

Applying Single Cell RNA-Seq to Dissect a Phenotype

BY

MAJD M ARISS

B.S., American University of Beirut 2013

M.S., Georgetown University 2015

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry and Molecular Genetics
in the Graduate College of the
University of Illinois at Chicago, 2020

Chicago, Illinois

Dissertation Committee:

Maxim Frolov, Chair and Advisor
Bradley Merrill
Nissim Hay
Alisa Katzen
Teresa Vales Orenic, Biological Sciences

DEDICATION

This thesis is dedicated to my family, without whom it would have never been accomplished.

ACKNOWLEDGEMENTS

I would like to acknowledge my thesis committee Drs. Bradley Merrill, Nissim Hay, Alisa Katzen, Teresa Vales Orenic, for their constant constructive guidance throughout my PhD training. I would like to thank Dr. Maxim Frolov for mentoring me over the last four years. He saw the potential to entrust me with a very challenging project, drove me to achieve excellence and precision while also being incredibly patient. Without his guidance, my next career steps would have not been possible, and I will forever be thankful for his mentorship.

I would like to thank Dr. Abul BMMK Islam for his extensive collaboration throughout my PhD which has been instrumental in the work in this thesis and teaching me to run the computational analyses. Current and previous lab members Alexandre Aissa, Maria Paula Zappia, James Kwon, Alexandra Rader, Meg Critcher, and Alice Rogers, provided technical troubleshooting.

I would like to thank GEMS students and close friends Isha, Tara, Sean, James, Josh, Meg, Afif, Ramzi, Hassib, Hasan, Alice, Alexandra, Alexandre, Milica, Katarina, Alexa, Faisal, Nader, Yves, and my cat Fish for keeping me sane and providing the emotional support during my time in Chicago. Finally, I thank my mother, sister, father and brother for all the love, support, and being my rock throughout my PhD. I would not have made it through without my friends, family, and cat.

CONTRIBUTION OF AUTHORS

Chapter II is from published work in *Nature Communications*:

- Ariss, M. M., Islam, A. B., Critcher, M., Zappia, M. P., & Frolov, M. V. (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nature communications*, 9(1), 1-13. Copyright: “Authors have the right to reuse their article’s Version of Record, in whole or in part, in their own thesis. Additionally, they may reproduce and make available their thesis, including Springer Nature content, as required by their awarding academic institution.” <https://www.nature.com/nature-research/reprints-and-permissions/permissions-requests>

Chapter III is work in review at *eLife* on the use of single cell RNA-seq to find Ama as a novel component of the receptor tyrosine kinase signaling pathway and the blood-brain barrier.

- Ariss, M. M., Terry A. R., Islam, A. B., Hay, N., & Frolov M. V. (submitted 2020). Amalgam regulates the receptor tyrosine kinase pathway through Sprouty in glial cell development. Copyright: eliFe grants permission to “copy, distribute, and reuse them (in part or in full), without needing to seek permission, as long as the author and original source are properly cited.” <https://reviewer.elifesciences.org/author-guide/journal-policies>

UIC Genome Research Core assisted in sequencing the single cell RNA-seq libraries throughout all this work with the help of Dr. Stefan J Green and Kevin Kunstman. Dr. Abul BMMK Islam established, aligned the sequencing data, and supervised the bioinformatics analyses. In **Chapter II**, Meg Critcher helped performed immunostainings and Dcp-1 quantifications. Dr. Maria Paula Zappia performed the ChIP-qPCR assay. In **Chapter III**, Alex Terry performed the human cell lines experiments, lentiviral production and transduction, and immunoblots.

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
I. INTRODUCTION.....	1
A. Cell-type specific responses to the Retinoblastoma pathway.....	1
B. The context dependent response of Sprouty.....	7
C. Goals of this thesis.....	10
D. Cited literatures.....	14
II. SINGLE CELL RNA-SEQUENCING IDENTIFIES A METABOLIC ASPECT OF APOPTOSIS IN <i>Rbf</i> MUTANT.....	19
A. Summary.....	19
B. Introduction.....	19
C. Results.....	21
1. A cell atlas of the wild-type third-instar larval eye disc.....	21
2. Developing photoreceptors undergo a transcriptional switch.....	31
3. scRNA-seq identified an <i>Rbf</i> mutant-specific cell cluster.....	35
4. Intracellular acidification increases apoptosis in <i>Rbf</i> ^{d20a}	39
D. Discussion.....	43
E. Materials and Methods.....	45
1. Fly stocks.....	45
2. Fluorescent in situ hybridization.....	47
3. DIG-labeled probes.....	47
4. Chromatin immunoprecipitation–quantitative polymerase chain reaction.....	48
5. Immunofluorescence.....	50
6. BrdU labeling.....	50
7. Cleaved Dcp-1 quantification.....	51
8. Live pH staining.....	51
9. Tissue dissociation for Drop-seq.....	52
10. Drop-seq.....	52
11. scRNA-seq data analysis.....	52
12. Cell clustering and discovery of cell types.....	53
13. Gene set enrichment analysis.....	54
14. Monocle 2 trajectory.....	55
F. Cited literatures.....	55
III. AMALGAM REGULATES THE RECEPTOR TYROSINE KINASE PATHWAY THROUGH SPROUTY IN GLIAL CELL DEVELOPMENT.....	59
A. Summary.....	59
B. Intriduction.....	59
C. Results.....	61
1. <i>Ama</i> is required for glial cell development.....	61
2. Profiling of <i>Ama</i> depleted glia by scRNA-seq.....	66

TABLE OF CONTENTS (continued)

<u>CHAPTER</u>	<u>PAGE</u>
3. Blood brain barrier is disrupted by Ama depletion.....	75
4. Ama depletion reduces RTK signaling.....	80
5. Ama affects RTK signaling through Sprouty.....	86
6. Ama knockdown suppresses neoplasia in a Drosophila glioma model.....	89
7. Knockdown of Ama ortholog, LSAMP, increases SPROUTY2 in human glioblastoma cell lines	91
D. Discussion.....	94
E. Materials and Methods.....	97
1. Fly stocks.....	97
2. Ama transgene.....	99
3. Fluorescent in situ hybridization.....	99
4. Immunofluorescence.....	100
5. Heat-shock treatment.....	101
6. Dextran labeling.....	101
7. Tissue dissociation and Drop-seq.....	101
8. scRNA-seq alignment.....	102
9. Cell clustering and single cell analysis.....	103
10. Human cell lines.....	103
11. Plasmids and lentiviral production and transduction.....	104
12. Immunoblotting.....	104
13. TCGA (cBioportal).....	105
F. Cited literatures.....	106
IV. CONCLUSIONS.....	112
A. The RB pathway in glycolysis and pH dysregulation.....	112
B. Placing Ama in the RTK pathway.....	116
C. scRNA-seq is a storm of information.....	119
1. Establishing a single cell wild type atlas.....	119
2. Analyzing the single cell perturbations between genotypes.....	121
D. Cited literatures.....	123
APPENDICES.....	125
A. Nature permission of inclusion.....	125
B. eLife permission of inclusion.....	126
VITA.....	127

LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE</u>
1. The Rb pathway.....	6
2. Mixed species barnyard plot.....	13
3. A cell atlas of the wild-type third-instar larval eye disc.....	27
4. A heat map showing markers across cell populations.....	28
5. Single-cell RNA-seq identifies cell-specific markers.....	29
6. Photoreceptors undergo a transcriptional switch at onset of axonogenesis.....	33
7. Wild-type and <i>Rbf</i> ^{d20a} scRNA-seq identifies a mutant-specific metabolic cluster..	37
8. Upregulation of <i>Ldh</i> , <i>Ald</i> and <i>HIF1A</i> make <i>Rbf</i> mutant prone to apoptosis.....	41
9. <i>Ama</i> is expressed in glia and is required in glial development.....	64
10. scRNA-seq identifies cellular perturbations following <i>Ama</i> knockdown.....	69
11. scRNA-seq supervised analysis in glial cells from <i>repo</i> >+ brains reveal heterogeneity in glia.....	71
12. Supervised scRNA-seq analysis using in <i>repo</i> >+ and <i>repo</i> > <i>Ama</i> ^{RNAi} glia identifies distinct cellular perturbations in gene expression.....	74
13. <i>Ama</i> depletion decreases glial cell proliferation and disrupts the BBB.....	78
14. <i>Ama</i> knockdown decreases RTK signaling.....	82
15. <i>Ama</i> affects Sty in the brain.....	84
16. <i>Ama</i> over expression in the eye affects RTK signaling.....	85
17. Sty knockdown largely rescues the phenotype of <i>Ama</i> depletion.....	88
18. <i>Ama</i> depletion suppresses neoplastic growth in Drosophila glioma model.....	90
19. The relationship between <i>Ama</i> and Sprouty is conserved in human glioblastoma cell lines.....	93
20. Illustration of the molecular mechanism of the metabolic aspect of apoptosis in <i>Rbf</i> mutant cells.....	114
21. ROS levels in <i>Rbf</i> mutant eye discs.....	115
22. <i>Ama</i> affects RTK signaling.....	118

LIST OF ABBREVIATIONS

Ald	Aldolase
Ama	Amalgam
BBB	Blood-Brain Barrier
ChIP-qPCR	Chromatin Immunoprecipitation Quantitative Polymerase Chain Reaction
Dcp-1	Death caspase-1
Dp110	Phosphatidylinositol 3-kinase 92E
E2F	E2-Promoter Binding Protein
EGFR	Epidermal Growth Factor Receptor
Elav	Embryonic Lethal Abnormal Vision
FLP	Flippase
GBM	Glioblastoma Multiforme
GMR	Glass Multiple Reporter
hid	head involution defective
HIF1A	Hypoxia Inducible Factor 1 Subunit Alpha
Ldh	Lactate Dehydrogenase
LSAMP	Limbic System Associated Membrane Protein
MF	Morphogenetic Furrow
P-ERK	Phospho-Extracellular Signal-Regulated Kinase
Pnt	Pointed
Rb	Retinoblastoma
Repo	Reversed Polarity
RNAi	RNA interference
RTK	Receptor Tyrosine Kinase
scRNA-seq	Single Cell RNA-seq
Sty	Sprouty
tSNE	t-Stochastic Neighbor Embedding
UAS	Upstream Activating Sequences
UMAP	Uniform Manifold Approximation and Projection

SUMMARY

Recent technological advances in single cell RNA-sequencing (scRNA-seq) have enabled the study of complex cellular dynamics, including the identification of accurate transcriptomic gene signatures in a multitude cell types. Its superior resolution has facilitated the discovery of novel sub-cell types, biomarkers, and intricate cellular differentiation processes which were undetectable using formerly conventional techniques such as RNA-seq. In this thesis, I employ scRNA-seq to create two comprehensive cell atlases of the *Drosophila* developing eye and glia during the late third instar larval stage. After resolving the cellular heterogeneity and dynamics that characterize wild type eye tissues, I subjected *Rbf* (Retinoblastoma-family protein) mutant eyes and *Amalgam* (*Ama*) depleted brains to scRNA-seq in order to identify the cellular transcriptomic perturbations in their respective contexts. These findings led to the discovery of novel players in the Retinoblastoma and Receptor Tyrosine Kinase (RTK) pathways. Finally, my analysis of the aforementioned mutant and knockdown phenotypes demonstrates the benefits of scRNA-seq and emphasizes its value in uncovering the molecular mechanisms that drive cellular transcriptomic discrepancies.

I. INTRODUCTION

A. Cell-type specific responses to the Retinoblastoma pathway

The retinoblastoma protein (pRB) was the first tumor suppressor, a protein whose inactivation indirectly promotes cancer formation, ever discovered. The RB family is comprised of three proteins pRb, p107, and p130, that are continuously post-translationally modified throughout the cell cycle. In senescent cells, unphosphorylated pRB binds to E2F transcription factors (TF) and represses the E2F/DP transcriptional activity (Dyson, 2016). When pRB is phosphorylated by CyclinD/CDK4 or CDK6, CyclinE/CDK2, and CyclinA/CDK2, it dissociates from E2F, resulting in the de-repression of gene transcription that promotes G1/S cell cycle progression such as of CCNE1, MYB, and CDK2 (Kitajima & Takahashi, 2020). Additionally, the E2F can promote expression of genes involved in DNA synthesis such as PCNA, MCM2-7, and CDC6 during the S cell cycle phase (Dyson, 2016). In the RB pathway, CDKN2A (p16^{INK4a}), CDKN1A (p21^{CIP1}), and CDKN1B (p27^{Kip1}), act as tumor suppressors by binding to Cyclin/CDK complexes to inhibit the catalytic domains of CDKs and to prevent RB phosphorylation (Figure 1). Consequently, the RB pathway regulates proliferation and its inactivation is a crucial event in cancer (Burkhart et al., 2019, Kitajima & Takahashi, 2020).

This basic approach to the RB pathways makes the assumption that pRB functions exclusively as an E2F/DP inhibitor and understates its versatility and complex activity depending on cellular context. It has, in fact, been known for some time that RB has cell and context specific functions. For example, E2F1 has different mechanisms of action in estrogen receptor-positive breast cancer cells and its expression is regulated in part by hormone signaling (Ngwenya & Safe, S. 2003). This shows that E2Fs can be regulated outside of the RB pathway in specific cell types. RB-E2F was also shown to act as a transducer of TGF β signaling by forming a complex with Smad cofactors, resulting in inhibition of the growth-promoting gene *c-myc*. This complex between RB and Smad explains why

TGF β presents different contextual phenotypes, acting as a tumor suppressor in some epithelial cells and as an oncogene in fibroblasts (Chen et al., 2002). Furthermore, early work in *Drosophila* shows that E2F1 can be both pro- and anti-apoptotic based on the position of the cells in the wing disc (Moon et al., 2005). The E2F-independent roles of RB are reflected in recent work in *RB* mutant mice. Endogenous mutation of RB that prevents its binding to the transactivation domain of E2F proteins (*RB1^G*) resulted in accelerated cell cycle entry with no effect on novel tumor formation or overgrowth in a *Cdkn1a* and *Kras^{G12D}* mutant background (Thwaites et al., 2019). However, dual mutation in *RB1^G* and *Trp53* or *Cdkn1b* resulted in enhanced cancer susceptibility and formation, respectively (Thwaites et al., 2019). As inactivation of RB occurs in many cancers, it is essential to study RB pathway activity in a variety of cell types and observe its function in different intra- and intercellular contexts.

Retinoblastoma is the most common pediatric intraocular cancer with most cases occurring under the age of 3 (Li & Zhang, 2020). Mutations in RB1 are sufficient to drive retinoblastoma pathogenesis and they display distinct contextual phenotypes due to cell specific molecular circuitry in different cell types. The Cobrinik group found that retinoblastoma originates from cone cells since they express cone cell markers RXR γ and TR β 2. They recently revealed that RB mutations affect maturing normal human cone cells expressing ARR3, since they display intrinsically high levels of MDM2, thus inhibiting the p53-induced apoptotic signal and driving pathogenesis (Singh et al., 2018). The Dyer group, on the other hand, found that retinoblastoma originates from mature horizontal cells since RB inactivation drives them to re-enter cell cycle. These cells displayed MDM4 amplifications which similarly inhibits the p53 pathways and apoptosis (Dyer, 2016). In light of these findings, histological studies classified retinoblastoma into two types: Homer-Wright (HW) well differentiated, and Flexner-Wintersteiner (FW) poorly differentiated (Singh & Kashyap, 2018). Differentiated HW tumors display over 50% normal rosette-like arrangement, relatively low

malignancy, small nuclear to cytoplasm ratios, and are insensitive to radiotherapy. Poorly differentiated FW on the other hand have less than 50% rosette features, a relatively higher malignancy, a barely noticeable cytoplasm, and are sensitive to radiotherapy (Li & Zhang, 2020). The tumors originating from cone cells are classified as FW (Singh et al., 2018), while horizontal cells re-entering cell cycle maintain their differentiation features (Dyer, 2016) suggesting that the horizontal cell of origin in retinoblastoma are likely associated with HW. These results further underly the complexity of the RB pathway in driving different cell-type specific responses in the same cancer.

Recently, high throughput methods such as RNA-seq, microarrays, and high resolution Single Nucleotide Polymorphism (SNP) arrays were used to study retinoblastoma tumors (Elchuri et al., 2018). Although these techniques were able to detect gains of specific chromosomal regions, changes in long non-coding RNA, and alternate splicing events, they failed to explain the large genetic variations that accompany different types of retinoblastoma (Elchuri et al., 2018). In summation, the findings and inadequacies of previous and current work implore the field to develop new methods for analyzing cell-specific phenotypes and to characterize the molecular mechanisms that are altered following RB inactivation.

Drosophila presents a robust and simplified organism in which to study the Rb pathways as it is highly conserved in mammals. Flies have less genetic redundancy than humans, so they provide a highly malleable and streamlined model, which enables a comprehensive approach to study the role of a protein or a gene family (Yamaguchi & Yoshida, 2018). For example, all three RB and eight E2F (E2F1-E2F8) human genes are simplified into two Rbf1 and Rbf2 and two dE2F1 and dE2F2 in *Drosophila* (Van Den Heuvel & Dyson, 2008).

The late third instar larval eye disc in *Drosophila* offers a unique framework for examining neuronal differentiation, as a physical indentation called the morphogenetic furrow (MF)

continuously specifies photoreceptors (PR) in a synchronized manner. The highly heterogeneous tissue is comprised of undifferentiated eye progenitor cells, transitioning cells anterior to the MF, differentiating cells in the MF, and differentiated photoreceptors in the posterior compartment (Moses, 2002). At this developmental stage, the six core transcription factors that establish the eye field are expressed in different compartments of the eye: *toy*, *ey*, and *optix* are expressed anterior to the MF, *dac* is upregulated around the MF, whereas *so* and *eya* initiate expression around the MF and remain upregulated in the posterior compartment of the eye disc. These eye specification genes form a gene regulatory network and regulate each other. The network is initially driven by Toy and Ey to eventually express So and Eya with the latter two forming a complex that activates multiple targets important for retinal development.

The MF is initiated following a short-range signal of Hedgehog which causes a long-range Dpp signal that induces PR differentiation in the posterior compartment. This Dpp signal is inhibited and kept in check in the anterior compartment by Wingless. Once initiated, the MF shapes multiple rosettes of proneural cells that express Atonal which drive neurogenesis. Each rosette cluster resolves to select out only one founder R8 photoreceptor cell. When the R8 founder photoreceptor is recruited, the remaining PR cells are recruited in a stepwise manner R2+R5->R3+R4->R1+R6->R7 through successive rounds of Notch and EGFR signaling (Cagan & Reh, 2010). During the pupal stage, each ommatidium's outer PRs (R1-R6) initiate expression of rhodopsin Rh1 which is sensitive to dim light. The inner PRs R7 and R8 are sensitive to UV and blue light respectively. R7 expresses rhodopsins Rh3 and Rh4 whereas R8 expresses rhodopsins Rh5 and Rh6.

Additionally, the eye disc possesses a squamous peripodial membrane in the apical compartment which gives rise to the head case and migrating glial cells in the basal compartment that differentiate to wrap and support PR axons.

Chapter II discusses a novel way to analyze the cell-type specific impact of an inactivated RB pathway in a heterogenous and complex eye disc tissue.

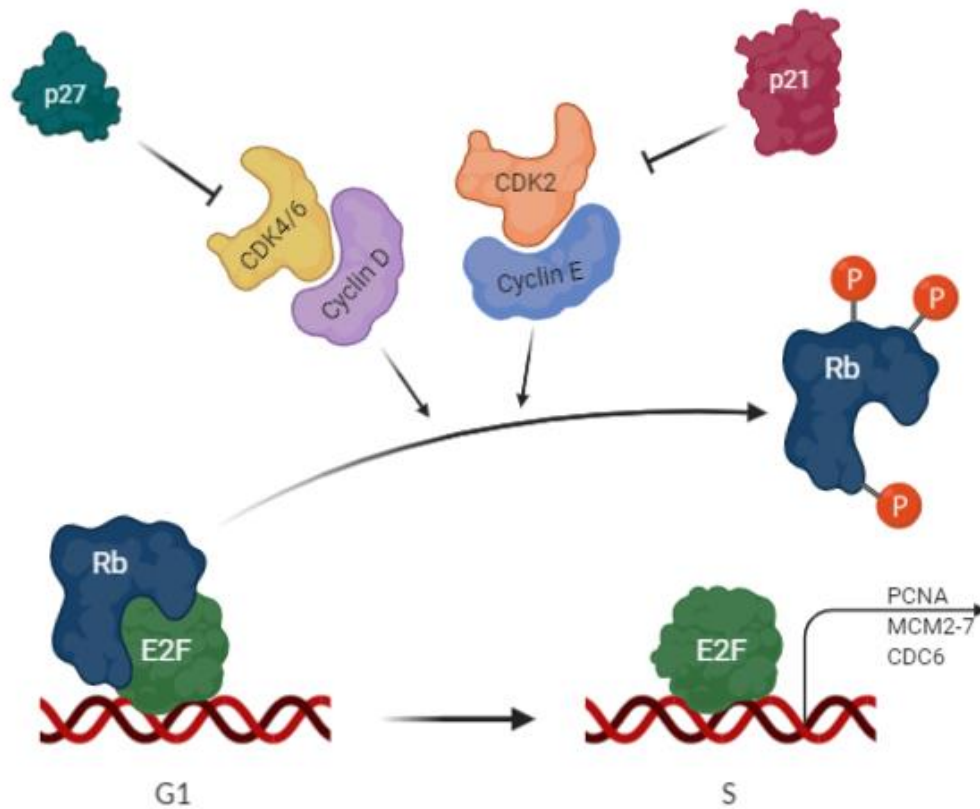


Figure 1. The RB pathway

During the G1 cell cycle phase, Rb binds to E2F transcription factors and inhibit their transcriptional activity. While transitioning to the S phase, Cyclin/CDKs phosphorylate Rb to de-repress genes involved in DNA synthesis.

B. The context dependent response of Sprouty

The receptor tyrosine kinase (RTK) pathway is among the most characterized signaling pathways and regulates multiple significant cellular processes such as proliferation, growth, differentiation, and survival during both cell and cancer development (Mele & Johnson, 2020). The pathway is comprised of 58 surface receptors which, when bound to extracellular ligands, dimerize and activate through auto-phosphorylation in their intracellular domain tyrosine residues. This in turn recruits the adaptor protein GRB2 and guanine nucleotide exchange factor SOS to the plasma membrane. SOS then induces conformational changes in the RAS protein by exchanging GDP to GTP, triggering a cascade of phosphorylation that activates downstream effector kinases RAF, MEK, and ERK (Regad, 2015). Mutations in the EGFR receptor tyrosine kinase gene are predominant in many cancers and cause pathogenesis in glioblastomas, non-small-cell lung, and colorectal cancers (Hatanpaa et al., 2010; Riely et al., 2006; McKay et al., 2002). This has led to the development of several cancer drugs that target the RTK signaling pathway such as Lapatinib, Sorafenib, Gefitinib, and Erlotinib (Johnston et al., 2009; Llovet et al., 2008; Mok et al., 2009; Zhou et al., 2011).

The RTK signaling pathway incorporates several effectors that bind and interact with different downstream components. Among them is the general RTK signaling inhibitor protein Sprouty, which was initially discovered in *Drosophila*. There are four mammalian counterparts for Sprouty (Sprouty1-4) and its function is evolutionary conserved from *Drosophila* to humans (Kawazoe & Taniguchi, 2019). Following growth factor stimulation, Sprouty is recruited to cellular membranes where it is activated by a series of tyrosine residue phosphorylations and serine residue dephosphorylations (Hausott & Klimaschewski 2019). Once active, the inhibitor binds to GRB2 or RAF, thereby disrupting the recruitment of downstream RTK effector kinases and inhibiting the signaling phosphorylation cascade (Hausott & Klimaschewski 2019). Sprouty proteins are

upregulated by growth factors and provide a negative feed-back loop to fine-tune and coordinate RTK signaling with the physiological environment (Masoumi-Moghaddam et al., 2014).

Even though Sprouty plays an antagonistic role in the RTK pathway, it has become evident through recent work that its function is ultimately highly complex and multifaceted. The general RTK inhibitor was shown to act as either a tumor suppressor or as an oncogene depending on cellular context. For example, Sprouty inhibits FGF signaling and decreases P-ERK which inhibits proliferation and migration in osteosarcoma cells (Rathmanner et al., 2013) and has been shown to act as a tumor suppressor in KRAS^{G12D} driven lung tumors (Cho et al., 2011). Additionally, recent work has revealed that Sprouty is directly regulated by various microRNAs (miRs) such as miR-23a and miR-592 which inhibit migration, proliferation, and invasion in gastric cancer (Li et al., 2018; He et al., 2018). Similarly, loss of Sprouty in endometrial carcinoma results in a decrease in E-cadherin and an increase in EGF signaling, both of which promote invasion (So et al., 2015). Conversely, in colon cancer, Sprouty over-expression increased cell proliferation and phosphorylation of c-Met, as well as ERK and Akt. This resulted in an increase in tumor size as well as metastasis (Holgren et al., 2010). Recent work also showed Sprouty to be an oncogene in testicular germ cell tumors since knockdown of the RTK inhibitor decreased P-Akt levels, cell growth, migration, and invasion (Das et al., 2018). In the context of prostate cancer, Sprouty deficiency arrests tumor growth, suggesting that it is contributing to tumorigenesis. However, loss of both the RTK inhibitor and PTEN loss results in enhanced tumorigenesis (Patel et al., 2013). This shows that Sprouty can act as an oncogene and only displays tumor suppressive capabilities in specific cellular contexts.

In 2016, two studies further outlined the context-dependent and complex role of Sprouty by showing that it can function either as a tumor suppressor or an oncogene within breast cancers. He et al., (2016) found that knockdown of Sprouty impaired Snail induction and suppressed mesenchymal

phenotype in triple-negative breast cancer cells (TNBC) MDA-MB-231 cells, preventing migration and invasion. They show that E-cadherin expression increases and that Sprouty functions as an oncogene in TNBC *in vitro* and *in vivo*. Six months later, Koledovaa et al., (2016) revealed that loss of Sprouty in stable breast cancer MCF7 cells and mammary fibroblasts resulted in increased cell invasion, branching, and promoted ECM remodeling. They show that Sprouty acts as a tumor suppressor in these contexts.

Some work attempted to explain these cell specific phenotypic discrepancies by showing that FGF-induced stimulation of ERK is inhibited by Sprouty, whereas EGF-induced stimulation of ERK is activated by Sprouty (Kawazoe & Taniguchi, 2019). One potential mechanism explains that EGF stimulation promotes Sprouty binding to the E3 ubiquitin ligase c-Cbl, inhibiting its role in endocytosis and degradation of RTK surface receptors. However, the complex cellular responses of Sprouty besides c-Cbl binding are not well understood, especially *in vivo*, where multiple growth factors concomitantly activate RTK signaling pathway (Masoumi-Moghaddam et al., 2014). Therefore, there is a considerable need to develop new means and methods for holistic study of the cell type specific roles of Sprouty.

Sprouty was initially discovered and characterized in *Drosophila* and was demonstrated to affect the number of glia during embryogenesis (Kramer et al., 2000). This initial pioneer study in glia led to the discovery that Sprouty's function is conserved in mammals as well (Impagnatiello et al., 2001). Furthermore, multiple RTK surface receptors such as PDGFR, FGFR, InR, and MEGF10 receptor are crucial in *Drosophila* glial cell development and have been shown to promote glia proliferation, differentiation, migration, and locomotion (Read 2018; Avet-Rochex et al., 2012; Franzdóttir et al., 2009; Ray et al., 2017). Therefore, the *Drosophila* developing glia present an ideal context to study the role of Sprouty as the inhibitor can affect all these different RTK receptors which

are vital for glia. Recent studies classified glial cells in the fruit fly into four major classes: surface, cortex, neuropil-associated, and wrapping glia (Yildirim et al., 2019). Surface glia surround the nervous system to form a physical layer called the blood-brain barrier that protects neurons from the high solute content of the hemolymph and subsequent neurodegeneration (Bainton et al., 2005). These glial cells support optimal neuronal growth by controlling the concentration of amino-acids and metabolites obtained from the blood stream (Volkenhoff et al., 2015). Cortex glia, located between surface glia and neuropils glia of the nervous system, form a meshwork that surround neuronal bodies, supplying physical and metabolic support to neurons (Melom & Littleton, 2013). Neuropil-associated glia, which develop in the neuropils, serve a similarly structural role by encasing dendrites, axons, and synapses. Additionally, they participate in neurotransmitter homeostasis and removal of neuronal debris in the central nervous system (Awasaki et al., 2006). Finally, wrapping glia have similar features to non-myelinating Schwann cells since they surround and provide support to axons in the peripheral nervous system (Matzat et al., 2015).

While RTK plays a crucial role in glia development and glioblastoma, Chapter III describes a new method to determine the cell specific impact following a decrease in RTK signaling in different glia, gliomas, and uncovers a novel gene to regulate Sprouty.

C. **Goals of this thesis**

scRNA-seq is a robust tool for studying complex heterogeneous tissues by isolating and identifying transcriptomic profiles in disparate cells that were otherwise not possible using conventional approaches, such as RNA-seq. This technique has enabled the discovery of new, rare cell types, sub-types, and novel biomarkers (Saliba et al., 2014). scRNA-seq has also been shown to uncover complex cellular dynamics within specific cell types such as differentiation, transcriptomic

responses to different extracellular stimuli, and spatial positioning of a cell within a tissue (Wagner et al., 2016; Papalexi & Satija, 2018; Liu & Trapnell, 2016).

Drop-seq is an ideal scRNA-seq platform that generates thousands of single cell cDNA libraries at an affordable cost. The platform consists of a microfluidic device with input for cells, beads, and oil, and an output which creates thousands of nanoliter droplets, each capturing one cell and one bead in a uniform flow (Macosko et al., 2015). To test if the platform is operational, Macosko et al. describe an initial “barnyard” experiment in which mouse 3T3 fibroblast cells were mixed with human HEK293 cells prior to performing Drop-seq (Macosko et al., 2015). The purpose of the experiment was to show that mouse and human cell transcripts are separated with the technology. Upon performing the experiment with the Drop-seq equipment in our hands, the results indicate that mouse cDNA libraries did not capture human transcripts and vice-versa, indicating that Drop-seq was successful in creating single cells libraries. The chance of doublets, capturing two cells in the same droplet, was as low as 1.1%, as shown in the barnyard plot (Figure 1).

The goal of this thesis is to test the applicability of scRNA-seq in dissecting a mutant and knockdown phenotype. The work here outlines a new methodology starting from: 1- A scRNA-seq experiment to analyze single cell clusters. 2- Using additional tools to validate these clusters *in vivo* during development. 3- Ascertaining biological features that uncover the molecular mechanisms and signaling pathways affected in different cell clusters. The aim of this dissertation is to provide the groundwork and layout for future studies that seek to analyze cell-specific transcriptional changes following a lesion event in *Drosophila* and mammalian tissues.

The Seurat R package is widely used to examine scRNA-seq datasets and perform an unsupervised unbiased cell clustering analysis (Stuart et al., 2019). The program initially identifies highly variable genes across the whole sample to create multiple principle components (PCs) which

separate cells into two groups that are distinct in their transcriptional program. The highly significant PCs are further selected to create a t-Distributed Stochastic Neighbor Embedding (tSNE) or Uniform Manifold Approximation and Projection (UMAP) two-dimensional plot. The tSNE/UMAP in turn reveals the heterogeneity of the sample by grouping cells similar in gene expression into cell clusters, each presenting a unique gene expression signature and biomarkers.

In this thesis, scRNA-seq was used to dissect phenotypes in *Drosophila* and characterize the cell-type specific molecular mechanisms following inactivation of the RB pathway in the eye disc, and a reduction in RTK signaling in glial cells.

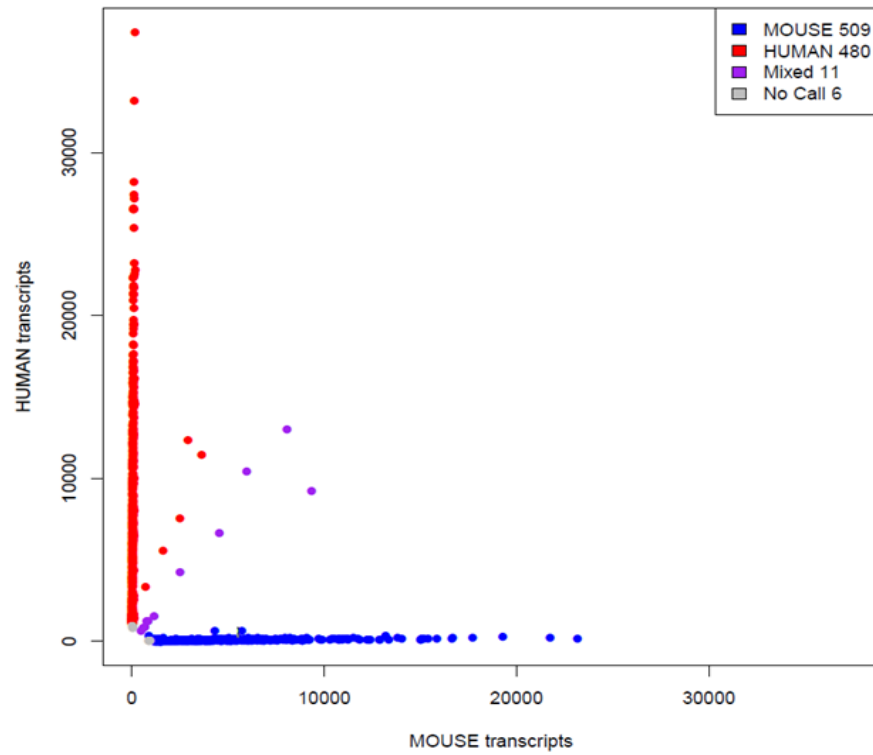


Figure 2. Mixed species barnyard plot.

Barnyard plot showing 509 and 480 barcodes having species specific mouse and human transcripts respectively whereas only 11 (1.1%) barcodes display mixed species transcript.

D. Cited literatures

Avet-Rochex, A., Kaul, A. K., Gatt, A. P., McNeill, H., & Bateman, J. M. (2012). Concerted control of gliogenesis by InR/TOR and FGF signalling in the *Drosophila* post-embryonic brain. *Development*, 139(15), 2763-2772.

Awasaki, T., Tatsumi, R., Takahashi, K., Arai, K., Nakanishi, Y., Ueda, R., & Ito, K. (2006). Essential role of the apoptotic cell engulfment genes *draper* and *ced-6* in programmed axon pruning during *Drosophila* metamorphosis. *Neuron*, 50(6), 855-867.

Bainton, R. J., Tsai, L. T. Y., Schwabe, T., DeSalvo, M., Gaul, U., & Heberlein, U. (2005). *moody* encodes two GPCRs that regulate cocaine behaviors and blood-brain barrier permeability in *Drosophila*. *Cell*, 123(1), 145-156.

Burkhart, D. L., Morel, K. L., Sheahan, A. V., Richards, Z. A., & Ellis, L. (2019). The Role of RB in Prostate Cancer Progression. In *Prostate Cancer* (pp. 301-318). Springer, Cham.

Cagan, R. L., & Reh, T. A. (2010). *Invertebrate and vertebrate eye development*. Academic Press.

Chen, C. R., Kang, Y., Siegel, P. M., & Massagué, J. (2002). E2F4/5 and p107 as Smad cofactors linking the TGF β receptor to c-myc repression. *Cell*, 110(1), 19-32.

Cho, H. C., Lai, C. Y., Shao, L. E., & Yu, J. (2011). Identification of tumorigenic cells in KrasG12D-induced lung adenocarcinoma. *Cancer research*, 71(23), 7250-7258.

Das, M. K., Furu, K., Evensen, H. F., Haugen, Ø. P., & Haugen, T. B. (2018). Knockdown of SPRY4 and SPRY4-IT1 inhibits cell growth and phosphorylation of Akt in human testicular germ cell tumours. *Scientific reports*, 8(1), 1-8.

Dyer, M. A. (2016). Lessons from retinoblastoma: implications for cancer, development, evolution, and regenerative medicine. *Trends in molecular medicine*, 22(10), 863-876. *Cancer cell*, 17(4), 376-387.

Dyson, N. J. (2016). RB1: a prototype tumor suppressor and an enigma. *Genes & development*, 30(13), 1492-1502.

Elchuri, S. V., Rajasekaran, S., & Miles, W. O. (2018). RNA-sequencing of primary Retinoblastoma tumors provides new insights and challenges into tumor development. *Frontiers in genetics*, 9, 170.

Franzdóttir, S. R., Engelen, D., Yuva-Aydemir, Y., Schmidt, I., Aho, A., & Klämbt, C. (2009). Switch in FGF signalling initiates glial differentiation in the *Drosophila* eye. *Nature*, 460(7256), 758-761.

- Hatanpaa, K. J., Burma, S., Zhao, D., & Habib, A. A. (2010). Epidermal growth factor receptor in glioma: signal transduction, neuropathology, imaging, and radioresistance. *Neoplasia*, 12(9), 675-684.
- He, Y., Ge, Y., Jiang, M., Zhou, J., Luo, D., Fan, H., ... & Yang, L. (2018). MiR-592 promotes gastric cancer proliferation, migration, and invasion through the PI3K/AKT and MAPK/ERK signaling pathways by targeting Spry2. *Cellular Physiology and Biochemistry*, 47(4), 1465-1481.
- He, Q., Jing, H., Liaw, L., Gower, L., Vary, C., Hua, S., & Yang, X. (2016). Suppression of Spry1 inhibits triple-negative breast cancer malignancy by decreasing EGF/EGFR mediated mesenchymal phenotype. *Scientific reports*, 6, 23216.
- Holgren, C., Dougherty, U., Edwin, F., Cerasi, D., Taylor, I., Fichera, A., ... & Khare, S. (2010). Sprouty-2 controls c-Met expression and metastatic potential of colon cancer cells: sprouty/c-Met upregulation in human colonic adenocarcinomas. *Oncogene*, 29(38), 5241-5253.
- Impagnatiello, M. A., Weitzer, S., Gannon, G., Compagni, A., Cotten, M., & Christofori, G. (2001). Mammalian sprouty-1 and-2 are membrane-anchored phosphoprotein inhibitors of growth factor signaling in endothelial cells. *The Journal of cell biology*, 152(5), 1087-1098.
- Johnston, S., Pippen Jr, J., Pivot, X., Lichinitser, M., Sadeghi, S., Dieras, V., ... & Press, M. F. (2009). Lapatinib combined with letrozole versus letrozole and placebo as first-line therapy for postmenopausal hormone receptor–positive metastatic breast cancer. *Journal of Clinical Oncology*, 27(33), 5538-5546.
- Kitajima, S., Li, F., & Takahashi, C. (2020). Tumor Milieu Controlled by RB Tumor Suppressor. *International Journal of Molecular Sciences*, 21(7), 2450.
- Koledova, Z., Zhang, X., Streuli, C., Clarke, R. B., Klein, O. D., Werb, Z., & Lu, P. (2016). SPRY1 regulates mammary epithelial morphogenesis by modulating EGFR-dependent stromal paracrine signaling and ECM remodeling. *Proceedings of the National Academy of Sciences*, 113(39), E5731-E5740.
- Kramer, S., Okabe, M., Hacohen, N., Krasnow, M. A., & Hiromi, Y. (1999). Sprouty: a common antagonist of FGF and EGF signaling pathways in *Drosophila*. *Development*, 126(11), 2515-2525.
- Liu, S., & Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Research*, 5.
- Li, Y., Chen, H., She, P., Chen, T., Chen, L., Yuan, J., & Jiang, B. (2018). microRNA-23a promotes cell growth and metastasis in gastric cancer via targeting SPRY2-mediated ERK signaling. *Oncology letters*, 15(6), 8433-8441.
- Li, B., & Zhang, X. (2020). A Review of Retinoblastoma from the Perspective of Integrative Medicine. In *Integrative Ophthalmology* (pp. 103-107). Springer, Singapore.

- Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., ... & Schwartz, M. (2008). Sorafenib in advanced hepatocellular carcinoma. *New England journal of medicine*, 359(4), 378-390.
- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., ... & Trombetta, J. J. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161(5), 1202-1214.
- Matzat, T., Sieglitz, F., Kottmeier, R., Babatz, F., Engelen, D., & Klämbt, C. (2015). Axonal wrapping in the *Drosophila* PNS is controlled by glia-derived neuregulin homolog *Vein*. *Development*, 142(7), 1336-1345.
- McKay, J. A., Murray, L. J., Curran, S., Ross, V. G., Clark, C., Murray, G. I., ... & McLeod, H. L. (2002). Evaluation of the epidermal growth factor receptor (EGFR) in colorectal tumours and lymph node metastases. *European Journal of Cancer*, 38(17), 2258-2264.
- Mele, S., & Johnson, T. K. (2020). Receptor Tyrosine Kinases in Development: Insights from *Drosophila*. *International Journal of Molecular Sciences*, 21(1), 188.
- Melom, J. E., & Littleton, J. T. (2013). Mutation of a NCKX eliminates glial microdomain calcium oscillations and enhances seizure susceptibility. *Journal of Neuroscience*, 33(3), 1169-1178.
- Mok, T. S., Wu, Y. L., Thongprasert, S., Yang, C. H., Chu, D. T., Saijo, N., ... & Nishiwaki, Y. (2009). Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *New England Journal of Medicine*, 361(10), 947-957.
- Moon, N. S., Frolov, M. V., Kwon, E. J., Di Stefano, L., Dimova, D. K., Morris, E. J., ... & Dyson, N. J. (2005). *Drosophila* E2F1 has context-specific pro-and antiapoptotic properties during development. *Developmental cell*, 9(4), 463-475.
- Moses, K. (2002). *Drosophila eye development* (Vol. 37). Springer Science & Business Media.
- Ngwenya, S., & Safe, S. (2003). Cell context-dependent differences in the induction of E2F-1 gene expression by 17 β -estradiol in MCF-7 and ZR-75 cells. *Endocrinology*, 144(5), 1675-1685.
- Papalexi, E., & Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature Reviews Immunology*, 18(1), 35.
- Patel, R., Gao, M., Ahmad, I., Fleming, J., Singh, L. B., Rai, T. S., ... & Sansom, O. J. (2013). Sprouty2, PTEN, and PP2A interact to regulate prostate cancer progression. *The Journal of clinical investigation*, 123(3), 1157-1175.
- Rathmanner, N., Haigl, B., Vanas, V., Doriguzzi, A., Gsur, A., & Sutterlüty-Fall, H. (2013). Sprouty2 but not Sprouty4 is a potent inhibitor of cell proliferation and migration of osteosarcoma cells. *FEBS letters*, 587(16), 2597-2605.

- Ray, A., Speese, S. D., & Logan, M. A. (2017). Glial Draper rescues A β toxicity in a *Drosophila* model of Alzheimer's Disease. *Journal of Neuroscience*, 37(49), 11881-11893.
- Read, R. D. (2018). Pvr receptor tyrosine kinase signaling promotes post-embryonic morphogenesis, and survival of glia and neural progenitor cells in *Drosophila*. *Development*, 145(23), dev164285.
- Regad, T. (2015). Targeting RTK signaling pathways in cancer. *Cancers*, 7(3), 1758-1784.
- Riely, G. J., Pao, W., Pham, D., Li, A. R., Rizvi, N., Venkatraman, E. S., ... & Miller, V. A. (2006). Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clinical cancer research*, 12(3), 839-844.
- Saliba, A. E., Westermann, A. J., Gorski, S. A., & Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic acids research*, 42(14), 8845-8860.
- Singh, L., & Kashyap, S. (2018). Update on pathology of retinoblastoma. *International journal of ophthalmology*, 11(12), 2011.
- Singh, H. P., Wang, S., Stachelek, K., Lee, S., Reid, M. W., Thornton, M. E., ... & Cobrinik, D. (2018). Developmental stage-specific proliferation and retinoblastoma genesis in RB-deficient human but not mouse cone precursors. *Proceedings of the National Academy of Sciences*, 115(40), E9391-E9400.
- So, W. K., Cheng, J. C., Fan, Q., Wong, A. S., Huntsman, D. G., Gilks, C. B., & Leung, P. C. (2015). Loss of Sprouty2 in human high-grade serous ovarian carcinomas promotes EGF-induced E-cadherin down-regulation and cell invasion. *FEBS letters*, 589(3), 302-309.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck III, W. M., ... & Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7), 1888-1902.
- Thwaites, M. J., Cecchini, M. J., Passos, D. T., Zakirova, K., & Dick, F. A. (2019). Context dependent roles for RB-E2F transcriptional regulation in tumor suppression. *PloS one*, 14(1).
- Van Den Heuvel, S., & Dyson, N. J. (2008). Conserved functions of the pRB and E2F families. *Nature reviews Molecular cell biology*, 9(9), 713-724.
- Volkenhoff, A., Weiler, A., Letzel, M., Stehling, M., Klämbt, C., & Schirmeier, S. (2015). Glial glycolysis is essential for neuronal survival in *Drosophila*. *Cell Metabolism*, 22(3), 437-447.
- Wagner, A., Regev, A., & Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. *Nature biotechnology*, 34(11), 1145.

Xu, X. L., Fang, Y., Lee, T. C., Forrest, D., Gregory-Evans, C., Almeida, D., ... & Cobrinik, D. (2009). Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell*, 137(6), 1018-1031.

Yamaguchi, M., & Yoshida, H. (2018). *Drosophila* as a model organism. In *Drosophila Models for Human Diseases* (pp. 1-10). Springer, Singapore.

Yildirim, K., Petri, J., Kottmeier, R., & Klämbt, C. (2019). *Drosophila* glia: Few cell types and many conserved functions. *Glia*, 67(1), 5-26.

Zhou, C., Wu, Y. L., Chen, G., Feng, J., Liu, X. Q., Wang, C., ... & Lu, S. (2011). Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *The lancet oncology*, 12(8), 735-742.

This chapter was previously published

II. SINGLE CELL RNA-SEQUENCING IDENTIFIES A METABOLIC ASPECT OF APOPTOSIS IN *Rbf* MUTANT

Citation: Ariss, M. M., Islam, A. B., Critcher, M., Zappia, M. P., & Frolov, M. V. (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in *Rbf* mutant. *Nature communications*, 9(1), 1-13.

A. Summary

The function of Retinoblastoma tumor suppressor (pRB) is greatly influenced by the cellular context, therefore the consequences of pRB inactivation are cell-type-specific. Here we employ single cell RNA-sequencing (scRNA-seq) to profile the impact of an *Rbf* mutation during *Drosophila* eye development. First, we build a catalogue of 11,500 wild type eye disc cells containing major known cell types. We find a transcriptional switch occurring in differentiating photoreceptors at the time of axonogenesis. Next, we map a cell landscape of *Rbf* mutant and identify a mutant-specific cell population that shows intracellular acidification due to increase in glycolytic activity. Genetic experiments demonstrate that such metabolic changes, restricted to this unique *Rbf* mutant population, sensitize cells to apoptosis and define the pattern of cell death in *Rbf* mutant eye disc. Thus, these results illustrate how scRNA-seq can be applied to dissect mutant phenotypes.

B. Introduction

Functional inactivation of the retinoblastoma protein (pRB) is considered an obligatory event in the development of human cancer and is usually attributed to its ability to block cell-cycle progression through negative regulation of the E2F transcription factor. Binding to pRB inhibits E2F transcriptional activity and halts cell cycle. Conversely, the inactivation of pRB releases E2F and

allows S-phase entry (Dick, Goodrich, Sage, & Dyson, 2018). Such a simplistic view is built on the assumption that pRB operates in the same way across different cell types. However, mouse models and clinical studies have revealed that the function of pRB is greatly influenced by the cellular context. The consequences of pRB inactivation are thought to be determined by a unique, cell-type-specific molecular circuitry around pRB. Such specific interactions may also help to explain why cancer originates in a specific cell type. For example, human retinoblastoma is believed to be derived from post-mitotic cone precursors. These cells are uniquely sensitive to Rb loss as they express cone lineage factors (TR β 2 and RXR γ) and the oncoproteins MYCN and MDM2 (Xu et al., 2014). Thus, it is important to understand how mutations in the RB pathway affect individual cell types. This point is especially relevant in interpreting the results of genome-wide studies, which have been extensively used to deduce how the RB pathway operates. However, averaging gene expression using bulk samples does not provide sufficient resolution to determine the impact of RB pathway mutations on individual cell types.

Recent advances in single-cell RNA-sequencing (scRNA-seq) offer an opportunity to detect variation at the cellular level and dissect heterogeneous tissues into unique cell clusters. Surprisingly, although scRNA-seq has been used to study tumor heterogeneity in cancer, this technology has yet to be adapted to dissect the mutant phenotypes in model organisms. *Drosophila* has a streamlined version of the mammalian RB pathway and proved to be invaluable in deciphering its role *in vivo* (Van Den Heuvel & Dyson, 2008). For example, investigating the mutant phenotype of *Rbf*, encoding the pRB ortholog, provided key insights into the function of *Rbf* in development. Like in mammalian cells, inactivation of *Rbf* in the larval eye imaginal disc results in mild cell-cycle defects and apoptosis. Increased sensitivity to apoptosis of *Rbf*-deficient cells is commonly attributed to elevated expression

of apoptotic E2F target genes such as *hid* in flies. Notably, despite *hid* being upregulated throughout almost the entire *Rbf* mutant eye disc, apoptosis is restricted to cells anterior to the morphogenetic furrow that show a transient reduction in epidermal growth factor receptor (EGFR) signaling (Moon, Di Stefano, & Dyson, 2006). Thus, the *Rbf* mutant eye disc represents an ideal setting to apply scRNA-seq methodology and identify a precise cellular context that makes *Rbf* mutant cells sensitive to apoptosis.

Here, we report an atlas of 11,500 wild-type eye disc cells with 1× cellular coverage that includes major cell types in the developing larval eye. We also find a transcriptional switch during photoreceptor differentiation. We then utilize this resource to examine the *Rbf* mutant phenotype and identify a specific population of cells with increased glycolysis that makes them sensitive to E2F-dependent apoptosis. Thus, our results illustrate the applicability of scRNA-seq to profile mutant phenotypes.

C. **Results**

1. **A cell atlas of the wild-type third-instar larval eye disc**

The *Drosophila* eye remains a preferable model to investigate the control and coordination of cell proliferation, differentiation and apoptosis. During the third-instar larval stage, the morphogenetic furrow (MF) sweeps across the eye disc from the posterior margin towards the anterior, demarcating the onset of neuronal differentiation. The asynchronously dividing cells of the anterior compartment arrest at G1 upon entry into the MF, specifying into photoreceptors as they exit posteriorly. However, not all cells commit to the differentiation program upon emerging from the MF. The remaining uncommitted interommatidial cells undergo one synchronous round of cell

division called the second mitotic wave (SMW) posterior to the MF and then remain quiescent (Figure 1A).

There are several transcriptional domains in the eye disc. The pre-proneural (PPN) domain located immediately anterior to the MF is defined by the expression of *hairy* (*h*) and *dachshund* (*dac*) (Figure 2A, 2B) (Bessa, Gebelein, Pichaud, Casares, & Mann, 2002). Cells anterior to the PPN express *homothorax* (*hth*) in the eye-antennal domain as well as the apical peripodial epithelium, a thin squamous membrane overlying the eye disc proper, which gives rise to the head capsule (Figure 2A, 2B). In the posterior compartment, photoreceptors are organized in repetitive units called ommatidia. Eight photoreceptors (R1–R8) are recruited into each ommatidium in a sequential manner and can be distinguished by specific markers: Senseless (Sens) marks R8 (Figure 1B), BarH1/2 are expressed in R1 and R6, while Rough (Ro) delineates R2 and R5 (Moses, 2002). Following specification, photoreceptors project their axons in the basal compartment to form connections with migratory glial cells from the brain.

We began by building a catalog of cells in the wild-type third-instar larval eye disc using Drop-seq, a microfluidic-based scRNA-seq platform (Macosko et al., 2015). The third-instar larval eye discs were dissected at the optic stalk and eye-antennal domain, then dissociated into a single-cell suspension. To control for potential batch effects, a total of 550 eye discs were dissected over 11 biological replicates and scRNA-seq was performed on each replicate. Following quality controls, sequencing data were collected from 11,416 individual cells at an average depth of 30,344 mapped reads per cell, with a total of 16,331 genes detected. To unbiasedly classify cell types, we performed a t-stochastic neighbor embedding (tSNE) analysis using Seurat (Satija, Farrell, Gennert, Schier, & Regev, 2015) and identified 15 distinct cell clusters (Figure 2C). Cells from each of 11 biological replicates were evenly distributed among the cell populations and contributed to each cluster, with

the exception of the rare cell-type HEMO shared by 10 replicates. This demonstrates that the results are highly reproducible and show no apparent batch effect between samples.

To assess the efficiency of cell dissociation in generating a single-cell suspension, we took advantage of known markers that are highly specific for individual photoreceptors. *Sens* is expressed only in R8 and therefore its expression is mutually exclusive with *Rough*, a marker for R2/5, or *BarH2*, a marker for R1/6. Since R8 is positioned in the center of each ommatidium, it is making physical contact with other photoreceptors. If tissues were not effectively dissociated, we expected to find co-expression of *Sens* with other photoreceptor-specific transcripts such as *Rough* or *BarH2* in single-cell libraries. Significantly, less than 4% of libraries contained mixed R cell-specific transcripts (i.e., *sens* and *ro* transcripts present in the same library), suggesting that >95% represent single-cell transcriptomes (Figure 2D). We calculated the expected co-expression values based on independent expression of the three genes and note that our observed results (1.7% and 3.6%) (Figure 2D) are significantly lower than expected values (p value = 2.3×10^{-7} using chi-squared test).

Unsupervised Seurat clustering identified 15 cell populations each having a set of differential expressed genes (Figure 3). Each population was assigned based on known markers computationally derived from the analysis (Figure 4A). The MF was distinguished by high levels of Notch signaling (Figure 1B) including detection of several *E(spl)* genes (Figure 4A) (Lai, Bodner, Kavalier, Freschi, & Posakony, 2000). The expression of *h* and *dac* was used to identify the PPN and, consistent with previous literature, the latter was also observed in the MF (Figure 4A) (Bessa et al., 2002). Undifferentiated cells (UND) located anterior to the PPN were identified by the presence of *hth* and *toy* expression.

The eye-antennal border (EAB) forms the boundary between these two tissues. EAB cells were assigned by the expression of *Lim1* and *cut* (Figure 4A) (Roignant, Legent, Janody, & Treisman,

2010). Since the dissection included part of the antenna (ANT), these cells were distinguished by the expression of *Dll* (Figure 4A) (Emerald, Curtiss, Mlodzik, & Cohen, 2003).

Four cell populations were mapped to the posterior compartment of the eye disc. Two of which (early photoreceptor (EPR) and late photoreceptor (LPR)) were identified as photoreceptors based on the presence of the classical neuronal markers *elav*, *futsch* and *Appl* (Figure 4A). The remaining two are interommatidial cells (INT and SMW) since they expressed posterior markers such as *ed* and *Mmp2* (Figure 4A) but lacked the expression of neuronal genes. As described above, interommatidial cells undergo one synchronous round of cell division in the second mitotic wave and exit the cell cycle in the posterior (Figure 2B) (Moses, 2002). Since cells in the SMW cluster exhibit high expression of cell-cycle genes such as *Claspin*, *PCNA*, *Mcm7* (Figure 4B), *dap* and *stg*, we concluded that these are cells of the second mitotic wave. Conversely, the expression of cell-cycle genes is low in INT cells and therefore they are the quiescent interommatidial cells. Interestingly, we did not find a distinct cell population corresponding to cone cells but their respective marker *cut* was detected throughout the LPR population. This is likely because cone precursors share similarities with photoreceptors and therefore Seurat analysis does not distinguish between them at this developmental stage. We identified several cell type-specific markers for the eye disc. Neuromusculin (encoded by *nrm*) was previously shown to be only expressed in photoreceptors (Kania, Han, Kim, & Bellen, 1993). However, *nrm* expression was also detected in interommatidial cells in our data set and this was validated by immunofluorescence (Figure 4A, 4C). Interestingly, computational analysis identified a muscle-specific gene *sls* (also known as *kettin*) (Lakey et al., 1993) as a posterior marker (Figure 4C). We confirmed this surprising observation by using a *sls* reporter that showed robust expression in the posterior domain of the eye disc (Figure 4C).

The tSNE analysis revealed two glial cell populations (PG and WG+SPG) based on *repo* expression (Figure 4A), a well-established glial marker (Moses, 2002). These clusters are distinct by their proliferation and differentiation states. Perineurial glia (PG) are undifferentiated, dividing cells that label with 5-bromo-2'-deoxyuridine (BrdU) (Franzdóttir et al., 2009) and can be detected by immunofluorescence in the posterior compartment of the eye disc (Figure 2B). Since PG cell population exhibited high expression of cell-cycle genes *Claspin*, *stg* and *PCNA* (Figure 4B), we assigned this population to perineurial glial cells. Consistently, the PG-specific gene *CG3168* (DeSalvo et al., 2014) was also exclusively detected in this population (Figure 4A). As PG cells migrate towards the eye and make contacts with photoreceptor axons, they exit cell cycle and differentiate into wrapping glia (WG). These two types of glial cells are physically separated by two large subperineurial, or carpet, glial cells (SPG). The presence of the known WG markers *Gli*, *nrv2* and *sty* and the SPG marker *moody* (Sasse, Neuert, & Klämbt, 2015) (Figure 4A) was used to assign the WG+SPG cluster to wrapping glia and subperineurial glial cells. Consistently, the expression of cell-cycle genes was low in these cells (Figure 4B) as they are no longer proliferating. Additionally, we identified two markers of wrapping glial cells, *NK7.1* (most highly expressed in WG+SPG) and *CG9336*, expressed exclusively in this cluster and was previously presumed to be expressed in axon projections (Firth & Baker, 2007). The glial-specific expression of *NK7.1* and *CG9336* was confirmed by a green fluorescent protein (GFP) reporter and fluorescence in situ hybridization (FISH), respectively (Figure 4D). Another gene is *couch potato* (*cpo*) that has been previously shown to be expressed throughout the posterior compartment (Harvie, Filippova, & Bryant, 1998). We found that its expression is restricted to photoreceptors and absent in interommatidial cells (Figure 4A, 4C). *cpo* is strongly expressed in wrapping glial cells of WG+SPG (Figure 4D) and this was validated with a *cpo-lacZ* reporter line (Figure 4B, 4D).

We discerned two cell clusters representing cells of the peripodial epithelium, a thin squamous membrane that overlies the eye disc proper and gives rise to head structures during metamorphosis. This membrane is composed of two distinct cell populations; the dorsal peripodial epithelium (DPE), and the ventral peripodial epithelium (VPE). These populations were distinguished by dorsal markers such as *mirror* (Figure 2B) (McNeill, Yang, Brodsky, Ungos, & Simon, 1997) that is present in DPE but absent in VPE. The VPE expresses *cv-c*, a top marker for the EAB (Pilgram, Potikanond, van der Plas, Fradkin, & Noordermeer, 2011). However, the VPE does not express *Lim1* which further distinguishes that population from the EAB. Cells of the ocellar complex (OCx), which form the ocelli on the adult head, were identified by the expression of *ocelliless (oc)* (Dominguez-Cejudo & Casares, 2015). Finally, we identified a cluster of hemocytes (HEMO) consisting of around 60 cells (0.53% of all cells in our dataset), based on the expression of the classical marker *Hemolectin (Hml)* (Sinenko, Mandal, Martinez-Agosto, & Banerjee, 2009) among others (Figure 4A). From our computational analysis we identified *pigs*, a gene that plays a role in ovary and muscle development (Pines, Housden, Bernard, Bray, & Roper, 2010), as a hemocyte marker and confirmed its expression in hemocytes using a *pigs* reporter (Figure 4A, 4D). In addition to hemocytes, *pigs* is expressed in PG cells as revealed by Seurat and immunofluorescence (Figure 4A, 4D). The identification of a distinct HEMO cluster within the 10,000 cells of the eye disc (Silies, Yuva-Aydemir, Franzdóttir, & Klämbt, 2010) is notable as only a small number of hemocytes are associated with each disc. We concluded that our atlas of 11,415 cells accurately represents known cell types of the larval eye disc and efficiently captures rare cell types (<1%). Additionally, our results identified unique transcriptional signatures within glial cells and two distinct photoreceptors populations.

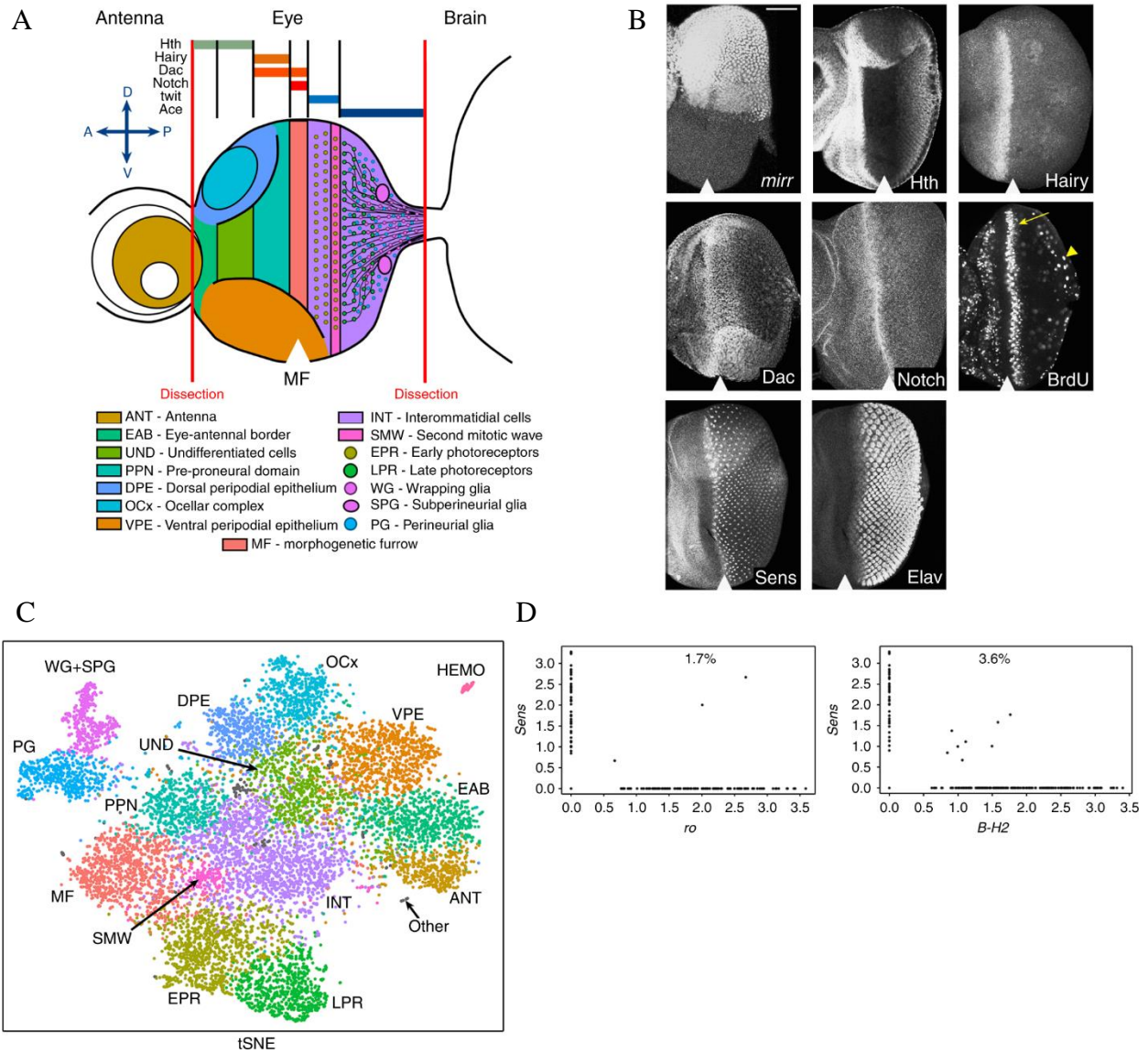


Figure 3. A cell atlas of the wild-type third-instar larval eye disc.

(A) A cartoon representation of the eye-antennal disc and known cell types that are color coded. The red vertical lines indicate where the dissections were made. The expression domains of genes above that were used to assign cell clusters. D dorsal, V ventral, P posterior to the MF, A anterior to the MF. (B) The pattern of the expression of cell-type markers in the eye as revealed by immunofluorescence. BrdU labeling shows the location of SMW (yellow arrow) and PG (yellow arrowhead) cells undergoing cell division. The scale bar is 50 μ m. (C) The tSNE representation of 11,416 wild-type eye disc cells. Each cluster represents a cell type/domain, characterized by a unique gene expression signature. Clusters are color coded as in (A). Other cells are labeled as gray. (D) Gene/gene plots generated on all photoreceptors for *sens* (R8) and *ro* (R2/5), *sens* and *B-H2* (R1/6). The percentage represents doublets, based on the detection of two R cell-specific markers in the same single-cell library. The position of the MF is shown by white arrowhead.

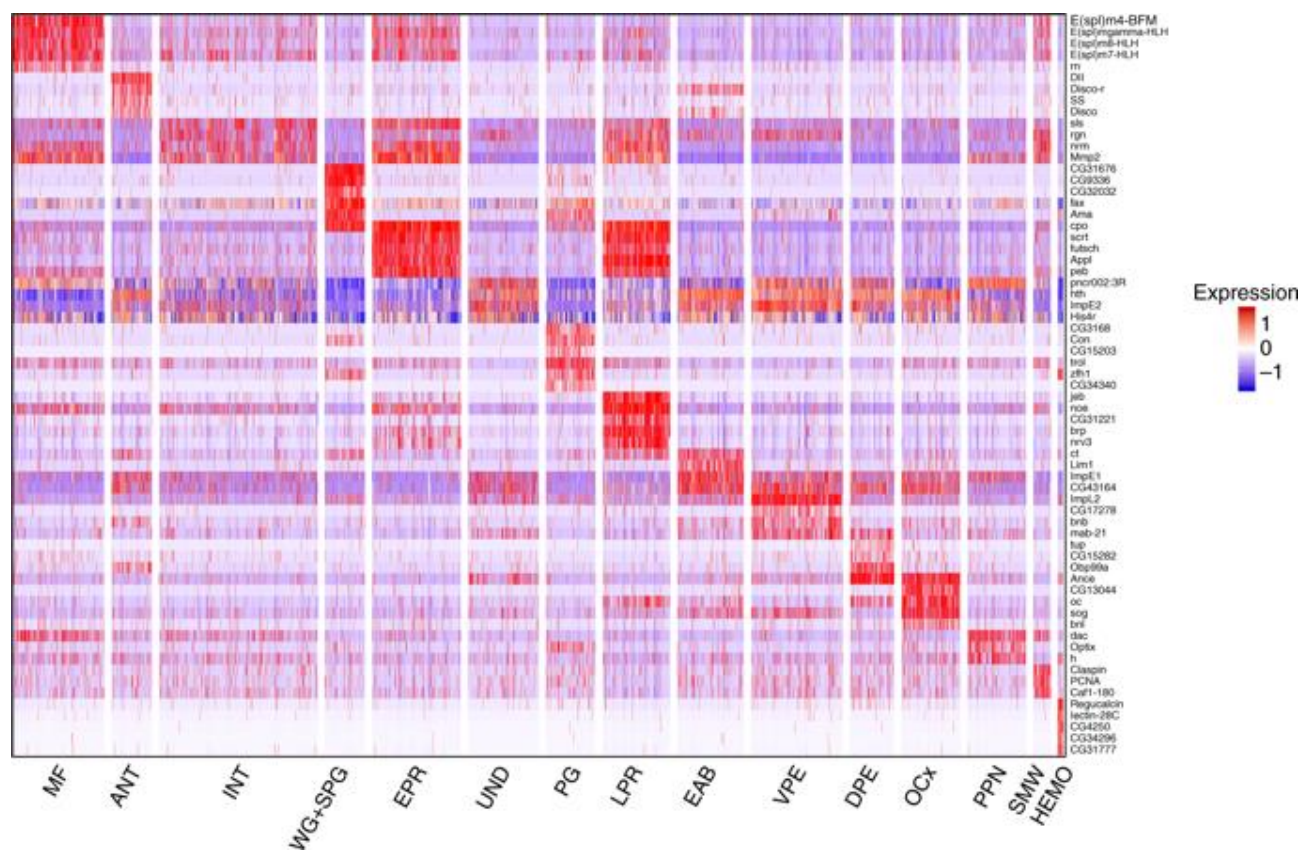
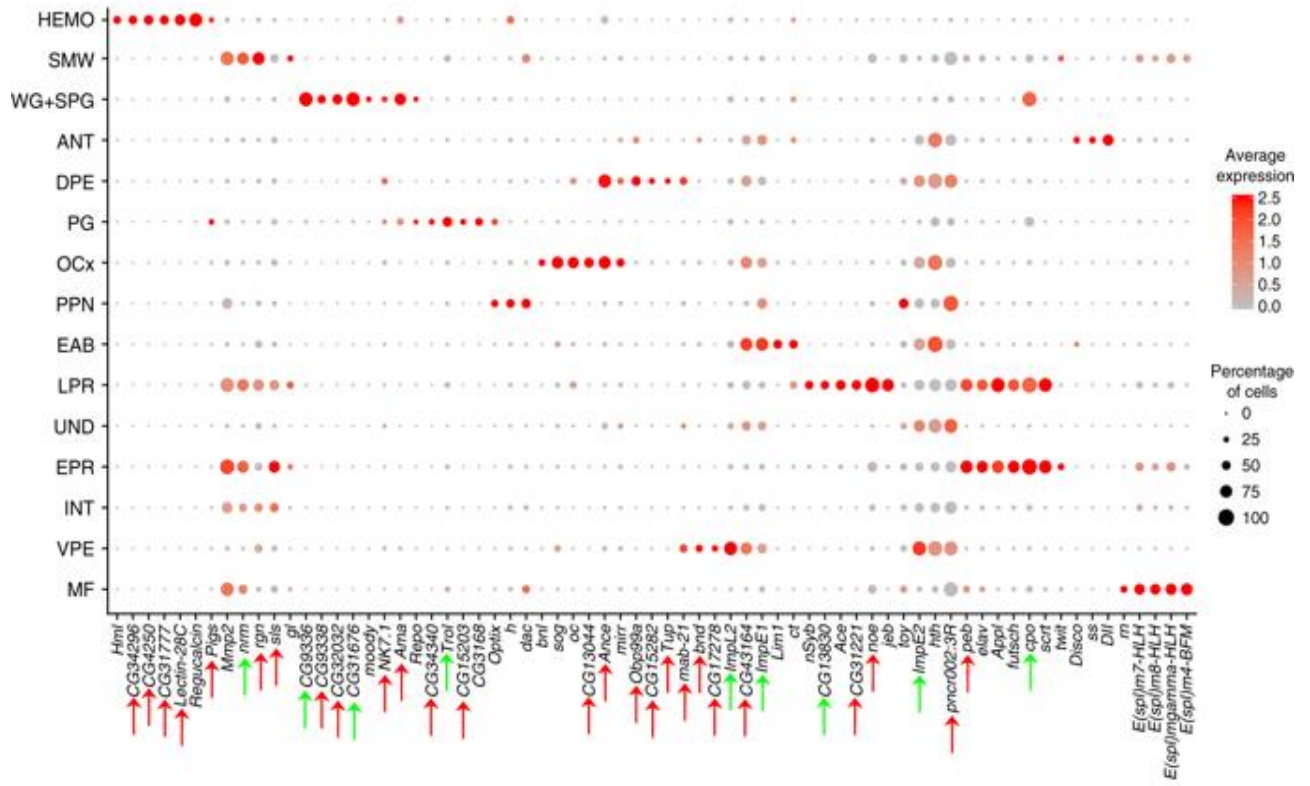


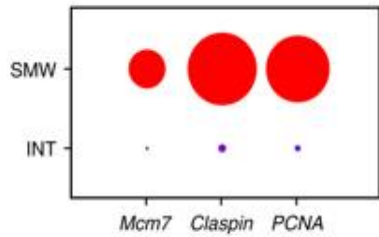
Figure 4. A heat map showing markers across cell populations.

Heat map displaying the expression of 65 top markers across 15 wild-type populations.

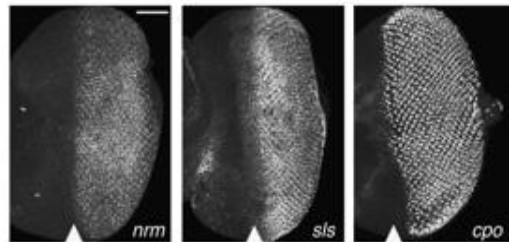
A



B



C



D

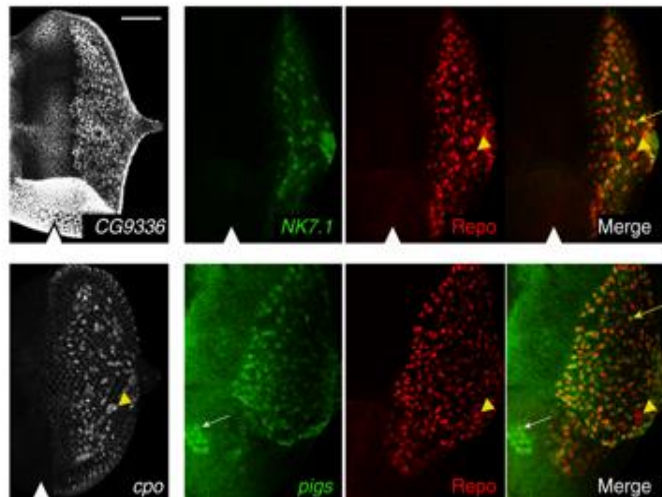


Figure 5. Single-cell RNA-seq identifies cell-specific markers.

(A) Dot plot of top markers for all populations. Red arrows point at previously unreported eye cell-type markers. Green arrows indicate genes previously studied in the eye that were not assigned to a specific cell type. The genes without arrows are known markers used to assign cell populations. (B) Dot plots, top panel shows interommatidial populations INT and SMW, showing expression of cell-cycle genes exclusively in SMW subpopulation. Bottom panel shows glial cell populations PG and WG+SPG showing expression of cell-cycle genes exclusively in PG. The scale bar is 50 μ m. (C) LacZ enhancer traps for posterior candidate markers *neuromusculin* (*nrm*), *sallimus* (*sls*, also known as *kettin*) and photoreceptor marker *couch potato* (*cpo*). (D) Wrapping glia markers *CG9336* and *NK7.1* of and their expression as revealed by FISH (*CG9336*) and a GFP reporter for *NK7.1* (green). Repo (red) is a glial marker. Merge image marks cells co-expressing *NK7.1* and Repo, also shown with yellow arrow. Markers of PG, *cpo* and *pigs* are revealed by LacZ enhancer trap and GFP reporter, respectively. SPG are indicated by a yellow arrowhead, overlap of Repo and *pigs* shown with yellow arrow. The scale bar is 50 μ m. The position of the MF is shown by white arrowhead.

2. Developing photoreceptors undergo a transcriptional switch

Given that there are eight photoreceptors, the tSNE analysis unexpectedly detected only two photoreceptor clusters, EPR and LPR. To determine whether EPR and LPR represent photoreceptors of different R types, we plotted the R-cell-specific markers *sens*, *ro* and *B-H2* against *elav*, a pan-neuronal marker expressed in all R cells. Surprisingly, cells expressing *sens*, *ro* and *B-H2* were found in both clusters (Figure 5A) revealing lack of bias towards a particular photoreceptor type between EPR and LPR populations. In order to determine whether there is an R-type segregation within each cluster, we selected the cells in EPR and performed a supervised Seurat analysis on these cells only. Yet, *sens*-, *ro*- and *B-H2*-positive cells remained randomly scattered throughout the tSNE feature plots, revealing lack of clustering based on R-type. Similar result was observed when Seurat analysis was done only on LPR cells.

The comparison of top marker gene lists for EPR and LPR revealed a significant overlap, as the majority of EPR markers were also found in LPR (Figure 5B). One of the rare EPR-specific markers is *twit*. We used *twit* probe for FISH to localize the position of EPR in the eye disc and detected its expression within a narrow band posterior to the MF, ending at column 4 (Figure 5C, 5D). Conversely, the expression of the LPR-specific top marker *Ace*, encoding the enzyme Acetylcholine esterase that degrades acetylcholine in synapses, was detected posterior to column 4, and in photoreceptor axons in the basal compartment (Figure 5C, 5D). Since 153 genes were upregulated in LPR in comparison to EPR (Figure 5B), we decided to further elucidate the differences between EPR and LPR. We performed gene set enrichment analysis for gene ontology biological processes (GOBP) on all computationally derived genes for EPR, LPR and UND. As expected, undifferentiated cells were dominated by cell cycle-related categories. In contrast, EPR and LPR showed enrichment for photoreceptor differentiation-related GO terms (Figure 5E). Interestingly,

LPR genes exhibited a statistically significant enrichment of axonogenesis and related categories compared to EPR. Accordingly, computationally identified top markers for LPR *jeb*, *brp*, *Dscam2* and *nSyb* have a known role in axon guidance (Bazigou et al., 2007; Bhattacharya et al., 2002; Blockus & Chédotal, 2016; Wagh et al., 2006).

Finally, we employed Monocle 2 (Qiu et al., 2017) to determine the temporal order of development of EPR and LPR along the pseudotime-derived neuronal differentiation axis. The progression of cellular differentiation was visualized on the plot; from UND progenitor cells to PPN in the anterior, through MF to INT in the posterior, and finally to EPR and LPR (Figure 5F). Notably, LPR clustered consecutively to EPR in the plot, with the latter located closer to MF cells. We concluded that photoreceptors cluster irrespective of R-cell type. However, after column 4, photoreceptors initiate axonogenesis (Moses, 2002) and this transcriptional switch distinguishes late (LPR) from early (EPR) photoreceptors.

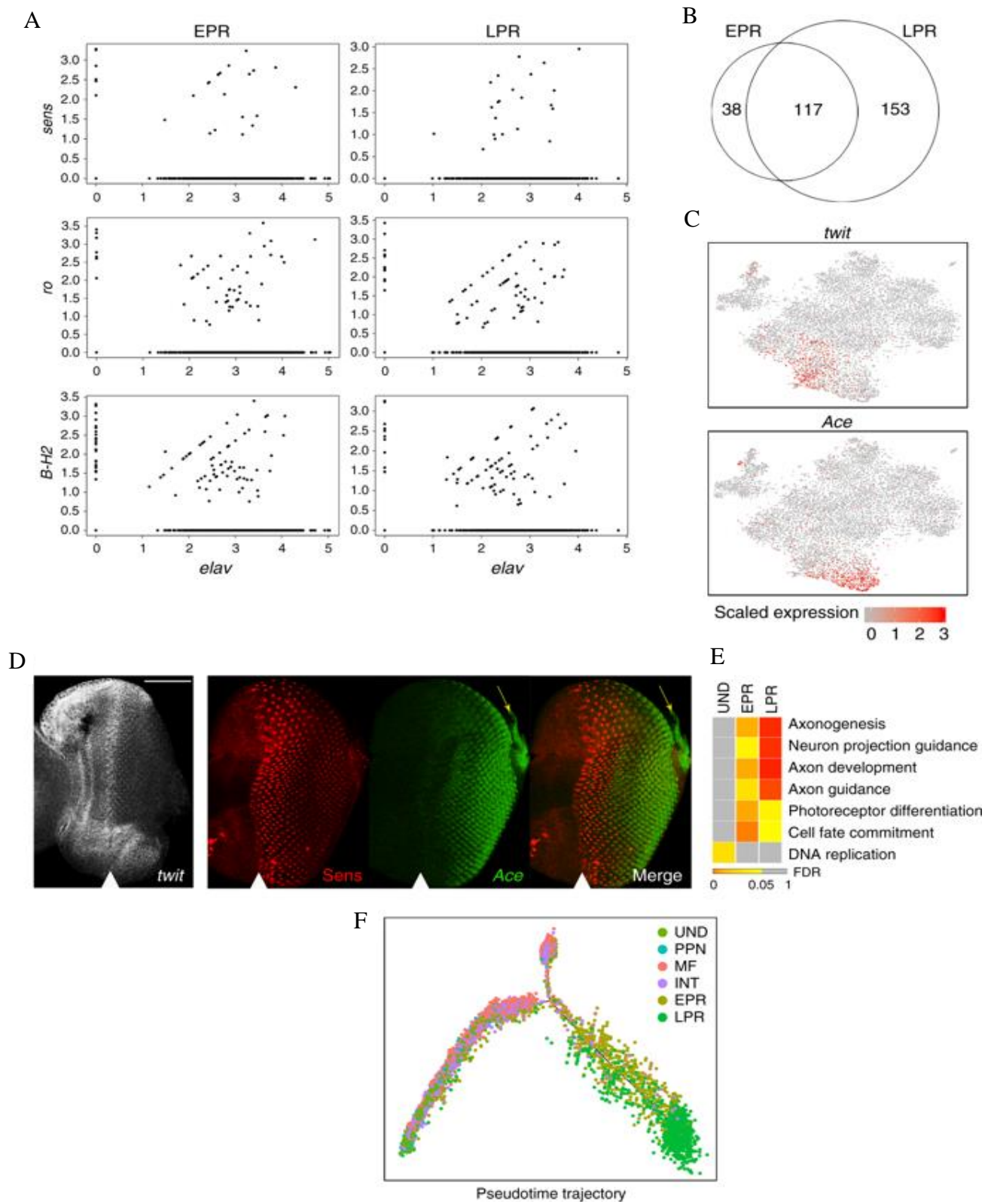


Figure 6. Photoreceptors undergo a transcriptional switch at onset of axonogenesis.

(A) Gene/gene plots showing expression of R type-specific markers *sens*, *ro* and *B-H2* in relation to *elav* in EPR (left panel) and LPR (right panel). (B) Venn diagram of the gene lists for EPR and LPR clusters. (C) Feature plots on tSNE for *twit* and *Ace* show predominant expression of each gene in EPR and LPR cells, respectively. (D) EPR and LPR marker gene localization. Left panel: FISH for *twit* mRNA, an EPR-specific gene, shows expression in the MF and posterior until column 4. Right panel: GFP enhancer trap for *Ace* (green), an LPR top marker, was costained with Sens (red). *Ace-GFP* was detected after column 4 and in axonal projections (yellow arrow). The scale bar is 50 μ m. (E) Gene enrichment analysis for biological processes in UND, EPR and LPR. (F) Monocle 2 trajectory using shown clusters. Stages of differentiation are visualized, with the split of INT and photoreceptor populations. LPR arise from the EPR population, clustering consecutively in pseudotime. The position of the MF is shown by white arrowhead

3. scRNA-seq identified an *Rbf* mutant-specific cell cluster

Having built a cell atlas of the wild-type third-instar larval eye disc, we proceeded to investigate the effect of an *Rbf* mutation on individual cell types in this tissue. We performed scRNA-seq on *Rbf*^{d20a} mutant eye discs⁴ in 3 replicates and generated a scRNA-seq dataset of 5203 *Rbf* mutant cells. We then selected 5591 wild-type cells and conducted a tSNE analysis on the combined dataset containing both wild-type and *Rbf* mutant cells.

The tSNE plot identified 15 distinct cell clusters (Figure 6A) that were readily assigned based on markers from the wild-type cell atlas (Figure 2A). De-repression of cell-cycle genes is a known hallmark of *Rbf* mutant (Moon et al., 2006). Accordingly, *PCNA*, *Mcm7* and others cell-cycle genes were upregulated in *Rbf* mutant cells (Figure 6B). Interestingly, Seurat analysis no longer groups cells of the SMW into a distinct cluster, likely because SMW cells are highly proliferative and therefore become indistinguishable from other *Rbf* mutant cells. tSNE analysis revealed that cells of both genotypes contributed to each cell population, albeit with small variations between the numbers of wild-type and *Rbf* mutant cells in several cell populations. For example, the number of photoreceptors in *Rbf* mutant appeared to be lower than in the wild type. However, these differences are likely to be due to a subtle developmental delay of *Rbf* mutants that affects the position of the morphogenetic furrow at the time of dissection and therefore influences the number of photoreceptors that are recruited in the posterior.

In striking contrast to subtle differences described above, cluster 13 consisted of almost exclusively *Rbf*^{d20a} cells (Figure 6A). The *Ldh* (*ImpL3*) gene encoding lactate dehydrogenase is the top ranked marker and its expression is restricted to cells of that cluster (Figure 6B, 6C). Therefore, to identify the spatial location of cluster 13 in the eye disc, we examined the pattern of *Ldh* expression in *Rbf* mutant discs by in situ hybridization. Notably, *Ldh* was highly expressed in a distinct stripe

preceding the MF in *Rbf* mutant eye discs, while being undetectable in wild-type eye discs (Figure 6D). This is consistent with previous observations of very low endogenous *Ldh* activity at this developmental stage (Wang, Purkayastha, Jones, Thaker, & Banerjee, 2016). Two other predicted markers for cluster 13 were *Ald*, encoding the glycolytic enzyme Aldolase, and *HIF1A* (*sim*), a transcription factor that directly regulates *Ldh* in flies and mammals (Figure 6C) (Wang et al., 2016; Zhao et al., 2013). Consistently, both *Ald* and *HIF1A* matched the pattern of *Ldh* expression (Figure 6C).

One consequence of upregulated glycolytic gene expression is increased glycolytic flux and subsequently excessive production of lactate which decreases intracellular pH. To determine whether this phenomenon occurs in *Rbf^{d20a}* mutant, we used a live intracellular probe that fluoresces only in low pH conditions. As expected, no signal was detected in control eye discs (Figure 6E) as the expression of glycolytic genes is low during late larval development (Tennessen, Baker, Lam, Evans, & Thummel, 2011). Strikingly, *Rbf* mutant exhibited a high level of intracellular acidification immediately anterior to the MF (Figure 6E) that largely corresponds to the region of elevated expression of *Ldh* and *Ald* (Figure 6D). This is due to increased glycolytic activity as downregulation of these genes by RNA interference (RNAi) in the *Rbf^{d20a}* background completely suppresses intracellular acidification (Figure 6E). We concluded that an *Rbf* mutation leads to upregulation of glycolytic gene expression that is restricted to a population of cells anterior to the MF. This consequently lowers intracellular pH through increased glycolytic activity.

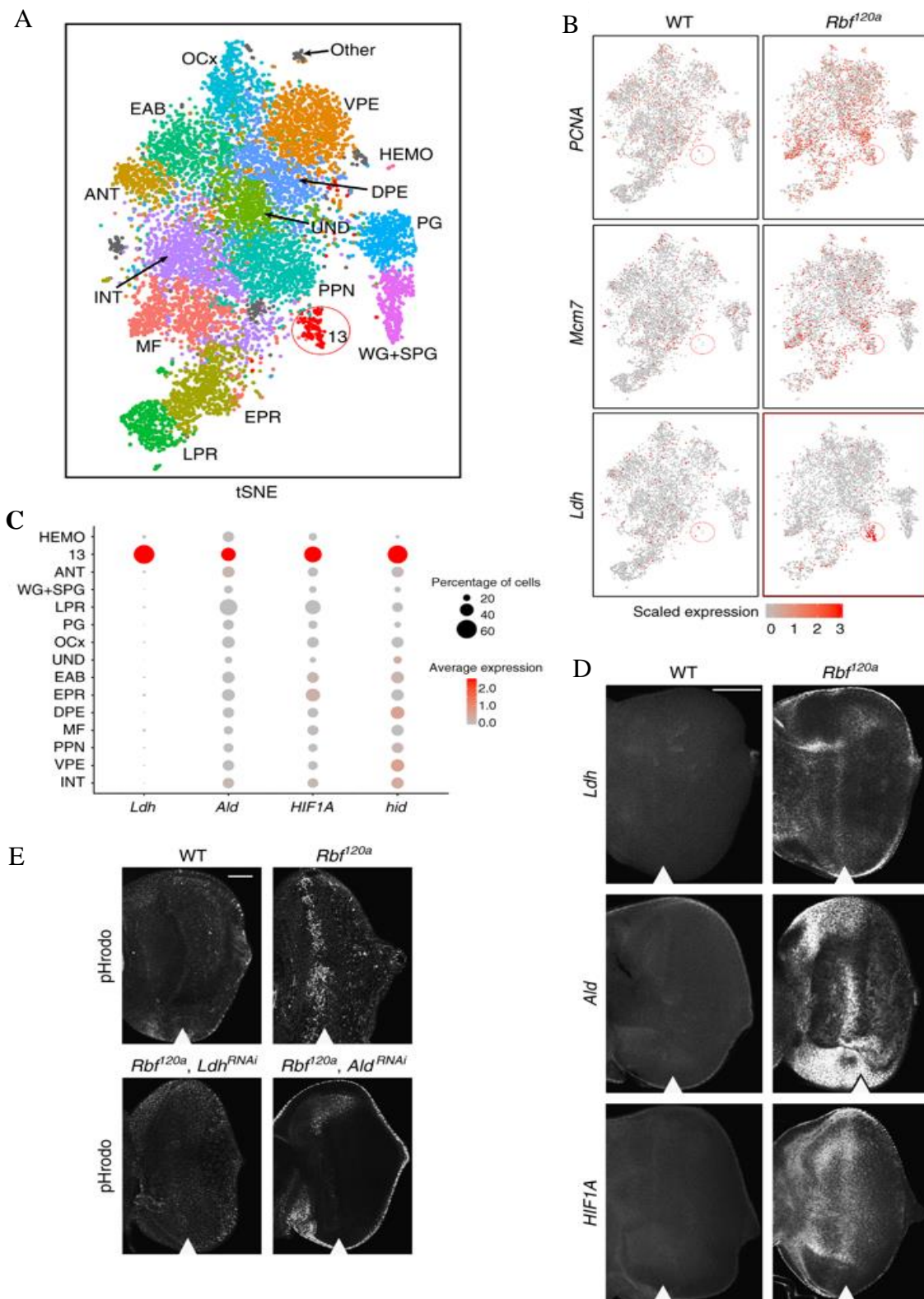


Figure 7. Wild-type and *Rbf^{120a}* scRNA-seq identifies a mutant-specific metabolic cluster.

(A) Seurat analysis using a combined wild-type and *Rbf* mutant single-cell transcriptomes. Cluster 13, outlined in red, is almost exclusively comprised of *Rbf*^{d20a} mutant cells. Clusters are color coded as in Figure 1A. (B) Feature plots on tSNE showing the expression of *PCNA*, *Mcm7* and *Ldh* in wild-type and *Rbf* mutant cells. Location of cluster 13 is outlined in red. (C) Dot plot showing expression of cluster 13 markers *Ldh*, *Ald*, *HIF1A* and *hid* throughout all clusters. (D) Expression of *Ldh*, *Ald* and *HIF1A* in the eye disc as revealed by FISH. The scale bar is 50 μ m. (E) Pattern of intracellular acidification as revealed by pHrodo. The scale bar is 50 μ m. The position of the MF is shown by white arrowhead.

4. Intracellular acidification increases apoptosis in *Rbf*^{120a}

Loss of *Rbf* results in upregulation of the apoptotic gene *hid* that sensitizes cells to apoptosis. Curiously, cell death does not occur throughout the entire *Rbf* mutant eye disc. Instead, apoptosis is restricted to a stripe of cells immediately anterior to the MF (Figure 7A) (Moon et al., 2006), which was attributed to a transient reduction in pro-survival EGFR signaling in this region. We noted that the pattern of apoptosis appears to match the location of the *Rbf* mutant-specific cluster 13 that also showed the expression of *hid* as a top marker (Figure 6C, 6D). Since upregulation of *Ald* and *Ldh* and increased intracellular acidification are the major features of this cluster, we examined their contribution to apoptosis in *Rbf* mutant. RNAi was used to downregulate expression of each *Ald*, *Ldh* and *HIF1A* in the *Rbf*^{120a} background and apoptotic cells were visualized by immunostaining for cleaved Dcp-1, the *Drosophila* homolog of effector Caspase-7. Strikingly, the downregulation of *Ldh*, *Ald* or *HIF1A* significantly suppressed cell death in *Rbf* mutant cells (Figure 7A, 7B). Collectively, these results support the interpretation that cluster 13 consists of *Rbf* mutant cells undergoing apoptosis. Importantly, intracellular acidification is not merely a consequence of apoptosis in *Rbf* mutants, since pH remains normal when cell death is induced by overexpression of *rpr* or *hid*. These results indicate that increased intracellular acidification occurs independently of the activation of pro-apoptotic signals *in vivo*.

In order to understand the regulation of *Ldh*, *Ald* and *HIF1A* in the *Rbf* mutant, we isolated chromatin from whole wild-type third-instar larvae and performed chromatin immunoprecipitation–quantitative PCR (ChIP-qPCR). Compared to a nonspecific antibody (immunoglobulin G (IgG)), the promoter regions of *Ald* and *HIF1A* were significantly enriched with Rbf, Dp, E2f1 and E2f2 (*p* value < 0.001 using a two-way analysis of variance (ANOVA) test) (Figure 7C). Since *HIF1A* is known to directly activate *Ldh* (Wang et al., 2016), its upregulation in *Rbf* mutant is likely to be

mediated by hypoxia-inducible factor-1 α (HIF1 α). These findings indicate that Rbf directly regulates the expression of *Ald* and *HIF1A* to prevent their inappropriate activation that leads to apoptosis.

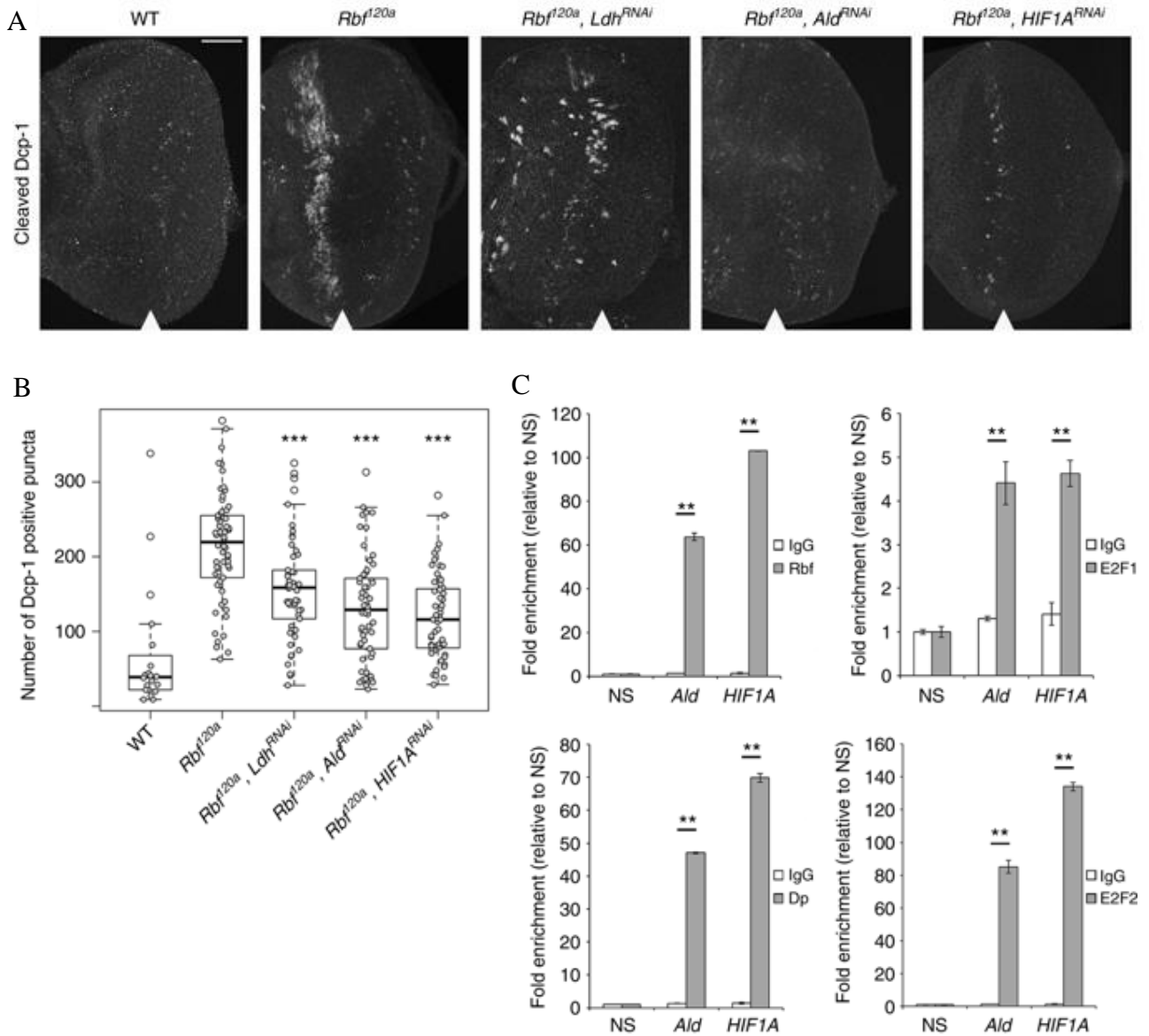


Figure 8. Upregulation of *Ldh*, *Ald* and *HIF1A* make *Rbf* mutant prone to apoptosis.

(A) Cleaved Dcp-1 immunofluorescence displays the distinct wave of apoptosis in *Rbf^{d20a}* that is rescued by downregulation of each *Ldh*, *Ald* and *HIF1A*. The scale bar is 50 μ m. (B) Quantification of cleaved Dcp-1-positive puncta in N (WT) = 20 eyes, N (*Rbf^{d20a}*) = 62 eyes, N (*Rbf^{d20a}, Ldh^{RNAi}*) = 54 eyes, N (*Rbf^{d20a}, Ald^{RNAi}*) = 58 eyes, and N (*Rbf^{d20a}, HIF1A^{RNAi}*) = 59 eyes. Center box line represents the median. The lower box limit is the first quartile. The upper box limit is the third quartile. The whiskers point at the variabilities outside the first and third quartiles. Values outside the whiskers are outliers. The hollow circles are outliers. One-way analysis of variance (ANOVA) test compares knockdowns versus *Rbf^{d20a}*, ***adjusted p value $< 5 \times 10^{-4}$. (C) ChIP-qPCR performed on

larval chromatin using Rbf, E2f1, E2f2 and Dp antibodies at the promoter regions of *Ald* and *HIF1A* relative to the negative site (NS). Data are representative of two experimental replicates and show the mean and error bars represent s.e.m., two-way ANOVA test was used, ** p value < 0.001.

D. Discussion

Here, we employed scRNA-seq to characterize a mutant phenotype. tSNE analysis on a combined scRNA-seq dataset of wild-type and *Rbf* mutant eye disc cells identified a small cluster that corresponds to *Rbf* mutant cells undergoing apoptosis. These cells exhibit elevated glycolytic activity and increased intracellular acidification that makes them uniquely sensitive to apoptosis. The results of our computational analysis were confirmed by genetic experiments. Thus, dissecting a mutant phenotype by scRNA-seq proved to be informative in unraveling acute cellular alterations that would remain undetected using conventional methods, such as whole tissue RNA-seq.

The proper assignment of computationally predicted clusters of cells inferred from scRNA-seq data to a particular cell type is dependent on prior information about cell type-specific markers. We therefore chose the *Drosophila* eye disc due to exceptional knowledge of the underlying biology. Our wild-type catalog contains ~11,500 cells and is comparable with the estimated count of 10,000 cells in the eye disc (Silies et al., 2010). This 1× cellular coverage proved sufficient to identify known domains, cell types and states. For example, progenitor cells in the anterior compartment were split into two clusters that matched previously described transcriptional domains (Bessa et al., 2002). Notably, additional markers were identified for each cell population (Figure 4A) and were often ranked higher than classical markers by computational analysis, raising the question of their importance in eye development.

Unexpectedly, photoreceptors did not segregate according to photoreceptor type (R1–8), which is usually defined by one or two markers. One possibility is that the 1× cellular coverage, or relatively low gene capture in Drop-seq, may be insufficient to detect transcriptional variations between R types. Alternatively, this may indicate that transcriptional variations between R types are minimal at the larval stage as photoreceptors are yet to express specific rhodopsins. Moreover, the

type of rhodopsin (rh) does not correlate with R type. R1 through R6 express the broad-spectrum *rh1*, while R7 and R8 express one of the four rhodopsins, *rh3*, *rh4*, *rh5* or *rh6*, responsible for color vision (Cook & Desplan, 2001). Instead, we find that large transcriptional changes occur later in photoreceptor development, at the onset of axonal projections. These axons establish the crucial connections with glial cells migrating from the brain required for photoreceptor survival (Campos, Fischbach, & Steller, 1992). Thus, scRNA-seq identified a distinct transcriptional switch occurring in developing photoreceptors, between EPR and LPR, that is important for their survival. This also illustrates a potential limitation when interpreting single-cell expression data, as previous findings may provide a biased expectation of how cells should cluster.

Finally, our results provide another perspective as to why the loss of *Rbf* sensitizes cells to apoptosis. Previous studies suggested that in *Rbf* mutant larval eye discs, E2F-dependent upregulation of *hid* and the transient reduction of pro-survival EGFR signaling leads to apoptosis anterior to the MF (Moon et al., 2006). We show that increased intracellular acidification is another key event to trigger apoptosis exclusively in this domain. Whether intracellular acidification is a cause or consequence of apoptosis remains a long-standing conundrum. The idea that cellular acidification precedes cell death was initially shown in neutrophils (Gottlieb, Giesing, Zhu, Engler, & Babior, 1995). Our results support and expand on this hypothesis by demonstrating the causative role of intracellular acidification in apoptosis *in vivo*. Since lactate is readily secreted, it is possible that glycolytic cells of population 13 induce apoptosis in adjacent cells. Although we cannot completely exclude this interpretation, we note that *hid* is highly expressed in cells of population 13 indicating that apoptosis occurs in a cell-autonomous manner. Importantly, pH remains neutral in cells undergoing *rpr*- and *hid*-induced cell death, indicating that intracellular acidification is not merely a consequence of apoptosis. Whether these effects are specific to *Rbf* mutants is uncertain as it remains to be determined if elevated expression of glycolytic genes is sufficient to induce cell death in the

wild-type discs. Nevertheless, our data suggest a simple model wherein loss of *Rbf* results in upregulation of the glycolytic genes such as *Ldh* and *Ald* immediately anterior to the MF. This in turn increases lactate production and lowers pH in these cells which promotes apoptosis. Although E2F directly elevates *Ald* expression, upregulation of *Ldh* is indirect and likely mediated by *HIF1A*, an E2F target. These findings underscore the sensitivity of scRNA-seq to detect transcriptional changes in a small cell population, which were previously missed in microarray and RNA-seq experiments using bulk samples.

Our work also highlights the value of a high-resolution wild-type cell atlas in mapping mutant scRNA-seq datasets. We and others (Karaikos et al., 2017; Macosko et al., 2015) note remarkable reproducibility of scRNA-seq among samples, as cells from almost a dozen replicates unbiasedly contributed to each population. We suggest that using single-cell genomics to profile mutant phenotypes may identify previously unappreciated biology by detecting restricted cellular perturbations.

E. **Materials and Methods**

1. **Fly stocks**

All stocks and crosses were maintained at 25 °C in vials containing standard cornmeal agar medium.

Wild-type stock used for scRNA-seq and genetic experiments: y, v; attp2 (Bloomington#36303)

Rbf120a/FM7, GFP (gift from Nick Dyson)

Mi{y[+mDint2]=MIC}NK7.1[MI02850] (BDSC),

Mi{PT-GFSTF.1}Ace[MI07345-GFSTF.1] (BDSC),

P{ry[+t7.2]=A92}nrm[A37] sr[1] e[s] ca[1] (BDSC),

P{w[+mC]=lacW}mirr[B1-12]/TM6B, P{w[+mC]=tub-QS.P}4A Tb[1] (BDSC),

P{w[+mC]=lacW}sls[j1D7] (BDSC)

P{ry[+t7.2]=lArB}cpo[2] ry[506](BDSC)

Mi{y[+mDint2]=MIC}pigs[MI11007] (BDSC)

Rbf120a ey-FLP / Y; act5C>FRT>stop>FRT>GAL4, UAS-GFP (gift from Nam-Sung Moon),

P{y[+t7.7] v[+t1.8]=TRiP.JF02071}attP2 (BDSC),

P{y[+t7.7] v[+t1.8]=TRiP.HMC06145}attP40 (BDSC),

UAS-LdhRNAi (VDRC-KK#110190),

P{y[+t7.7] v[+t1.8]=TRiP.HMS00832}attP2 (BDSC),

P{y[+t7.7] v[+t1.8]=TRiP.HMS00833}attP2 (BDSC),

P{w[+mC]=GAL4-ninaE.GMR}12 (BDSC),

P{w[+mC]=GMR-rpr.H}S (BDSC),

P{w[+mC]=GMR-hid}SS1, y[1] w[*] P{ry[+t7.2]=neoFRT}19A; P{w[+m*]=GAL4-ey.H}SS5,

P{w[+mC]=UAS-FLP.D}JD2 (BDSC)

P{y[+t7.7] v[+t1.8]=VALIUM20-EGFP.shRNA.3}attP40 (BDSC)

Final genotypes:

Rbf120a, ey-FLP / Y; act5C>FRT>stop>FRT>GAL4 / UAS-AldRNAi

Rbf120a, ey-FLP / Y; act5C>FRT>stop>FRT>GAL4 / UAS-LdhRNAi

Rbf120a, ey-FLP / Y; act5C>FRT>stop>FRT>GAL4 / UAS-simaRNAi

GMR-GAL4 / UAS-EGFPRNAi.

2. Fluorescent in situ hybridization

Third instar wandering larval eye discs were dissected in phosphate-buffered saline (PBS) and fixed in 4% formaldehyde+1× PBS for 15 min on ice, then in 4% formaldehyde+1× PBS + 0.1% DOC + 0.1% Triton X-100 for 15 min at room temperature (RT). Proteinase K (New England Biolabs) was used at a final concentration of 0.5 µg/mL in PBT (1× PBS+0.1% Tween-20). Digoxigenin (DIG)-labeled RNA probes were denatured at 90 °C for 4 min, using 3–5 µL probe in 25 µL hybridization solution and hybridization was carried out following ThermoFisher FISH Tag RNA Kit suggested hybridization protocol (<https://tools.thermofisher.com/content/sfs/manuals/mp32952.pdf>). The protocol was modified after Post-Hybridization step 13.7. Mouse anti-DIG (Sigma, #11333062910 (1:250) was added in PBT and samples were incubated overnight at 4 °C. After overnight incubation, samples were washed 5 times in PBT for 5 min. Anti-mouse secondary Cy3 antibody (Jackson ImmunoResearch (1:300)) was added in PBT and incubated for 1 h. Finally, samples were washed three times in PBT for 5 min then mounted in FluorSave (EMD Millipore, #345789) on a glass slide. All steps carried out with gentle rocking on nutator, unless specified. Samples were imaged using Zeiss Confocal microscope.

3. DIG-labeled probes

PCR primers were created for genes, with an average fragment length of 500 bp. Whole third-instar larvae complementary DNA (cDNA) was amplified by PCR using Q5 High-Fidelity DNA Polymerase (New England Biolabs, #M0491S). DNA purification was performed with QiaGen QIAEX II 20051. A second round of PCR was carried out using 1 ng of cDNA from first PCR, and reverse primers containing the T7 primer; AGGGATCCTAATACGACTCACTATAGGGCCCGGGGC. T7 RNA polymerase (Sigma, #10881767001) along with DIG RNA labeling mix (Sigma #11277073910) were used for in vitro transcription and incubated for 5–6 h at 37 °C. RNA was then purified using NucleoSpin RNA clean-up (Macherey-Nagel, #740948). Gene primers were as follows:

CG9336: CCTGAAGTTCGAGGCTGATG, CAACCATTCGTCGTGCATTT,

twit: TGAAGCCCATCCATCCAACC, GGCAAGTGTCCCGAAACTA,

Ldh (ImpL3): ACGGCTCCAACCTTCTGAAG, TTGTTCAACTTCGGTGGGAG,

Ald: CTCTGAGGATGAGGTCACCA, ATGTTCTCCTTCTTGCCAGC,

HIF1A (sima): GAGGGTGGCTTCGAGTTTAG, GGTTTCCACTCTCCTCTGC.

4. Chromatin immunoprecipitation–quantitative polymerase chain reaction

Fifteen third-instar larvae were collected and homogenized using a tissue grinder with 60 mM KCl, 15 mM NaCl, 4 mM MgCl₂, 15 mM HEPES (pH 7.6), 0.5% Triton X-100, 0.5 mM dithiothreitol (DTT), 10 mM sodium butyrate and protease inhibitor cocktail (Complete, Roche). Crosslinking was done in 1.8% formaldehyde for 15 min followed by 225 mM Glycine. Lysis buffer is 15 mM HEPES

at pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 1% Triton X-100, 0.5 mM DTT, 0.1% SDS, 0.5% lauroylsarcosine and 10 mM sodium butyrate with protease inhibitor cocktail (Complete, Roche). Chromatin was sheared using a Branson 450 Sonifier. Antibodies used for immunoprecipitation are rabbit anti-DP (#212, gift from Nick Dyson), mouse anti-Rbf (DX3/DX5, ratio 1:1 gift from Nick Dyson), rabbit anti-E2f1 (#210 gift from Nick Dyson and anti-E2f1 from Gunter Reuter, ratio 1:1), rabbit anti-E2f2 antibodies (#79 gift from Nick Dyson), and rabbit IgG (Sigma) as nonspecific antibody. Samples were pulled down with Protein G Dynabeads (Invitrogen), then washed with lysis buffer four times, with TE (pH 8) twice and eluted. Decrosslinking was done overnight at 65 °C. Samples were treated with RNase A (Sigma) for 1 h at 37 °C followed by proteinase K for 2 h at 50 °C. Then, DNA was purified by phenol–chloroform extraction and overnight ethanol precipitation.

The immunoprecipitated DNA along with the input genomic DNA (collected before precipitation) was quantified by qPCR. SensiFast SYBR No-ROX Mix (Bioline) was used for qPCR. Reactions were run on a LightCycler 480 (Roche). Primer sequences are: Aldolase forward: TTTACCGCCCAAACGAAAGC and reverse: GCCAAGCGCTTTAAATTCCC; sima (HIF1A) forward: AAACGACCAACGCACATACG and reverse: TTTGGTTCGCGTGCAATACC. A negative sequence site that does not contain any predicted E2F-binding sites is forward: TGTGTATGCCTTGCTTGAC and reverse TCTATGCACACGCTCTACTGAG. The protein enrichment was calculated as the percentage of immunoprecipitated DNA relative to input DNA (prior DNA precipitation) for each antibody. Data presented are relative to the negative binding site for each CHIP. Each sample was measured twice. The two-way ANOVA was used to calculate the p values.

5. Immunofluorescence

Wandering third-instar larval eye discs were dissected in 1× PBS and fixed in 4% formaldehyde+1× PBS for 15 min, permeabilized in 0.3% PBS-T (1× PBS, 0.3% Triton X-100) two times for 10 min and then incubated with antibodies overnight at 4 °C in 1× PBS+10% normal donkey serum (NDS, Jackson ImmunoResearch)+0.1% Triton X-100 blocking serum. The following day, samples were washed in 0.1% PBS-T (1× PBS+0.1% Triton X-100) three times for 5 min. Samples were then incubated with appropriate fluorescently labeled secondary antibodies (Jackson ImmunoResearch) for 1 h in 1× PBS+10% NDS+0.1% Triton X-100 followed by 4',6-diamidino-2-phenylindole (DAPI) for 5 min. Finally, samples were washed five times for 5 min then mounted in FluorSave (EMD Millipore) on glass slides. All steps were carried out at RT and with gentle rocking, unless specified otherwise. Whenever fluorescent images have been compared, they have been obtained with the same acquisition and display settings.

Primary antibodies used were: anti-β-gal: DSHB 40-1a (1:200), anti-cleaved Drosophila Dcp-1: Cell Signaling Asp216 (1:500), anti-Dac DSHB mAbdac1-1 (1:100), anti-Elav: DSHB 7E8A10 (1:200), anti-GFP (FITC): Abcam ab6662 (1/1000), anti-Hairy: from T. Orenic (1:4), anti-Hth: from Richard Mann (1:2000), anti-Notch DSHB C458.2H (1:100), anti-Repo: DSHB 8D13 (1:500), anti-Senseless: from H. Bellen (1:100).

6. BrdU labeling

Dissected eye discs were incubated with Schneiders media (Fisher, 21720024) containing BrdU at a final concentration of 0.2 mg/mL for 2 h at RT. The tissues were fixed in 1.5% formaldehyde+0.01% Tween-20+1× PBS overnight at 4 °C. Tissues were washed twice with 1× PBS

twice for 30 min. The eye discs were digested using 5 μ L DNase (Promega RQ1, #M6101) in 1 \times PBS at a total volume of 100 μ L and incubated for 40 min at 37 °C, then washed twice for 5 min with 1 \times PBS+0.3% Triton. Tissues were blocked for an hour (see Immunofluorescence) and then the BrdU antibody was added (1:50 (BD Biosciences, 347580)). The Immunofluorescence protocol was followed from this step, see above.

7. Cleaved Dcp-1 quantification

Confocal images were inverted and cropped in Adobe Photoshop to remove antennal discs, reducing background staining. The following filters were applied: Brightness +125, Contrast +50, levels: 0, 1.0, 220. Dcp-1-positive cells were then counted using ImageJ software using conservative parameters in order to minimize detection of background; Threshold, BlackBackground is false, Convert to Mask, Fill Holes, Watershed, Analyze Particles (size = 50–Infinity). The same parameters were used for all eye discs in each genotype. Quantification was done using two RNAi lines for Ald (BL#26301 and BL#65883) and counting 18 and 42 eye discs respectively, and HIF1A/sima (BL#33894 (48 eye discs), BL#33895 (11 eye discs)) and the datasets combined. A total number of 62 Rbf120a and 54 Ldh knockdown eyes were used. Pooled data were analyzed in R using one-way ANOVA followed by Tukey's HSD post-hoc test of significance. A box plot was created, individual eyes plotted as gray circles and outliers as white circles.

8. Live pH staining

Third-instar larval eye discs were dissected and washed twice with 1 \times PBS (pH 7.4) Then, it was incubated with 1 μ L of pHrodo Red AM Intracellular pH (Life Technologies cat. no. P35372)

Indicator probe+98 μ L 1 \times PBS (pH 7.4)+1 μ L of the buffer included in pHrodo for 25 min. The eye discs were mounted in FluorSave (EMD Millipore #345789) on a glass slide then imaged using Zeiss Confocal microscope.

9. Tissue dissociation for Drop-seq

Eye discs were dissected within 1 h and then transferred to a microcentrifuge tube where they were dissociated in a final concentration of 2.5 mg/mL Collagenase (Sigma #C9891) and 1 \times trypsin (Sigma #59418C) in Rinaldini solution. The microcentrifuge tube was horizontally positioned on a 225 rpm shaker for 20 min at RT. This dissociation protocol resulted in healthy single cells and less than 5% clumps.

10. Drop-seq

We followed the Drop-seq protocol⁷ Online-Drop-seq-Protocol-v.-3.1-Dec-2015.pdf at <http://mccarrolllab.com/dropseq/> while having four modifications. the final concentration of Sarkosyl in the lysis buffer was 0.4%. The cycles in the PCR step post exonuclease are 4 cycles, and then 12. The cDNA Post PCR was purified twice with 0.6 \times AMPure beads. The tagmented DNA for sequencing was purified twice: first using 0.6 \times AMPure beads and the second time using 1 \times AMPure beads.

11. scRNA-seq data analysis

Illumina paired end raw sequences (FastQ file) were processed for read alignment and gene expression quantification. Drop-seq single-cell data were analyzed using the data analysis protocol described in Drop-seq cook-book (version 1.2 Jan 2016)⁷ (<http://mccarrolllab.com/dropseq/>) and used the Drop-seq_tools-1.13. We used STAR aligner to align the reads against *Drosophila melanogaster* genome version BDGP6 (Ensembl) and corresponding gene model was extracted from Ensembl version 90. Quality of reads and mapping were checked using the program FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Digital Gene Expression (DGE) matrix data obtained from an aligned library is done using the Drop-seq program DigitalExpression (integrated in Drop-seq_tools-1.13). Number of cells that were extracted from aligned BAM file is based on knee plot which extracts the number of reads per cell, then plot the cumulative distribution of reads and select the knee of the distribution.

Drop-seq alignment was performed as mentioned in Drop-seqAlignmentCookbookv1.2Jan2016 at <http://mccarrolllab.com/dropseq/>. Alignment was performed using the *Drosophila melanogaster* BDGP6 genome while excluding mitochondria encoded genes. Quality checks were performed then matrices were created for downstream bioinformatic analyses.

12. Cell clustering and discovery of cell types

The following computational figures: dotplots, feature plots, tSNE, heatmaps and gene/gene plots were generated using Seurat.

Following alignment, the extracted gene matrices were subjected to unsupervised identify probable cell types using R package Seurat (R version 3.3.2, Seurat version v2.2.1). In Seurat analysis we performed initial quality control analysis and low-quality cells were filtered out using 200 and

3000 gene/cell as a low and top cutoff, respectively, $\text{min.cell} = 3$ and $\text{total.expr} = 1e4$. Statistically significant principal components were determined using JackStraw procedure. The first 20 principle components were selected to run non-linear dimensional reduction (tSNE) while a resolution of 1.6 was applied in both WT (11,416 cells) and WT+Rbf (5591+5203) analyses. The WT analysis was performed only on the cell barcodes as listed in WT.xlsx whereas the WT+Rbf analysis was performed on WT_Rbf.txt barcodes listed in text file WT_Rbf.xlsx. These files exclude cells having a predominant ribosomal gene expression signature coming from each sample. Following a resolution of 1.6, the interommatidial cells (INT) split into 3 clusters in the WT analysis. Two of these have the same top markers indicating an artifact from Seurat analysis and, as a result, were both labeled as INT. The third population has a cell-cycle signature with top markers such as Claspin, stg and PCNA. This population was therefore labeled as SMW. Populations having heat shock proteins as top markers and any remaining ribosomal proteins were grouped and labeled as other in dark gray.

13. Gene set enrichment analysis

Gene set enrichment analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states. We applied this statistical analysis to find which Gene Ontology (GOBP) shows statistically significant, concordant differences in each cell cluster.

Using average expression of genes in each cell cluster, we selected UND, EPR and LPR populations for GSEA analysis. We calculated the \log_2 fold change ($\log_2\text{FC} = \log_2(B) - \log_2(A)$); where B is the average expression value of a specific gene in the cluster of interest and A is the median expression value of that gene in the other clusters. The $\log_2\text{FC}$ values were then ranked and ranked genes matched to categories of GOBP (<http://www.go2msig.org/cgi-bin/prebuilt.cgi?taxid=7227>).

The normalized enrichment score for each process was calculated based on the Broad Institute's GSEA directions after 10,000 permutations to obtain accurate false discovery rate (FDR) values (<http://software.broadinstitute.org/gsea/index.jsp>). We used gitools v1.8.4 to create the heatmap while using $\text{FDR} < 0.05$ as the cutoff for significantly enriched, upregulated biological processes. If the processes are down-regulated, they were labeled as $\text{FDR} = 1$.

14. Monocle 2 trajectory

Single-cell pseudotime trajectory was constructed using Monocle 229 (monocle version 2.8.0, with R version 3.5.0) using UND, PPN, MF, INT, EPR and LPR cells from the WT scRNA-seq data. The following parameters were used to generate the plot: top 1500 genes were selected, $\text{Rho} = 20$ and $\text{Delta} = 20$, $\text{qval} < 0.01$ and DDRTree dimension reduction method.

F. Cited literatures

Bazigou, E., Apitz, H., Johansson, J., Lorén, C. E., Hirst, E. M. A., Chen, P. L., ... Salecker, I. (2007). Anterograde Jelly belly and Alk Receptor Tyrosine Kinase Signaling Mediates Retinal Axon Targeting in *Drosophila*. *Cell*, 128(5), 961–975. <https://doi.org/10.1016/j.cell.2007.02.024>

Bessa, J., Gebelein, B., Pichaud, F., Casares, F., & Mann, R. S. (2002). Combinatorial control of *Drosophila* eye development by Eyeless, Homothorax and Teashirt. *Genes & Development*, 2415–2427. <https://doi.org/10.1101/gad.1009002.terior>

Bhattacharya, S., Stewart, B. A., Niemeyer, B. A., Burgess, R. W., McCabe, B. D., Lin, P., ... Schwarz, T. L. (2002). Members of the synaptobrevin/vesicle-associated membrane protein (VAMP) family in *Drosophila* are functionally interchangeable in vivo for neurotransmitter release and cell viability. *Proceedings of the National Academy of Sciences*, 99(21), 13867–13872. <https://doi.org/10.1073/pnas.202335999>

Blockus, H., & Chédotal, A. (2016). Fly Dscams Can Also Help You Find the Right Partners. *Neuron*, 89(3), 423–425. <https://doi.org/10.1016/j.neuron.2016.01.021>

- Campos, a R., Fischbach, K. F., & Steller, H. (1992). Survival of photoreceptor neurons in the compound eye of *Drosophila* depends on connections with the optic ganglia. *Development* (Cambridge, England), 114, 355–366.
- Cook, T., & Desplan, C. (2001). Photoreceptor subtype specification: From flies to humans. *Seminars in Cell and Developmental Biology*, 12(6), 509–518. <https://doi.org/10.1006/scdb.2001.0275>
- DeSalvo, M. K., Hindle, S. J., Rusan, Z. M., Orng, S., Eddison, M., Halliwill, K., & Bainton, R. J. (2014). The *Drosophila* surface glia transcriptome: Evolutionary conserved blood-brain barrier processes. *Frontiers in Neuroscience*, 8(OCT), 1–22. <https://doi.org/10.3389/fnins.2014.00346>
- Dick, F. A., Goodrich, D. W., Sage, J., & Dyson, N. J. (2018). Non-canonical functions of the RB protein in cancer. *Nature Reviews Cancer*, 18(7), 442–451. <https://doi.org/10.1038/s41568-018-0008-5>
- Dominguez-Cejudo, M. A., & Casares, F. (2015). Antero-posterior patterning of *Drosophila* ocelli requires an anti-repressor mechanism within the hh-pathway mediated by the Six3 gene Optix. *Development*, 142(16), 2801–2809. <https://doi.org/10.1242/dev.125179>
- Emerald, B. S., Curtiss, J., Mlodzik, M., & Cohen, S. M. (2003). Distal antenna and distal antenna related encode nuclear proteins containing pipsqueak motifs involved in antenna development in *Drosophila*. *Development* (Cambridge, England), 130(6), 1171–1180. <https://doi.org/10.1242/dev.00323>
- Firth, L. C., & Baker, N. E. (2007). Spitz from the retina regulates genes transcribed in the second mitotic wave, peripodial epithelium, glia and plasmatocytes of the *Drosophila* eye imaginal disc. *Developmental Biology*, 307(2), 521–538. <https://doi.org/10.1016/j.ydbio.2007.04.037>
- Franzdóttir, S. R., Engelen, D., Yuva-Aydemir, Y., Schmidt, I., Aho, A., & Klämbt, C. (2009). Switch in FGF signalling initiates glial differentiation in the *Drosophila* eye. *Nature*, 460(7256), 758–761. <https://doi.org/10.1038/nature08167>
- Gottlieb, R. A., Giesing, H. A., Zhu, J. Y., Engler, R. L., & Babior, B. M. (1995). Cell acidification in apoptosis: granulocyte colony-stimulating factor delays programmed cell death in neutrophils by up-regulating the vacuolar H(+)-ATPase. *Proceedings of the National Academy of Sciences of the United States of America*, 92(13), 5965–5968. <https://doi.org/10.1073/pnas.92.13.5965>
- Harvie, P. D., Filippova, M., & Bryant, P. J. (1998). Genes expressed in the ring gland, the major endocrine organ of *Drosophila melanogaster*. *Genetics*, 149(1), 217–231.
- Kania, A., Han, P. L., Kim, Y. T., & Bellen, H. (1993). Neuromusculin, a *Drosophila* Gene Expressed in Peripheral Neuronal Precursors and Muscles, Encodes a Cell Adhesion Molecule. *Neuron*, 11(4), 673–687. [https://doi.org/10.1016/0896-6273\(93\)90078-6](https://doi.org/10.1016/0896-6273(93)90078-6)
- Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., ... Zinzen, R. P. (2017). The *Drosophila* embryo at single-cell transcriptome resolution. *Science*, 358(6360), 194–199. <https://doi.org/10.1126/science.aan3235>

- Lai, E. C., Bodner, R., Kavalier, J., Freschi, G., & Posakony, J. W. (2000). Antagonism of notch signaling activity by members of a novel protein family encoded by the bearded and enhancer of split gene complexes. *Development (Cambridge, England)*, 127, 291–306.
- Lakey, a, Labeit, S., Gautel, M., Ferguson, C., Barlow, D. P., Leonard, K., & Bullard, B. (1993). Kettin, a large modular protein in the Z-disc of insect muscles. *The EMBO Journal*, 12(7), 2863–2871. <https://doi.org/10.1002/j.1460-2075.1993.tb05948.x>
- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., ... McCarroll, S. A. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161(5), 1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>
- McNeill, H., Yang, C. H., Brodsky, M., Ungos, J., & Simon, M. A. (1997). Mirror encodes a novel PBX-class homeoprotein that functions in the definition of the dorsal-ventral border in the *Drosophila* eye. *Genes and Development*, 11(8), 1073–1082. <https://doi.org/10.1101/gad.11.8.1073>
- Moon, N.-S., Di Stefano, L., & Dyson, N. (2006). A Gradient of Epidermal Growth Factor Receptor Signaling Determines the Sensitivity of rbf1 Mutant Cells to E2F-Dependent Apoptosis. *Molecular and Cellular Biology*, 26(20), 7601–7615. <https://doi.org/10.1128/MCB.00836-06>
- Moses, K. (2002). *Drosophila Eye Development*. (K. Moses, Ed.). Springer, Berlin, Heidelberg.
- Pilgram, G. S. K., Potikanond, S., van der Plas, M. C., Fradkin, L. G., & Noordermeer, J. N. (2011). The RhoGAP crossveinless-c Interacts with Dystrophin and Is Required for Synaptic Homeostasis at the *Drosophila* Neuromuscular Junction. *Journal of Neuroscience*, 31(2), 492–500. <https://doi.org/10.1523/JNEUROSCI.4732-10.2011>
- Pines, M. K., Housden, B. E., Bernard, F., Bray, S. J., & Roper, K. (2010). The cytolinker Pigs is a direct target and a negative regulator of Notch signalling. *Development*, 137(6), 913–922. <https://doi.org/10.1242/dev.043224>
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, 14(10), 979–982. <https://doi.org/10.1038/nmeth.4402>
- Roignant, J.-Y., Legent, K., Janody, F., & Treisman, J. E. (2010). The transcriptional co-factor Chip acts with LIM-homeodomain proteins to set the boundary of the eye field in *Drosophila*. *Development*, 137(2), 273–281. <https://doi.org/10.1242/dev.041244>
- Sasse, S., Neuert, H., & Klämbt, C. (2015). Differentiation of *Drosophila* glial cells. *Wiley Interdisciplinary Reviews: Developmental Biology*, 4(6), 623–636. <https://doi.org/10.1002/wdev.198>
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, 33(5), 495–502. <https://doi.org/10.1038/nbt.3192>

- Silies, M., Yuva-Aydemir, Y., Franzdóttir, S. R., & Klämbt, C. (2010). The eye imaginal disc as a model to study the coordination of neuronal and glial development. *Fly*, 4(1), 71–79. <https://doi.org/10.4161/fly.4.1.11312>
- Sinenko, S. A., Mandal, L., Martinez-Agosto, J. A., & Banerjee, U. (2009). Dual Role of Wingless Signaling in Stem-like Hematopoietic Precursor Maintenance in *Drosophila*. *Developmental Cell*, 16(5), 756–763. <https://doi.org/10.1016/j.devcel.2009.03.003>
- Tennessen, J. M., Baker, K. D., Lam, G., Evans, J., & Thummel, C. S. (2011). The *Drosophila* estrogen-related receptor directs a metabolic switch that supports developmental growth. *Cell Metabolism*, 13(2), 139–148. <https://doi.org/10.1016/j.cmet.2011.01.005>
- Van Den Heuvel, S., & Dyson, N. J. (2008). Conserved functions of the pRB and E2F families. *Nature Reviews Molecular Cell Biology*, 9(9), 713–724. <https://doi.org/10.1038/nrm2469>
- Wagh, D. A., Rasse, T. M., Asan, E., Hofbauer, A., Schwenkert, I., Dürrbeck, H., ... Buchner, E. (2006). Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in *Drosophila*. *Neuron*, 49(6), 833–844. <https://doi.org/10.1016/j.neuron.2006.02.008>
- Wang, C. W., Purkayastha, A., Jones, K. T., Thaker, S. K., & Banerjee, U. (2016). In vivo genetic dissection of tumor growth and the Warburg effect. *ELife*, 5(September), 1–21. <https://doi.org/10.7554/eLife.18126>
- Xu, X. L., Singh, H. P., Wang, L., Qi, D. L., Poulos, B. K., Abramson, D. H., ... Cobrinik, D. (2014). Rb suppresses human cone-precursor-derived retinoblastoma tumours. *Nature*, 514(7522), 385–388. <https://doi.org/10.1038/nature13813>
- Zhao, D., Zou, S. W., Liu, Y., Zhou, X., Mo, Y., Wang, P., ... Guan, K. L. (2013). Lysine-5 acetylation negatively regulates lactate dehydrogenase a and is decreased in pancreatic cancer. *Cancer Cell*, 23(4), 464–476. <https://doi.org/10.1016/j.ccr.2013.02.005>

III. AMALGAM REGULATES THE RECEPTOR TYROSINE KINASE PATHWAY THROUGH SPROUTY IN GLIAL CELL DEVELOPMENT

A. Summary

The receptor tyrosine kinase (RTK) pathway plays an essential role in development and disease by controlling cell proliferation and differentiation. Here, we profile the *Drosophila* larval brain by single cell RNA-sequencing and identify *Amalgam* (*Ama*), encoding a cell adhesion protein of the immunoglobulin IgLON family, that regulates the RTK pathway activity during glial cell development. Depletion of *Ama* reduces cell proliferation, affects glial cell type composition and disrupts the blood-brain barrier (BBB) that leads to hemocyte infiltration and neuronal death. We show that *Ama* depletion lowers RTK activity by upregulating Sprouty (*Sty*), a negative regulator of RTK pathway. Knockdown of *Ama* blocks oncogenic RTK signaling activation in the *Drosophila* glioma model and halts malignant transformation. Finally, knockdown of a human ortholog of *Ama*, LSAMP, results in upregulation of SPOUTY2 in glioblastoma cell lines suggesting that the relationship between *Ama* and *Sty* is conserved.

B. Introduction

The receptor tyrosine kinase (RTK) pathway regulates growth, cell proliferation, differentiation, and survival, and therefore has a prominent role in development and in cancer (*Regad, 2015*). It is comprised of surface receptors that are activated and dimerized upon binding by growth factors and propagate signals through the RAS-RAF-MEK-ERK axis. The RTK pathway is tightly regulated at multiple levels including inhibition by members of the Sprouty gene family. This type of

inhibition, measured by a decrease in ERK phosphorylation, occurs when Sprouty binds to either GRB2 or RAF and disrupts the propagation of the signal (*Masoumi-Moghaddam et al., 2014*). However, beyond this classical role of binding to different components of the RTK pathway, the mechanism of Sprouty repression and its regulation is not fully understood (*Masoumi-Moghaddam et al., 2014*). Notably, Sprouty is a potential therapeutic target in many neurological diseases (*Hausott and Klimaschewski, 2019*) and was shown to act either as a tumor suppressor or as an oncogene in different cancers (*Masoumi-Moghaddam et al., 2014*). The latter suggests that Sprouty's function and its impact on RTK signaling is likely cell type dependent, which further complicates investigation of Sprouty's role in cancer given high heterogeneity in tumors.

Recent advances in single cell RNA-sequencing (scRNA-seq) allow characterization of cell diversity at a high resolution and help to dissect cellular heterogeneity. This technology enabled identification of novel biomarkers in cell types as well as the discovery of rare cell subtypes that would have otherwise been missed in bulk RNA-seq (*Papalexi and Satija, 2018*). Additionally, scRNA-seq has been shown to uncover dynamic spatiotemporal processes such as differentiation (*Wagner et al, 2016*) and complex cell-to-cell responses following stimuli (*Liu and Trapnell, 2016*).

Drosophila has proven to be a highly advantageous model to study signaling pathways given its amenability to genetic analysis and remarkable conservation of signal transduction pathways between humans and flies (*Chatterjee and Deng, 2019*). This is best illustrated by pioneering studies of the RTK signaling in *Drosophila* (*Simon et al., 1991*) leading to discovery of several members of the pathway that are preserved across species (*Perrimon, 1994*). For instance, Sprouty was initially discovered in *Drosophila* (*Hacohen et al., 1998*) and it was later found to have a similar role in mammals (*Impagnatiello et al., 2001*). *Drosophila* has also been instrumental in studying cancers as several fly tumor models of the lung (*Levine and Cagan, 2016*), eye (*Pagliarini and Xu, 2003*), blood

(Osman *et al.*, 2009), glia (Read *et al.*, 2009), and colon (Bang *et al.*, 2016) cancers have recently been established.

Here, we employed scRNA-seq to identify *Amalgam* (*Ama*), a member of the IgLON family encoding a cell adhesion immunoglobulin, as a new regulator of the RTK pathway that acts through *sprouty* (*sty*). Depletion of *Ama* decreases glial cell proliferation, disrupts the blood-brain barrier (BBB), and results in a dramatic increase of hemocytes infiltration in the brain. We show that knockdown of *Ama* increases *Sty* which reduces RTK signaling pathway in glia during development and in a *Drosophila* glioma model. Notably, the impact of knockdown of *Ama* on *Sprouty* is conserved in human glioblastoma cell lines suggesting a functional conservation across species.

C. **Results**

1. ***Ama* is required for glial cell development**

RTK signaling has been extensively studied in the *Drosophila* eye primarily during photoreceptor differentiation but less so in glial cells, called wrapping glia (WG), that envelope axonal projections. This glial cell type differentiates from perineurial glia (PG) as PG are migrating from the brain towards the developing eye disc during the third instar larval stage. It has been shown that a transient increase in RTK signaling is required during the transition from perineurial to wrapping glia (Franzdóttir *et al.*, 2009).

In order to identify novel genes important in this process, we examined the scRNA-seq dataset of the larval eye imaginal disc (Ariss *et al.*, 2018) to isolate genes that are specifically expressed in the perineurial and wrapping glia. *Amalgam* (*Ama*) was found to be one of the top markers in glial cells (Figure 8A) with a higher expression in wrapping glia than in perineurial glia, which parallels

the transient increase in RTK signaling during this transition. *Ama* is an adhesion protein of the immunoglobulin super-family (Seeger et al., 1988) and affects axon pathfinding in *Drosophila* embryos (Frémion et al., 2000). To determine whether *Ama* is important in glia, we employed the *UAS-Gal4* system to knockdown *Ama* by RNAi using a *UAS-Ama^{RNAi}* transgene driven by a pan-glial *repo-Gal4* driver or a subperineurial glia (SPG) specific *moody-Gal4* driver. Depletion of *Ama* in all glial cells *repo-Gal4* resulted in early pupal lethality (~24h after pupa formation) indicating that *Ama* may have an essential role in development. This result was validated with another *UAS-Ama^{RNAi}* transgene, thus, confirming the specificity of RNAi knockdown.

We began the investigation of *Ama*'s function in glia by confirming its glial specific expression. The *repo-Gal4 UAS-mCD8-GFP* eye discs that express GFP in glial cell membranes were dissected and subjected to fluorescent in situ hybridization (FISH) protocol using the *Ama* specific Stellaris probes. As shown in Figure 8B, *Ama* transcripts were detected exclusively in GFP positive glial cells. In a complementary approach, *Ama*'s expression was examined by immunofluorescence using an *Ama-Gal4* enhancer trap line crossed to *UAS-GFP* while glial cells were visualized with pan-glial Repo antibody. In agreement with FISH, *Ama* (GFP positive) was detected only in glial cells (Figure 8C). Notably, intensity of the GFP signal was higher in wrapping glia than in perineurial glia. Indeed, the more intense GFP signal colocalizes with Cut, a wrapping glia marker that is not expressed in perineurial glia. We also noticed that *Ama* is expressed in a subset of cells of the ventral peripodial membrane epithelium (VPE), as also revealed in the scRNA-seq dataset. These results suggest that *Ama*'s expression is primarily restricted to glial cells, with wrapping glia expressing a higher level of *Ama* than perineurial glia.

Since perineurial glial cells originate in the brain before migrating towards the eye disc, we examined the expression of *Ama* in the third instar larval brain using a G-TRACE cell lineage tracing

system (Evans et al., 2009). This technique provides a visual representation of temporal gene expression, labeling cells that expressed *Ama* in the past with GFP, while RFP labels cells that currently express *Ama*. With the exception of a few cells, *Ama* was almost exclusively detected in glia based on the colocalization with Repo (Figure 8D). Notably, every glial cell that expressed RFP also expressed GFP indicating that *Ama* remains expressed in the same cell throughout development and is not turned off (Figure 8D). Individual confocal sections show that *Ama* is expressed in most types of glial cells but was largely absent in the medulla region of the optic lobe. Interestingly, only half of glial cells expressed *Ama* (Figure 8E). The glial specific *Ama* expression was further confirmed by FISH (Figure 8F). Importantly, the signal was lost when *Ama* was depleted by RNAi using *repo-Gal4* (Figure 8F).

To characterize the consequences of glia specific *Ama* knockdown, the *repo>Ama^{RNAi}* eye discs were stained with Repo antibody to visualize glial cells. Strikingly, no glial cells were found in *repo>Ama^{RNAi}* (Figure 8G). Since glial cells play a crucial role in photoreceptor axon guidance (Xie et al., 2014), we examined axons in the brain using a 24B10 antibody. In the wild type, axonal projections are found in the lamina and medulla of the optic lobe, however, axons fail to land in their respective compartments in the brains when glial cells are depleted of *Ama* (Figure 8H). Furthermore, there was a severe reduction in the number of glial cells in *repo>Ama^{RNAi}* brain and its size of was significantly smaller (Figure 8I).

Thus, we concluded that *Ama* is essential in glial cell development. *Ama* depletion results in severe reduction of glia in the brain and prevents migration of glial cells from the brain towards the eye disc.

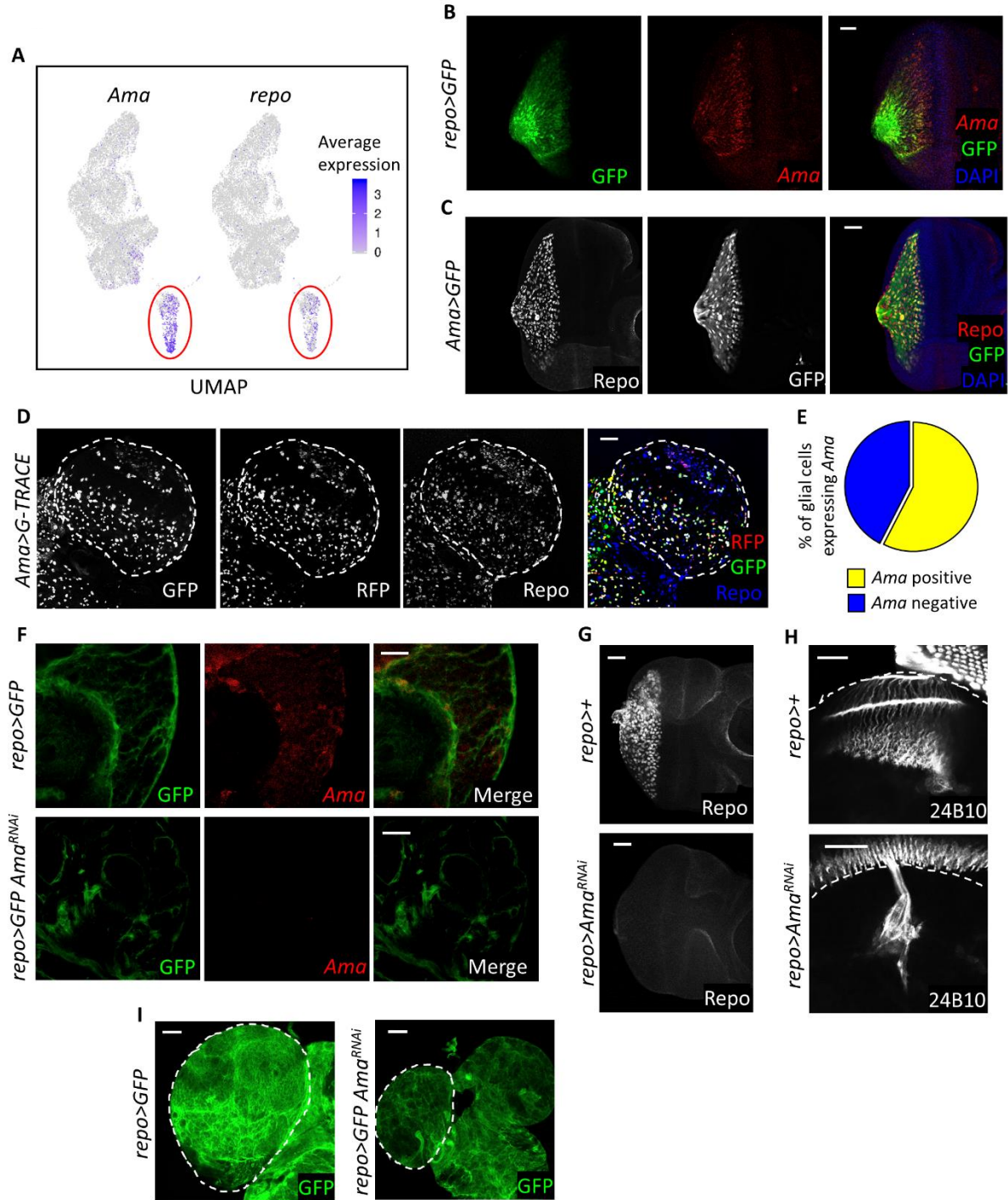


Figure 9. *Ama* is expressed in glia and is required in glial development.

(A) Feature plots on third instar eye disc UMAP for *Ama* and *repo* indicate both genes to be co-expressed in the same cluster. (B) FISH for *Ama* mRNA in the eye disc shows expression in glial cells with glial membranes labelled by GFP. Final genotype: *repo>mCD8GFP*. (C) Repo

immunofluorescence in the eye disc display expression in *Ama* positive cells labeled by GFP. (D) Repo immunofluorescence in the third instar brain of *Ama>G-TRACE* showing the lineage and real-time expression of *Ama-Gal4*. GFP labels the lineage whereas RFP represents the real-time expression. (E) Pie chart outlying the percent of glial cells expressing *Ama*. Z-stacks of *Ama>G-TRACE* brains were counted. Data represent means of three experiments. (F) FISH for *Ama* mRNA in the brain shows expression in glial cells with glial membranes labelled by GFP. Knockdown of *Ama* in glia results in loss of the FISH signal. Final genotypes: *repo>mCD8GFP* (top panel) *repo>mCD8GFP Ama^{RNAi}* (bottom panel). (G) Repo immunofluorescence shows the lack of glia in the eye disc following *Ama* depletion. (H) Immunofluorescence using 24B10 antibody to label photoreceptor axons indicate defects in axons guidance in *repo>Ama^{RNAi}*. (I) Brains are smaller *repo>Ama^{RNAi}* than *repo>GFP*. GFP labels glial cell membranes. Final genotypes: *repo>mCD8GFP* (left panel) *repo>mCD8GFP Ama^{RNAi}* (right panel). Scale bar is 20 μ m.

2. Profiling of *Ama* depleted glia by scRNA-seq

To precisely characterize the impact of *Ama* depletion on different glial cell types in the larval brain we employed scRNA-seq. A total of 45 brains across four biological replicates in each of the *repo*>+ and *repo*>*Ama*^{RNAi} larva were dissected and then dissociated into a single cell suspension to perform scRNA-seq by Drop-seq (Macosko et al., 2015) (Figure 9A). Following sequencing alignment and digital expression matrix generation, Seurat v3.0 package (Stuart et al., 2019) was used to normalize and filter out the low-quality cells. In wild type, 16,553 high quality cells were retained. These cells were grouped into 17 distinct clusters based on similarity in gene expression and the clustering was visualized by a Uniform Manifold Approximation and Projection (UMAP) algorithm. Cells in each cluster are characterized by specific list of expressed genes or biomarkers. The neuronal cell clusters were identified based on the neuronal marker *elav*, whereas the glial clusters were identified by *repo* expression. Glial cell clusters were selected for further analysis.

There are multiple types of glial cells in the developing third instar larval brain. Surface glia (SG), consisting of perineurial glia and subperineurial glial cells, are situated at the surface of the third instar larval brain and form the blood-brain barrier (BBB) (Bainton et al., 2005). The barrier protects neurons from degeneration and the acidic environment of hemolymph. Neuropil associated glia (NP) are comprised of two glial cells types, astrocyte like glia (AG) and ensheathing glia (EG) that infiltrate and cover neuropils respectively in the central brain and optic lobe (Omoto et al., 2015). Cortex glia (CG) encase neurons and are located between SG and the neuropils (Pereanu et al., 2007). Finally, glia progenitor cells (GPC) are situated between the lamina neuropil and SG of the optic lobe and give rise to glia in the optic lobe (Yoshida et al., 2005).

The supervised analysis of glial cells selected based on *repo* expression encompassed 760 cells across 9 clusters. The distinct marker list in each cluster was used to assign and label the different

glial cell types based on previously published work (Figure 10A, 10B, 10C). Thus, cells of GPC, SG, NP, WG cell types were readily identified. Interestingly, cells of the Fasciclin (Fas) cluster express markers of multiple glial cell types. For instance, *Fas2* and *alrm*, which are highly expressed in the Fas cluster, are surface and neuropil glia markers respectively (DeSalvo et al., 2014; Doherty et al., 2009). Moreover, *Ama* is also expressed in multiple glial cell types (Figure 8C) and is a top marker in that cluster (Figure 10B). One of the prominent features of the Fas cluster is the expression of cell adhesion proteins (*Fas2*, *Fas3*, *Ama*), which is consistent with the published findings underlying that cell-cell contact is a hallmark of glial cells (DeSalvo et al., 2014; Sasse et al., 2015). Thus, the Fas cluster does not appear to correspond to a specific cell type and therefore may represent a cell state.

Next, we performed Seurat analysis of a combined dataset of 9,147 *repo>Ama^{RNAi}* cells and the 16,553 control cells to characterize glial cells differences between the knockdown and control brains. Cells of non-neuronal clusters were selected as above (Figure 9B) and used to perform a supervised Seurat analysis. Reassuringly, the expression of *Ama* was strongly reduced in *repo>Ama^{RNAi}* cells (Figure 9C), thus, indicating the efficiency of *Ama* depletion.

UMAP revealed several major differences in the presentation of *Ama* depleted cells (Figure 9D, 9E). First, there was a complete loss of the Fas cluster in *repo>Ama^{RNAi}* cells as it was represented exclusively by wild type cells. Interestingly, this cluster displays the highest level of *Ama* expression (Figure 9D, 9E). Second, *repo>Ama^{RNAi}* cells were missing in the SG cluster (Figure 9D, 9E). Instead, *Ama* depleted cells formed a distinct cluster that displayed SG markers *Tret1-1*, *trol*, *CG3168* (Volkenhoff et al., 2015; Ariss et al., 2018) that was labeled SG-like (SGL) (Figure 9F). Third, the UMAP reveals a new *Ama* knockdown specific cluster (ASC) (Figure 9D, 9E) that expresses *repo* but none of other known glial cell markers (Figure 9F). Instead *sprouty* (*sty*) is one of its top markers (Figure 11). Finally, a hemocyte (HEMO) cluster that was identified by the expression of blood cell

markers such as *Hml* and *He* (Figure 11) (Goto et al., 2003; Kurucz et al., 2007) comprised predominantly of cells of *repo>Ama^{RNAi}* than of wild type brains (Figure 9D, 9E). These results indicate that depletion of *Ama* affects multiple clusters at the single cell level.

(A) Illustration of scRNA-seq pipeline. Third instar larval brains from Control (*repo>+*) and *Ama^{RNAi}* (*repo>Ama^{RNAi}*) were dissected then dissociated into a single cell suspension. Drop-seq was performed to capture single cells and generate cDNA libraries. Following alignment and generating a single cell gene expression matrix, the samples were analyzed using Seurat to unbiasedly find cell clusters having distinct gene expression profiles. (B) Feature plot of 25,700 cells outlying the neuronal and glial cells in the UMAP from the combined *repo>+* (16,553 cells) and *repo>Ama^{RNAi}* (9,147 cells) scRNA-seq brains. (C) Dot plot in *repo>+* and *repo>Ama^{RNAi}* brains displaying the depletion of *Ama* expression using scRNA-seq. (D) UMAP from the supervised analysis on glial cells from *repo>+* and *repo>Ama^{RNAi}* brains displaying 10 distinct cluster. GPC/CG = Glia Precursor Cells/Cortex Glia. ASC = *Ama^{RNAi}* Specific Cluster. SG = Surface Glia. SGL = Surface Glia-Like. NP1 = Neuropil Glia 1. NP2 = Neuropil Glia 2. NP3 = Neuropil Glia 3. Fas = Fasciclin cluster. (E) UMAP from D showing the genotype of each cell. (F) Dot plot displaying the top markers in different clusters from the analysis in D. ASC glia do not share markers of other control glial cluster. SGL share markers with SG cells. Clusters labeled in red are observed in the *repo>+* supervised glial analysis. Clusters labeled in blue are additional clusters that appeared after pooling *repo>+* cells with *repo>Ama^{RNAi}* brains. HEMO and WG are excluded as they originate from the hemolymph and the peripheral nervous system respectively and not the brain. In this analysis GPC and CG were pooled.

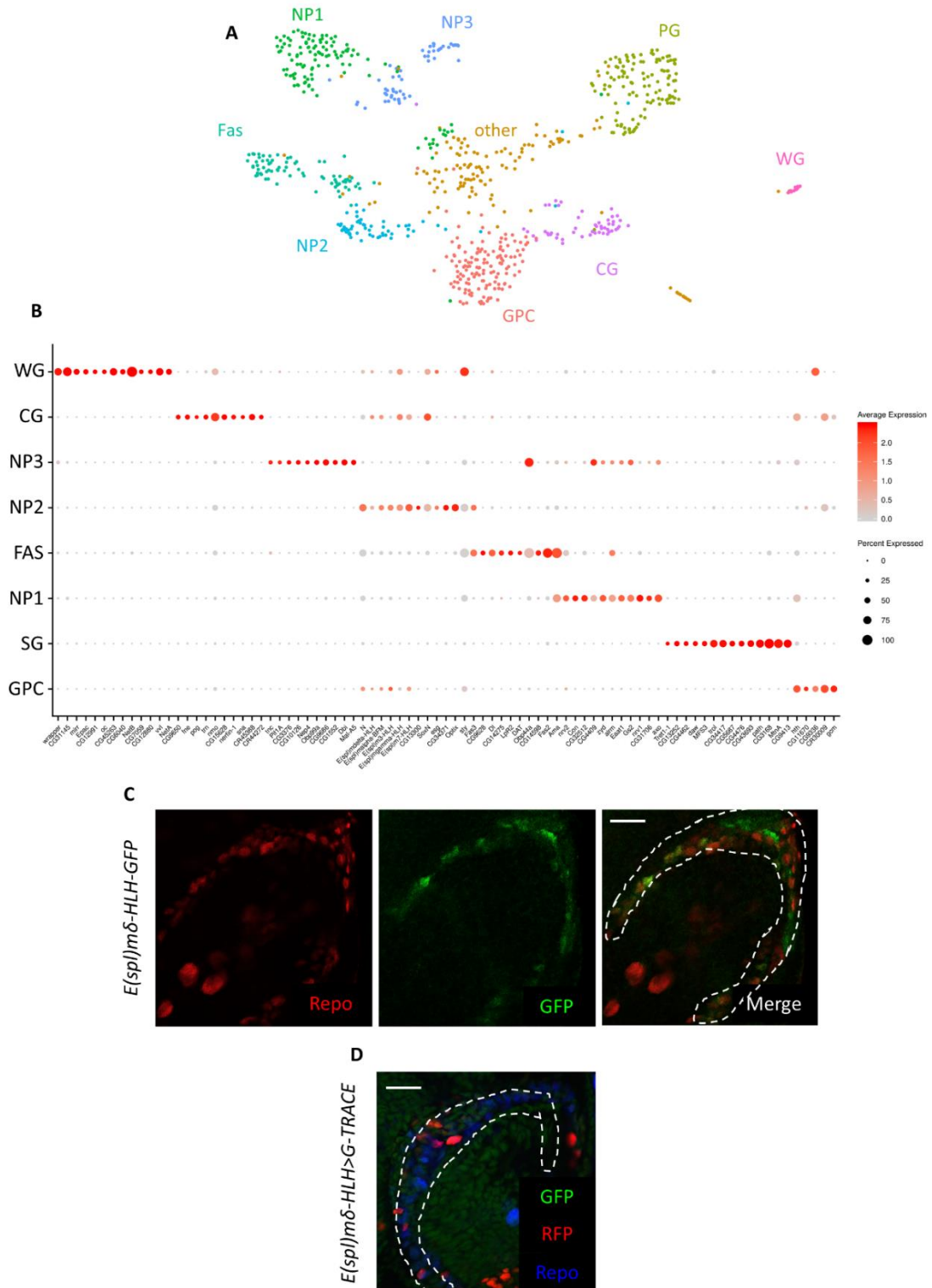


Figure 11. scRNA-seq supervised analysis in glial cells from *repo*>+ brains reveal heterogeneity in glia.

(A) UMAP of 760 *repo*>+ glial identifies 8 distinct glial clusters:

GPC: Glia Progenitor Cells, display *gcm* as a top marker which was shown to be expressed in these progenitor cells (Chotard 2005 and Soustelle 2007)

SG: Surface glia, have top markers such as *CG3168*, *trol* (Ariss et al., 2018), and *Tret1-1* (Volkenhoff et al., 2015) shown to be expressed at the surface of the brain

NP1: Neuropil associated glia 1, express neuropil glia specific markers such as *Eaat1* and *Gs2*, and *alrm* (Luna et al., 2017; Yildirim et al., 2019). This cluster displays a high expression level of genes involved in cell oxidation-reduction homeostasis such as NADH/NADPH production and recycling. For instance, transketolase (*CG8036*) is an oxireductase gene in the pentose phosphate pathway, which is important for NADPH production, is highly expressed in NP1. *Nmdmc* is also a top marker that functions to regenerate NADH. *Mtacd1* is another marker in NP1 which recycles NAD and the thioredoxin peroxidase, *Jafrac1*, is a gene highly expressed in this cluster which utilizes NADPH to neutralize reactive oxygen species. This is interesting since glial cells are involved in multiple neuronal homeostatic processes (Yildirim et al., 2019). NP1 are likely clustering by this biological function.

NP2: Neuropil associated glia 2, display *Notch* (*N*) as a top marker and *E(spl)* genes which are expressed when Notch signaling is activated. Two *E(spl)mδ-HLH* reporters show that NP2 represents neuropil associated glia in the optic lobe.

NP3: Neuropil associated glia 3, also expresses the neuropil glia *Eaat1* and *Gs2*, and *alrm* like NP2. This cluster however expresses a multitude of genes involved in the enzymatic activity of macromolecule and peptide breakdown. For instance, Acetylcholine esterase (*Ace*), metalloproteinase *Nep4*, Sphingomyelin phosphodiesterase (*CG3376*), Alpha-glucosidase (*Mal-A5*), and serine protease rhomboid (*rho*) are top markers in this cluster. This is interesting since glial cells have been shown to clear neuronal debris (Doherty et al., 2009). Not only does NP3 have markers that promote macromolecule breakdown, but the cluster also expresses *crq*, *Arc1* (Ashley et al., 2018), *myo* (Awasaki et al., 2011), *cathD* (Eissenberg et al., 2011) which are genes involved in neuronal clearance, remodeling, and lysosomal peptidolysis. These findings suggest that NP3 groups glia involved in neural clearance.

We interestingly observe that *zyd*, a specific marker of CG (Melom and Littleton 2013), to be specifically expressed in NP1 and NP3 suggesting that some CG function in either in cellular homeostasis or neural clearance.

CG: Cortex glia, displays *nerfin-1* and *trn* as top markers which have been shown to be expressed near the medulla cortical region in the optic lobe between the brain surface and neuropil (Xu et al., 2017; Chang et al., 1993). Hence, this cluster was labeled as CG.

Fas: Fasciclin cluster, reveals to have top markers of different cell types such as *Fas2*, *alrm*, and *Ama*. For example, *Fas2* is highly expressed in surface glia (DeSalvo et al., 2014), *alrm* is a neuropil associated glia marker (Doherty et al., 2009), and *Ama* is expressed in multiple glial cell types. Since *Fas2*, *Fas3*, and *Ama* are all cell adhesion genes, the Fas cluster is likely pooling glia by that biological function.

WG: Wrapping Glia, originate from the eye disc (and optic stalk) or the peripheral nervous system (PNS). Since the dissection included some part of the optic stalk and PNS, Drop-seq captured these glial cells. WG displays top makers such as *wrapper* (Noordermeer et al., 1998), *NetA*, *NetB* (Harris

et al., 1996), *vvl* (Wheeler et al., 2006) and *Epac* (Kearney et al., 2004) that have been shown to be expressed in midline glia. Midline glia are a category of wrapping glia that surround neuronal axons (Yildirim et al., 2019). This cluster also shares common top markers with recently published scRNA-seq data on WG in the third instar eye-disc (Ariss et al., 2018) such as *CG9336*, *Clic*, and *nuf*.

The “other” cluster represents low quality cells that predominantly express ribosomal genes.

(B) Dot plot showing top markers in each glial cluster in *repo*>+ brains. (C) *E(spl)mδ-HLH-GFP* reporter shows GFP expression in neuropil glia in the optic lobe indicating that NP2 are neuropil associated glia in that region of the brain. Repo labels glia. (D) An *E(spl)mδ-HLH-Gal4* reporter was crossed to *UAS-G-TRACE* shows RFP (current expression of *E(spl)mδ-HLH* to be in neuropil glial cells in the optic lobe. Glia are labeled with Repo. GFP labels lineage expression of *E(spl)mδ-HLH*. Scale bar is 20 μm.

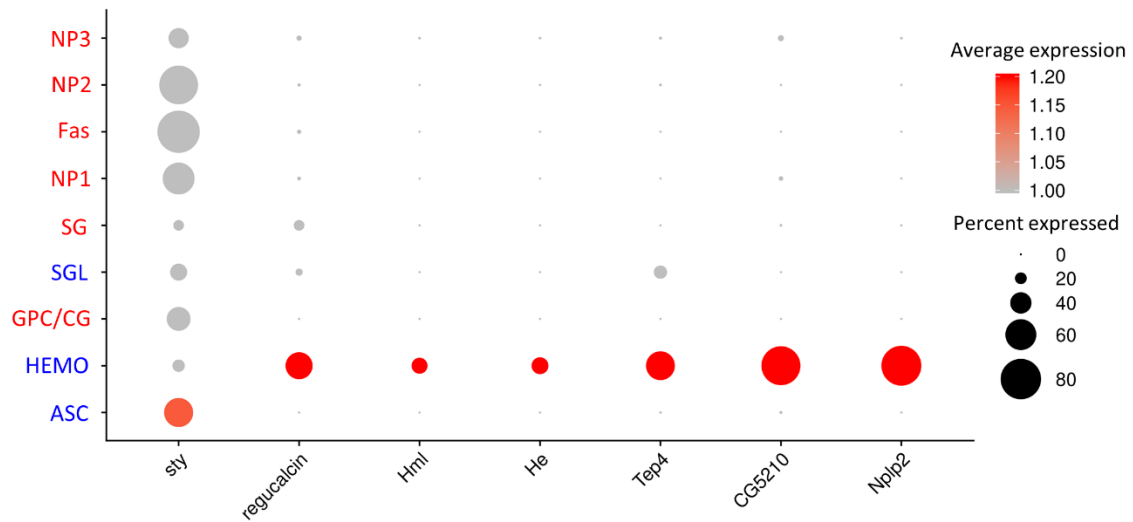


Figure 12. Supervised scRNA-seq analysis using in *repo>+* and *repo>Ama^{RNAi}* glia identifies distinct cellular perturbations in gene expression.

Dot plot showing expression of hemocyte markers in the HEMO cluster as well as sty highly expressed in ASC. Blue clusters are additional clusters that appeared after pooling control cells with *repo>Ama^{RNAi}* brains.

3. Blood brain barrier is disrupted by Ama depletion

scRNA-seq profiling revealed that *Ama* is highly expressed in the Fas cell cluster and that this cluster is lost in *repo>Ama^{RNAi}* brains. To validate these results *in vivo*, we selected *Fas2*, a top marker of the Fas cluster, and examined its expression by immunofluorescence with Fas2 antibody. Glial cells were labeled with membrane targeted GFP using *repo>mCD8-GFP*. In the wild type brain, Fas2 was observed in multiple glial cell types but with strongest expression in surface glia (Figure 12A). This is consistent with published data that Fas2 is predominantly expressed in surface glia (DeSalvo et al., 2014). Notably, there was a striking reduction in Fas2 at the surface of the brain in SG cells of *repo>mCD8-GFP Ama^{RNAi}* brains (Figure 12A). Such defect in surface glia of *repo>Ama^{RNAi}* is particularly interesting given that SGL cells are clustering distinctly from wild type SG cells on UMAP plot (Figure 9D, 9E) suggesting that the gene expression profile of surface glial cells changes dramatically following *Ama* knockdown. We therefore decided to investigate the effect of *Ama* depletion in surface glia.

SG cells proliferate extensively during larval stages (Yildirim et al., 2019) in order to fully cover the rapidly increasing brain mass. This is essential to maintain the blood brain barrier (BBB), a physical barrier that shields neurons from the blood stream solutes and protects them from neurodegeneration (Yeh et al., 2018). The barrier is a vital component of the nervous system and it is damaged in multiple neurodegenerative diseases such as Parkinson's and Alzheimer's (Sweeney et al., 2018).

In order to determine whether *Ama* affects glial cell proliferation, the expression of cell cycle genes from the scRNA-seq dataset was examined. In comparison to wild type controls, *repo>Ama^{RNAi}* glia strongly downregulate *PCNA*, *CycD*, *CycB*, *CycA*, and *dpa* (Figure 12B). Accordingly, there was

a striking reduction of glial cells undergoing mitosis as revealed by immunostaining with an anti-PH3 antibody (Figure 12C, 12D). These findings suggest that Ama plays a role in glial cell proliferation.

To examine the morphology of surface glia in *Ama* depletion these cells were visualized with a C494 antibody. In wild type brains, C494 reveals a continuous membrane layer that is encompassing the brain. In contrast, surface glia membranes were discontinuous in *repo>Ama^{RNAi}* indicating that the brains were not completely covered by SG cells (Figure 12E). This was confirmed with a different Gal4 driver *moody-Gal4* that is expressed in subperineurial glia (Figure 12E). Discontinuity of SG membranes is a hallmark of a damaged BBB. To directly confirm this, *repo>Ama^{RNAi}* and *moody>Ama^{RNAi}* brains were incubated with a fluorescent dextran dye that only penetrates the brain when the BBB is broken (Bainton et al., 2005). Although no fluorescent signal was detected in control brains, both *repo>Ama^{RNAi}* and *moody>Ama^{RNAi}* brains displayed penetration of the fluorescent dye (Figure 12F). Intriguingly, while analyzing scRNA-seq dataset we noted an elevated number of hemocytes in *repo>Ama^{RNAi}* relative the control brain (Figure 9D, 9E), which would be in agreement with a defective BBB. This was confirmed by staining with a pan hemocyte antibody (H2) (Kurucz et al., 2007) that showed a dramatic increase of infiltrating hemocytes in *repo>Ama^{RNAi}* brain lobes (Figure 12G).

One of the consequences of disruption of BBB is the exposure of neurons to the high concentration of solutes in the hemolymph that may cause neuronal degeneration and death. To determine whether *Ama* depletion in glia affects neurons, we examined the expression of apoptotic genes using scRNA-seq dataset. Interestingly, *hid*, a known pro-apoptotic gene in flies, is highly expressed in *repo>Ama^{RNAi}* cells but not in the wild type cells of the neuronal cluster 6. Accordingly, there was a significant level of neuronal death in *repo>Ama^{RNAi}* brains as revealed by the staining with a cleaved DCP-1 antibody that labels the apoptotic cells (Figure 12H).

We concluded that *Ama* has an essential function in glia because its glia specific depletion severely impairs the BBB, leads to hemocytes infiltration into the brain and extensive neuronal cell death. The latter and severe proliferative defects of *Ama* depleted glia may help to explain the small brain size in *repo>Ama^{RNAi}* animals.

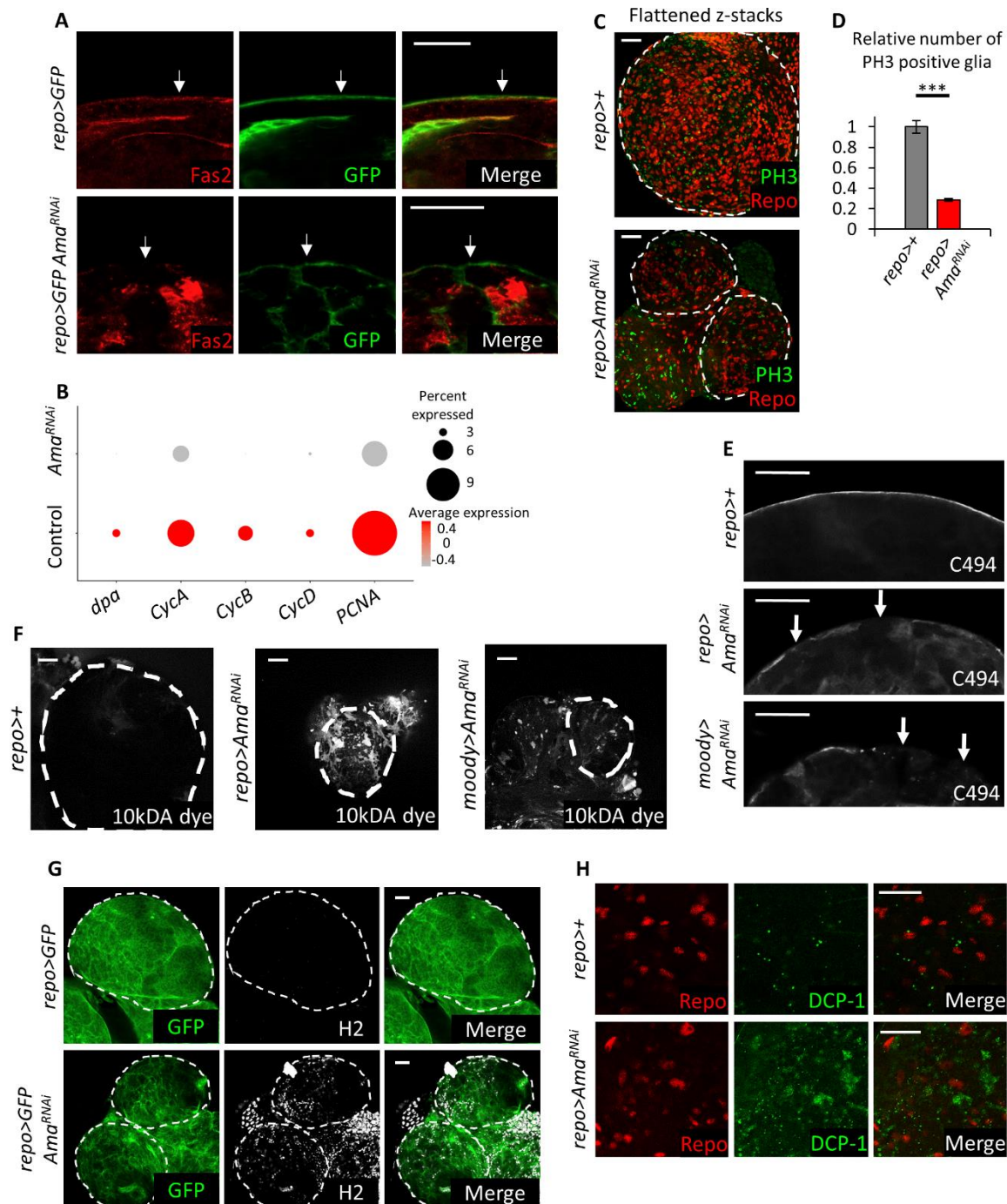


Figure 13. Ama depletion decreases glial cell proliferation and disrupts the BBB.

(A) Immunofluorescence shows a decrease in Fas2 in surface glia following Amd knockdown. Final genotype: *repo>mCD8GFP* (top panel) *repo>mCD8GFP Amd^{RNAi}* (bottom panel). White arrow points at surface glia. (B) Dot plot indicating a decrease in expression level of cell cycle genes *dpa*, *CycA*, *CycB*, *CycD*, and *PCNA* using scRNA-seq in *repo>+* and *repo>Amd^{RNAi}* glia. (C) Flattened z-

stacks of *repo>+* and *repo>Ama^{RNAi}* brains immunofluorescence with REPO and PH3 show that Ama knockdown decreases overall Repo and PH3 colocalization. (D) Quantification of Repo and PH3 colocalization in C counted reveal a significant reduction in mitotically active glia following Ama depletion. Z-stacks from *repo>+* (n=5) and *repo>Ama^{RNAi}* (n=6) brains were counted and normalized to the average count in *repo>+*. Data represents the mean after normalization \pm standard error of the mean (\pm s.e.m), and ***p<0.001 using a one-way ANOVA test. (E) C494 Immunofluorescence labeling SPG of third instar larval brains shows discontinuous surface glia membranes following Ama knockdown. White arrow points at discontinuous membranes. (F) Brains labeled with 10kDA dextran dye indicate that Ama knockdown increases dye penetration in the tissue implying a damaged BBB. (G) Immunofluorescence using pan hemocyte H2 antibody shows a strong signal in the brain in *repo>Ama^{RNAi}*. (H) Cleaved DCP-1 immunofluorescence shows a neuronal apoptotic signal in *repo>Ama^{RNAi}*. Scale bar is 20 μ m.

4. *Ama* depletion reduces RTK signaling

The other major impact of *Ama* knockdown is the appearance of the Amalgam specific cluster (ASC) (Figure 9D, 9E, 9F). Seurat analysis revealed that *sty*, a general receptor tyrosine kinase (RTK) inhibitor (Hacohen et al., 1998), is an ASC top marker (Figure 11). *sty* is highly upregulated in *Ama* depleted glia cells, while the expression of a downstream effector and a read-out of RTK signaling in glia *pointed* (*pnt*) (O'Neill et al., 1994) is reduced (Figure 12A). Interestingly, RTK signaling through FGFR (Avet-Rochex et al., 2012) and PDGFR (Read, 2018) was shown to be important for glial cell proliferation in the third instar brain as well as for glial migration (Franzdóttir et al., 2009). Since *Ama* deficient glia proliferate poorly and fail to migrate from brain to the eye disc, we investigated RTK signaling in *repo>Ama^{RNAi}*.

We began by examining the Sty expression in surface glia cells of *repo>Ama^{RNAi}* brains by immunofluorescence using a Sprouty antibody that has been previously validated (Hacohen et al., 1998). In control, Sty is present at a low level uniformly throughout the confocal plane. In *Ama* depleted glia, Sty was markedly elevated especially at plasma and nuclear membranes and this result was confirmed with two different *UAS-Ama^{RNAi}* lines (Figure 13B, Figure 14). Conversely, Sty staining was moderately reduced when *Ama* was overexpressed using a *repo-Gal4* driver (Figure 13B). To determine whether changes in Sty levels in *Ama* depleted glia are consequential, we examined the phospho-ERK (P-ERK) staining, a commonly used marker of RTK signaling, in surface glia. Strikingly, the P-ERK signal was largely lost in *repo>Ama^{RNAi}* while overexpression of *Ama* resulted in a modest increase in P-ERK (Figure 13C). Another read-out of RTK signaling, Pointed (PntP1), was strongly reduced in *Ama* depleted glia and slightly elevated in *repo>Ama* (Figure 13D), which is in agreement with scRNA-seq data (Figure 13A).

Since Ama is found at the membrane and is also secreted (Fremion et al., 2000), we asked whether it could exert a non-cell-autonomous effect on P-ERK. Clones of glia cells overexpressing Ama were generated by heat shock in *hs-FLP Act>>Ama GFP*, stained with P-ERK and distinguished by GFP. Notably, an elevated P-ERK was observed within the clone of cells expressing Ama (GFP positive) as well as in cells adjacent to the clonal boundary (Figure 13E). This suggests that Ama can affect RTK signaling in non-cell-autonomous manner. Additionally, ectopic expression of Ama in the posterior of the eye disc results in an elevated P-ERK and rough eye phenotype (Figure 15A, 15B, 15C), while clonal analysis demonstrates a non-cell-autonomous effect (Figure 15D). We also note that an increase in Sty occurs in PG and SPG even Ama was depleted using an SPG driver (Figure 14). This suggests that the effect of Ama on Sty can occur in a non-cell-autonomously, which may, in turn, explain non-cell-autonomous impact of Ama on RTK activity described above.

These findings suggest that Ama affects RTK in different cell types and can act in a non-cell-autonomous manner. This is consistent with previous biochemical studies of Ama showing that it is present at the plasma membrane and is also secreted (Fremion et al., 2000).

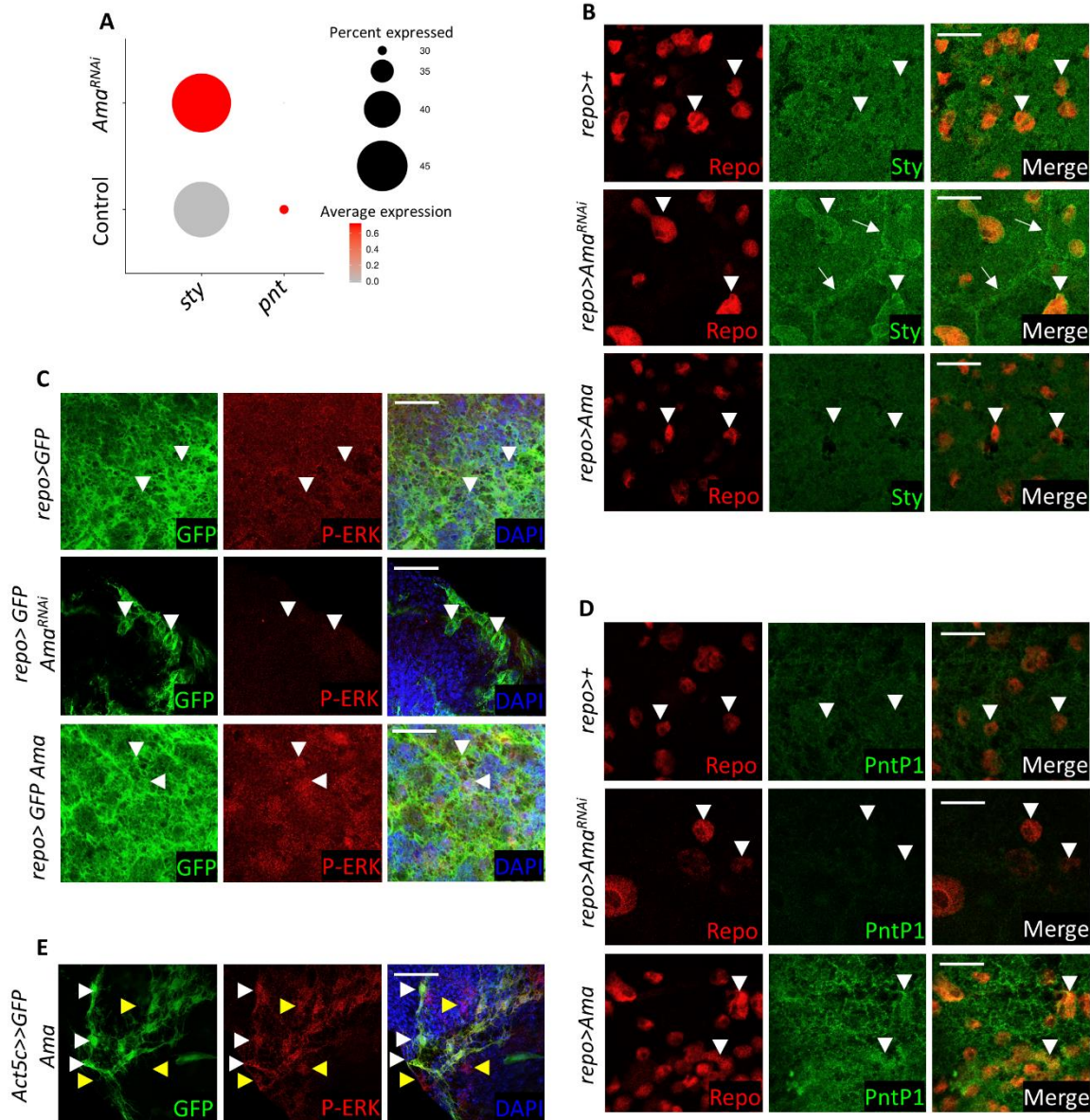


Figure 14. Ama knockdown decreases RTK signaling.

(A) Dot plot showing an increase in *sty* and a decrease in *pnt* using scRNA-seq in *repo>Ama^{RNAi}* relative to *repo>+*. (B) Sty immunofluorescence in SG shows an increase in Sty especially at the membrane in *repo>Ama^{RNAi}* and a modest decrease in basal Sty in *repo>Ama*. White arrowheads point at Repo positive glial cells whereas white arrows point at SPG membranes. (C) P-ERK immunofluorescence in SG shows a decrease in P-ERK signal in *repo>Ama^{RNAi}* and a modest increase in P-ERK in *repo>Ama*. White arrowheads point at glia nuclei. Final genotypes: *repo>mCD8GFP* (top panel), *repo>mCD8GFP Ama* (middle panel), *repo>mCD8GFP Ama^{RNAi}* (bottom panel). (D) Immunofluorescence in SG using PntP1 antibodies reveal a decrease in signal in *repo>Ama^{RNAi}* and a modest increase in PntP1 in *repo>Ama*. White arrowheads point at Repo positive glia nuclei. Final

genotypes: *repo>mCD8GFP* (top panel), *repo>mCD8GFP Ama* (middle panel), *repo>mCD8GFP Ama^{RNAi}* (bottom panel). (E) P-ERK immunofluorescence in a clone of cells using FLP-Out to over express Ama. GFP labels PG cells in the clone on the surface of the brain that over express Ama. White arrowheads point at cell autonomous increase of P-ERK. Yellow arrowheads point at cell non-autonomous increase of P-ERK. Scale bar is 20 μ m.

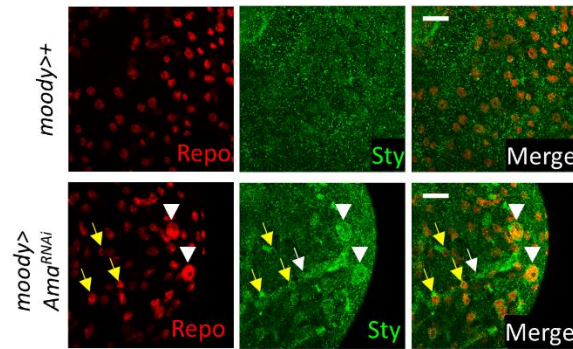


Figure 15. Ama affects Sty in the brain.

Sty immunofluorescence in SG shows an increase in Sty especially at the membranes in *moody>Ama^{RNAi}* (GD22945). White arrowheads point at SPG cells whereas white arrow points at SPG membranes. Yellow arrows point at PG cells. Repo labels all glia.

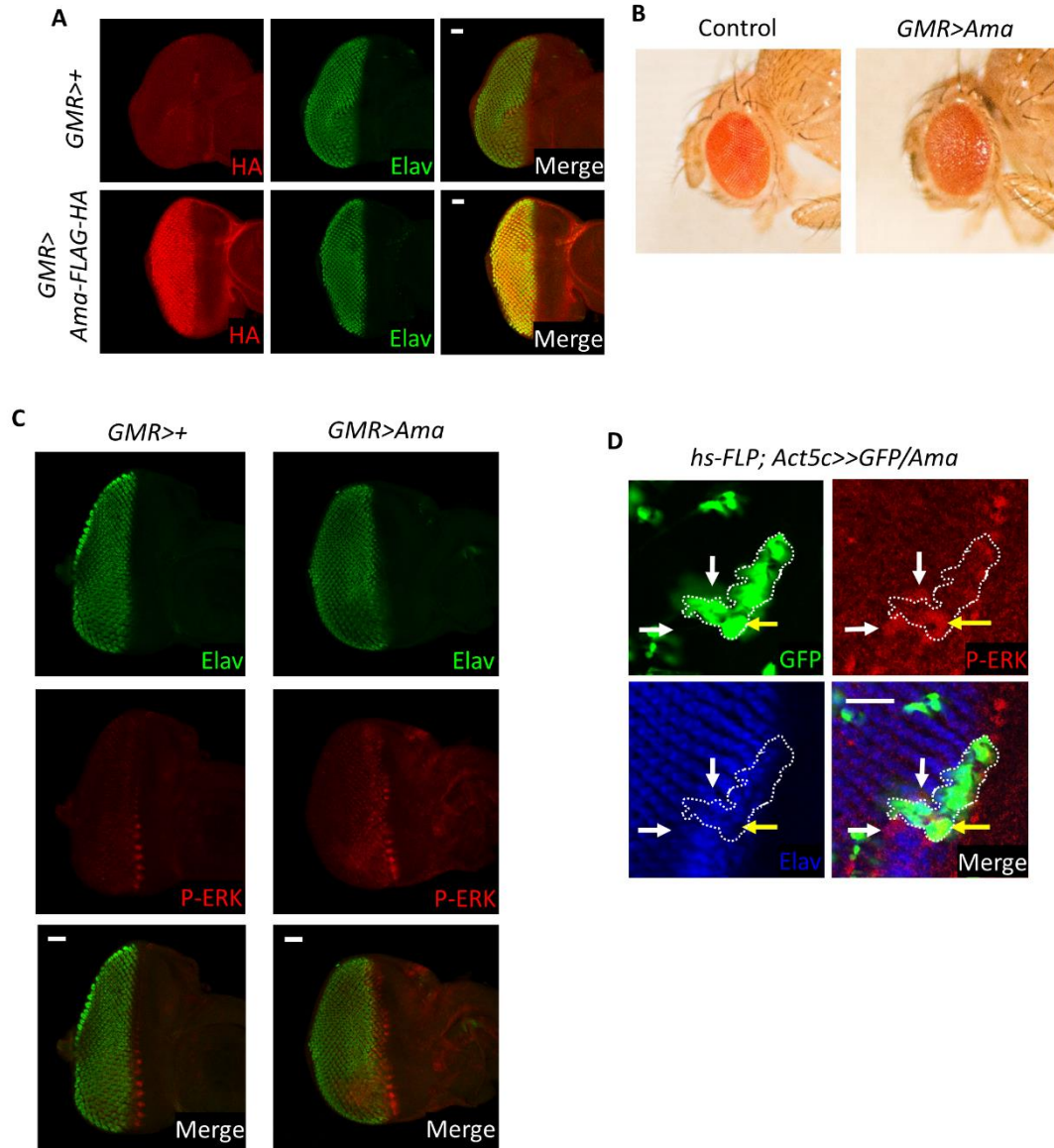


Figure 16. Ama over expression in the eye affects RTK signaling.

(A) HA immunofluorescence shows HA signal after over expressing HA-tagged Ama in the eye disc. Elav labels photoreceptors. (B) Ama over expression in the eye disc results in a mild rough eye phenotype. (C) Immunofluorescence showing that Ama over expression increases P-ERK levels in the eye disc. Elav labels photoreceptors. (D) P-ERK immunofluorescence in a clone of cells using FLP-Out to over express Ama. GFP labels the clone of cells in the eye disc that over express Ama. Yellow arrow points at cell autonomous increase of P-ERK. White arrows point at cell non-autonomous increase of P-ERK. Elav labels photoreceptors. Scale bar is 20 μ m.

5. *Ama* affects RTK signaling through Sprouty

As described above, *repo>Ama^{RNAi}* animals result in reduced glial cell proliferation, a small brain size and defects in glial migration. Previous studies highlight the importance of RTK signaling in proliferation and migration of glial cells (Avet-Rochex et al., 2012; Read, 2018; Franzdóttir et al., 2009). Our results show that a major consequence of *Ama* knockdown is upregulation of Sty, a negative regulator of RTK, and reduced RTK signaling. These observations suggest a simple model where high level of Sty in *Ama* depleted cells lowers RTK signaling in these cells and may account for the *Ama* phenotype.

This model was tested in two ways. First, we asked whether upregulation of Sty is an important event in *Ama* depleted glia. This was determined by a genetic test in which Sty was downregulated by RNAi in glial cells after which brains of *repo>sty^{RNAi}Ama^{RNAi}* and *repo>Ama^{RNAi}* were compared. The efficiency of *sty* RNAi was confirmed by staining with the Sty antibody, and showed that Sty was no longer upregulated in *repo>sty^{RNAi}Ama^{RNAi}*. Downregulation of Sty partially rescued the small brain size of *repo>Ama^{RNAi}* as the *repo>sty^{RNAi}Ama^{RNAi}* brains were larger in size than in *repo>Ama^{RNAi}* (Figure 16A). Accordingly, cell counting showed a significant increase in glial cells in *repo>sty^{RNAi}Ama^{RNAi}* compared to *repo>Ama^{RNAi}* (Figure 16B). However, Sty depletion was insufficient to rescue the neuronal cell death as there was no significant difference in the DCP-1 staining between *repo>Ama^{RNAi}* and *repo>sty^{RNAi}Ama^{RNAi}*.

As another test of the model, we asked whether Erk, which is downstream of Sty and therefore insensitive to Sty upregulation, rescues the phenotype of *Ama* depletion. To do this, a constitutively active Erk transgene (*UAS-rl^{sem}*) was overexpressed in *Ama* depleted glial cells. As shown in Figure 16A, brains of *repo>rl^{sem}Ama^{RNAi}* animals were significantly larger than *repo>Ama^{RNAi}* and there was a significant increase in glial cell number relative to *repo>Ama^{RNAi}* (Figure 16B). Interestingly, glial

migration on the eye disc was partially rescued in *repo>rl^{sem} Ama^{RNAi}* (Figure 16C), further underscoring the importance of RTK and ERK in this process.

From these results we concluded that the phenotype of glial specific Ama knockdown is, at least partially, due to reduced RTK signaling since elevating RTK signaling by activated ERK largely rescues glial cell migration and the small brain size in *repo>Ama^{RNAi}*. Our results suggest that Ama is upstream of ERK and acts through Sty to alter RTK signaling. Thus, upregulation of Sty is functionally important in Ama depleted cells.

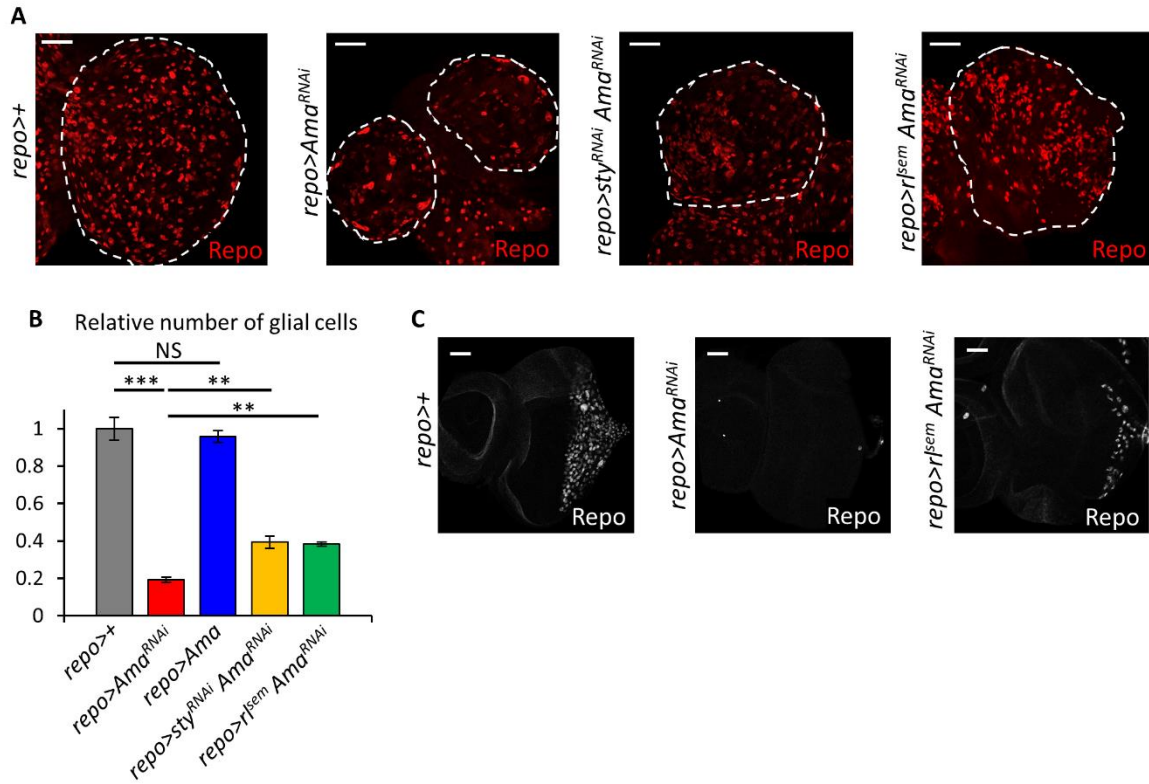


Figure 17. Sty knockdown largely rescues the phenotype of Ama depletion.

(A) Repo immunofluorescence of third instar larval brains show a partial increase in size in *repo>sty^{RNAi} Ama^{RNAi}*, and *repo>rl^{sem} Ama^{RNAi}* relative to *repo>Ama^{RNAi}*. (B) Quantification of Repo in brains in A reveal that *sty* knockdown or over expressing activated ERK in Ama depleted brain partially rescues the number of glia *repo>Ama^{RNAi}*. Z-stacks from *repo>+* (n=4), *repo>Ama^{RNAi}* (n=4), *repo>Ama* (n=3), *repo>Ama^{RNAi}* (n=4), *repo>sty^{RNAi} Ama^{RNAi}* (n=4), *repo>rl^{sem} Ama^{RNAi}* (n=4) brains were counted and normalized to the average count in *repo>+*. Data represents the mean after normalization \pm s.e.m. and **p<0.01, ***p<0.001, NS = not significant, using one-way ANOVA tests to compare each genotype to *repo>+*. (C) Repo immunofluorescence of eye discs indicates that activated ERK over expression in Ama depleted brains partially rescues glial cell migration in *repo>Ama^{RNAi}* eye discs. Scale bar is 20 μ m.

6. Ama knockdown suppresses neoplasia in a *Drosophila* glioma model

One of the best cancer models developed in *Drosophila* is a glioma model in which co-activation of EGFR-Ras and PI3K pathways faithfully recapitulates many hallmarks of human glioma (Furnari et al., 2007). In the glioma fly model, co-expression of *dEGFR^λ* and *dp110^{CAAX}* transgenes in glial cells gives rise to highly proliferative, transplantable and invasive neoplastic tumor-like growth (Read et al., 2009). Since Ama is important in regulation of normal RTK signaling during development, we asked whether Ama's role is conserved in the oncogenic context of glioma model.

As expected, glial specific expression of *dEGFR^λ* and *dp110^{CAAX}* resulted in activation of EGFR-Ras pathway that was visualized by a dramatic increase in the P-ERK and PntP1 staining relative to the control (Figure 17A, 17B) (Read et al., 2009). This led to excessive glia proliferation and enlargement of the brain (Figure 17C) (Read et al., 2009). Remarkably, concomitant depletion of Ama blocked hyperactivation of the pathway as the levels of P-ERK and PntP1 were reduced below the wild type control (Figure 17A, 17B). Accordingly, Ama depletion reduced the abnormally large brain size of *repo>dEGFR^λ dp110^{CAAX}* (Figure 17C). Since upregulation of Sty is a key event of Ama downregulation, we compared Sty expression between *repo>dEGFR^λ dp110^{CAAX}* and *repo>dEGFR^λ dp110^{CAAX} Ama*. In *repo>dEGFR^λ dp110^{CAAX}* brains, Sty was uniformly upregulated throughout the glial plane (Figure 17D). This is expected as Sty was shown to be directly regulated by EGFR signaling (Butchar et al., 2012) and fine-tunes RTK signaling (Rubin et al., 2003). Strikingly, in *repo>dEGFR^λ dp110^{CAAX} Ama^{RNAi}* glial cells, Sty was upregulated to a much higher level that was especially evident at the nuclear membrane (Figure 17D). Thus, Ama depletion results in increase of Sty, reduces RTK activation caused by activated EGFR and blocks neoplasia in a glioma fly model.

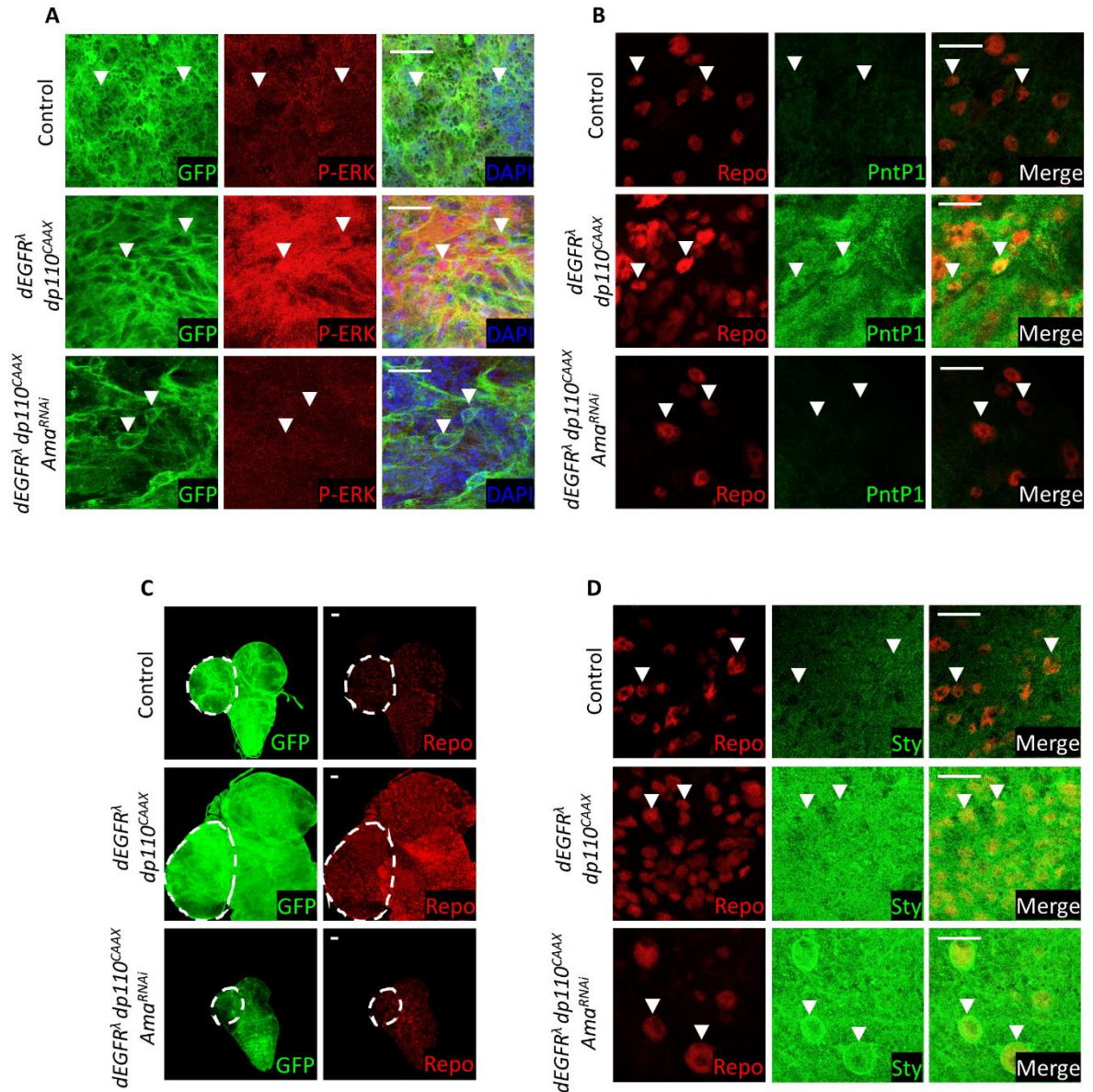


Figure 18. Ama depletion suppresses neoplastic growth in *Drosophila* glioma model.

(A) P-ERK immunofluorescence in SG reveals that knockdown of Ama in *Drosophila* glioma model a striking decrease in P-ERK. White arrowheads point at glial cells. (B) PntP1 immunofluorescence in SG indicates a drastic decrease in PntP1 following Ama depletion in *Drosophila* glioma model. White arrowheads point at Repo positive glial cells. (C) Repo immunofluorescence in SG shows that Ama depletion in a *Drosophila* glioma model model drastically decreases the brain size and number of glia. GFP labels glial cell membranes. (D) Immunofluorescence reveals an increase in Sty following Ama knockdown in *Drosophila* glioma model in SG membranes. White arrowheads point at Repo positive glial cells. Scale bar is 20 μ m. Final genotypes in (A-D): *repo>mCD8GFP* (top panel), *repo>dEGFR^λ dp110^{CAAX} mCD8GFP* (middle panel), *repo>dEGFR^λ dp110^{CAAX} mCD8GFP Ama^{RNAi}* (bottom panel).

7. Knockdown of Ama ortholog, LSAMP, increases SPROUTY2 in human glioblastoma cell lines

Ama has several human orthologs that belong to the IgLON immunoglobulin cell adhesion family (flybase.org) (Thurmond et al., 2019), with one member of this family LSAMP showing the highest sequence similarity. Interestingly, IgLONs family members are expressed in astrocyte glia and have been shown to promote astrocyte proliferation through FGF signaling (Sugimoto et al., 2012). The experiments described above suggest that upregulation of Sty is the major consequence of Ama depletion in glia during normal development and in glioma model. Therefore, we asked whether the relationship between Ama and Sty is conserved in human cells.

To address this question, we selected two glioblastoma cell lines, U251 and T98G, and examined the expression of SPROUTY 2 (SPRY2) following LSAMP knockdown by western blot. Endogenous LSAMP was detected in U251 and T98G glioblastoma cell lines by western blot (Figure 18A, 18B). U251 cells expressing *LSAMP* shRNA were generated using a lentiviral construct, and western blot confirmed that LSAMP was successfully knocked down by the *LSAMPsh* (Figure 18A). A doxycycline (Dox) inducible knockdown was generated in T98G cell line, since conventional lentiviral shRNA resulted in reduced cell viability. A dose-dependent decrease in LSAMP was observed with increasing concentration of Dox (Figure 18B). Having established the efficiency of LSAMP knockdown we examined SPRY2 levels. As shown in Figure 18A, 18B, in both U251 and T98G cell lines, the *LSAMPsh* led to an increase in SPRY2 expression as revealed by western blot.

To explore the relationship between LSAMP and SPRY2 in human glioblastoma, we analyzed publicly available Glioblastoma Multiforme (GBM) TCGA PanCancer Atlas database (cBioportal.org). We found a significant inverse correlation in mRNA levels between *LSAMP* and *SPRY2* (Figure 18C). GBM patients with *EGFR* mutations and amplifications were selected to

perform Kaplan-Meier survival analyses based on *LSAMP* expression. The results revealed a significant decrease in survival in patients with high expression of *LSAMP* (Figure 18D). These findings suggest that the relationship between Ama and Sty is conserved in human glioblastoma cell lines.

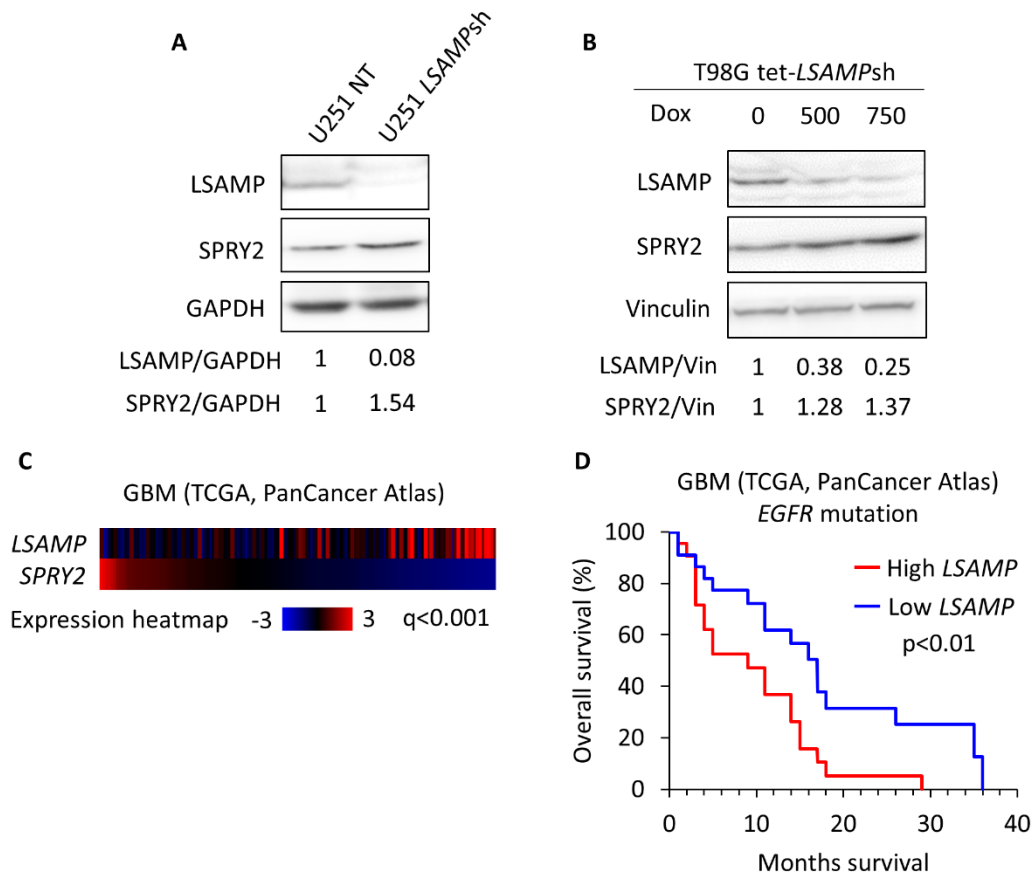


Figure 19. The relationship between Ama and Sprouty is conserved in human glioblastoma cell lines.

(A) Immunoblot showing LSAMP, SPRY2 and GAPDH levels in U251 cells stably expressing non-target shRNA (NT) or *LSAMP* shRNA (*LSAMPsh*). Knockdown of LSAMP increases SPRY2 with relative densitometry levels for each target labeled below. (B) Immunoblot showing LSAMP, SPRY2 and Vinculin levels in T98G cells expressing a doxycycline-inducible shRNA. Cells were treated for three days with the corresponding amounts of doxycycline (ng/ml in 0.15% DMSO) before harvesting. Dose dependent knockdown of LSAMP gradually increases SPRY2. Relative densitometry levels for each target are labeled below. (C) Heatmap displaying mRNA expression z-score (RNA Seq V2 RSEM) of *LSAMP* and *SPRY2* in each row in GBM human patients (TCGA PanCancer Atlas) shows a significant inverse correlation between those genes. Each column represents a tumor tissue sample. q -value were generated from a two-sample t-test. (D) Kaplan-Meier plot reveals that GBM patients (TCGA PanCancer Atlas) with EGFR mutations have a significant decrease in survivability with high levels of *LSAMP* mRNA. p -value was generated from a logrank test.

D. Discussion

RTK signaling pathway is critical in a plethora of glial cell functions such as migration, differentiation, proliferation, neurodegeneration and locomotion (Yildirim et al., 2019; Ray et al., 2017). Here, we identify *Ama* as a regulator of RTK signaling in glial cells and uncover an essential function of *Ama* in the maintenance of the BBB. Our study underscores the power of scRNA-seq profiling to explore a knockdown phenotype that led to identification of Sprouty as a major *Ama* target in regulation of RTK signaling.

Profiling cells by scRNA-seq identified distinct changes in gene expression profiles across multiple glial cell types in *Ama* depleted brain that otherwise would have not been possible using conventional approaches. First, we find that *Ama* depletion affects gene expression in surface glia, which in turn leads to a disruption in the BBB. Second, the increase in hemocyte numbers in *repo>Ama^{RNAi}* scRNA-seq dataset enabled the discovery of infiltrating blood cells in *Ama* depleted brains. Third, the increase in *Sty* and decrease in *PntP1* levels following *Ama* knockdown in glia guided us to explore the impact of *Ama* in RTK signaling.

scRNA-seq identifies and clusters cells based on similarity in gene expression profile to uncover the cellular heterogeneity in a tissue. In this study, scRNA-seq of the normal brains identified all the glial cell types in addition to a Fas cluster that appear to encompass multiple glial cell types. This cell cluster displays high expression levels of cell adhesion proteins which is a hallmark of glial cells (DeSalvo et al., 2014; Sasse et al., 2015), while also supporting the observation that scRNA-seq not only clusters cells by type but also by similar biological features (Ariss et al., 2018). This underlies the robustness of scRNA-seq since it uncovers complex cellular dynamics related to certain stimuli, continuous temporal differentiation processes (Liu and Trapnell, 2016), as well as the spatial arrangement of cells (Wagner et al, 2016). Notably, the cellular perturbations in clustering that we

observed following Ama depletion were validated experimentally by genetic analysis, immunofluorescence, and other assays. This type of single cell data validation is essential as it addresses the concern of a batch effect between scRNA-seq samples (Haghverdi et al., 2018). Thus, our work highlights the power of scRNA-seq to profile a knockdown or mutant phenotype.

Our results revealed that Ama is critical for maintaining the BBB as its depletion results in discontinuous SG membranes, suggesting a lack of tight junctions or organization that leads to disruption of the barrier (Babatz et al., 2018). This in turn exposes the larval brains to the high potassium content of hemolymph that damages neurons (Yildirim et al., 2019). Ama knockdown decreases overall glial cell proliferation which can affect SG cells and the BBB in two additional ways. First, lack of proliferation in PG cells can potentially affect the secretion of metabolites, which is important to prevent neurodegeneration (Volkenhoff et al., 2015). Second, Ama depletion can alter SPG growth by reducing endoreplication and endomitosis, as evident by reduced expression of cell cycle genes in *repo>Ama^{RNAi}*, and, therefore, hinder the ability of SPG to accommodate the growing brain during late larval stages (Unhavaithaya and Orr-Weaver 2012).

The conventional way to determine the intactness of the BBB function is by labeling brains with a fluorescent dextran dye. If the barrier is permissible to large molecules such as the dye, then the BBB is considered to be broken. Here, we present an alternative approach to monitor a disruption in the BBB by measuring the increase in infiltrating hemocytes in larval brains both by scRNA-seq and through staining with a hemocyte specific antibody. Although SG protect neurons from the hemolymph (Yeh et al., 2018), penetration of hemocytes into the brain through the damaged BBB has not been previously reported. Whether infiltrating hemocytes have a role in inflicting the damage in the brain is unknown but raises the possibility that they might have a function in that context.

Ama and Lachesin are part of the IgLON family, with Lachesin also shown to be required for the BBB maintenance (Strigini et al., 2006). Since IgLONs are also expressed in the BBB in mammals, this suggests that the immunoglobulin superfamily may indeed have an evolutionarily conserved BBB function in humans as well (Kubick et al., 2018). These findings highlight an important role of IgLONs in neurodegenerative diseases that result in BBB breakdown (Sweeney et al., 2018).

Through scRNA-seq and measuring P-ERK and PntP1 levels, we found a drastic reduction in the level of RTK signaling in *Ama* depleted brains. We suggest that *Ama* regulates the RTK pathway since its depletion increases Sty levels, a general inhibitor of the pathway, and, conversely, overexpression of *Ama* has an opposite effect. Sty in *repo>Ama^{RNAi}* brains predominantly localizes to the nuclear and plasma membranes. Intriguingly, in mammalian cells, SPRY2 localization to the membrane was shown to be crucial for its phosphorylation and inhibitory effect (Hanafusa et al., 2002). Our results therefore suggest that, while depletion of *Ama* increases Sty levels, there might additionally be an effect in the cellular localization or post-translational modification of Sty (Jarvis et al., 2006). Genetic experiments indicate that Sty is the key target of *Ama* and that activated ERK can partially rescue the *repo>Ama^{RNAi}* phenotype whereas activated EGFR in the glioma fly model cannot. These epistatic interactions support the model that *Ama*, acts through Sty, and is downstream of EGFR but upstream of ERK in RTK signaling.

Although knockdown of Sty in *Ama* depleted brains partially rescues the glial cell numbers and brain size, it fails to suppress the neuronal apoptosis in *repo>Ama^{RNAi}* brains. Since *Ama* can act non-cell-autonomously, we cannot exclude the possibility that *Ama* might also affect neurons non-cell-autonomously and, therefore, the expression of Sty in glial cells fails to rescue neurons. This is noteworthy since LSAMP, the human counterpart of *Ama*, was reported to control neurite growth

(Akeel et al., 2010). The precise mechanism of how Ama affects Sty is unknown. However, SPRY2 and EGFR in human cell lines have been shown to compete with c-Cbl (ubiquitin ligase) binding. In this context, the ubiquitin ligase attenuates SPRY2's inhibitory effect and vice-versa (Rubin et al., 2003). Since Ama and Sty localize to cell membranes, one possible explanation is that they may interact in a competitive manner, whereby loss of Ama results in Sty binding to RTK receptors and the pathway inhibition.

Finally, we show that Ama's function to control Sty levels is conserved in human glioblastoma (GBM) cell lines. Additionally, GBM patients with EGFR mutations display a significant increase in survival when they have low expression levels of LSAMP, suggesting a potential role for the IgLON family member in this cancer. Previous work showed LSAMP either promotes growth or acts as a tumor suppressor (Kresse et al., 2009). Intriguingly, SPRY2 inhibits or activates RTK signaling based on the context and cell type and also acts either as an oncogene or tumor-suppressor (Masoumi-Moghaddam et al., 2014). Our results suggest a potential connection between LSAMP and SPRY2 that may help to explain the role of LSAMP in tumorigenesis and, thus, point to *LSAMP* as a potential therapeutic target.

E. **Materials and Methods**

1. **Fly stocks**

Drosophila stocks and crosses were kept at 25 °C using standard cornmeal agar medium unless mentioned otherwise. Stocks were obtained from the Bloomington *Drosophila* Stock Center (BDSC) the Kyoto *Drosophila* Genetic Resource Center (DGRC), and the Vienna *Drosophila* Resource Center (VDRC).

The stocks used for scRNA-seq are *w[1118]; P{w[+m*]=GAL4}repo/TM3, Sb[1]* (BDSC *P{GAL4}repo*) and *y[1] sc[*] v[1] sev[21]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00297}attP2* (BDSC *HMS00297, UAS-Ama^{RNAi}*)

y[1] v[1]; P{y[+t7.7] v[+t1.8]=UAS-GFP.VALIUM10}attP2 (BDSC)

w[]; P{y[+t7.7] w[+mC]=10XUAS-IVS-mCD8::GFP}attP2* (BDSC)

y[] w[*]; P{w[+mW.hs]=GawB}Ama[NP1297] / TM6, P{w[-]=UAS-lacZ.UW23-1}UW23-1* (Kyoto DGRC)

moody-Gal4 (gift from Vicki Losick)

UAS-dEGFR^Δ, UAS-dp110^{CAAX}; repo-Gal4, UAS-mCD8-GFP/TM6C, Tb, tub-Gal80 (gift from Renee Read)

repo-Gal4, UAS-mCD8-GFP/TM6C, Tb (gift from Renee Read)

UAS-sty^{RNAi} (VDRC shRNA330208)

y[1] w[]; P{w[+mC]=UAS-rl[Sem].S}2* (BDSC)

w[1118]; P{y[+t7.7] w[+mC]=GMR57E08-GAL4}attP2 (BDSC)

y[1] w[]; P{y[+t7.7] w[+mC]=E(spl)mdelta-HLH-GFP.FPTB}attP40* (BDSC)

hs-FLP; act5c-FRT-stop-FRT-Gal4, UAS-GFP (gift from Teresa Orenic)

UAS-Ama^{RNAi} (VDRC GD22945)

The *moody>Ama^{RNAi}* (VDRC GD22945) cross was performed at 25C for two days then the vial was transferred to 30C for the rest of development.

All the Ama knockdown experiments were performed using BDSC HMS00297 unless the VDRC GD22945 UAS-Ama^{RNAi} line is mentioned in figure and figure legend.

2. Ama transgene

A UAS-Ama C-terminal FLAG-HA tagged clone UFO01101 having a *white* gene selectable marker and an *attB* site was used from the Universal Proteomics Resource and part of the Berkeley Drosophila Genome Project (BDGP) (<https://www.fruitfly.org/EST/proteomics.shtml>). This clone was received from the Drosophila Genomics Resource Center (DGRC) (<https://dgrc.bio.indiana.edu/Home>) stock number 1621050.

PhiC31 transformation services in BestGene (<http://www.thebestgene.com/>) were performed on the *y[1] w[1118]; PBac{y[+]-attP-9A}VK00023* line (BDSC 9741) with an estimated CytoSite 70A2 (Chromosome 3) using the UAS-Ama construct above.

3. Fluorescent *in situ* hybridization

38 RNA probes labeled with Quasar 670 fluorophores were generated with the Stellaris custom probe designer v4.2 that target the Ama-RA coding sequence (Oligo length = 18; min. spacing length = 2; masking level = 5)

(https://biosearchassets.blob.core.windows.net/assets/bti_custom_stellaris_drosophila_protocol.pdf)

Third instar larval brains were dissected and the Stellaris imaginal disc fluorescent *in situ* hybridization protocol was followed

(https://biosearchassets.blob.core.windows.net/assets/bti_custom_stellaris_drosophila_protocol.pdf)

Samples were imaged using Zeiss Confocal microscope.

4. Immunofluorescence

Third instar larval brain were dissected in 1X PBS then fixed for 30 minutes in 4% formaldehyde, 1X PBS. The brains were permeabilized and washed twice with 0.3% PBST for 10 minutes. Primary antibodies were incubated overnight in 10% NDS. The samples were washed three times the second day with 0.1% PBST and incubated with the fluorescent labeled secondary antibody in 10% NDS for one hour. Brains were then washed five times with 0.1% PBST and mounted on a glass slide using FluorSave (EMD Millipore).

Wandering third-instar larval eye discs were dissected in 1× PBS and fixed in 4% formaldehyde+1× PBS for 15 min, permeabilized in 0.3% PBS-T (1× PBS, 0.3% Triton X-100) two times for 10 min and then incubated with antibodies overnight at 4 °C in 1× PBS+10% normal donkey serum (NDS, Jackson ImmunoResearch)+0.1% Triton X-100 blocking serum. The following day, samples were washed in 0.1% PBS-T (1× PBS+0.1% Triton X-100) three times for 5 min. Samples were then incubated with appropriate fluorescently labeled secondary antibodies (Jackson ImmunoResearch) for 1 h in 1× PBS+10% NDS+0.1% Triton X-100 followed by 4',6-diamidino-2-phenylindole (DAPI) for 5 min. Finally, samples were washed five times for 5 min then mounted in FluorSave (EMD Millipore) on glass slides. All steps were carried out at RT and with gentle rocking, unless specified otherwise. Whenever fluorescent images have been compared, they have been obtained with the same acquisition and display settings. The slides were imaged using Zeiss Confocal microscope.

Primary antibodies were used: anti-Repo: DSHB 8D13 (1:50), anti-cleaved Drosophila Dcp-1 Cell Signaling Asp216 (1:500), anti-GFP (FITC) Abcam ab6662 (1/1000), anti-P-ERK Sigma M8159 (1:200), anti-Sty-C-terminal (gift from Mark Krasnow) (1:900), anti-P-Glycoprotein, C494 (recommended by Roland Bainton) ThermoFisher Scientific MA1-26529 (1:100), anti-Fas2 DSHB 1D4 (1:50), anti-PH3, Millipore Sigma (1:1000), 24B10 antibody DSHB (1:100), anti-PntP1 (gift from James Skeath) (1:200), H2 antibody (gift from István Ando) (1:100).

5. Heat-shock treatment

To generate third instar larvae with clones of cells that over express Ama, heat shock was induced 72 hours after egg deposition at 37C for 1 hour then transferred to 25C. Third instar larvae were collected around 55 hours after heat-shock for immunostaining.

6. Dextran labeling

Larval brains were dissected and rinsed twice with 1xPBS then incubated for 10 to 15 minutes in 2.5mM of tetramethylrhodamine labeled 10kDa dextran dye (ThermoFisher D1816) in 1xPBS. The brains were rinsed again twice with 1xPBS and mounted on a glass slide using FluorSave (EMD Millipore). Zeiss Confocal microscope was quickly used to image the slides.

7. Tissue dissociation and Drop-seq

The dissociation protocol uses both Collagenase (Sigma C9891) resuspended in 10X Krebs-Ringer with Calcium solution and 10X Trypsin-EDTA (Sigma 59418C). The EDTA partially inhibits

the Collagenase to minimize tissue damage. Ten to fifteen third instar brains were dissected in 30 minutes in 1xPBS on ice then transferred to a 1.5 mL microcentrifuge tube and spun down at 5,000 rpm (~2,000 g) 4C for 5 min. The brains were washed with Rinaldini solution and spun down at 5,000 rpm (~2,000 g) 4C for 5 min. While washing, the pellet remained carefully submerged to prevent the tissues from being pipetted up and sticking to the pipette tip. A digestion mix consisting of 1x Trypsin and 2.5mg/ml Collagenase in Rinaldini (500 uL) was added to resuspend the pelleted tissues. The microcentrifuge tube was oriented horizontally (to maximize the mechanical digestion and dissociation mix movement in the tube) on a shaker at 225 rpm for 45 minutes to an hour. The microcentrifuge tube was flicked every 10 minutes during the digestion step. The tube was spun down at 5,000 rpm (~2,000 g) 4C for 5 min before and between each following step: rinsing with 1xPBS-0.01% BSA, rinsing with 1xPBS, resuspending cells in 60-100ul 1xPBS. This protocol results in single cells and less than 5% clumps. Drop-seq protocol was performed on the single cell suspension as previously described (Ariss et al., 2018).

8. scRNA-seq alignment

Drop-seq data was analyzed using the pipeline wrapped in DropSeqPipe (version 0.5) (<https://github.com/Hoohm/dropSeqPipe>). This is based on Drop-seq core computational pipeline described in “Drop-seq_Alignment_Cookbook.pdf” version 2.0.0 on 9-28-18 (<http://mccarrolllab.com/dropseq/>) and the Drop-seq tools (version-2.3.0) was used to process the single cell RNA-seq Illumina paired end raw sequences (FastQ files). STAR aligner was used to align the raw sequences against the *Drosophila melanogaster* genome version BDGP6 (Ensembl gene model version 90). The quality of reads and mapping were examined with the FastQC program

(v0.11.8) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and poor quality reads (below 20) were eliminated before mapping.

The Digital Gene Expression (DGE) was generated using the Drop-seq alignment protocol. The number of cells extracted from aligned BAM file is based on knee plot which extracts the number of reads per cell, then plot the cumulative distribution of reads and select the knee of the distribution.

9. Cell clustering and single cell analysis

The DGE files were subjected to Seurat (version 3.0.0) single cell analysis to generate computational figures such as dot plots, feature plots, UMAP, and violin plots in R (version 3.5.3). The standard Seurat workflow was followed, and we filtered the low-quality cells and outlier cells with high gene number (min. gene cutoff = 200 and max. gene cutoff = 3000). The first 20 principle components were selected to run the UMAP with a resolution of 0.5 for the brains analyses and a resolution 1.2 for the supervised glial cell analyses. The *repo*>+, *repo*>*Ama*^{RNAi} combined supervised glia analysis was performed using the specific barcodes. The population labeled as “other” display high levels of ribosomal genes and, therefore, considered to represent low quality cells.

10. Human cell lines

U251 cells were gifted by Dr. Nissim Hay and T98G cells were gifted by Dr. Elizaveta Benevolenskaya. 293FT cells were acquired from Thermo Fisher. Both GBM cell lines were grown in MEM (Corning 10-010-CV) and 293FT were grown in DMEM (Corning 10-017-CV). All media was supplemented with 10% FBS (Gemini) and 1% penicillin/streptomycin (Corning 30-002-CI). U251 cells were transduced with non-target (NT) or LSAMP shRNAs that were stably expressed. For

immunoblot experiments, 800k cells were split onto 6cm plates and harvested 24 hours later. T98G cells were transduced with a doxycycline inducible shRNA targeting LSAMP. Following transduction, T98G cells were grown in media with 10% tet-free FBS (Gemini). For immunoblot experiments, 300k cells were plated onto 6cm plates with the corresponding amount of doxycycline (sigma) added and grown for 72 hours before harvesting. All conditions had 0.15% DMSO.

11. Plasmids and lentiviral production and transduction.

Stably expressed shRNA plasmids, plko.1 empty vector and non-target (NT), were purchased from Sigma. pTIP doxycycline inducible shRNA plasmid was a gift from Dr. Marcus Peter. The LSAMP-specific shRNA cloned into each plasmid was: CCGGCAAGTTTACTTGATCGTACAA. For lentiviral production, 6×10^6 293FT were reverse transfected with 9ug Virapower (Thermo Fisher) and 3ug of the corresponding shRNA vector in 36ul of lipofectamine. Media was changed 16 hours after transfection, and 6mL of viral particles was collected twice before filtering through a 0.45um filter. Cells were transduced with 2mL of viral supernatant diluted in 2mL of complete media supplemented with 8ug/ml of polybrene (Sigma 107689). 48 hours after transduction, cells were selected in 1ug/ml of puromycin (ACROS organics 227420100) for 5 days. (Putzbach et al., 2017)

12. Immunoblotting

At the indicated time point, cells were washed 2x with cold PBS on ice. 100ul of 1x RIPA (Cell Signaling #9806S), supplemented with phosphatase and protease inhibitor tablets (Peirce A32957 and A32953), was added to each plate for 5 minutes before scraping and collecting into 1.5mL Eppendorf tubes. Lysates were vortexed on high for 10s every 10min for 30min before

centrifuging at 13,000 x g for 10 minutes. The lysate supernatant was collected, and the protein concentration was determined with the Bradford Method (BioRad). 20-30ug of protein was loaded for each sample and ran via standard electrophoresis on a 10% polyacrylamide gel and transferred for 75 minutes at 115v onto PVDF membranes. Membranes were then blocked in 5% milk in 0.1% TBS-tween for one hour. Primary antibodies were incubated on a shaker over night at 4 degrees Celsius in 2.5% milk in TBS-T. Membranes were washed 3x in TBS-T before the corresponding secondary antibody (1:3000) was added for 1 hour at room temperature. Membranes were washed 3x in TBS-T before developing on an Azure cSeries with chemiluminescence. Primary antibodies used: LSAMP 1:500 (Abcam ab64427), SPRY2 1:1000 (Milipore #07-524), GAPDH 1:5000 (Cell signaling D16H11), and Vinculin 1:5000 (Sigma V9131). The relative densitometries were determined using Image Studio Lite Ver 5.2.

13. TCGA (cBioportal)

Glioblastoma TCGA PanCancer Atlas patient samples were selected in cBioportal (<https://www.cbioportal.org/>) (Gao et al., 2013; Cerami et al., 2012). The *LSAMP* and *SPRY2* heatmap was generated using the “OncoPrint” tab in cBioportal.

Using the “Custom Selection” in cBioportal, GBM patients with EGFR mutations were split into two groups based on the median in *LSAMP* expression (*LSAMP* high versus *LSAMP* low) in the “Groups” tab. The survivability was then plotted. The same was achieved using GBM patients with EGFR amplifications that were split into two groups based on the first and fourth quartiles in *LSAMP* expression (*LSAMP* high versus *LSAMP* low).

F. Cited literatures

- Akeel, M., McNamee, C. J., Youssef, S., & Moss, D. (2011). DIgLONs inhibit initiation of neurite outgrowth from forebrain neurons via an IgLON-containing receptor complex. *Brain research*, 1374, 27-35. doi: 10.1016/j.brainres.2010.12.028.
- Ariss, M. M., Islam, A. B., Critcher, M., Zappia, M. P., & Frolov, M. V. (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nature communications*, 9(1), 1-13. doi: 10.1038/s41467-018-07540-z.
- Ashley, J., Cordy, B., Lucia, D., Fradkin, L. G., Budnik, V., & Thomson, T. (2018). Retrovirus-like Gag protein Arc1 binds RNA and traffics across synaptic boutons. *Cell*, 172(1-2), 262-274. doi: 10.1016/j.cell.2017.
- Avet-Rochex, A., Kaul, A. K., Gatt, A. P., McNeill, H., & Bateman, J. M. (2012). Concerted control of gliogenesis by InR/TOR and FGF signalling in the *Drosophila* post-embryonic brain. *Development*, 139(15), 2763-2772. doi: 10.1242/dev.074179.
- Awasaki, T., Huang, Y., O'connor, M. B., & Lee, T. (2011). Glia instruct developmental neuronal remodeling through TGF- β signaling. *Nature neuroscience*, 14(7), 821. doi: 10.1038/nn.2833.
- Babatz, F., Naffin, E., & Klämbt, C. (2018). The *Drosophila* blood-brain barrier adapts to cell growth by unfolding of pre-existing septate junctions. *Developmental cell*, 47(6), 697-710. doi: 10.1016/j.devcel.2018.10.002.
- Bainton, R. J., Tsai, L. T. Y., Schwabe, T., DeSalvo, M., Gaul, U., & Heberlein, U. (2005). moody encodes two GPCRs that regulate cocaine behaviors and blood-brain barrier permeability in *Drosophila*. *Cell*, 123(1), 145-156. doi: 10.1016/j.cell.2005.07.029.
- Bangi, E., Murgia, C., Teague, A. G., Sansom, O. J., & Cagan, R. L. (2016). Functional exploration of colorectal cancer genomes using *Drosophila*. *Nature communications*, 7(1), 1-16. doi: 10.1038/ncomms13615.
- Butchar, J. P., Cain, D., Manivannan, S. N., McCue, A. D., Bonanno, L., Halula, S., Truesdell, S., Austin, C. L., Jacobsen, T. L., & Simcox, A., (2012). New negative feedback regulators of Egfr signaling in *Drosophila*. *Genetics*, 191(4), 1213-1226. doi: 10.1534/genetics.112.141093.
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E. and Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., & Schultz, N. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. doi: 10.1158/2159-8290.CD-12-0095.
- Chang, Z., Price, B. D., Bockheim, S., Boedigheimer, M. J., Smith, R., & Laughon, A. (1993). Molecular and genetic characterization of the *Drosophila* tartan gene. *Developmental biology*, 160(2), 315-332. doi: 10.1006/dbio.1993.1310

- Chatterjee, D., & Deng, W. M. (2019). *Drosophila Model in Cancer: An Introduction*. In *The Drosophila Model in Cancer* (pp. 1-14). Springer, Cham. doi: 10.1007/978-3-030-23629-8_1.
- DeSalvo, M. K., Hindle, S. J., Rusan, Z. M., Orng, S., Eddison, M., Halliwill, K., & Bainton, R. J. (2014). The *Drosophila* surface glia transcriptome: evolutionary conserved blood-brain barrier processes. *Frontiers in neuroscience*, 8, 346. doi: 10.3389/fnins.2014.00346.
- Doherty, J., Logan, M. A., Taşdemir, Ö. E., & Freeman, M. R. (2009). Ensheathing glia function as phagocytes in the adult *Drosophila* brain. *Journal of Neuroscience*, 29(15), 4768-4781. doi: 10.1523/JNEUROSCI.5951-08.2009.
- Eissenberg, J. C., Ilvarsonn, A. M., Sly, W. S., Waheed, A., Krzyzanek, V., Pohlmann, R., Waschkau, D., Kretzschmar, D., & Dennes, A. C. (2011). *Drosophila* GGA model: an ultimate gateway to GGA analysis. *Traffic*, 12(12), 1821-1838. doi: 10.1111/j.1600-0854.2011.01285.x.
- Evans, C. J., Olson, J. M., Ngo, K. T., Kim, E., Lee, N. E., Kuoy, E., E., Patananan, A. N., Sitz, D., Tran, P., Do, M. T., Yackle, K., Cespedes, A., Hartenstein, V., Call, G. B., & Banerjee, U. (2009). G-TRACE: rapid Gal4-based cell lineage analysis in *Drosophila*. *Nature methods*, 6(8), 603. doi: 10.1038/nmeth.1356.
- Franzdóttir, S. R., Engelen, D., Yuva-Aydemir, Y., Schmidt, I., Aho, A., & Klämbt, C. (2009). Switch in FGF signalling initiates glial differentiation in the *Drosophila* eye. *Nature*, 460(7256), 758-761. doi: 10.1038/nature08167.
- Fremion, F., Darboux, I., Diano, M., Hipeau-Jacquotte, R., Seeger, M. A., & Piovant, M. (2000). Amalgam is a ligand for the transmembrane receptor neurotactin and is required for neurotactin-mediated cell adhesion and axon fasciculation in *Drosophila*. *The EMBO journal*, 19(17), 4463-4472. doi: 10.1093/emboj/19.17.4463.
- Furnari, F. B., Fenton, T., Bachoo, R. M., Mukasa, A., Stommel, J. M., Stegh, A., Hahn, W. C., Ligon, K. L., Louis, D. N., Brennan, C. and Chin, L., DePinho, R. A., & Caveness, W. K. (2007). Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes & development*, 21(21), 2683-2710. doi: 10.1101/gad.1596707.
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E. and Cerami, E., Sander, C., & Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.*, 6(269), p11-p11. doi: 10.1126/scisignal.2004088.
- Goto, A., Kadowaki, T., & Kitagawa, Y. (2003). *Drosophila* hemolymph gene is expressed in embryonic and larval hemocytes and its knock down causes bleeding defects. *Developmental biology*, 264(2), 582-591. doi: 10.1016/j.ydbio.2003.06.001.
- Hacohen, N., Kramer, S., Sutherland, D., Hiromi, Y., & Krasnow, M. A. (1998). sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the *Drosophila* airways. *Cell*, 92(2), 253-263. doi: 10.1016/s0092-8674(00)80919-8.

- Haghverdi, L., Lun, A. T., Morgan, M. D., & Marioni, J. C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature biotechnology*, 36(5), 421-427. doi: 10.1038/nbt.4091.
- Hanafusa, H., Torii, S., Yasunaga, T., & Nishida, E. (2002). Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nature cell biology*, 4(11), 850-858. doi: 10.1038/ncb867.
- Harris, R., Sabatelli, L. M., & Seeger, M. A. (1996). Guidance cues at the Drosophila CNS midline: identification and characterization of two Drosophila Netrin/UNC-6 homologs. *Neuron*, 17(2), 217-228. doi: 10.1016/s0896-6273(00)80154-3.
- Hausott, B., & Klimaschewski, L. (2019). Sprouty2—A novel therapeutic target in the nervous system?. *Molecular neurobiology*, 56(6), 3897-3903. doi: 10.1007/s12035-018-1338-8.
- Impagnatiello, M. A., Weitzer, S., Gannon, G., Compagni, A., Cotten, M., & Christofori, G. (2001). Mammalian sprouty-1 and -2 are membrane-anchored phosphoprotein inhibitors of growth factor signaling in endothelial cells. *The Journal of cell biology*, 152(5), 1087-1098. doi: 10.1083/jcb.152.5.1087.
- Jarvis, L. A., Toering, S. J., Simon, M. A., Krasnow, M. A., & Smith-Bolton, R. K. (2006). Sprouty proteins are in vivo targets of Corkscrew/SHP-2 tyrosine phosphatases. *Development*, 133(6), 1133-1142. doi: 10.1242/dev.02255.
- Kearney, J. B., Wheeler, S. R., Estes, P., Parente, B., & Crews, S. T. (2004). Gene expression profiling of the developing Drosophila CNS midline cells. *Developmental biology*, 275(2), 473-492. doi: 10.1016/j.ydbio.2004.08.047.
- Kresse, S. H., Ohnstad, H. O., Paulsen, E. B., Bjerkehagen, B., Szuhai, K., Serra, M., Schaefer, K. L., Myklebost, O., & Meza-Zepeda, L. A. (2009). LSAMP, a novel candidate tumor suppressor gene in human osteosarcomas, identified by array comparative genomic hybridization. *Genes, Chromosomes and Cancer*, 48(8), 679-693. doi: 10.1002/gcc.20675.
- Kubick, N., Brösamle, D., & Mickael, M. E. (2018). Molecular evolution and functional divergence of the IgLON family. *Evolutionary Bioinformatics*, 14, 1176934318775081. doi: 10.1177/1176934318775081.
- Kurucz, É., Váczi, B., Márkus, R., Laurinyecz, B., Vilmos, P., Zsámboki, J., Csorba, K., Gateff, E., Hultmark, D., & Andó, I. (2007). Definition of Drosophila hemocyte subsets by cell-type specific antigens. *Acta Biologica Hungarica*, 58(1), 95-111. doi: 10.1556/ABiol.58.2007.
- Levine, B. D., & Cagan, R. L. (2016). Drosophila lung cancer models identify trametinib plus statin as candidate therapeutic. *Cell reports*, 14(6), 1477-1487. doi: 10.1016/j.celrep.2015.12.105.
- Liu, S., & Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Research*, 5. doi:10.12688/f1000research.7223.1.

- Luna, A. J. F., Perier, M., & Seugnet, L. (2017). Amyloid precursor protein in *Drosophila* glia regulates sleep and genes involved in glutamate recycling. *Journal of Neuroscience*, 37(16), 4289-4300. doi: 10.1523/JNEUROSCI.2826-16.2017.
- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M. and Trombetta, J.J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161(5), 1202-1214. doi: 10.1016/j.cell.2015.05.002.
- Masoumi-Moghaddam, S., Amini, A., & Morris, D. L. (2014). The developing story of Sprouty and cancer. *Cancer and Metastasis Reviews*, 33(2-3), 695-720. doi: 10.1007/s10555-014-9497-1.
- Melom, J. E., & Littleton, J. T. (2013). Mutation of a NCKX eliminates glial microdomain calcium oscillations and enhances seizure susceptibility. *Journal of Neuroscience*, 33(3), 1169-1178. doi: 10.1523/JNEUROSCI.3920-12.2013.
- Noordermeer, J. N., Kopczynski, C. C., Fetter, R. D., Bland, K. S., Chen, W. Y., & Goodman, C. S. (1998). Wrapper, a novel member of the Ig superfamily, is expressed by midline glia and is required for them to ensheath commissural axons in *Drosophila*. *Neuron*, 21(5), 991-1001. doi: 10.1016/s0896-6273(00)80618-2.
- O'Neill, E. M., Rebay, I., Tjian, R., & Rubin, G. M. (1994). The activities of two Ets-related transcription factors required for *Drosophila* eye development are modulated by the Ras/MAPK pathway. *Cell*, 78(1), 137-147. doi: 10.1016/0092-8674(94)90580-0.
- Omoto, J. J., Yogi, P., & Hartenstein, V. (2015). Origin and development of neuropil glia of the *Drosophila* larval and adult brain: two distinct glial populations derived from separate progenitors. *Developmental biology*, 404(2), 2-20. doi: 10.1016/j.ydbio.2015.03.004.
- Osman, D., Gobert, V., Ponthan, F., Heidenreich, O., Haenlin, M., & Waltzer, L. (2009). A *Drosophila* model identifies calpains as modulators of the human leukemogenic fusion protein AML1-ETO. *Proceedings of the National Academy of Sciences*, 106(29), 12043-12048. doi: 10.1073/pnas.0902449106.
- Pagliarini, R. A., & Xu, T. (2003). A genetic screen in *Drosophila* for metastatic behavior. *Science*, 302(5648), 1227-1231. doi: 10.1126/science.1088474.
- Papalexi, E., & Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature Reviews Immunology*, 18(1), 35. doi: 10.1038/nri.2017.76.
- Pereanu, W., Spindler, S., Cruz, L., & Hartenstein, V. (2007). Tracheal development in the *Drosophila* brain is constrained by glial cells. *Developmental biology*, 302(1), 169-180. doi: 10.1016/j.ydbio.2006.09.022
- Perrimon, N. (1994). Signalling pathways initiated by receptor protein tyrosine kinases in *Drosophila*. *Current opinion in cell biology*, 6(2), 260-266. doi: 10.1016/0955-0674(94)90145-7.

- Putzbach, W., Gao, Q. Q., Patel, M., Van Dongen, S., Haluck-Kangas, A., Sarshad, A. A., ... & Zhao, J. C. (2017). Many si/shRNAs can kill cancer cells by targeting multiple survival genes through an off-target mechanism. *Elife*, 6, e29702. doi: 10.7554/eLife.29702.
- Ray, A., Speese, S. D., & Logan, M. A. (2017). Glial Draper rescues A β toxicity in a *Drosophila* model of Alzheimer's Disease. *Journal of Neuroscience*, 37(49), 11881-11893. doi: 10.1523/JNEUROSCI.0862-17.2017.
- Read, R. D. (2018). Pvr receptor tyrosine kinase signaling promotes post-embryonic morphogenesis, and survival of glia and neural progenitor cells in *Drosophila*. *Development*, 145(23), dev164285. doi: 10.1242/dev.164285.
- Read, R. D., Cavenee, W. K., Furnari, F. B., & Thomas, J. B. (2009). A *drosophila* model for EGFR-Ras and PI3K-dependent human glioma. *PLoS genetics*, 5(2). doi: 10.1371/journal.pgen.1000374.
- Regad, T. (2015). Targeting RTK signaling pathways in cancer. *Cancers*, 7(3), 1758-1784. doi:10.3390/cancers7030860.
- Rival, T., Soustelle, L., Strambi, C., Besson, M. T., Iché, M., & Birman, S. (2004). Decreasing glutamate buffering capacity triggers oxidative stress and neuropil degeneration in the *Drosophila* brain. *Current biology*, 14(7), 599-605. doi: 10.1016/j.cub.2004.03.039.
- Rubin, C., Litvak, V., Medvedovsky, H., Zwang, Y., Lev, S., & Yarden, Y. (2003). Sprouty fine-tunes EGF signaling through interlinked positive and negative feedback loops. *Current biology*, 13(4), 297-307. doi: 10.1016/s0960-9822(03)00053-8.
- Sasse, S., Neuert, H., & Klämbt, C. (2015). Differentiation of *Drosophila* glial cells. *Wiley Interdisciplinary Reviews: Developmental Biology*, 4(6), 623-636. doi: 10.1002/wdev.198.
- Seeger, M. A., Haffley, L., & Kaufman, T. C. (1988). Characterization of amalgam: a member of the immunoglobulin superfamily from *Drosophila*. *Cell*, 55(4), 589-600. doi: 10.1016/0092-8674(88)90217-6.
- Simon, M. A., Bowtell, D. D., Dodson, G. S., Lavery, T. R., & Rubin, G. M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. *Cell*, 67(4), 701-716. doi: 10.1016/0092-8674(91)90065-7.
- Strigini, M., Cantera, R., Morin, X., Bastiani, M. J., Bate, M., & Karagogeos, D. (2006). The IgLON protein Lachesin is required for the blood–brain barrier in *Drosophila*. *Molecular and Cellular Neuroscience*, 32(1-2), 91-101. doi: 10.1101/gad.177436.111.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck III, W. M., Hao, Y., Stoeckius, M., Smibert, P. and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7), 1888-1902. doi: 10.1016/j.cell.2019.05.031.

- Sugimoto, C., Morita, S., & Miyata, S. (2012). Overexpression of IgLON cell adhesion molecules changes proliferation and cell size of cortical astrocytes. *Cell biochemistry and function*, 30(5), 400-405. doi: 10.1002/cbf.2813.
- Sweeney, M. D., Sagare, A. P., & Zlokovic, B. V. (2018). Blood–brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nature Reviews Neurology*, 14(3), 133. doi: 10.1038/nrneuro.2017.188.
- Thurmond, J., Goodman, J. L., Strelets, V. B., Attrill, H., Gramates, L. S., Marygold, S. J., Matthews, B. B., Millburn, G., Antonazzo, G., Trovisco, V., & Kaufman, T.C. (2019). FlyBase 2.0: the next generation. *Nucleic acids research*, 47(D1), D759-D765. doi: 10.1093/nar/gky1003.
- Unhavaithaya, Y., & Orr-Weaver, T. L. (2012). Polyploidization of glia in neural development links tissue growth to blood–brain barrier integrity. *Genes & development*, 26(1), 31-36. doi: 10.1101/gad.177436.111.
- Volkenhoff, A., Weiler, A., Letzel, M., Stehling, M., Klämbt, C., & Schirmeier, S. (2015). Glial glycolysis is essential for neuronal survival in *Drosophila*. *Cell Metabolism*, 22(3), 437-447. doi: 10.1016/j.cmet.2015.07.006.
- Wagner, A., Regev, A., & Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. *Nature biotechnology*, 34(11), 1145. doi: 10.1038/nbt.3711.
- Wheeler, S. R., Kearney, J. B., Guardiola, A. R., & Crews, S. T. (2006). Single-cell mapping of neural and glial gene expression in the developing *Drosophila* CNS midline cells. *Developmental biology*, 294(2), 509-524. doi: 10.1016/j.ydbio.2006.03.016.
- Xie, X., Gilbert, M., Petley-Ragan, L., & Auld, V. J. (2014). Loss of focal adhesions in glia disrupts both glial and photoreceptor axon migration in the *Drosophila* visual system. *Development*, 141(15), 3072-3083. doi: 10.1242/dev.101972.
- Xu, J., Hao, X., Yin, M. X., Lu, Y., Jin, Y., Xu, J., Ge, L., Wu, W., Ho, M., Yang, Y. and Zhao, Y., & Zhang, L. (2017). Prevention of medulla neuron dedifferentiation by Nerfin-1 requires inhibition of Notch activity. *Development*, 144(8), 1510-1517. doi: 10.1242/dev.141341.
- Yeh, P. A., Liu, Y. H., Chu, W. C., Liu, J. Y., & Sun, Y. H. (2018). Glial expression of disease-associated poly-glutamine proteins impairs the blood–brain barrier in *Drosophila*. *Human molecular genetics*, 27(14), 2546-2562.
- Yildirim, K., Petri, J., Kottmeier, R., & Klämbt, C. (2019). *Drosophila* glia: Few cell types and many conserved functions. *Glia*, 67(1), 5-26. doi: 10.1002/glia.23459.
- Yoshida, S., Soustelle, L., Giangrande, A., Umetsu, D., Murakami, S., Yasugi, T., Awasaki, T., Ito, K., Sato, M. and Tabata, T. (2005). DPP signaling controls development of the lamina glia required for retinal axon targeting in the visual system of *Drosophila*. *Development*, 132(20), 4587-4598. doi: 10.1242/dev.02040.

IV. CONCLUSIONS

A. The RB pathway in glycolysis and pH dysregulation

Chapter II outlines the chain-of-events for a Drop-seq experiment, from tissue isolation to identification of the molecular mechanism associated with metabolic dysregulation in E2F induced cell death (Figure 19). These findings provide a crucial piece of information based on previously published work outlying the importance of the RB pathway in metabolic reprogramming (Burkhart et al 2019). It was shown that *Rbf* knockdown reprograms the glutamine flux towards glutathione synthesis to minimize oxidative stress in the wing disc (Nicolay et al., 2013). Since PPN cells have an increase in intracellular acidification, I tested whether they also displayed stress upon *Rbf* mutation using a general reactive oxygen species (ROS) dye. As expected, *Rbf* mutant eye discs show an increase in ROS near the furrow compared to WT eye discs (Figure 20). This result, along with the findings in Chapter II, suggests that *Rbf* depleted tissues likely reprogram glutamine flux in favor of glutathione in order to cope with the increase in acidification, ROS, and apoptosis.

Recently published work revealed that *Rbl* depleted lung tumors resulted in a direct upregulation of glycolytic genes and an increase in lactate production (Conroy et al., 2020). Even though apoptosis was not measured in this context, the connection between *Rbf* and an increase in glycolytic flux causing lactate production seems to be conserved in mice lung cancers. Moreover, *Rbl* null human embryonic stem cells (hESC) were shown to have an increase in extracellular acidification, likely due to lactate production (Avior et al., 2017). Although this also occurs in a different context, there seems to be a unifying theme that RB plays an important role in glycolysis, lactate production, and acidification. My results unravel a direct link between the RB pathway and glycolytic genes expression levels as well as acidification, and I associate this connection to a biological process since these glycolytic *Rbf* mutant cells are sensitized to apoptosis.

It was shown that RB loss in mouse embryonic fibroblasts increases ROS levels and re-wires cells to metabolize glutamine as a source of energy for the TCA cycle. The latter is likely through a direct upregulation of the ASCT2 glutamine transporter in RB mutants (Burkhart et al 2019). In hindsight, the connection between loss of RB and lactate production and secretion supports the hypothesis that RB mutant cells switch their carbon energy source from glucose to glutamine (Reynolds et al., 2013).

Dysregulation of pH has recently been shown to be an emerging hallmark of cancer (Webb et al., 2011). Cancer cells readily secrete lactate and acidify the environment to escape the immune system (Kato et al., 2013) and to promote metastasis (Webb et al., 2011). Recent work also shows that targeting lactate transporters in pancreatic cancer present a promising therapeutic strategy (Choi et al., 2018). Since RB mutant hESC increases glycolytic activity and extracellular acidification (Avior et al., 2017), it is important to consider the impact of the RB pathway when designing drugs to target pH regulators (Izumi et al., 2003) and lactate transporters (Choi et al., 2018). Moreover, since RB inactivation is a crucial event in cancer (Kitajima & Takahashi, 2020), employing combination therapy could improve patient outcome.

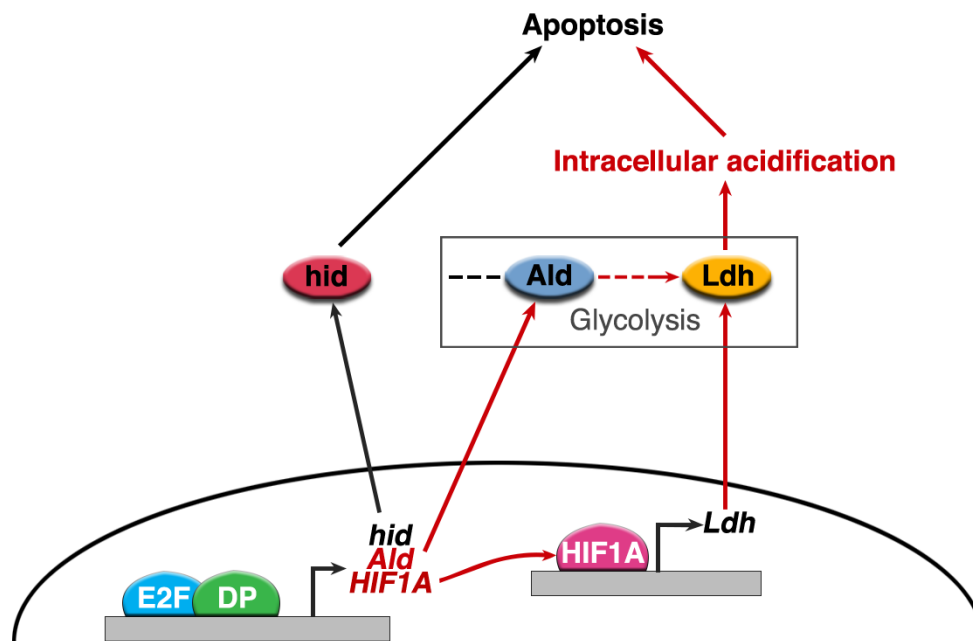


Figure 20. Illustration of the molecular mechanism of the metabolic aspect of apoptosis in *Rbf* mutant cells.

E2F1/DP directly bind to *Ald* and *HIF1A* promoters to increase glycolytic activity, decrease intracellular pH and sensitize cells to apoptosis.

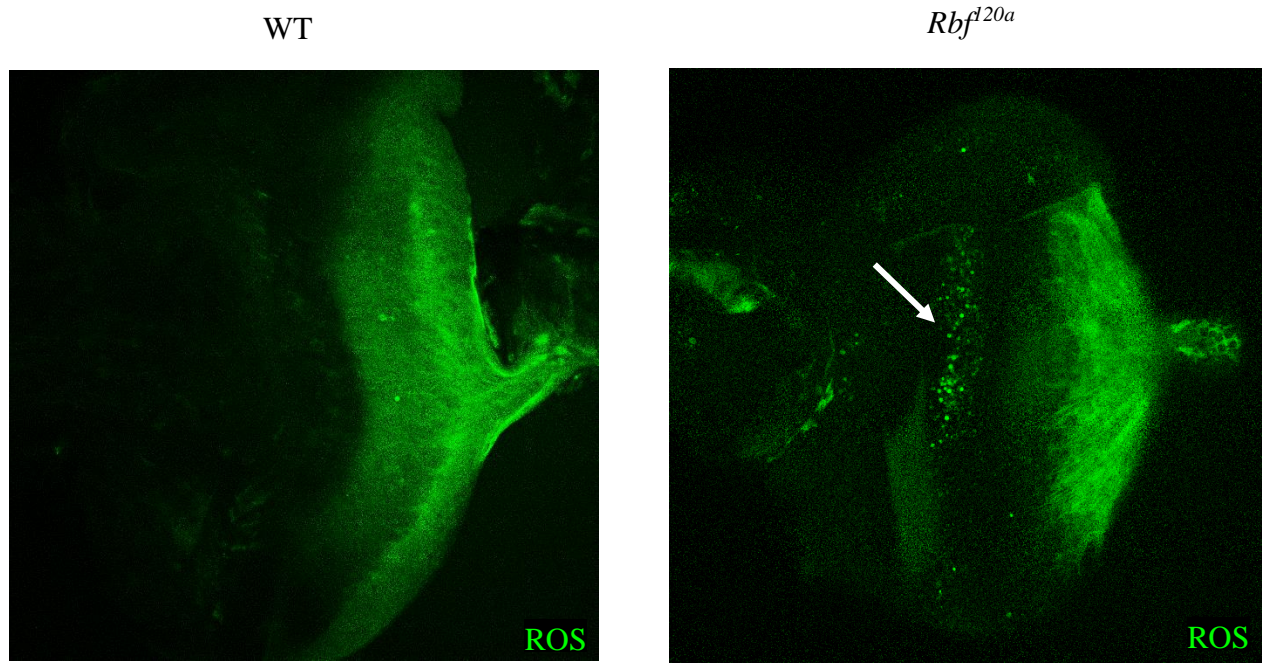


Figure 21. ROS levels in *Rbf* mutant eye discs.

Reactive Oxygen Species (ROS) was measured with a CM-H₂DCFDA dye in the WT and *Rbf* mutant eye discs shows an increase in ROS near the furrow as indicated by the white arrow.

B. Placing Ama in the RTK pathway

There are multiple extracellular canonical ligands that bind and activate RTK receptors such as growth factors, neurotrophic factors, cytokines, and hormones (Regad, 2015). Chapter III shows that scRNA-seq identified Ama, a cell adhesion IgLON extracellular and membrane protein, to affect RTK signaling by acting through Sprouty (Figure 20).

Although knockdown and over-expression of Ama respectively decreases and increases P-ERK levels, one might think that Ama acts similarly to any growth factor (GF) (Liu et al., 2009). However, GFs have been shown to upregulate Sprouty in mammals (Rubin et al., 2005) and *Drosophila* (Butchar et al., 2012) whereas this thesis shows that Ama and Sprouty are inversely correlated across species. Moreover, GFs act upstream of surface receptor tyrosine kinases, whereas epistatic interactions place Ama downstream of EGFR. These results suggest that the cell adhesion protein functions differently than GFs in the RTK signaling pathway. Further biochemical studies need to be done to accurately place Ama in the RTK pathway and determine whether it binds to surface receptors or whether there is competitive model with Sprouty, as mentioned in the Chapter III conclusion.

It was noted in the *Drosophila* tracheal system and in mice neuronal precursor cells that Sprouty can function in a non-cell autonomous manner (Kim & Bar-Sagi 2004; Hausott & Klimaschewski 2019). This has puzzled the field for some time since Sprouty is located in the cytoplasm and recruited to cellular membranes following growth factor stimulation, making a non-cell autonomous effect difficult to grasp as Sprouty functions in the intracellular compartment. Ama, on the other hand, is expressed at the membrane as well as the extracellular compartment and its knockdown results in a non-cell autonomous increase in Sprouty. If these two proteins physically interact with each other, then in Sprouty mutant cells, Ama would be freed-up to interact with nearby

cells. This, in turn, would decrease Sprouty levels in the neighboring cells, thereby relieving inhibition of RTK signaling in a non-cell autonomous manner. Therefore, chapter III provides a possible explanation to the puzzling non-cell autonomous effect of Sprouty, which is that it could occur through Ama. This is a strong possibility since it was shown that Ama can physically affect and aggregate cells (Frémion et al., 2000) while I also demonstrate that it can increase P-ERK in nearby cells.

This thesis shows that Sprouty can be affected by components outside the RTK pathway and opens the possibility for other cell adhesion or other IgLON proteins to have a similar effect on Sprouty as well. Ama has 2 main human counterparts, LSAMP and OPCML, as well as 3 other orthologs showing less similarity NTM, IgLON5, NEGR1. Although the effect of Ama on Sprouty is functionally conserved through LSAMP and SPRY2, it remains unknown whether this function is redundant through OPCML or other IgLON proteins. One way to test this would be to perform a concomitant double knockdown of LSAMP and OPCML and determine whether SPRY2 is further upregulated compared to the LSAMP (and OPCML) single knockdown. Conversely, although SPRY2 displays the highest similarity with the *Drosophila* Sprouty, it remains unclear if LSAMP affects other SPRY orthologs (SPRY1, 3, and 4).

Sprouty is a promising target in neurological disease. However, it is difficult to create compounds that will specifically target and alter its function because it is an intracellular protein that lacks enzymatic activity (Hausott & Klimaschewski 2019). This thesis shows that Ama, an extracellular and membrane protein can modulate Sprouty, providing an alternative RTK inhibitor target.

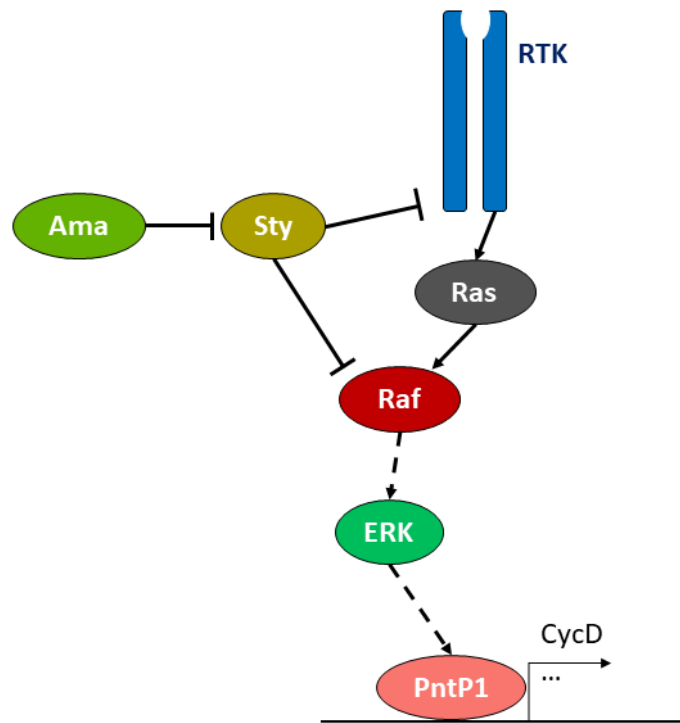


Figure 22. Ama affects RTK signaling.

Illustration of the molecular mechanism of Ama acting through Sty to inhibit RTK signaling.

C. scRNA-seq is a storm of information

This thesis outlines the benefits and depth achieved when employing scRNA-seq to dissect a mutant or knockdown phenotype in *Drosophila*. Use of this technique has allowed us to answer several challenging questions that were otherwise difficult to address such as: What is the amount of heterogeneity in a tissue? Does a mutation or knockdown affect one or multiple cell types and what are the transcriptomic variations at the single cell level? What different methods can be used to validate or disprove those cellular alterations?

Chapter II and III delineate the methodology starting from a scRNA-seq experiment to identifying the molecular features in different cell clusters following a tumorigenic lesion. The thorough description of the workflow is outlined below.

1. **Establishing a single cell wild type atlas**

Before tackling a mutant or knockdown tissue, it is vital to finalize controls and characterize the different cell clusters by generating a single cell wild type atlas. Once the tSNE/UMAP specifies the different clusters in the wild type, one must assign and label these clusters based on known and published literature. Using a top-down approach allows one to identify the clusters expressing known classical markers and assign such clusters to the various cell types captured by the technique. This often fails to provide a clear image of what the cluster represents since some markers are not captured due to reasons mentioned later in this section. However, each cluster is provided with a set of markers that are unbiasedly generated by the analysis, thus providing an alternative bottom-up approach to assign the cluster. For example, looking up the top 10-20 markers in each cluster by searching previously published literatures offers an unbiased method for cluster labeling. Ideally, assignment should be accomplished using multiple markers, thereby reducing the possibility of mislabeling. Gene set enrichment analysis using the top markers in each cluster offers an alternative useful method to

label a cluster by determining whether a process is enriched. This type of analysis reveals whether the cells are clustering based on biological function. It is also useful to determine whether a gene transcriptional profile is similar between different clusters, as they could represent different differentiation states of a single cell type. Finally, if these methods fail to provide insight on a cluster, lowering the tSNE/UMAP resolution granularity (k-means) provides a simpler view of the tissue. This is important to consider since artificial clusters appear in the analysis when the resolution is too high. It is also possible that these represent lower quality cells with low gene count and high ribosomal RNA that were not completely filtered out during the computational analysis.

More often than not, classical markers don't meet the minimal parameters and are not included in the unbiased marker list during the scRNA-seq analysis. In this scenario, the matrix displaying the average expression of all the genes for each cluster provides crucial additional information. For example, although the glial cells clustered separately than the eye disc or neuronal cells in the brain, the pan-glial marker *repo* did not meet all the parameters and was therefore not included in the marker list. However, the average expression matrix clearly showed *repo* to be expressed, although at low levels, only in the glial cell clusters.

This also addresses a widely known limitation of scRNA-seq: the technique does not to capture all the transcripts in a cell. Therefore, one should be cautious when determining if a classical marker was not captured or is expressed at low levels in the dataset. This more likely represents a flaw within the technique (or dissociation) rather than the actual expression of the gene and emphasizes the importance of labeling a cluster using on a variety of markers.

Recently, the Aerts group used the 10X scRNA-seq platform, which has a higher gene per cell count than Drop-seq, and showed that the technique was also incapable of clustering individual photoreceptors in the third instar eye disc based on their types (R1-8). They did, interestingly, observe

the same transcriptional switch in photoreceptors (EPR, LPR) as in Chapter II (González-Blas et al., 2019). These limitations in scRNA-seq will eventually become obsolete once the dissociation and single cell RNA capture are refined in the coming years. Nevertheless, more work still needs to be done to perfect these techniques, as they are still in their early developmental stages.

When the clusters are accurately labeled, Chapter II and III reveal that scRNA-seq cluster cells based on: 1- cell types (CG, NP, PG, ANT), 2- domains (PPN, VPE, DPE), 3- differentiation patterns (PG->WG, EPR->LPR, GPC->NP2), 4- biological functions (Fas, NP1, NP3), 5- signaling pathways signatures (MF and NP2). Is it however notable to mention that a central recurring challenge in scRNA-seq is the accurate assignment of clusters to a cell type or cell state (Lähnemann et al., 2020). One can argue that a cell state represents a cell type (or vice versa), or that they are two completely distinct features. The line separating these two aspects of a cell remains unclear and up for discussion. Nonetheless, it is crucial not to let the expected heterogeneity of a specific tissue impact the correct assignment and labelling of a single cell cluster (Lähnemann et al., 2020).

2. Analyzing the single cell perturbations between genotypes

Once labeled, the wild type single cell atlas provides a backbone upon which one can superimpose and compare mutant (or knockdown) cells to identify the individual transcriptional changes in each cluster. scRNA-seq notoriously displays a significant amount of batch effect between different samples. One way to overcome this is to perform multiple biological replicates in each genotype and determine whether the same cluster variations are observed in all replicates. For example, the *Rbf* mutant specific “cluster 13” and the ASC knockdown clusters were shared by all three and four biological replicates respectively, making it unlikely to be the result of batch effect. On the other hand, the fifth, sixth, and eleventh biological replicates in the wild type eye disc captured

either brain cells, or showed clusters displaying a high ribosomal gene signature. Since these differences were not observed across all the other biological replicates, they were excluded from the analysis and counted as batch effect variations.

Analyzing and comparing different genotypes can result in four different outcomes: 1- gain of a cluster, 2- loss of a cluster, 3- gain or loss of markers within a cluster, 4- changes in the ratio and distribution of cells in the clusters. Although a combination of these specific outcomes is very likely, it is important to validate them through the means described below.

Since the first three outcomes are followed by changes in gene expression, the most direct method of validation is to perform a chromogenic or fluorescent in situ hybridization (ISH) experiment to test whether there are variations in mRNA. Since ISH experiments are tedious, the use of enhancer trap and reporter lines have proven useful in accurately identifying which cells are most affected since they provide a better single cell resolution. Finally, performing immunostainings using antibodies against a specific marker can also provide a clear image as to which cells either completely lose or gained expression of a gene in different genotypes. Although this technique recognizes the protein instead of the mRNA, it has proven to be incredibly useful in validating novel specific markers in different cell types.

In the scenario where there is a change in the distribution of cells along the clusters, enhancer traps, reporters, and antibody immunostainings can be used to accurately count individual cells and validate the changes in cell-type ratios between genotypes.

Using the workflow described above, scRNA-seq proved to be a holistic and unbiased tool for dissecting the cell-type specific molecular mechanisms behind two phenotypes in *Drosophila*. When analyzed thoroughly, one can identify hidden processes that are otherwise undetectable using conventional techniques, thereby introducing a new form of research in the molecular biology and

genetics fields. I also outline the importance to overcome the limitations of scRNA-seq by using it in combination with other techniques and genetic approaches. This thesis provides the groundwork and outline for future mammalian work when using scRNA-seq to answer complicated questions such as elucidating the cell-type and contextual phenotypic variations of an RB or Sprouty mutation in various cancers.

D. Cited literatures

Avior, Y., Lezmi, E., Yanuka, D., & Benvenisty, N. (2017). Modeling developmental and tumorigenic aspects of trilateral retinoblastoma via human embryonic stem cells. *Stem cell reports*, 8(5), 1354-1365.

Burkhart, D. L., Morel, K. L., Sheahan, A. V., Richards, Z. A., & Ellis, L. (2019). The Role of RB in Prostate Cancer Progression. In *Prostate Cancer* (pp. 301-318). Springer, Cham.

Butchar, J. P., Cain, D., Manivannan, S. N., McCue, A. D., Bonanno, L., Halula, S., Truesdell, S., Austin, C. L., Jacobsen, T. L., & Simcox, A., (2012). New negative feedback regulators of Egfr signaling in *Drosophila*. *Genetics*, 191(4), 1213-1226.

Choi, S. Y. C., Ettinger, S. L., Lin, D., Xue, H., Ci, X., Nabavi, N., ... & Gleave, M. E. (2018). Targeting MCT 4 to reduce lactic acid secretion and glycolysis for treatment of neuroendocrine prostate cancer. *Cancer medicine*, 7(7), 3385-3392.

Cobrinik, D. (2005). Pocket proteins and cell cycle control. *Oncogene*, 24(17), 2796-2809.

Conroy, L. R., Dougherty, S., Kruer, T., Metcalf, S., Lorkiewicz, P., He, L., ... & Sun, R. C. (2020). Loss of Rb1 Enhances Glycolytic Metabolism in Kras-Driven Lung Tumors In Vivo. *Cancers*, 12(1), 237.

Fremion, F., Darboux, I., Diano, M., Hipeau-Jacquotte, R., Seeger, M. A., & Piovant, M. (2000). Amalgam is a ligand for the transmembrane receptor neurotactin and is required for neurotactin-mediated cell adhesion and axon fasciculation in *Drosophila*. *The EMBO journal*, 19(17), 4463-4472.

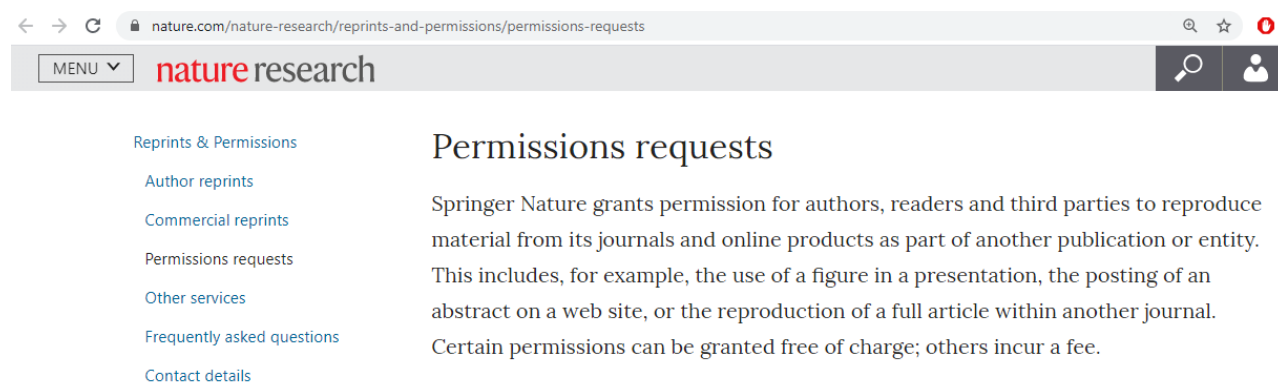
González-Blas, C. B., Quan, X. J., Duran-Romaña, R., Taskiran, I. I., Koldere, D., Davie, K., ... & Mauduit, D. (2019). Identification of genomic enhancers through spatial integration of single-cell transcriptomics and epigenomics. *bioRxiv*.

Hausott, B., & Klimaschewski, L. (2019). Sprouty2—A novel therapeutic target in the nervous system?. *Molecular neurobiology*, 56(6), 3897-3903.


- Izumi, H., Torigoe, T., Ishiguchi, H., Uramoto, H., Yoshida, Y., Tanabe, M., ... & Kohno, K. (2003). Cellular pH regulators: potentially promising molecular targets for cancer chemotherapy. *Cancer treatment reviews*, 29(6), 541-549.
- Kato, Y., Ozawa, S., Miyamoto, C., Maehata, Y., Suzuki, A., Maeda, T., & Baba, Y. (2013). Acidic extracellular microenvironment and cancer. *Cancer cell international*, 13(1), 89.
- Kim, H. J., & Bar-Sagi, D. (2004). Modulation of signalling by Sprouty: a developing story. *Nature reviews Molecular cell biology*, 5(6), 441-450.
- Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D. J., Hicks, S. C., Robinson, M. D., ... & Pinello, L. (2020). Eleven grand challenges in single-cell data science. *Genome biology*, 21(1), 1-35.
- Liu, R., Zheng, H. Q., Zhou, Z., Dong, J. T., & Chen, C. (2009). KLF5 promotes breast cell survival partially through fibroblast growth factor-binding protein 1-pERK-mediated dual specificity MKP-1 protein phosphorylation and stabilization. *Journal of Biological Chemistry*, 284(25), 16791-16798.
- Masoumi-Moghaddam, S., Amini, A., & Morris, D. L. (2014). The developing story of Sprouty and cancer. *Cancer and Metastasis Reviews*, 33(2-3), 695-720.
- Nicolay, B. N., Gameiro, P. A., Tschöp, K., Korenjak, M., Heilmann, A. M., Asara, J. M., ... & Dyson, N. J. (2013). Loss of RBF1 changes glutamine catabolism. *Genes & development*, 27(2), 182-196.
- Regad, T. (2015). Targeting RTK signaling pathways in cancer. *Cancers*, 7(3), 1758-1784.
- Reynolds, M. R., Lane, A. N., Robertson, B., Kemp, S., Liu, Y., Hill, B. G., ... & Clem, B. F. (2014). Control of glutamine metabolism by the tumor suppressor Rb. *Oncogene*, 33(5), 556-566.
- Rubin, C., Zwang, Y., Vaisman, N., Ron, D., & Yarden, Y. (2005). Phosphorylation of carboxyl-terminal tyrosines modulates the specificity of Sprouty-2 inhibition of different signaling pathways. *Journal of Biological Chemistry*, 280(10), 9735-9744.
- Webb, B. A., Chimenti, M., Jacobson, M. P., & Barber, D. L. (2011). Dysregulated pH: a perfect storm for cancer progression. *Nature Reviews Cancer*, 11(9), 671-677.
- Xu, X. L., Fang, Y., Lee, T. C., Forrest, D., Gregory-Evans, C., Almeida, D., ... & Cobrinik, D. (2009). Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell*, 137(6), 1018-1031.

APPENDICES

A. Nature permission of inclusion



B. eLife permission of inclusion

 reviewer.elifesciences.org/author-guide/journal-policies

Licensing

Because articles published by eLife are licensed under a [Creative Commons Attribution license](#), others are free to copy, distribute, and reuse them (in part or in full), without needing to seek permission, as long as the author and original source are properly cited.

VITA

Majd M. Ariss

EDUCATION

- | | |
|--|-----------|
| University of Illinois at Chicago , Chicago, IL | 2015-2020 |
| • Ph.D. in Biochemistry and Molecular Genetics | |
| Georgetown University , Washington, DC | 2014-2015 |
| • M.S. in Biochemistry and Molecular Biology | |
| American University of Beirut , Beirut, Lebanon | 2010-2013 |
| • B.S. Biology, minor in Chemistry | |

PUBLICATIONS

- Ariss M. M.** et al., (2018). Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant. *Nature communications*, 9(1), 5024
- Ariss M. M.** et al., (2020). Amalgam regulates the receptor tyrosine kinase pathway through Sprouty in glial cell development. *eLife*, in review
- Chen X., **Ariss M. M.** et al., (2020). Cell autonomous versus systemic Akt isoform deletions uncovered new roles for Akt1 and Akt2 in breast cancer. in review
- Zappia M. P., Castro L. de, **Ariss M. M.** et al., (2020) A single-cell transcriptome atlas of the adult muscle precursors uncovers early events in fiber-type divergence in *Drosophila*. *EMBO reports*, in review
- Aissa A. F., Islam A. B. M. M. K., **Ariss M. M.** et al., (2020) Drug resistant changes at the level of single cells inform combination therapy. In preparation

RESEARCH EXPERIENCE

- | | |
|--|-----------|
| University of Illinois at Chicago , Dr. Maxim V. Frolov lab | 2016-2020 |
|--|-----------|

Project 1: Single cell RNA-sequencing identifies a metabolic aspect of apoptosis in Rbf mutant

- Set up a single cell RNA-sequencing Drop-seq platform and analyzed datasets using Seurat and Monocle using R
- Created the first single cell atlas of the *Drosophila* third instar eye disc Identified novel markers for each cell type in the tissue, as well as a transcriptional switch in photoreceptors
- Uncovered glycolytic cells in *Rbf* mutant eye discs that are highly acidic which in turn are sensitized to apoptosis.

Project 2: Amalgam regulates the receptor tyrosine kinase pathway through Sprouty in glial cell development

- Generated the first single cell atlas of the *Drosophila* developing third instar larval glia
- Found that Amalgam is crucial for glial cell proliferation and maintaining the blood-brain barrier

- Revealed that Amalgam regulates the receptor tyrosine kinase signaling pathway through Sprouty in glia

Georgetown University, Dr. Leonard Bazar and Dr. Jack Chirikjian labs 2014-2015

- Cloned the human *KRAS* gene along with multiple mutated forms of the gene.
- Purified KRAS protein along with the mutated versions and studied the GTPase activity of each protein

Virginia Commonwealth University, Dr. Darrell Peterson 2013-2014

- Transformed and molecularly cloned different components of the Hepatitis B virus
- Purified multiple Hepatitis B proteins using column chromatography

FELLOWSHIP and AWARDS

UIC Pre-doctoral Education for Clinical and Translational Scientists (PECTS) Fellowship Award (2018)

Won the “Best Poster Award” at the Biochemistry and Molecular Genetics Retreat (2018). Single cell RNA-sequencing reveals a metabolic aspect of apoptosis in Rbf mutant. Lake Geneva, WI

Georgetown University “Internship Excellency Award” (2015)

French Baccalaureate, with Honors “mention bien” (2009)

PRESENTATIONS

Talk, Chicago Cancer Biology Retreat, Chicago, IL 2018

Talk, UIC GEMS Research Symposium, Chicago, IL 2018

Poster, Chicago Cancer Biology Retreat, Chicago, IL 2018

Poster, Biochemistry and Molecular Genetics Retreat, Lake Geneva, WI 2018

Poster, Single Cell Analysis Meeting, Cold Spring Harbor, NY 2017

Poster, Chicago Cancer Biology Retreat, Chicago, IL 2016

Poster, Fourth International Rb Meeting, Columbus, OH 2015