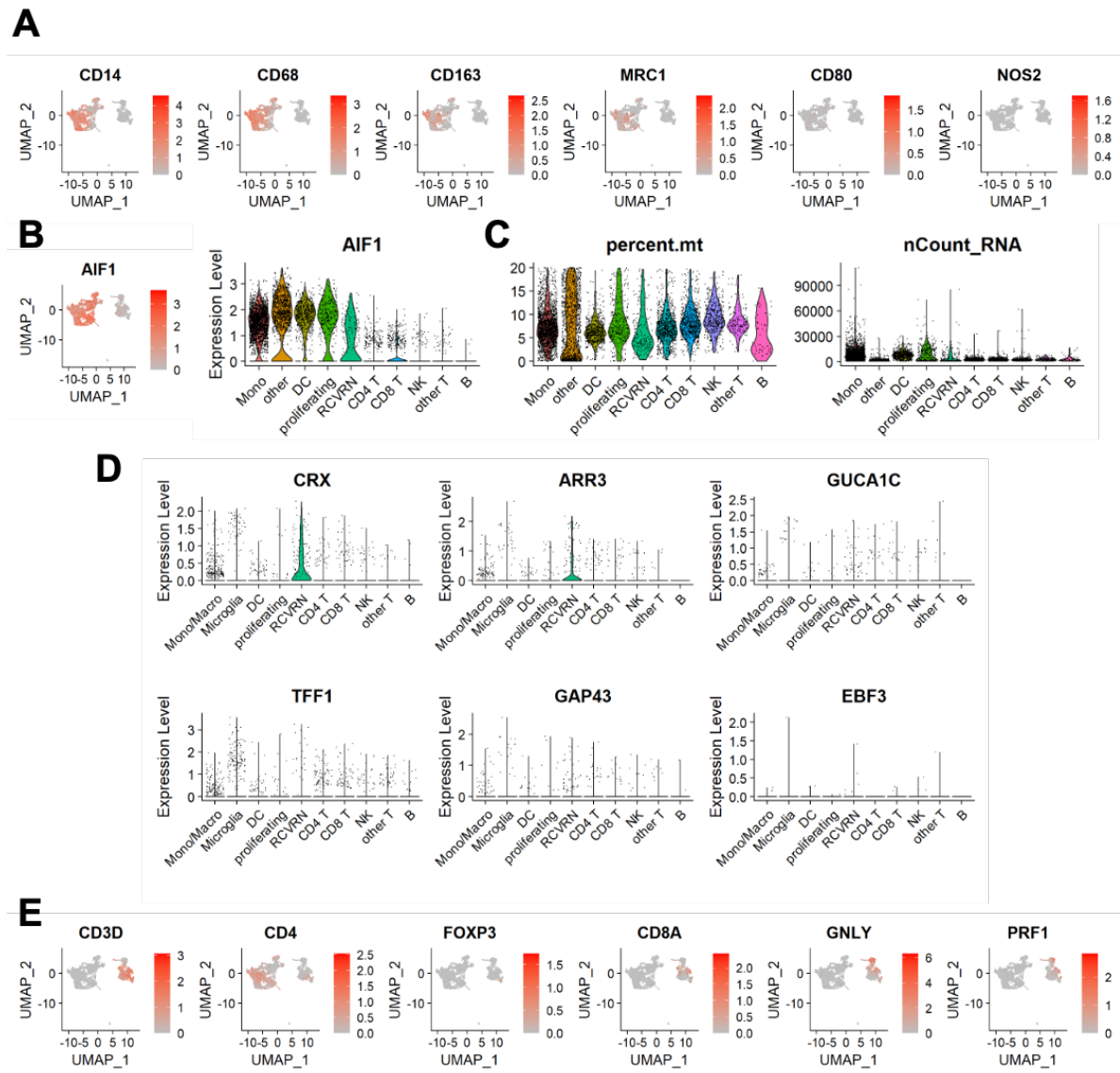


Figure 8. Expression of selected genes in retinoblastoma-infiltrated immune cells.

- A, Expression of markers of monocytes, macrophages, M2 and M1 macrophages presented in UMAP.
- B, Expression of microglia marker in UMAP or by Violin plot.
- C, Violin plot showing the mitochondria gene percentage and the number of count in each cell.
- D, Expression of cone photoreceptor markers, neuronal/ganglion markers, and aberrant retinoblastoma related gene in each immune cell type.
- E, Expression of T cell markers in UMAP.

Ligand-receptor analysis between retinoblastoma and tumor infiltrated macrophages

To investigate the possible mechanisms through which the M2 macrophages were activated, we performed ligand-receptor interaction prediction using CellPhoneDB (29). We randomly sampled 2000 cells in tumor, monocytes/macrophages or T cells, and iterated the predictions for 100 times. We kept the interactions that were predicted in more than 80 times out of the 100 iterations. CD74-MIF appeared as the top prediction between monocytes/macrophages and tumor cells (Figure 9), same as what we previously predicted using the data of one tumor (13). Other top predictions included C5AR1-RPS19, CD74-COPA, CD74-APP, GRN-SORT1, HLA-C-FAM3C, NRP1-VEGFB, LGALS9-CD47 (Figure 9). Among them, several genes were reported to be involved in the immune-suppressive microenvironment and M2 macrophage polarization. Complement C5a Receptor 1, encoded by *C5AR1* gene, is a complement receptor that plays a role in promoting M2 macrophage phenotype in squamous cell carcinoma (30) and breast cancer (31). Inhibition of the C5AR1 protein or blockade of its interaction with RPS19 can revert the protumor phenotype. The gene *LGALS9* encodes galectin 9, which participates in T cell exhaustion (32) and binds to CD206 in M2 macrophages (33). Neuropilin-1 (NRP1) was also shown to promote M2 macrophage infiltration in hypoxic environments (34). High expression of SPP1 and CD44 were reported in glioma (35) and in hepatocellular cancer (36). Co-culture of SPP1 knock down cell lines with macrophages had reduced levels of M2 macrophage marker expression, suggesting that SPP1 may play a role in M2 macrophage polarization (36).

CHAPTER 4

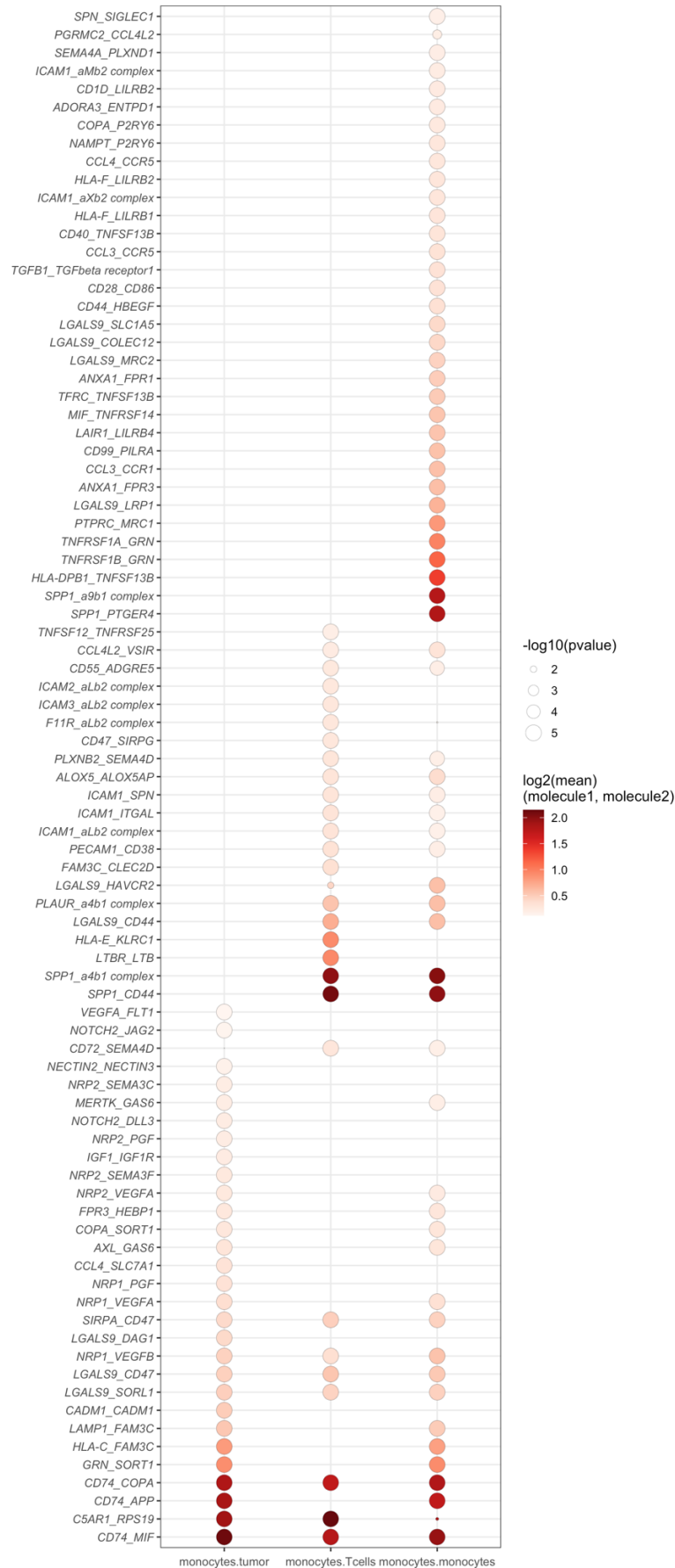


Figure 9. Inferred ligand-receptor pairs between monocytes and tumor cells, monocytes and T cells, or monocytes and monocytes.

DISCUSSION

Using a single-cell transcriptomic study of 14 primary retinoblastoma samples, we characterized the retinoblastoma intra- and inter- heterogeneity, as well as the tumor-infiltrating immune cells.

We reanalyzed a patient tumor alias as RBSC11 with marked intra-heterogeneity and refined the two possible progression pathways in this tumor that we proposed before (7). We uncovered, in part of the RBSC11, a dedifferentiation process in tumor progression, which was accompanied by the acquisition of additional genomic alterations. Some photoreceptor cone precursors that expressed *CRX*, *ARR3* and *GUCA1C*, after the losses of *RB1* genes and the gains of chromosome 10q, were transformed to cancer cells; with the addition of chromosome 2p gain, they decrease in their cone maturity, which was reflected by the loss of expression of a more mature cone marker *GUCA1C*; they may continue to lose maturing cone marker *ARR3*, and to express a neuronal stemness marker *CD24* and a pro-proliferation hormone *PRL*. In other part of RBSC11, *RB1* loss and accumulation of multiple chromosome alterations led to tumor progression. The gene *PRL* was also expressed by other retinoblastomas from patients who are diagnosed older than 18 months. It is not only the hormone produced by the pituitary gland for lactation, but also is endogenously expressed by the mammalian retina (16). It is a trophic factor in the retina that protects retinal cells from degeneration (17) and regulates retinal pigmented epithelia homeostasis (37). *PRL* is also involved in tumorigenesis and metastasis of multiple hormone responsive cancers (38).

We validated by single-cell RNA-seq in 14 patient samples that all tumor cells expressed *CRX*, an early photoreceptor marker, consistent with the hypothesis that retinoblastoma originate from cone photoreceptor precursors (39,40). Tumors from younger patients expressed maturing cone marker *ARR3*, while tumors from older patients expressed *EBF3*, a neuronal/ganglion marker, and *TFF1*, an aberrantly-expressed gene in retinoblastoma. This is in agreement of our previous findings of two molecular subtypes, that subtype 1 tumors expressed more matured cone signatures and manifested in younger patients, and subtype 2 tumors found in older patients

expressed less differentiated cone signatures but displayed neuronal/ganglion features and are associated with metastasis (7).

Our study portrayed comprehensively the immune landscape in the retinoblastoma tumor microenvironment. We identified tumor-associated macrophages/microglia, dendritic cells, CD4 T cells including regulatory T cells, CD8 T cells, NK cells and B cells in the retinoblastoma ecosystem. We showed that M2 macrophages were activated in retinoblastoma. Macrophages can be grouped into M1 (classically activated) and M2 (alternatively activated) phenotypes. M1 macrophages are often considered pro-inflammatory and anti-tumoral, while M2 macrophages tend to promote immune-suppressive status, angiogenesis, proliferation and metastasis in cancer (41). We inferred ligand-receptor interactions between tumor and immune cells and found several genes involved in the polarization of M2 macrophages. One of the predicted interactions between macrophages and tumor cells was MIF-CD74. MIF was shown to be one of the most secreted cytokines in retinoblastoma. Exposure of PBMC-derived macrophages to retinoblastoma-conditioned medium significantly increased M2 macrophage activation (13). Several other genes predicted in ligand-receptor interactions were also reported as factors to promote M2 phenotypes, such as *C5AR1*, *RPS19*, *LGALS9*, *NRP1*, and *SPP1*. Pharmacological inhibition of some of these factors could revert the M2 phenotype into the M1 phenotype (31) or improve anti-tumor activity synergistically through regulatory T cell depletion (32). Immunohistochemistry analysis of those proteins in retinoblastoma and functional experiments should be performed to validate their functions in their M2 polarization in retinoblastoma.

Taken together, our results depicted a comprehensive picture of retinoblastoma heterogeneity, including its immune microenvironment, which open new directions to targeted therapy and immune therapy in this rare pediatric cancer. We proposed several genes that may contribute to immune-suppressive environments that should be validated in future work.

METHODS

Retinoblastoma patients and ethics statement

We included 14 retinoblastomas in this study. All patients came from Institut Curie (Paris, France) from 2018 to 2021 and were diagnosed as unilateral and non-hereditary diseases. Median age at diagnosis was 20 months (range 9 to 63 months). All patients underwent primary enucleation without prior treatments.

The study was performed in accordance with the Declaration of Helsinki and French legislation. The study was approved by the Institut Curie Review Board. Written informed consent was obtained from parents or legal guardians of retinoblastoma patients, in accordance with the current guidelines and legislations of the country.

Single-cell RNA-sequencing of retinoblastoma

Within 1 hour after enucleation, a tumor specimen was obtained by needle aspiration through the anterior chamber of the eye and immediately placed in ice-cold CO₂ independent medium (Thermo Fisher). To obtain a high-quality sample, density gradient centrifugation was executed to remove debris, dead cells and erythrocytes using Histopaque-1077 (Sigma-Aldrich). Cell count and viability were determined by trypan blue exclusion on a Vi-CELL XR (Beckman Coulter Life Sciences). Samples with cell viability less than 60% underwent a second or a third round of density gradient centrifugation to further improve sample quality. All samples passed on for single-cell RNA-seq had viability more than 60%. Cells were then mechanically dissociated, washed and resuspended in phosphate-buffered saline supplemented with 0.04% bovine serum for single-cell processing.

With the aim of recovering 3000 cells, 6000 cells were loaded onto the Chromium controller (10X Genomics) using Chromium single cell 3' reagent kits v2 or v3 chemistry (10X Genomics). Each single cell was incorporated in one droplet together with one barcoded gel bead containing reagents. Following in-droplet lysis, cDNA incorporating cell barcode and UMI (unique molecular identifier) was synthesized and amplified. Amplified cDNA was fragmented and an Illumina sequencing library was constructed (Illumina). For quality control and quantification of cDNA and library, BioAnalyzer (Agilent BioAnalyzer High Sensitivity chip) was used. Libraries were

loaded at 400pM and pair-end sequenced on Novaseq 6000 using NovaSeq 6000 S1 Reagent Kit (Illumina). Cells were sequenced at a mean depth of 100000 reads/cell.

Single-cell transcriptome analysis

Sample demultiplexing, alignment to the reference genome (GRCh38, Ensembl 84, pre-built Cell Ranger reference version 1.2.0), quantification and initial quality control (QC) were performed using the Cell Ranger software (version 3.0.0 or 3.1.0, 10X Genomics).

For reanalysis of RBSC11, genes which were expressed in more than 3 cells, and cells that expressed more than 500 genes and less than 5% of mitochondria genes, were retained. For analysis of 14 retinoblastomas, genes which were expressed in more than 3 cells and cells which expressed more than 500 genes and less than 20% of mitochondria genes were retained. Normalization and clustering were performed using Seurat package version 3. UMI counts were normalized by `NormalizeData` function with `logNormalize` method, by a scaling factor of the median UMI count. UMI counts were then scaled to regress out the effect of UMI counts and the difference of G1S phase score minus S phase score. Variable genes were found with `FindVariableGenes` function with `logVMR` function. Genes with average expression more than 0.0125 and less than 8 and with dispersion more than 0.5 were considered as variable genes for principal component analysis (PCA).

Cell clusters were identified by Seurat V3 `FindNeighbors` and `FindClusters` functions that implemented K-nearest neighbor (KNN) graph-based method and modularity optimization with Louvain algorithm (15,42–44). The first 20 to 30 principal components were used. The parameter `Resolution` in the `FindClusters` was set between 0.1 to 1.4 in order to find an optimal resolution for biological interpretation.

Cluster markers were identified by `FindAllMarkers` function. Briefly, expression of genes that expressed in more than 10% of cells in one cluster were compared with expression of these genes in all other clusters, using Wilcoxon rank-sum test and corrected with a Benjamini & Hochberg (BH) correction. The procedure was repeated for all clusters. Genes upregulated in each cluster with more than 0.2 fold were considered as cluster markers. Pathway analysis of cluster markers were performed

by R clusterProfiler package (45). Genesets tested were from the Molecular Signatures Database v7.0 (46).

Copy number variations (CNVs) were inferred from the single-cell gene expression by InferCNV package 1.0.0, using normal retinal cells derived from hiPSCs as reference (47).

Cell type prediction was made using Seurat V4 reference mapping, where the reference were the single-cell gene expression data with 162000 human peripheral blood mononuclear cells sequenced by CITE-seq and normalized by SCTransform (variance-stabilizing transformations for single-cell UMI data) (24). Expression data of immune cells identified from our previous analysis were also normalized using SCTransform for query to keep consistent with the reference data. Anchors between the reference and query were identified using FindTransferAnchors function with the first 50 dimensions from a supervised PCA on normalized expression data. Cell type labels were transferred from the reference to the query with MapQuery function.

CellPhoneDB v2.1.7 was used to predict ligand-receptor interactions between immune cells and tumor cells (29). Briefly, 2000 cells were randomly selected from each cell type and the prediction function was applied with option 'statistical_analysis' and 1000 iterations. Analyses were repeated 100 times and the stable ligand-receptor pairs appeared more than 80 times were retained in the results.

REFERENCES

1. Dimaras H, Corson TW, Cobrinik D, White A, Zhao J, Munier FL, et al. Retinoblastoma. *Nat Rev Dis Prim.* 2015;(August):15021.
2. Rushlow DE, Mol BM, Kennett JY, Yee S, Pajovic S, Thériault BL, et al. Characterisation of retinoblastomas without RB1 mutations: Genomic, gene expression, and clinical studies. *Lancet Oncol.* 2013;14(4):327–34.
3. Vasef MA, Auerbach A, Czuchlewski DR, Bocklage T, Chabot-Richards D, Aguilera N, et al., editors. *Retinoblastoma*. In: *Diagnostic Pathology*. Philadelphia: Elsevier; 2016. p. 8–419.
4. McEvoy J, Flores-Otero J, Zhang J, Nemeth K, Brennan R, Bradley C, et al. Coexpression of Normally Incompatible Developmental Pathways in Retinoblastoma Genesis. *Cancer Cell.* 2011;20(2):260–75.
5. Kapatai G, Brundler M-A, Jenkinson H, Kearns P, Parulekar M, Peet AC, et al. Gene expression profiling identifies different sub-types of retinoblastoma. *Br J Cancer.* 2013 Jul;109(2):512–25.
6. Kooi IE, Mol BM, Moll AC, van der Valk P, de Jong MC, de Graaf P, et al. Loss of photoreceptor and gain of genomic alterations in retinoblastoma reveal tumor progression. *EBioMedicine.* 2015;2(7):660–70.
7. Liu J, Ottaviani D, Sefta M, Desbrousses C, Chapeaublanc E, Aschero R, et al. A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression. *Nat Commun.* 2021;12(1):5578.
8. Collin J, Queen R, Zerti D, Steel DH, Bowen C, Parulekar M, et al. Dissecting the Transcriptional and Chromatin Accessibility Heterogeneity of Proliferating Cone Precursors in Human Retinoblastoma Tumors by Single Cell Sequencing-Opening Pathways to New Therapeutic Strategies? *Invest Ophthalmol Vis Sci.* 2021 May;62(6):18.
9. Yang J, Li Y, Han Y, Feng Y, Zhou M, Zong C, et al. Single-cell transcriptome profiling reveals intratumoral heterogeneity and malignant progression in retinoblastoma. *Cell Death Dis.* 2021;12(12):1100.
10. Singh L, Singh MK, Rizvi MA, Bakhshi S, Meel R, Lomi N, et al. Clinical relevance of the comparative expression of immune checkpoint markers with the clinicopathological findings in patients with primary and chemoreduced retinoblastoma. *Cancer Immunol Immunother.* 2020;69(6):1087–99.
11. Ganesan B, Parameswaran S, Sharma A, Krishnakumar S. Clinical relevance of B7H3 expression in retinoblastoma. *Sci Rep.* 2020;10(1):10185.
12. Miracco C, Toti P, Gelmi MC, Aversa S, Baldino G, Galluzzi P, et al. Retinoblastoma Is Characterized by a Cold, CD8+ Cell Poor, PD-L1- Microenvironment, Which Turns Into Hot, CD8+ Cell Rich, PD-L1+ After Chemotherapy. *Invest Ophthalmol Vis Sci.* 2021 Feb 1;62(2):6.
13. Cuadrado-Vilanova M, Liu J, Paco S, Burgueño V, Sirab N, Pascual-Pasto G, et al. Identification of immunosuppressive factors in retinoblastoma cell secretomes and aqueous humor from patients. *J Pathol.* 2022 Mar 7.
14. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, Marioni JC, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* 2019;20(1):63.
15. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive Integration of Single-Cell Data. *Cell.* 2019 Jun;177(7):1888–1902.e21.
16. Rivera JC, Aranda J, Riesgo J, Nava G, Thebault S, López-Barrera F, et al. Expression and cellular localization of prolactin and the prolactin receptor in mammalian retina. *Exp Eye Res.* 2008 Feb;86(2):314–21.
17. Arnold E, Thebault S, Baeza-Cruz G, Arredondo Zamarripa D, Adán N, Quintanar-Stéphano A, et al. The hormone prolactin is a novel, endogenous trophic factor able to regulate reactive glia and to limit retinal degeneration. *J Neurosci.* 2014 Jan;34(5):1868–78.
18. Gilliam DT, Menon V, Bretz NP, Pruszk J. The CD24 surface antigen in neural development and disease. *Neurobiol Dis.* 2017 Mar;99:133–44.
19. Guo K, Searfoss G, Krolkowski D, Pagnoni M, Franks C, Clark K, et al. Hypoxia induces the expression of the pro-apoptotic gene BNIP3. *Cell Death Differ.* 2001 Apr;8(4):367–76.
20. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* (80-). 2014 Jun 20;344(6190):1396 LP – 1401.
21. Sung CH, Chuang JZ. The cell biology of vision. *J Cell Biol.* 2010;190(6):953–63.
22. Salman A, McClements ME, MacLaren RE. Insights on the Regeneration Potential of Müller Glia in the Mammalian Retina. *Cells.* 2021 Jul;10(8).
23. Xu XL, Fang Y, Lee TC, Forrest D, Gregory-Evans C, Almeida D, et al. Retinoblastoma has

- properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell*. 2009 Jun;137(6):1018–31.
24. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021 Jun;184(13):3573-3587.e29.
 25. Hu Z-W, Chen L, Ma R-Q, Wei F-Q, Wen Y-H, Zeng X-L, et al. Comprehensive analysis of ferritin subunits expression and positive correlations with tumor-associated macrophages and T regulatory cells infiltration in most solid tumors. *Aging (Albany NY)*. 2021 Apr;13(8):11491–506.
 26. Jézéquel P, Campion L, Spyrtatos F, Loussouarn D, Campone M, Guérin-Charbonnel C, et al. Validation of tumor-associated macrophage ferritin light chain as a prognostic biomarker in node-negative breast cancer tumors: A multicentric 2004 national PHRC study. *Int J cancer*. 2012 Jul;131(2):426–37.
 27. Naito Y, Takagi T, Higashimura Y. Heme oxygenase-1 and anti-inflammatory M2 macrophages. *Arch Biochem Biophys*. 2014 Dec;564:83–8.
 28. Lu Y, Shiao F, Yi W, Lu S, Wu Q, Pearson JD, et al. Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. *Dev Cell*. 2020;53(4):473-491.e9.
 29. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc*. 2020;15(4):1484–506.
 30. Medler TR, Murugan D, Horton W, Kumar S, Cotecchini T, Forsyth AM, et al. Complement C5a Fosters Squamous Carcinogenesis and Limits T Cell Response to Chemotherapy. *Cancer Cell*. 2018 Oct;34(4):561-578.e6.
 31. Markiewski MM, Vadrevu SK, Sharma SK, Chintala NK, Ghouse S, Cho J-H, et al. The Ribosomal Protein S19 Suppresses Antitumor Immune Responses via the Complement C5a Receptor 1. *J Immunol*. 2017 Apr;198(7):2989–99.
 32. Yang R, Sun L, Li C-F, Wang Y-H, Yao J, Li H, et al. Galectin-9 interacts with PD-1 and TIM-3 to regulate T cell death and is a target for cancer immunotherapy. *Nat Commun*. 2021;12(1):832.
 33. Enninga EAL, Chatzopoulos K, Butterfield JT, Sutor SL, Leontovich AA, Nevala WK, et al. CD206-positive myeloid cells bind galectin-9 and promote a tumor-supportive microenvironment. *J Pathol*. 2018 Aug;245(4):468–77.
 34. Casazza A, Laoui D, Wenes M, Rizzolio S, Bassani N, Mambretti M, et al. Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor immunity. *Cancer Cell*. 2013 Dec;24(6):695–709.
 35. He C, Sheng L, Pan D, Jiang S, Ding L, Ma X, et al. Single-Cell Transcriptomic Analysis Revealed a Critical Role of SPP1/CD44-Mediated Crosstalk Between Macrophages and Cancer Cells in Glioma. *Frontiers in Cell and Developmental Biology*. 2021 Vol 9.
 36. Liu L, Zhang R, Deng J, Dai X, Zhu X, Fu Q, et al. Construction of TME and Identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. *Cancer Immunol Immunother*. 2022 Jan;71(1):121–36.
 37. Meléndez García R, Arredondo Zamarripa D, Arnold E, Ruiz-Herrera X, Noguez Imm R, Baeza Cruz G, et al. Prolactin protects retinal pigment epithelium by inhibiting sirtuin 2-dependent cell death. *EBioMedicine*. 2016 May;7:35–49.
 38. Karayazi Atıcı Ö, Govindarajan N, Lopetegui-González I, Shemanko CS. Prolactin: A hormone with diverse functions from mammary gland development to cancer metastasis. *Semin Cell Dev Biol*. 2021 Jun;114:159–70.
 39. Singh HP, Wang S, Stachelek K, Lee S, Reid MW, Thornton ME, et al. Developmental stage-specific proliferation and retinoblastoma genesis in RB-deficient human but not mouse cone precursors. *Proc Natl Acad Sci*. 2018 Oct 2;115(40):E9391 LP-E9400.
 40. Xu XL, Fang Y, Lee TC, Forrest D, Gregory-Evans C, Almeida D, et al. Retinoblastoma Has Properties of a Cone Precursor Tumor and Depends Upon Cone-Specific MDM2 Signaling. *Cell*. 2009;137(6):1018–31.
 41. Liu J, Geng X, Hou J, Wu G. New insights into M1/M2 macrophages: key modulators in cancer progression. *Cancer Cell Int*. 2021;21(1):1–7.
 42. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 2015 May 21;161(5):1202–14.
 43. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir ED, Tadmor MD, et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*. 2015 Jul;162(1):184–97.
 44. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large

CHAPTER 4

- networks. 2008 Oct;2008(10):P10008.
45. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012 May;16(5):284–7.
 46. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005 Oct;102(43):15545–50.
 47. Slembrouck-Brec A, Nanteau C, Sahel J-A, Goureau O, Reichman S. Defined Xeno-free and Feeder-free Culture Conditions for the Generation of Human iPSC-derived Retinal Cell Models. *J Vis Exp*. 2018 Sep;(139).

CHAPTER 5.

CONCLUDING REMARKS

Summary

During my doctoral study, I sought to better understand retinoblastoma heterogeneity. Our team previously collected multi-omics retinoblastoma datasets and identified two molecular subtypes, one showing a photoreceptor cone signature and another showing a mixture of cone and ganglion signatures. I continued in this direction, comparing and characterizing the two subtypes. We showed that subtype 1 tumors were often manifested by younger patients, harbored fewer genomic alterations, exhibited more-mature cone and immune signatures; and that subtype 2 tumors were usually diagnosed in older patients, displayed a CpG Island methylator phenotype, expressed less-differentiated cone markers, as well as neuronal/ganglion and stemness signatures. We collaborated with Dr. Guillermo's team, and showed in an independent series of retinoblastoma, that all metastatic cases were subtype 2 tumors. With the help of Dr. Sirab, a biologist in the team, we collected single-cell RNA-sequencing data from 14 retinoblastoma samples. We validated in the single-cell data that the molecular profiles of tumors from younger patients were similar to the subtype 1 tumors that expressing more-mature cone signature, while the tumors from older patients expressed neuronal/ganglion markers and the genes aberrantly upregulated in subtype 2 tumors. We uncovered that intra-tumoral heterogeneity in retinoblastoma can exist at both genomic and phenotypic level through the analysis of one retinoblastoma at single-cell level. We proposed a progression model for part of this tumor in which the accumulation of genomic alterations were accompanied by the loss of cone features and the gain of stemness feature and a retinal trophic factor. From the single-cell transcriptomic data, we identified the infiltrations of immune cells and normal retinal cells in retinoblastoma microenvironment. We made a comprehensive portrait of the immune cell landscape in retinoblastoma and showed an activation of protumoral M2 macrophages maybe due to several candidate genes. In collaboration with Dr. Carcaboso's team, we validated with cytokine array and functional experiments that Macrophage Migration Inhibitory Factor participated in the shaping

of the M2 phenotype and may contribute to immune-suppressive microenvironment in retinoblastoma.

Future directions

1) *MYCN*-amplified tumors as a third subtype

MYCN-amplified tumors were classified as subtype 2 retinoblastoma in our study, they expressed neuronal/ganglion markers and aberrant genes upregulated in subtype 2 tumors. On the other hand, they harbored several unique features, including their relative low proliferation rate and overall hypomethylation. *MYCN*-amplified retinoblastoma is found in 10% of all cases and in our cohort we only had 5 samples with transcriptomic or DNA methylation data. With more cases, we may be able to identify them as a third subtype or a subcluster of the subtype 2 retinoblastoma.

2) Retinoblastoma cell of origin

We showed that all retinoblastoma expressed early photoreceptor markers, along with several evidences from the literature, suggesting that retinoblastoma are derived from the early photoreceptor cone cells. However, there is still controversy on the retinoblastoma cell of origin as some researchers, based on the expression of some retinal progenitor markers, claimed that they are transformed from retinal progenitors. Analysis of the single-cell transcriptomic data in the fetal retina combined with our data from the tumors could provide more insights into this debated issue.

3) Aberrant pathways in retinoblastoma

Bulk data is a mixture of signals from both tumor and cells from the tumor microenvironment. Using single-cell data, we would be able to compare the tumor cells and normal cells to identify tumor-specific aberrant genes and pathways. Gene regulatory inference analysis can be applied to both bulk and single-cell data. It extracts pairwise gene information and can provide insights into gene regulations in tumorigenesis.

4) Immune therapies for retinoblastoma

Emerging evidence showed that immune therapies may also work in immune-cold pediatric cancers. We proposed several genes that may function in M2 macrophage polarization and tumor promotion. This need to be examined by *in vitro* and *in vivo* phenotypic and functional work. Our collection of immune cells at single-cell level also provides a rich resource for future investigations into retinoblastoma immune therapies.

Conclusion

Taken together, our comprehensive characterization of the two molecular subtypes of retinoblastoma, the intra-tumoral heterogeneity, and the immune landscape open up new directions for retinoblastoma studies and shed light on future target therapies and immune therapies for retinoblastoma.

ANNEX.


IDENTIFICATION OF IMMUNOSUPPRESSIVE FACTORS IN RETINOBLASTOMA CELL SECRETOMES AND AQUEOUS HUMOR FROM PATIENTS

In this annex, an article on work done in collaboration with Dr. Carcaboso and Dr. Cuadrado-Vilanova is enclosed. Utilizing cytokine array and immunohistochemistry, we identified cytokines secreted by 11 primary retinoblastoma derived cell models involved in the activation of protumoral M2 macrophages, including Macrophage Migration Inhibitory Factor (MIF). This is reminiscent of our findings of MIF-CD74 being one of the top predicted ligand-receptor pairs between tumor and monocytes/macrophages/microglia by single-cell analysis. We then made a ligand-receptor prediction using bulk gene expression data from tumors and from purified immune cells, and found MIF-CD74 still being the top hit. We showed that MIF was upregulated in tumor cells comparing to normal retina using bulk gene expression data. We further demonstrated that retinoblastoma conditioned medium or recombinant MIF could polarize macrophages to protumoral M2 type.

This manuscript is published in the Journal of Pathology (2022), available through the link: <https://doi.org/10.1002/path.5893>.

I am the second author of the article, and participated in the bioinformatics and statistical analysis, data visualization, manuscript writing of the relevant part and reviewing.

Identification of immunosuppressive factors in retinoblastoma cell secretomes and aqueous humor from patients

Maria Cuadrado-Vilanova^{1,2}, Jing Liu^{3,4}, Sonia Paco^{1,2}, Rosario Aschero^{1,2}, Victor Burgueño^{1,2}, Nanor Sirab^{3,4} , Guillem Pascual-Pasto^{1,2}, Genoveva Correa^{1,2}, Leire Balaguer-Lluna^{1,2}, Helena Castillo-Ecija^{1,2}, Sara Perez-Jaume^{1,2}, Oscar Muñoz-Aznar^{1,2}, Monica Roldan^{1,5}, Mariona Suñol^{1,6}, Paula Schaiquevich^{7,8}, Isabelle Aerts³, François Doz^{3,9}, Nathalie Cassoux^{9,10}, Fabiana Lubieniecki¹¹, Daniel Benítez-Ribas¹², Cinzia Lavarino^{1,2}, Jaume Mora^{1,2}, Guillermo L Chantada^{1,2,8,13}, Jaume Catala-Mora¹⁴, François Radvanyi^{3,4} and Angel M Carcaboso^{1,2*} 

¹ Institut de Recerca Sant Joan de Deu, Barcelona, Spain

² Pediatric Oncology, Hospital Sant Joan de Deu, Barcelona, Spain

³ Institut Curie, CNRS, UMR I 44, SIREDO Oncology Center, Paris, France

⁴ Institut Curie, PSL Research University, Paris, France

⁵ Genetic and Molecular Medicine - IPER, Hospital Sant Joan de Deu, Esplugues de Llobregat, Barcelona, Spain

⁶ Pathology, Hospital Sant Joan de Deu, Barcelona, Spain

⁷ Precision Medicine, Hospital de Pediatría JP Garrahan, Buenos Aires, Argentina

⁸ CONICET, Buenos Aires, Argentina

⁹ University of Paris, Paris, France

¹⁰ Institut Curie, Ophthalmic Oncology, Paris, France

¹¹ Pathology, Hospital de Pediatría JP Garrahan, Buenos Aires, Argentina

¹² Institut d'Investigacions Biomediques August Pi Sunyer, Barcelona, Spain

¹³ Universidad Austral-CONICET, Institute for Research in Translational Medicine (IIMT), Pilar, Argentina

¹⁴ Ophthalmology, Hospital Sant Joan de Deu, Barcelona, Spain

*Correspondence to: AM Carcaboso, Institut de Recerca Sant Joan de Deu, Santa Rosa 39-57, Esplugues de Llobregat, Barcelona, 08950, Spain.

E-mail: angel.montero@sjd.es

Abstract

The microenvironment of retinoblastoma, the solid malignancy of the developing retina, is immunosuppressive. To study the interactions between tumor-associated microglia/macrophages (TAMs) and tumor cells in retinoblastomas, we analyzed immunohistochemistry markers in 23 patient samples and characterized 105 secreted cytokines of 11 retinoblastoma cell models in culture. We detected profuse infiltration of CD163⁺ protumoral M2-like polarized TAMs in eyes enucleated due to cancer progression. Previous treatment of patients increased the number of TAMs but did not affect M2-like polarization. M2-like microglia/macrophages were almost absent in five eyes obtained from children enucleated due to nontumoral causes. CD8⁺ tumor-infiltrating lymphocytes (TILs) were moderately abundant in tumor eyes and very scarce in nontumoral ones. The expression of the immune checkpoint molecule PD-L1 was absent in 95% of the tumor samples, which is concordant with the finding of FOXP3⁺ Tregs infiltrating tumors. We confirmed the pathology results using single-cell transcriptome analysis of one tumor. We identified the cytokines extracellular matrix metalloproteinase inducer (EMMPRIN) and macrophage migration inhibitory factor (MIF), both with reported immunosuppressive activity, secreted at high levels in retinoblastoma primary cell cultures. Gene expression analysis of a large retinoblastoma cohort and single-cell transcriptome analysis confirmed that MIF and EMMPRIN were significantly upregulated in retinoblastomas, which led us to quantify both proteins by immunoassays in liquid biopsies (aqueous humor obtained from more than 20 retinoblastoma patients). We found a significant increase in the concentration of MIF and EMMPRIN in cancer patients, compared to 12 noncancer ones. Finally, we showed that macrophages derived from peripheral blood mononuclear cells increased the expression of markers of M2-like polarization upon exposure to retinoblastoma-conditioned medium or recombinant MIF. Overall, our findings suggest that retinoblastoma cell secretions induce the protumoral phenotype of this tumor. Our results might have clinical impact in the fields of biomarkers and treatment.

© 2022 The Pathological Society of Great Britain and Ireland.

Keywords: retinoblastoma; protumoral M2-like tumor associated microglia/macrophages (TAMs); macrophage migration inhibitory factor (MIF); extracellular matrix metalloproteinase inducer (EMMPRIN); liquid biopsy

Received 8 September 2021; Revised 10 February 2022; Accepted 3 March 2022

No conflicts of interest were declared.

INTRODUCTION

Retinoblastoma is a solid malignancy of the retina of small children, generally originated by biallelic alterations of the tumor suppressor gene *RB1* [1,2]. Conventional chemotherapy agents are effective for tumor chemoreduction in the intraocular form of the disease, but they can produce off-target effects and chemoresistance [1]. More selective treatments targeting specific genetic backbones of the tumor, such as the upregulation of the genes *SYK*, *MDM2*, and *MYCN* [3,4], or the overexpression of the transcription factor *E2F1* in an *RB1*-deregulated background [5], are still in the initial phases of preclinical and clinical research. A third group of treatments could emerge from the understanding of the microenvironmental and immunologic properties of the tumor [6,7]. Like most pediatric solid embryonal cancers, retinoblastomas are poor in stroma and in tumor-infiltrating lymphocytes (TILs) and show low expression of the programmed cell death ligand-1 (PD-L1) and other immune checkpoint molecules [8–10]. Such properties are consistent with an immunosuppressive or protumoral environment [7].

The identification of the causes for the cold tumor environment of retinoblastoma might help unveil actionable targets for future treatments. In this context, the interaction between tumor and stromal cells (including cells of the innate immune system, such as the retinal microglia) can lead to the activation of immunosuppressive signaling pathways [11]. Tumor cells may secrete antiinflammatory chemokines and cytokines, such as the macrophage migration inhibitory factor (MIF). This 12 kDa protein reportedly increases the infiltration of tumors by immunosuppressive regulatory T cells (Tregs) and tumor-associated microglia/macrophages (TAMs) polarized to the protumoral type M2-like [12–14]. Other tumor cell mechanisms that might suppress the immune system are the decreased expression of human leukocyte antigen (HLA) class I molecules [15], the production of transforming growth factor beta (TGF- β) [16], the activation of tryptophan degradation by the enzyme indoleamine 2,3-dioxygenase (IDO) [17], and the activation of the cyclooxygenase-2 prostaglandin E2 pathway [18].

A unique factor involved in the retinoblastoma microenvironment is the tumor confinement into the eye compartments in its early stages. This is especially relevant in the case of retinoblastomas with seeding, i.e. tumor cell clusters floating in the vitreous humor, or in the sub-retinal fluid, or, less frequently, in the aqueous humor [19]. These ocular fluids might contain protein factors appropriate for retinoblastoma growth, or secretions of retinoblastoma tumors, including exosomes, cytokines, and circulating cell-free DNA [20,21]. Aqueous humor is now safely accessible for liquid biopsy during the procedure of the intravitreal treatment injection [22]. The analysis of such aqueous humor samples might help characterize the microenvironment of the tumor in patients. In the experimental laboratory setting, primary cultures of patient-derived retinoblastoma cells offer a dynamic tool to identify secreted proteins or cellular

immune escape mechanisms. These cell models grow well in conditions that enrich stem cell-like cells [23].

In the present study we characterized histologic markers of the immune microenvironment in retinoblastoma samples (enucleated eyes) and compared them with the retinal microenvironment of nontumoral pediatric eyes, finding a strikingly high amount of M2-like polarized TAMs in retinoblastomas, independently of their pretreatment status. Then we hypothesized that retinoblastoma cells secrete soluble factors that induce such a protumoral environment. To address this question, we used the largest reported platform of patient-derived retinoblastoma cell cultures, validated by previous work [5,24–27]. We identified cytokines, such as MIF and extracellular matrix metalloproteinase inducer (EMMPRIN), that were abundantly and homogeneously secreted by all the cell models. Retinoblastoma-conditioned medium induced M2-like polarization and enhanced migration of human blood-derived macrophages. Finally, we analyzed liquid biopsies (aqueous humor) to address the clinical significance of the laboratory findings.

Materials and methods

Cell lines and cell culture

We used the retinoblastoma cell line Y79 (Sigma-Aldrich, St. Louis, MO, USA) and 11 primary retinoblastoma cell cultures established from the tumors of 11 patients. See Supplementary materials and methods and supplementary material, Table S1 for details. The Institutional Review Boards at the hospital Sant Joan de Deu (HSJD, Barcelona, Spain; protocol M-1608-C) and Hospital JP Garrahan (HPG, Buenos Aires, Argentina; protocols 838 and 904) approved the collection of tumor tissues, under informed consent.

Immunostaining, image acquisition and analysis

We carried out immunohistochemistry (IHC) staining on four-micron serial sections of formalin-fixed paraffin-embedded human tissue samples. See Supplementary materials and methods and supplementary material, Tables S2 and S3 for details.

Expression of common immune escape mechanisms by tumor cells

We assessed gene expression related to five different immune escape mechanisms in retinoblastoma cells by reverse-transcribed quantitative polymerase chain reaction (RT-qPCR). See Supplementary materials and methods and supplementary material, Table S4 for details.

Retinoblastoma secretome analysis

We cultured cells in 1 ml of supplemented retinoblastoma medium in 6-well plates, at a density of 10^6 cells/well. After 48 h, we collected supernatants by centrifugation at $400 \times g$ for 4 min and froze them at -20°C

until analysis. We performed immunodetection of cytokines with the Proteome Profiler Human XL Cytokine Array kit (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's instructions. This kit detects 105 proteins (in duplicate) simultaneously and includes negative and positive controls. See Supplementary materials and methods for details.

Gene expression of top-secreted cytokines in primary retinoblastomas and fetal retina

We analyzed the RNA of 59 retinoblastoma samples without prior treatments, six primary retinoblastoma cell cultures, two retinoblastoma cell lines, and three fetal retinas using the microarray Affymetrix U133 GeneChip (Affymetrix, Santa Clara, CA, USA). See Supplementary materials and methods for details.

Single-cell gene expression of top-secreted cytokines and immune markers in one retinoblastoma

We obtained single-cell RNA sequencing data from one retinoblastoma sample after primary enucleation of a patient older than 18 months with unilateral nonhereditary *RB1*^{-/-} disease, identified as RBSC11. See Supplementary materials and methods for details.

Quantification of secreted MIF and EMMPRIN in retinoblastoma cultures and liquid biopsies (aqueous humor)

We cultured tumor cells in 4 ml of supplemented retinoblastoma medium at a cell density of 2×10^6 cells/well in six-well plates. After 48 h, we collected supernatants by centrifugation at $400 \times g$ for 4 min. We obtained liquid biopsies (aqueous humor; 100 μ l) from patients before the procedures of intravitreal chemotherapy (retinoblastoma patients), intravitreal bevacizumab (patients with retinal disease), or cataract surgery. Samples were stored at -20°C until analysis. We diluted all samples (1:50) and performed Human MIF and Human EMMPRIN Quantikine ELISA Kits (R&D Systems) according to the manufacturer's instructions. We measured the assays with an Infinite M Nano (Tecan, Männedorf, Switzerland) microplate reader.

Differentiation of macrophages from peripheral blood mononuclear cells (PBMCs)

We isolated PBMCs from blood buffy coats of eight adult blood donors. See Supplementary materials and methods for details.

Macrophage polarization assay

We plated HSJD-RBT-2 cells at a density of 2×10^6 cells/well in six-well plates, in 2 ml of nonsupplemented retinoblastoma medium. After 72 h, we collected cell culture supernatants (i.e. retinoblastoma-conditioned media) and transferred them to macrophage cultures

containing 2×10^6 cells/well in complete macrophage culture medium six-well plates. The volume proportion of macrophage and retinoblastoma culture media was 1:1 in all the experimental conditions. To evaluate the effect of MIF on macrophage polarization, we prepared recombinant MIF (rMIF, Peprotech; Rocky Hill, NJ, USA) in nonsupplemented retinoblastoma medium, to achieve 500 ng/ml rMIF in the final macrophage cultures. We prepared M1-like (proinflammatory) polarization controls by exposing macrophages to lipopolysaccharides (LPS; Sigma-Aldrich), IFN- γ and tumor necrosis factor- α (TNF- α ; Peprotech), all dissolved in nonsupplemented retinoblastoma medium to achieve final concentrations of 100, 50, and 20 ng/ml, respectively. We prepared M2-like (protumoral) polarization controls by exposing macrophages to interleukin-10 (IL-10) and TGF- β 1 (Peprotech), dissolved in nonsupplemented retinoblastoma medium to achieve final concentrations of 10 ng/ml each. As an internal experimental control, we exposed macrophages to plain nonsupplemented retinoblastoma culture medium. Cells were maintained for 72 h at 37°C in a 5% CO_2 atmosphere until flow cytometry analysis for M1-like and M2-like markers. To correct for interpatient variability in basal levels of macrophage polarization, we expressed the polarization results as the ratio of CD163⁺ cells (M2-like) to CD80⁺ cells (M1-like).

Flow cytometry

See Supplementary materials and methods for details.

Invasion and migration assays

We plated 5×10^4 nonpolarized macrophages suspended in 250 μ l of complete macrophage culture medium in Transwell chambers (Falcon Permeable Support for 24-well Plate with 8.0 μ m Transparent PET Membrane; Corning, Glendale, AZ, USA) coated with Geltrex (Thermo Fisher Scientific, Waltham, MA, USA). In the lower chamber we deposited a 1:1 mixture of complete macrophage medium and retinoblastoma-conditioned medium or 500 ng/ml rMIF, both prepared using nonsupplemented retinoblastoma medium. We used 1:1 mixtures of complete macrophage culture medium and nonsupplemented retinoblastoma culture medium containing 10% iFBS or 0.1% BSA as the positive and negative migration controls, respectively. As an internal experimental control, we used a 1:1 mixture of complete macrophage culture medium and plain nonsupplemented retinoblastoma medium. After 72 h incubation, we fixed migratory cells with 4% paraformaldehyde and stained them with 1% crystal violet (Sigma-Aldrich) in 2% ethanol. We counted the number of the stained cells using a standard bright-field microscope. We represented the results relative to the migration of the negative control.

Statistical analysis

We used GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA) for statistical analysis. See Supplementary materials and methods for details.

Results

TILs and TAMs in retinoblastoma samples

We included in the study 28 enucleated eyes, 23 from retinoblastoma patients and five from nonretinoblastoma. Supplementary material, Table S5 shows the clinical and anatomic-pathological features of retinoblastomas, including germline *RB1* mutations. Median age at diagnosis was 29 months (range 5.4–111). Patients who received chemotherapy before enucleation and nontreated retinoblastoma patients (naïve) were equally represented in the cohort (supplementary material, Table S5). Clinical information of nonretinoblastoma enucleated eyes is in the supplementary material, Table S6; median patient age at enucleation was 55 months (range 28–134).

T and B lymphocyte counts in tumors and conserved retinal areas of most retinoblastoma eyes were moderate and not significantly different from counts in control eye retinas from nonretinoblastoma pediatric patients (Figure 1A). The supplementary material, Table S7 shows mean cell counts in the whole sample set. CD8⁺

lymphocytes were the most abundant TILs in tumors (Figure 1B). We found FOXP3⁺ cells (Tregs) at low density in around 83% of retinoblastoma tumors and 46% of conserved retinas in tumor eyes, but we did not find them in control eye retinas (supplementary material, Table S7). Infiltration of FOXP3⁺ and CD20⁺ cells in tumor areas achieved counts significantly higher than those of conserved retinas of tumor eyes and control eye retinas (Figure 1B). We did not detect significant differences in TIL counts between treated and naïve tumor samples. Enucleated eyes with prelaminar and retrolaminar optic nerve invasion showed a significantly higher number of CD4⁺ counts compared to eyes without signs of invasion (supplementary material, Table S8). Other anatomic-pathological properties or genetic background of the tumors were not related to TILs infiltration (supplementary material, Table S8).

All retinoblastoma tumors showed moderate infiltration of CD68⁺ macrophages and abundant Iba1⁺ microglial cells of ramified and amoeboid morphology, predominantly of the M2-like type (CD163⁺) (Figure 2A and supplementary material, Table S7). Expression of CD163 was

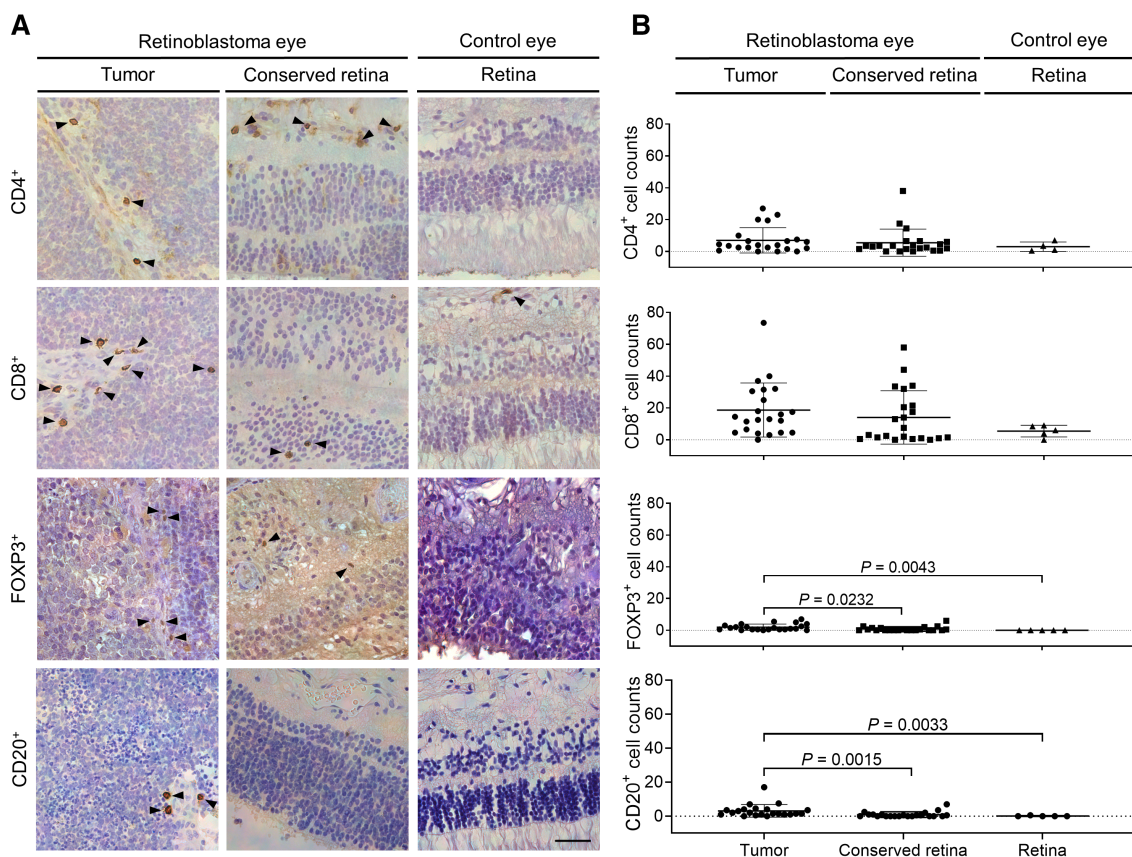


Figure 1. Infiltrating lymphocytes in retinoblastoma eyes (tumor and conserved retinal areas) and in retinas of nonretinoblastoma (control) eyes. (A) Representative examples of immunostaining for CD4, CD8, FOXP3, and CD20. Arrowheads indicate positive staining of the markers. All specimens are shown at 20× objective magnification with a 40-μm scale bar. Retinas are oriented so that the vitreous chamber is in the upper part of the image. (B), Total cell counts in tumor tissue ($n = 23$), evaluable conserved retinas ($n = 22$) and control eyes retinas ($n = 5$). Individual data (represented with dots) are means of the measures by two researchers (dots) and mean \pm SD are represented with lines.

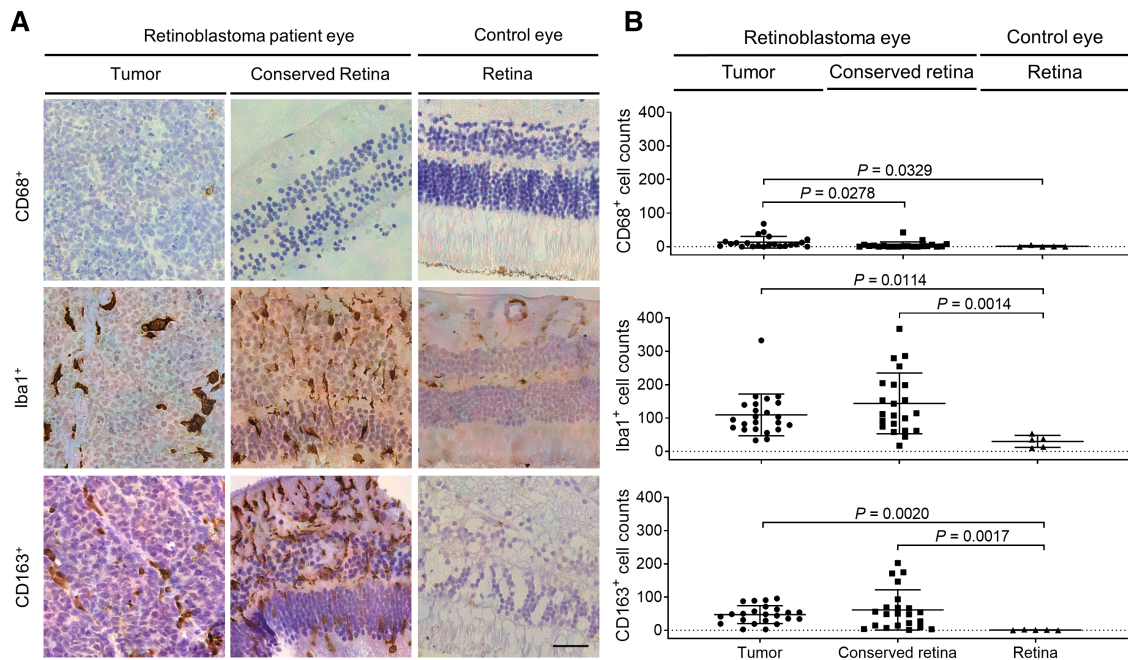


Figure 2. Infiltrating macrophages and microglia in retinoblastoma eyes (tumor and conserved retinal areas) and in retinas of nonretinoblastoma (control) eyes. (A) Representative examples of CD68, Iba1, and CD163 immunostaining. All specimens are shown at 20 \times objective magnification with a 40- μ m scale bar. Retinas are oriented so that the vitreous chamber is in the upper part of the image. (B) Total cell counts in tumor tissue ($n = 23$), evaluable conserved retinas ($n = 22$), and control eyes retinas ($n = 5$). Individual data (represented with dots) are means of the measures by two researchers (dots) and mean \pm SD are shown with lines.

exclusive of nonretinoblastoma cells (supplementary material, Figure S1). Conserved retinal areas in retinoblastoma eyes showed similar infiltration of these cells. Control eyes retinas, in contrast, showed no infiltration of CD68 $^{+}$ macrophages, a lower number of Iba1 $^{+}$ microglial cells than retinoblastoma eyes, and almost absence of M2-like positive cells (Figure 2A). Counts of CD68 $^{+}$, Iba1 $^{+}$, and CD163 $^{+}$ cells were significantly higher in retinoblastoma eyes compared to control eyes (Figure 2B).

Previous treatment of enucleated eyes significantly increased the abundance of Iba1 $^{+}$ TAMs in tumors, compared to naïve eyes, but did not affect the number of M2-like CD163 $^{+}$ TAMs (supplementary material, Table S9). Overall, our findings suggest that retinoblastoma promotes a microenvironment rich in TAMs of the M2-like type from the early development of the disease.

Immune escape mechanisms in retinoblastoma

Because our initial findings confirmed the immunosuppressive microenvironment in retinoblastoma tumors, we assessed possible immune escape mechanisms of tumor cells. Protein expression of the immune checkpoint molecule PD-L1 in tumors was negative in 96% of the samples (Figure 3A). The only positive sample corresponded to a treatment-naïve eye that presented 5% of the tumor cells positive for membrane PD-L1 staining (Figure 3A). We did not find positivity for the receptor PD-1 in any of the 23 analyzed tumors. Expression of B7-H3 was negative in 70% of the analyzed

samples. We identified focal B7-H3 staining in five samples and only two were positive in at least 5% of the tumor cells (Figure 3B). Microarray data confirmed low expression of genes *PDCD1* (PD-1) and *CD274* (PD-L1), and higher expression of *CD276* (B7-H3) (supplementary material, Figure S2). The expression of the gene encoding PD-L1, *CD274*, was not detectable by RT-qPCR in six retinoblastoma cell models cultured in basal conditions (Figure 3C). Upon incubation with IFN- γ , *CD274* expression became detectable in all cell models (Figure 3C). In contrast, the expression of the gene encoding B7-H3 was detectable by RT-qPCR in basal conditions and its expression levels did not increase upon IFN- γ exposure (Figure 3C). In similar experiments, we detected a significant increase in the membrane expression of PD-L1 protein when we exposed cells to IFN- γ (Figure 3D,E). Taken together, these results suggest that the immune escape mechanism of PD-L1 is functional in retinoblastomas, although not activated due to the immunosuppressed environment in the human tumors.

The expression of genes encoding for proteins of the HLA class I complex was low in retinoblastoma cells (supplementary material, Figure S3A). Concordantly, expression of HLA-ABC class I proteins was very low in retinoblastoma cells (supplementary material, Figure S3B). We detected a high increase in these membrane proteins upon exposure of tumor cells to IFN- γ (supplementary material, Figure S3B,C). Genes of the *TGF- β* family were at low expression levels in most cell

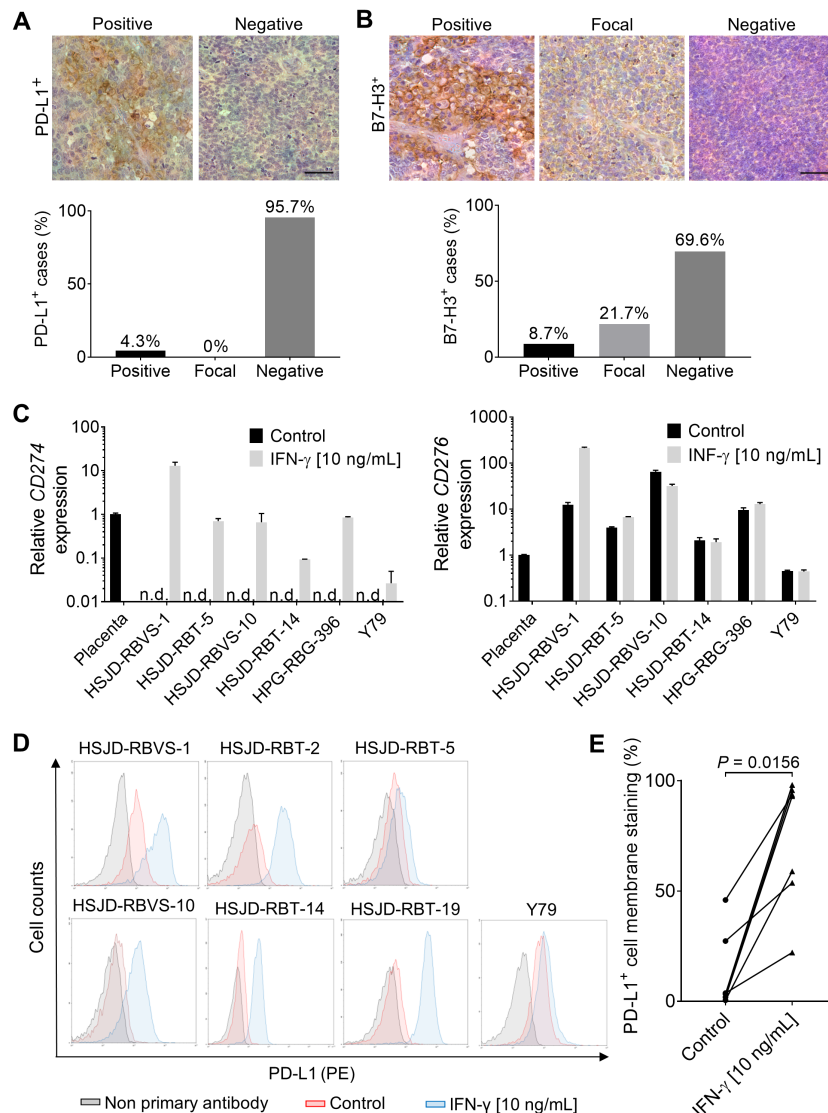


Figure 3. Expression of immune checkpoint molecules in retinoblastoma samples. (A) Top: representative examples of positive (brown staining in cell membranes) and negative PD-L1 immunostaining (20× objective magnification and 40-μm scale bar). Bottom: percentage of PD-L1 positivity in retinoblastomas ($n = 23$). (B) Top: representative examples of positive (brown membranes), focal, and negative B7-H3 immunostaining (20× objective magnification and 40-μm scale bar). Bottom: percentage of B7-H3 positivity in retinoblastomas ($n = 23$). (C) mRNA expression of *CD274* and *CD276* in six retinoblastoma cell models upon exposure to artificial proinflammatory microenvironment (10 ng/ml IFN-γ, 48 h). n.d., not detected. Data are shown as mean ± SD of two replicates. (D) Membrane PD-L1 staining by flow cytometry phycoerythrin (PE) fluorescence emission in retinoblastoma cells exposed to IFN-γ (10 ng/ml; 48 h). (E) Quantification of the flow cytometry experiment and paired analysis of the increase of PD-L1 membrane expression upon exposure to 10 ng/ml IFN-γ. Lines link paired samples.

models (supplementary material, Figure S4A). We did not detect mRNA of genes *IDO1*, *IDO2*, and *TDO*, of the kynurenine pathway, in any of the cells (supplementary material, Figure S4B). Cyclooxygenase-2 prostaglandin E2 pathway genes, *PTGES* and *PTGS2*, were not expressed in most of the cells (supplementary material, Figure S4C). Overall, these results suggest that reduced expression of HLA class I molecules may be an important immune evasion mechanism for retinoblastoma cells, while immunosuppressive pathways such as TGF-β, kynurenine, and cyclooxygenase-2 prostaglandin E2 are not activated.

Discovery of highly secreted macrophage inhibitory cytokines in retinoblastoma

We identified a homogeneous pattern of cytokine secretion in the supernatant of 11 retinoblastoma cell models (Figure 4A). The top-five secreted proteins were EMMPRIN, MIF, insulin-like growth factor-binding protein 2 (IGFBP-2), growth differentiation factor 15 (GDF-15), and pentraxin-3. All these molecules inhibit macrophages or cause immune evasion in solid tumors [14,28–30]. In the gene expression microarray,

Immunosuppressive factors in the microenvironment of retinoblastoma

7

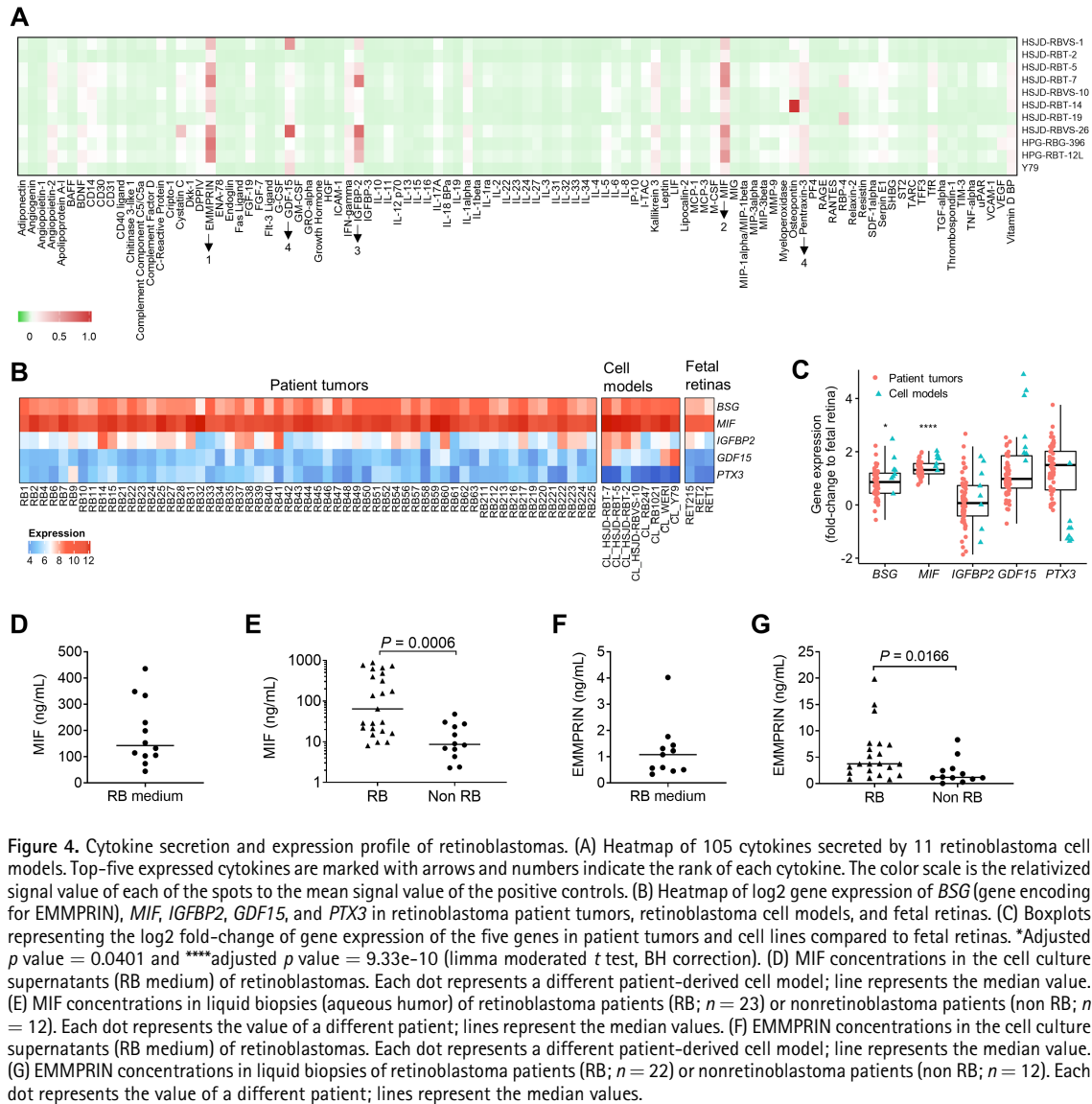


Figure 4. Cytokine secretion and expression profile of retinoblastomas. (A) Heatmap of 105 cytokines secreted by 11 retinoblastoma cell models. Top-five expressed cytokines are marked with arrows and numbers indicate the rank of each cytokine. The color scale is the relativized signal value of each of the spots to the mean signal value of the positive controls. (B) Heatmap of log₂ gene expression of *BSG* (gene encoding for EMMPRIN), *MIF*, *IGFBP2*, *GDF15*, and *PTX3* in retinoblastoma patient tumors, retinoblastoma cell models, and fetal retinas. (C) Boxplots representing the log₂ fold-change of gene expression of the five genes in patient tumors and cell lines compared to fetal retinas. *Adjusted *p* value = 0.0401 and ***adjusted *p* value = 9.33e-10 (limma moderated *t* test, BH correction). (D) MIF concentrations in the cell culture supernatants (RB medium) of retinoblastomas. Each dot represents a different patient-derived cell model; line represents the median value. (E) MIF concentrations in liquid biopsies (aqueous humor) of retinoblastoma patients (RB; *n* = 23) or nonretinoblastoma patients (non RB; *n* = 12). Each dot represents the value of a different patient; lines represent the median values. (F) EMMPRIN concentrations in the cell culture supernatants (RB medium) of retinoblastomas. Each dot represents a different patient-derived cell model; line represents the median value. (G) EMMPRIN concentrations in liquid biopsies of retinoblastoma patients (RB; *n* = 22) or nonretinoblastoma patients (non RB; *n* = 12). Each dot represents the value of a different patient; lines represent the median values.

patient tumors and cell models expressed significantly higher levels of genes *MIF* and *BSG* (gene encoding for EMMPRIN) compared to fetal retinas, with *MIF* being upregulated in all samples and *BSG* being upregulated in 56/59 tumors and in all cell models, compared to fetal retinas (Figure 4B,C). In cell culture supernatants, MIF concentration was in the range 44.9–436 ng/ml (*n* = 12 cell models) (Figure 4D). Median MIF concentration in liquid biopsies (aqueous humor) of retinoblastoma patients was 64.1 ng/ml (range 8.03–877; *n* = 23), significantly higher than in retinoblastoma-free aqueous humors, with median 8.61 ng/ml (range 2.29–47.6; *n* = 12) (Figure 4E). EMMPRIN concentration in cell culture supernatants was in the range of 0.34–4.02 ng/ml (*n* = 11 cell models) (Figure 4F). The median EMMPRIN concentration in liquid biopsies of retinoblastoma patients was 3.73 ng/ml (range 0.79–19.9; *n* = 22), significantly

higher than in retinoblastoma-free samples, with median 1.17 ng/ml (range 0.04–8.31; *n* = 12) (Figure 4G).

Single cell transcriptome analysis of selected immune markers and cytokines

We used single-cell RNA-sequencing data from one treatment-naïve retinoblastoma identified as RBSC11 [2]. We annotated tumor or immune cells based on estimated copy number variation status, expression of key markers, and pathways and the reference-based annotation [31]. Among 1,198 analyzed cells, 89.2% were tumor cells, 6.3% TAMs, and 4.4% TILs (Figure 5A). Most cells overexpressed *MIF*, independently of the cell type (Figure 5B). Among the genes encoding for MIF receptors, TAMs expressed *CD74*, while both TAMs and TILs expressed *CD44* (Figure 5B). Our mapping of potential ligand–receptor

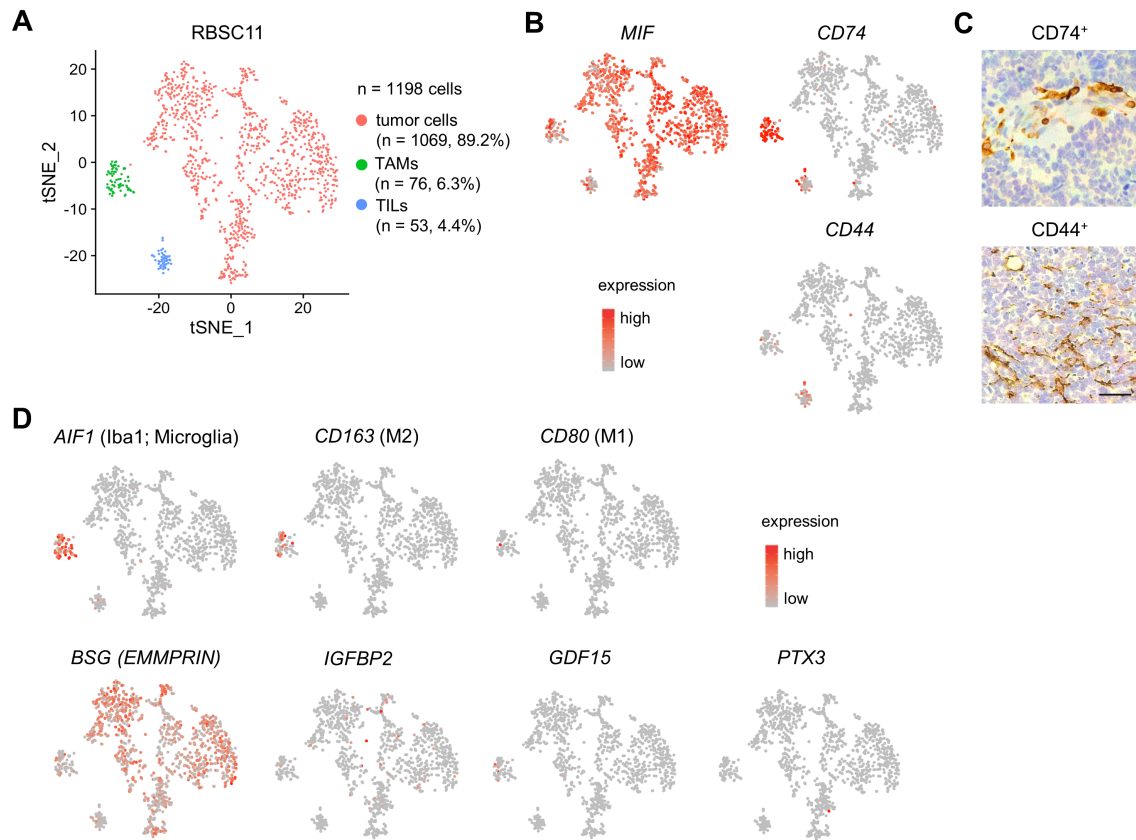


Figure 5. Single-cell sequencing of immune markers and cytokines in one retinoblastoma tumor. (A) 2D t-SNE plot of 1,198 single retinoblastoma cells from the tumor sample. Each dot represents one cell. (B) Expression of *MIF* and genes encoding for MIF receptors, *CD74* and *CD44*. (C) Immunostaining of *CD44* and *CD74* in one retinoblastoma sample. Pictures were taken at 20 \times objective magnification. Scale bar is 40 μ m. (D) Expression of genes encoding for immune markers *AIF1* (Iba1), *CD163*, and *CD80*, and secreted proteins *BSG* (EMMPRIN), *IGFBP2*, *GDF15*, and *PTX3*.

interactions at both the single-cell and bulk levels predicted that MIF/CD74 interactions are among the most significant pairs in tumor cells/macrophages (supplementary material, Figure S5A,B). Immunostaining for CD74 and CD44 in human retinoblastoma samples confirmed the gene expression profile, with only microglial-shaped cells expressing CD74 and with a more diffused staining of CD44 (Figure 5C). In the normal retinal tissue, we found CD44 diffusely expressed, while CD74 was restricted to monocytes (supplementary material, Figure S6).

The cluster of TAMs expressed the microglial marker Iba1 (gene *AIF1*) and the M2-like marker *CD163* (Figure 5D). TAMs did not express the M1-like marker *CD80* (Figure 5D). Among the other cytokines secreted by cell models, *BSG* (EMMPRIN) was expressed by a majority of tumor and immune cells, *IGFBP2* and *PTX3* were expressed by few tumor cells, and *GDF15* was restricted to a minority of TAMs (Figure 5D).

Polarizing activity of retinoblastoma-secreted factors on PBMC-derived macrophages

PBMC-derived macrophages used for polarization experiments expressed MIF receptors *CD44* and *CD74*

(supplementary material, Figure S7A) which colocalized at the cell membrane (supplementary material, Figure S7B, C). We used the macrophage marker *CD11b* staining as a quality control. We obtained a mean *CD11b* positivity of 99.3% in all the experiments. In preliminary experiments, we observed that *CD206* and *CD163* were positive in the same proportion of cells, and subsequently we used only *CD163*. In the macrophage polarization experiments, *CD163*⁺ cells were more abundant than *CD80*⁺ cells in all the studied conditions (Figure 6A). We used the proportion of both cell populations to represent data as M2-to-M1-like (i.e. *CD163*⁺-to-*CD80*⁺) ratios. The lowest M2-to-M1-like ratio corresponded to the M1-like control condition, which we used as the reference of the experiment. We observed that macrophages exposed to retinoblastoma-conditioned medium or rMIF increased their M2-to-M1-like ratio significantly, compared to the M1-like reference (Figure 6B).

Migration of macrophages through geltrex-covered transwells was significantly increased by the addition of retinoblastoma-conditioned medium to the acceptor chamber (Figure 6C,D). The addition of rMIF to the acceptor did not affect macrophage migration significantly ($p = 0.0708$).

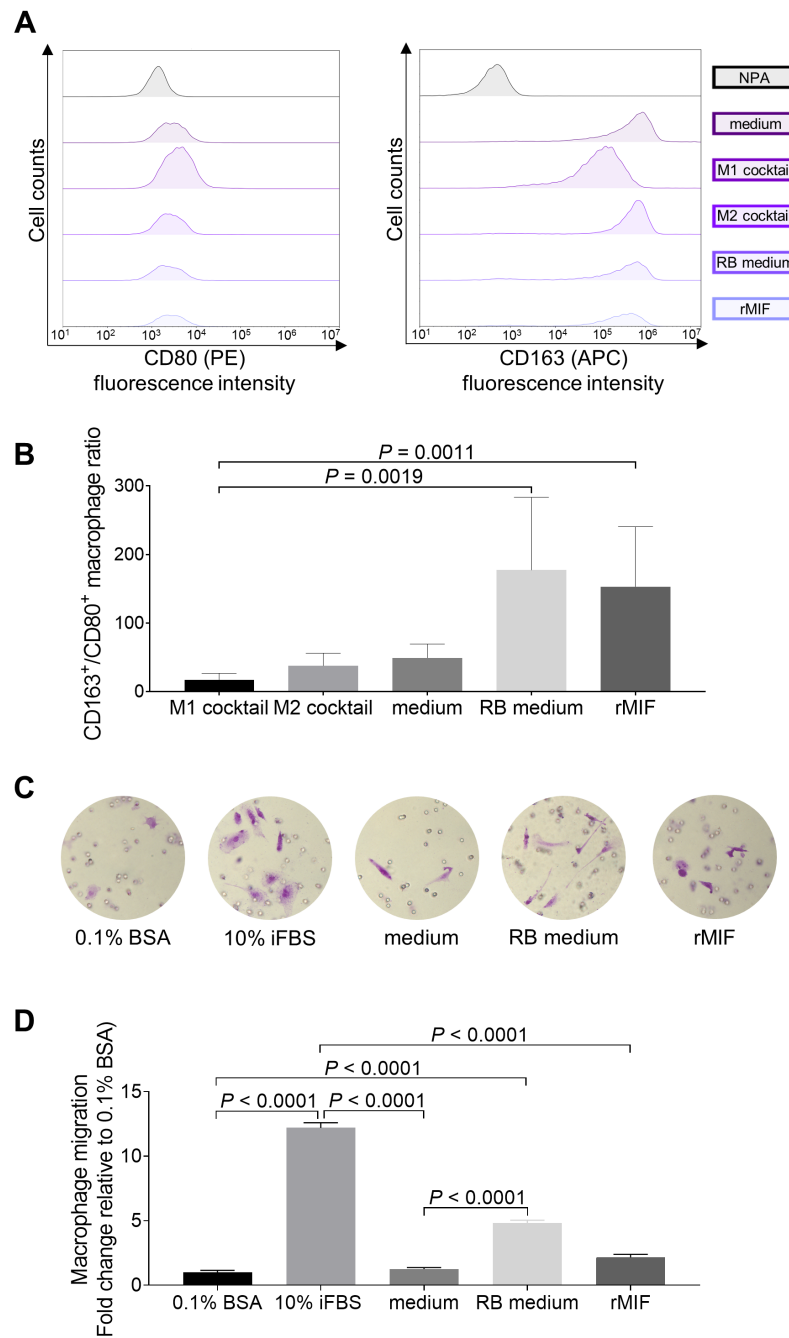


Figure 6. Polarization and migration of PBMC-derived macrophages stimulated with retinoblastoma-conditioned medium and rMIF. (A) Representative experiment showing cell counts of CD80⁺ (M1-like) (PE fluorescence emission), and CD163⁺ (M2-like), (allophycocyanin – APC – fluorescence emission), analyzed by flow cytometry. (B) CD163⁺/CD80⁺ ratios (mean and SD of eight independent experiments). (C) Representative experiment of macrophage invasion and migration in transwell assays. Cells were stained with crystal violet. (D) Migration of macrophages in each experimental condition, relativized to the migration of negative control macrophages (0.1% BSA). Data are mean and SD of 24 photographs, obtained from two independent experiments run in duplicate. NPA, nonprimary antibody; medium, non-supplemented retinoblastoma medium; M1 cocktail, medium containing M1-like polarizing cytokines; M2-like cocktail, medium containing M2-like polarizing cytokines; RB medium, retinoblastoma-conditioned medium; rMIF, recombinant MIF; 10% iFBS, inactivated fetal bovine serum; 0.1 BSA, bovine serum albumin.

Discussion

Our study demonstrates that retinoblastomas secrete soluble factors that could explain the protumoral M2-like type of retinoblastoma TAMs. We identified a homogeneous pattern in the secretome of a large set of these tumors, enriched in the proteins EMMPRIN, MIF, IGFBP-2, GDF-15, and pentraxin-3, all of them with reported activity to induce a protumoral microenvironment. For instance, MIF and IGFBP-2 have M2-like polarizing functions and recruit Tregs to solid tumors [12,14,29]; EMMPRIN increases the proportion of Tregs in T-cell cultures [28] and it is expressed in invasive retinoblastomas [32]; GDF-15 is a member of the TGF- β superfamily and inhibits macrophage surveillance during early tumor development [33]; and pentraxin-3 activates the TGF- β pathway and reduces HLA-DR and CD86 expression in human macrophages [30]. Therefore, we propose that the abundance of M2-like TAMs and the presence of Tregs in retinoblastoma, as shown in our patient cohort, is likely the result of the action of these or similar tumor cytokines in the microenvironment. Contrasting findings (lack of M2-like polarization and Tregs) in the enucleated eyes of five pediatric patients without retinoblastoma further substantiated our results.

Transcriptome analyses of patient samples, extended to the single-cell level, led us to select MIF and EMMPRIN as the leading candidates for further analyses. The significant increase of these two molecules in the aqueous humors of cancer patients might have clinical implications as biomarkers or treatment targets. The use of aqueous humor as liquid biopsy of retinoblastoma is well justified by the impossibility to perform tumor biopsy, due to the risks of extraocular dissemination associated with the surgical procedure [1]. Aqueous humor is a suitable source to obtain cell-free DNA from patients with retinoblastoma, allowing not only the diagnosis of *RB1* mutations but also identifying potential genomic markers for the risk of treatment failure [34]. The use of hybridization capture and deep sequencing of around 20,000X raw coverage further improves the sensitivity of this technique, allowing the analysis of *RB1* aberrations in cell-free DNA in blood samples [35]. However, deep sequencing techniques are generally not available outside large world-referral institutions. The selection of protein biomarkers, alternative or additional to genetic biomarkers, will increase the access to liquid biopsy analyses through simple immunoassays.

Our model prediction of MIF/CD74 as one of the top ligand/receptor pair interactions in retinoblastomas led us to study the interaction of MIF and human macrophages at a concentration within the range found in the analyzed liquid biopsies. The conclusion of our study, of M2-like polarization of macrophages by MIF, is robust because we obtained it based on eight independent experiments from eight blood donors. Conclusions regarding retinoblastoma-conditioned medium should be taken cautiously, however, because we evaluated only one tumor type, which was selected upon the analysis of the secretome heatmap and our previous preclinical work [5].

Many biological mediators regulate CD163 expression in monocytes and macrophages [36], but whether MIF-promoted expression is an actionable target with clinical meaning in retinoblastoma needs further preclinical work that was not addressed by our study. In melanoma, for instance, pharmacological inhibition of MIF signaling reduces tumor immunosuppression in tumor-bearing mice and restores the immune response to tumor cells *in vitro* [37,38]. The availability of highly selective administration routes to increase drug concentrations in the confined ocular compartment in which retinoblastoma arises might provide additional opportunities for the success of these treatments.

We observed that previous treatments did not modify the M2-like polarized CD163⁺ environment of retinoblastoma, which is in contrast with a recent study of the retinoblastoma immune environment by Miracco *et al* [8]. They reported a significantly higher number of CD163⁺ M2-like TAMs in previously treated human retinoblastoma eyes, compared to eyes enucleated without previous treatment [8]. The methodology of that study differed from ours in that they did not include ocular specimens without retinoblastoma, and we applied the multiple comparison correction to the *p* values for statistical analyses.

Retinoblastoma treatments can induce moderate inflammation, even the accumulation of multinucleated macrophages in the choroid and retina, as we previously reported in patients receiving intra-arterial chemotherapy [39]. Thus, we expected the enrichment of TAMs in treated tumors, compared to treatment-naïve tumors. However, PD-L1 expression did not change in our cohort of treated patients.

Immunostaining for PD-L1 in large collections of retinoblastoma samples shows that only a minority of tumors express PD-L1 [8,9]. In our patient cohort, we found almost all samples were PD-L1-negative, which was in agreement with our experiment showing that retinoblastoma cells express PD-L1 at low or undetectable levels that increase upon stimulation by proinflammatory molecules such as IFN- γ . Usui *et al* reported a similar finding for the Y79 cell line [40]. Because IFN- γ -secreting cells, and specifically CD8⁺ lymphocytes, were moderately abundant in a high proportion of the retinoblastoma samples, we propose that tumor microenvironment inhibits such CD8⁺ cells, leading to a decreased level of IFN- γ secretion [41]. The work by Miracco *et al*, in contrast, suggests that chemotherapy treatments induce a proinflammatory environment, leading to higher expression of PD-L1 [8], but our results and others do not confirm their observation [9]. We suggest that M2-like TAMs and Tregs suppress the activity of CD4⁺ and CD8⁺ TILs in retinoblastomas, because they inhibit the production of several proinflammatory cytokines such as IFN- γ and IL-2 [42]. Treg infiltration in our study was not homogeneous among the samples, but our IHC method might have underestimated the real number of these cells. In fact, Zhang *et al* found that around 10% of retinoblastoma TILs are FOXP3⁺ in a large set of 50 patient samples [43].

The low expression of PD-L1 in retinoblastoma suggests that few patients with confirmed PD-L1-positive

tumors would be eligible for anti-PD-L1 treatments. However, retinoblastoma biopsies for IHC remain impracticable before enucleation [1]. In addition, the use of immune checkpoint inhibitors in ocular tumors might cause difficult-to-manage inflammation due to pseudoprogession [44]. A similar concern with regard to inflammation would be raised by the use of anti-B7H3 antibodies or CAR-T cells in retinoblastomas [45,46].

Taken together, our results demonstrate that the interaction of retinoblastoma cells and TAMs through soluble factors secreted by tumor cells explains, at least partially, the cold tumor environment of retinoblastoma. Whether such tumor-secreted proteins are potential candidates for biomarkers or therapeutic modulation will be addressed in prospective work.

Acknowledgements

AMC acknowledges funding from Retos MINECO (Cure4RB project RTC-2015-4319-1) and ISCIII-FEDER (CPI13/00189, CPII18/00009, PI18/01970 and PI21/00216). We thank Xarxa de Bancs de Tumors de Catalunya (XBTC) sponsored by Pla Director d'Oncologia de Catalunya and Pediatric Tumor Bank of the Pathology Service of Garrahan Hospital (btp@garrahan.gov.ar).

Author contributions statement

MC-V and AMC conceived the study. GC, IA, FD, NC, FL, DB-R, GLC, JCM, FR and AMC provided patient samples and/or patient data. MC-V, JL, SP, RA, VB, NS, GP-P, LB-L, HC-E, OM-A and MR carried out experiments. MC-V, JL, SP, RA, NS, SP-J, FR and AMC analyzed data. PS provided key investigation materials. CL, JM and AMC provided investigation resources. MC-V and JL elaborated the figures. AMC acquired funding and supervised the study. MC-V and AMC wrote the original manuscript draft. All authors were involved in article review and editing and approved the submitted and published versions.

Data availability statement

Our array data for gene expression of top-secreted cytokines in primary retinoblastomas and fetal retina has been deposited in GEO (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58785>). The raw data for retinoblastoma RBSC11 has been deposited in the EGA database (EGAS00001005178).

References

- Munier FL, Beck-Popovic M, Chantada GL, *et al.* Conservative management of retinoblastoma: challenging orthodoxy without compromising the state of metastatic grace. "Alive, with good vision and no comorbidity". *Prog Retin Eye Res* 2019; **73**: 100764.
- Liu J, Ottaviani D, Sefta M, *et al.* A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression. *Nat Commun* 2021; **12**: 5578.
- Xu XL, Fang Y, Lee TC, *et al.* Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell* 2009; **137**: 1018–1031.
- Zhang J, Benavente CA, McEvoy J, *et al.* A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature* 2012; **481**: 329–334.
- Pascual-Pasto G, Bazan-Peregrino M, Olaciregui NG, *et al.* Therapeutic targeting of the RB1 pathway in retinoblastoma with the oncolytic adenovirus VCN-01. *Sci Transl Med* 2019; **11**: eaat9321.
- Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* 2013; **19**: 1423–1437.
- Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov* 2019; **18**: 197–218.
- Miracco C, Toti P, Gelmi MC, *et al.* Retinoblastoma is characterized by a cold, CD8+ cell poor, PD-L1- microenvironment, which turns into hot, CD8+ cell rich, PD-L1+ after chemotherapy. *Invest Ophthalmol Vis Sci* 2021; **62**: 6.
- Singh L, Singh MK, Rizvi MA, *et al.* Clinical relevance of the comparative expression of immune checkpoint markers with the clinicopathological findings in patients with primary and chemoreduced retinoblastoma. *Cancer Immunol Immunother* 2020; **69**: 1087–1099.
- Ganesan B, Parameswaran S, Sharma A, *et al.* Clinical relevance of B7H3 expression in retinoblastoma. *Sci Rep* 2020; **10**: 10185.
- Ge Z, Ding S. The crosstalk between tumor-associated macrophages (TAMs) and tumor cells and the corresponding targeted therapy. *Front Oncol* 2020; **10**: 590941.
- Choi S, Kim HR, Leng L, *et al.* Role of macrophage migration inhibitory factor in the regulatory T cell response of tumor-bearing mice. *J Immunol* 2012; **189**: 3905–3913.
- Abe R, Peng T, Sailors J, *et al.* Regulation of the CTL response by macrophage migration inhibitory factor. *J Immunol* 2001; **166**: 747–753.
- Ghoochani A, Schwarz MA, Yakubov E, *et al.* MIF-CD74 signaling impedes microglial M1 polarization and facilitates brain tumorigenesis. *Oncogene* 2016; **35**: 6246–6261.
- Ladányi A, Tímár J. Immunologic and immunogenomic aspects of tumor progression. *Semin Cancer Biol* 2020; **60**: 249–261.
- Thomas DA, Massagué J. TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell* 2005; **8**: 369–380.
- Uytendhoeve C, Pilotte L, Théate I, *et al.* Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2,3-dioxygenase. *Nat Med* 2003; **9**: 1269–1274.
- Porta C, Consonni FM, Morlacchi S, *et al.* Tumor-derived prostaglandin E2 promotes p50 NF-κB-dependent differentiation of monocytic MDSCs. *Cancer Res* 2020; **80**: 2874–2888.
- Francis JH, Abramson DH, Gaillard MC, *et al.* The classification of vitreous seeds in retinoblastoma and response to intravitreal melphalan. *Ophthalmology* 2015; **122**: 1173–1179.
- Galardi A, Colletti M, Lavarello C, *et al.* Proteomic profiling of retinoblastoma-derived exosomes reveals potential biomarkers of vitreous seeding. *Cancers (Basel)* 2020; **12**: 1555.
- Berry JL, Xu L, Murphree AL, *et al.* Potential of aqueous humor as a surrogate tumor biopsy for retinoblastoma. *JAMA Ophthalmol* 2017; **135**: 1221–1230.
- Munier FL, Gaillard MC, Balmer A, *et al.* Intravitreal chemotherapy for vitreous disease in retinoblastoma revisited: from prohibition to conditional indications. *Br J Ophthalmol* 2012; **96**: 1078–1083.
- Ma B, Lei X, Guan Y, *et al.* Maintenance of retinal cancer stem cell-like properties through long-term serum-free culture from human retinoblastoma. *Oncol Rep* 2011; **26**: 135–143.

24. Pascual-Pasto G, Olaciregui NG, Vila-Ubach M, et al. Preclinical platform of retinoblastoma xenografts recapitulating human disease and molecular markers of dissemination. *Cancer Lett* 2016; **380**: 10–19.
25. Pascual-Pasto G, Olaciregui NG, Opezzo JAW, et al. Increased delivery of chemotherapy to the vitreous by inhibition of the blood-retinal barrier. *J Control Release* 2017; **264**: 34–44.
26. Asnaghi L, White DT, Key N, et al. ACVR1C/SMAD2 signaling promotes invasion and growth in retinoblastoma. *Oncogene* 2019; **38**: 2056–2075.
27. Winter U, Ganiewich D, Ottaviani D, et al. Genomic and transcriptomic tumor heterogeneity in bilateral retinoblastoma. *JAMA Ophthalmol* 2020; **138**: 569–574.
28. Li X, Zhang Y, Ma W, et al. Enhanced glucose metabolism mediated by CD147 contributes to immunosuppression in hepatocellular carcinoma. *Cancer Immunol Immunother* 2020; **69**: 535–548.
29. Sun L, Zhang X, Song Q, et al. IGFBP2 promotes tumor progression by inducing alternative polarization of macrophages in pancreatic ductal adenocarcinoma through the STAT3 pathway. *Cancer Lett* 2021; **500**: 132–146.
30. Shiraki A, Kotooka N, Komoda H, et al. Pentraxin-3 regulates the inflammatory activity of macrophages. *Biochem Biophys Res* 2016; **5**: 290–295.
31. Aran D, Looney AP, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 2019; **20**: 163–172.
32. Adithi M, Nalini V, Kandam M, et al. Expression of matrix metalloproteinases and their inhibitors in retinoblastoma. *J Pediatr Hematol Oncol* 2007; **29**: 399–405.
33. Ratnam NM, Peterson JM, Talbert EE, et al. NF- κ B regulates GDF-15 to suppress macrophage surveillance during early tumor development. *J Clin Invest* 2017; **127**: 3796–3809.
34. Berry JL, Xu L, Kooi I, et al. Genomic cfDNA analysis of aqueous humor in retinoblastoma predicts eye salvage: the surrogate tumor biopsy for retinoblastoma. *Mol Cancer Res* 2018; **16**: 1701–1712.
35. Abramson DH, Mandelker D, Francis JH, et al. Retrospective evaluation of somatic alterations in cell-free DNA from blood in retinoblastoma. *Ophthalmol Sci* 2021; **1**: 100015.
36. Buechler C, Ritter M, Ors   E, et al. Regulation of scavenger receptor CD163 expression in human monocytes and macrophages by pro- and antiinflammatory stimuli. *J Leukoc Biol* 2000; **67**: 97–103.
37. Yaddanapudi K, Putty K, Rendon BE, et al. Control of tumor-associated macrophage alternative activation by macrophage migration inhibitory factor. *J Immunol* 2013; **190**: 2984–2993.
38. Figueiredo CR, Azevedo RA, Mousdell S, et al. Blockade of MIF-CD74 signalling on macrophages and dendritic cells restores the anti-tumour immune response against metastatic melanoma. *Front Immunol* 2018; **9**: 1132.
39. Parareda A, Catal   J, Carcaboso AM, et al. Intra-arterial chemotherapy for retinoblastoma. Challenges of a prospective study. *Acta Ophthalmol* 2014; **92**: 209–215.
40. Usui Y, Okunuki Y, Hattori T, et al. Expression of costimulatory molecules on human retinoblastoma cells Y-79: functional expression of CD40 and B7H1. *Invest Ophthalmol Vis Sci* 2006; **47**: 4607–4613.
41. Jiang Y, Li Y, Zhu B. T-cell exhaustion in the tumor microenvironment. *Cell Death Dis* 2015; **6**: e1792.
42. Curiel TJ, Coukos G, Zou L, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med* 2004; **10**: 942–949.
43. Zhang R, Song Y-N, Duo X, et al. Retinoblastoma cell-derived Twist protein promotes regulatory T cell development. *Cancer Immunol Immunother* 2021; **70**: 1037–1048.
44. Garcia GA, Topping KL, Mruthyunjaya P, et al. Pseudoprogression of metastatic melanoma to the orbit with pembrolizumab. *Ophthalmol Plast Reconstr Surg* 2020; **36**: e36–e40.
45. Majzner RG, Theruvath JL, Nellan A, et al. CAR T cells targeting B7-H3, a pan-cancer antigen, demonstrate potent preclinical activity against pediatric solid tumors and brain tumors. *Clin Cancer Res* 2019; **25**: 2560–2574.
46. Du H, Hirabayashi K, Ahn S, et al. Antitumor responses in the absence of toxicity in solid tumors by targeting B7-H3 via chimeric antigen receptor T cells. *Cancer Cell* 2019; **35**: 221, .e8–237.
47. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} method. *Methods* 2001; **25**: 402–408.
48. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; **43**: e47.
49. Efremova M, Vento-Tormo M, Teichmann SA, et al. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020; **15**: 1484–1506.
50. No  l F, Massenet-Regad L, Carmi-Levy I, et al. Dissection of inter-cellular communication using the transcriptome-based framework ICELLNET. *Nat Commun* 2021; **12**: 1089.

References 47–50 are cited only in the supplementary material.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods

Figure S1. Immunostaining for CRX (tumor marker) and CD163 (M2-like marker) in retinoblastoma cell lines

Figure S2. Boxplots representing average log₂ gene expression levels of *PDCD1* (PD-1), *CD274* (PD-L1) and *CD276* (B7-H3) in retinoblastoma tumors, cell models and fetal retina analyzed in the gene expression microarray

Figure S3. Expression of HLA class I molecules in retinoblastomas

Figure S4. mRNA expression levels of different immune escape mechanism in retinoblastoma primary cell cultures and Y79 commercial cell line

Figure S5. Prediction of ligand-receptor interactions in retinoblastoma tumors

Figure S6. Immunostaining of CD44 and CD74 in conserved retinal areas of one retinoblastoma patient (Retinoblastoma patient eye) and in the retina of one control eye

Figure S7. Characterization of CD74 and CD44 (MIF receptors) in PBMC-derived macrophages

Table S1. Short tandem repeat analysis of retinoblastoma primary cell lines (provided as an Excel file)

Table S2. Primary antibodies used in IHC studies: host species, commercial references, antigen retrieval procedures, antibody dilutions, staining methods and positive control samples

Table S3. Primary and secondary antibodies used for immunofluorescence studies: commercial references and antibody dilutions

Table S4. Primers used for RT-qPCR assays, including references of the TaqMan probes

Table S5. Clinical and anatomopathological features of the retinoblastoma samples

Table S6. Clinicopathological data of the control samples

Table S7. Infiltration of TILs, TAMs and microglial cells in retinoblastoma eyes, in both tumor and conserved retinal areas, and in control eyes (non-retinoblastoma) retinas

Table S8. Infiltration of TILs analyzed according to the clinical and anatomopathological features

Table S9. Infiltration of TAMs and microglia cells analyzed according to the clinical and anatomopathological features

Titre : Hétérogénéité du rétinoblastome : une analyse multi-omique et transcriptomique de cellules uniques

Mots clés : Rétinoblastome, hétérogénéité, microenvironnement, bioinformatique, multi-omique, transcriptome de cellules uniques

Résumé : Le rétinoblastome est un cancer pédiatrique dérivé de la rétine. Bien que rare, c'est la tumeur maligne intraoculaire la plus fréquente chez l'enfant. Le rétinoblastome est traité par une thérapie locale aux premiers stades, mais aux stades ultérieurs, il nécessite une chirurgie d'énucleation et une chimiothérapie systémique. Sans un diagnostic en temps utile et un traitement approprié, des métastases peuvent se développer et entraîner le décès de l'enfant.

Le rétinoblastome est une maladie hétérogène. Les cellules tumorales de différents patients ou différentes cellules d'un même patient peuvent présenter des caractéristiques moléculaires et phénotypiques distinctes. D'un point de vue histopathologique, les cellules tumorales peuvent présenter différents degrés de différenciation, ainsi qu'une croissance exophytique, endophytique ou mixte. En ce qui concerne la génomique, le rétinoblastome survient principalement après une inactivation bi-allélique de *RB1* et, dans de rares cas, une amplification de *MYCN* peut également déclencher la maladie. En termes de transcriptomique, quelques études ont été réalisées et ont révélé que le rétinoblastome peut présenter des degrés variés dans les signatures de différenciation des photorécepteurs.

Dans notre travail, nous avons identifié deux sous-types moléculaires basés sur l'analyse de 102 rétinoblastomes en utilisant le séquençage de l'exome entier, les SNP array, les puces pour mesurer l'expression des gènes et la méthylation de l'ADN. Nous avons réalisé le clustering en utilisant une stratégie combinant le clustering hiérarchique consensuel et le clustering basé sur les centroïdes. Nous avons démontré que les tumeurs de sous-type 1 présentaient une signature plus mature de différenciation de photorécepteurs cône et se manifestaient généralement chez les enfants de moins de 18 mois, tandis qu'un niveau plus faible de différenciation de photorécepteurs

cône est corrélé à un niveau élevé de caractéristiques ganglionnaires/neurales, de signatures de type souche et d'instabilité génomique dans les tumeurs de sous-type 2, que l'on retrouve davantage chez les enfants de plus de 18 mois. L'analyse d'une série indépendante de 112 rétinoblastomes présentant des facteurs pathologiques à haut risque a révélé que les tumeurs qui métastasaient expriment toutes la protéine TFF1, correspondant au gène le plus surexprimé dans les tumeurs de sous-type 2 par rapport aux tumeurs de sous-type 1. Nous avons analysé 14 rétinoblastomes en utilisant le séquençage de l'ARN de cellules uniques (scRNA-seq), et avons confirmé cette hétérogénéité inter-tumorale. L'une des tumeurs analysées présentait une hétérogénéité intra-tumorale au niveau phénotypique et génomique : certaines cellules tumorales présentaient un degré plus élevé de différenciation des cônes et la perte de différenciation des cônes s'accompagnait d'une accumulation d'altérations génomiques, tandis que d'autres cellules présentaient des altérations génomiques et des phénotypes totalement différents. Nous avons décrit le paysage immunitaire du rétinoblastome grâce au scRNA-seq, et avons découvert que différents types de cellules immunitaires sont présents dans le microenvironnement tumoral, notamment de multiples populations de cellules de la lignée monocyttaire telles que les macrophages M2 protumoraux et les cellules présentatrices d'antigènes, de multiples populations de cellules T dont les cellules T régulatrices CD4+ et les cellules T cytotoxiques CD8+, ainsi que les cellules NK. L'infiltration des macrophages M2 a été validée par immunohistochimie et a été associée à un niveau élevé d'expression du MIF par une analyse sur puces à cytokines et d'une prédiction *in silico* de ligand-récepteur. En conclusion, notre analyse multi-omique et transcriptomique de cellules uniques a permis de caractériser de manière détaillée l'hétérogénéité inter- et intra-tumorale et de décrire le paysage immunitaire du rétinoblastome.

Title : Retinoblastoma heterogeneity: a comprehensive multi-omic and single-cell transcriptomic analysis

Keywords : Retinoblastoma, heterogeneity, microenvironment, bioinformatics, multi-omics, single-cell transcriptome

Abstract : Retinoblastoma is a pediatric cancer derived from the retina. The global annual incidence is around 8000 new cases. Although rare, it is the most common intraocular malignancy in children. Retinoblastoma is treated with local therapy at early stages, but at later stages, it requires enucleation surgery and systemic chemotherapy. Without timely diagnosis and proper treatment, metastasis can develop and lead to lethal diseases.

Retinoblastoma is a heterogeneous disease. Tumor cells in different patients or different cells in one patient can exhibit distinct molecular and phenotypic characteristics. From the histopathological perspective, tumor cells can present differentiation at diverse levels, and exhibit exophytic, endophytic or mixed growth. With regard to genomics, retinoblastoma arises predominantly after bi-allelic *RB1* inactivation and in rare cases, *MYCN* amplification may also initiate the disease. Other mutations and genomic alterations may also contribute to the progression of the retinoblastoma such as *BCOR* mutations, chromosome 1q gain, 6p gain, or 16q loss. In terms of transcriptomics, a few studies were performed and revealed that retinoblastoma can display varied degrees in photoreceptor differentiation signatures.

In our work, we identified two molecular subtypes based on the analyses of a series of 102 retinoblastoma using whole-exome sequencing, SNP array, gene expression microarray and DNA methylation array. We made the clustering using a method combining consensus hierarchical clustering and centroid based clustering. We demonstrated that subtype 1 tumors exhibited a more matured cone differentiation signature and were usually manifested in children less than 18 months, while subtype 2

tumors were found more frequent in children more than 18 months, exhibited lower level of cone differentiation, elevated level of ganglion/neuronal features, stemness signatures, and genomic instability. Analysis of an independent series of 112 retinoblastomas with high risk pathological factors uncovered that metastatic tumors all expressed TFF1 protein at their primary sites, the gene that is most upregulated in subtype 2 tumors as compared to subtype 1 tumors. We confirmed this inter-tumoral heterogeneity in 14 additional retinoblastomas using single-cell RNA sequencing. One of the tumors analyzed displayed intra-tumoral heterogeneity at both phenotypic and genomic level, that some tumor cells exhibited higher grade of cone differentiation, the loss of cone differentiation was accompanied with the accumulation of genomic alterations, while other cells presented entirely different genomic alterations and phenotypes. We characterized the immune landscape of retinoblastoma through single-cell transcriptomics, and reported that various immune cells types were presented in the tumor microenvironment, including multiple populations of monocytic lineage cells such as the protumoral alternative M2 macrophages and antigen presenting cells, multiple populations of T cells such as CD4+ regulatory T cells and CD8+ cytotoxic T cells, as well as NK cells. The infiltration of M2 macrophages was validated by immunohistochemistry, and was associated with high-level of MIF expression through cytokine array and *in silico* ligand-receptor prediction.

Taken together, our multi-omic and single-cell transcriptomic analysis comprehensively characterized the inter- and intra- tumoral heterogeneity and characterized the immune landscape in retinoblastoma.