Aus dem Institut für Physiologie und Pathophysiologie

des Fachbereichs Medizin der Philipps-Universität Marburg

Mechanisms of niche to stem cell conversion exemplified by the follicle stem cell lineage in *Drosophila melanogaster*

Inaugural-Dissertation zur Erlangung des Doktorgrades der gesamten Zahnmedizin

dem Fachbereich Medizin der Philipps-Universität Marburg

vorgelegt von

Selma Beganović aus Gießen

Marburg, 2022

Angenommen vom Fachbereich Medizin
der Philipps-Universität Marburg
am

19.12.2022

Gedruckt mit Genehmigung des Fachbereichs Medizin

Dekanin: Prof. Dr. D. Hilfiker-Kleiner
Referent: Prof. Dr. S. Bogdan / Dr. K. Rust
1. Korreferent/in: Prof. Dr. A. Brehm

I dedicate my work to my mother, who dedicated her life to the struggle of giving us a better life.

Table of contents

Abbreviations

#1	Recombinant fly line 1
#3	Recombinant fly line 3
#5	Recombinant fly line 5
%	percent
*	p value under 0,05
**	p value under 0,01
***	p value under 0,001
::	Promoter
°C	Degree Celsius
aEC	Anterior escort cell
AMP	Antimicrobial peptide
AMPK	AMP activated protein kinase
ASC	Adult stem cell
bic	Bicaudal
BPA	Bisphenol A
BSA	Bovine serum albumin
bulkRNA-Seq	Bulk RNA Sequencing
CB	Cystoblast
CC	Cap cell
CD8	Cluster of differentiation 8
cDNA	Complementary DNA
cEC	Central escort cell
CG9220	Chondroitin sulfate synthase
CrebA	Cyclic-AMP response element binding protein A
CyO	Curly of Oster
dally	division abnormally delayed
DAPI	4',6-Diamidino-2-phenylindol
DD	Death domain
ddH2O	Double distilled water
dERR	Drosophila estrogen related receptor
dl	Dorsal
DNA	Deoxyribonucleic acid
DOR	Diabetes and obesity regulated
dpr17	defective proboscis extension response 17
dsRNA	Double stranded RNA
EC	Escort cell
EDC	Endocrine disrupting chemical
EGFP	Enhanced green fluorescent protein
Egfr	Epidermal growth factor receptor
e.g.	*exempli gratia* (for example)
en	Engrailed
ER	Estrogen receptor
ESC	Embryonic stem cell
et al.	*et alii* (and colleagues)
F1	Filial generation
FACS	Fluorescence activated cell sorting
Fas III	Fasciclin III
fax	failed axon connections

FC	Follicle cell
FCA	Fly Cell Atlas
FCA EC	Escort cell dataset of the Fly Cell Atlas
FCP	Follicle cell progenitor
Fer2LCH	Ferritin 2 light chain homologue
FLP	Flippase recombinase
foxo	Forkhead box, sub-group O e
FRT	Flippase recognition target
FSC	Follicle stem cell
g	gram
Gal	Galactokinase synthesis
Gal80ts	Temperature sensitive Gal80
GFP	Green fluorescent protein
gnu	Giant nuclei
GSC	Germline stem cell
G-TRACE	Gal4 technique for real-time and clonal expression
HCA	Human Cell Atlas
HCl	Hydrochloric acid
Idgf6	Imaginal disc growth factor 6
i.e.	*id est* (that is)
IGS	Inner germarial sheath cell
IκB	Inhibitor of Kappa-light-chain-enhancer of activated B cells
IL	Insulin-like peptide
InRK1409A	Dominant negative allele of the InR (Insulin Receptor)
Irc	Immune-regulated catalase
KCl	Potassium Chloride
KH2PO4	Monopotassium Phosphate
L	liter
lacZ	Beta-galactosidase
mamo	maternal gene required for meiosis
MBFC	Main body follicle cell
mEC	Middle escort cell
Mhc	Myosin heavy chain
ml	milliliter
mM	millimole
MMLV	Moloney Murine Leukemia Virus
mRNA	Messenger RNA
msk	Moleskin
mtrm	Matrimony
n	Sample Size
n.s.	Not significant
Na2HPO	Disodium Phosphate
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NF-κB	Nuclear Factor Kappa-light-chain-enhancer of activated B cells
NMDRC	Non-monotonic dose reponse curve
Nox	NADPH oxidase
p	Probability value

PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Triton X-100
PC	Polar cell
PCA	Principle component analysis
PCR	Polymerase chain reaction
pcs	Parcas
Pdk1	Phosphoinositide-dependent kinase 1
pEC	Posterior escort cell
PFA	Paraformaldehyde
pFC	Pre follicle cell
pH	Potential Hydrogen
PI3K	Phosphoinositide 3-kinase
ptc	Patched
QC	Quality Control
qPCR	Real-time quantitative PCR
RFP	Red fluorescent protein
Rheb	Ras homolog enriched in brain
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RT	Room temperature
S.E.M	Standard error of the mean
santa-maria	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
scRNA-Seq	Single cell RNA sequencing
siRNA	Small interfering RNA
SMART	Switching Mechanism At the end of the 5′-end of the RNA Transcript
stet	Stem cell tumor
TAG	Triacylglycerol
TF	Terminal filament cell
tl	Toll
TLR	Toll-like receptor
TM2	Third multiple 2
TM6B	Third multiple 6B
TOR	Target of rapamycin
T-SNE	T-distributed stochastic neighbor embedding
TSO	Template Switching Oligo
tub	Tubulin
UAS	Upstream activating sequence
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier
unc84	UNCoordinated 84
upd1	Unpaired 1
Wnt	Wingless and Int-1
wun2	Wunen-2

List of Figures

List of Tables

1. Introduction

1.1 *Drosophila melanogaster* as a model organism

Drosophila melanogaster, generally known as the fruit fly, serves as one of the most commonly used organisms for biological and biomedical research (Tolwinski, 2017). Its rapid generation time, low- cost cultivation, easily accessible genetic tools and the fact that the *Drosophila* genome is 60% homologous to the human genome, have made it crucial to studies and understanding physiological mechanisms (Mirzoyan et al., 2019; Tolwinski, 2017). These features yet also enable us to mimic and study human pathogenesis as approximately 75% of known human disease genes have matching orthologs in the genome of *Drosophila* (Reiter et al., 2001). Although non-vertebrate *Drosophila* and humans may appear severely different with respect to morphogenic features, it is well known and established that fundamental characteristics towards development and survival are conserved among both species (Jennings, 2011; Ugur et al., 2016).

There are many different tissue and cell types in *Drosophila* which can be successfully used for investigations. In this study, I used the *Drosophila* ovarian follicle stem cell as a model to study mechanisms of stem cell replacement.

1.1.1 Genetic tools in *Drosophila melanogaster*

In *Drosophila,* a wide spectrum of genetic tools can be utilized to study gene function and assess cellular behavior. In the following, I will describe genetic tools applied in this study.

1.1.1.1 The Gal4/UAS-System

The GAL4/UAS-System has been introduced for targeted gene expression in *Drosophila* and has been commonly used ever since (Brand & Perrimon, 1993) (Figure 1 A). Gal4 represents a transcription factor isolated from yeast which binds to a Gal4 dependent DNA sequence called UAS (Upstream Activating Sequence) and activates the expression of downstream genes. This system functions

bipartite and therefore ensures that the transcription of targeted genes remains silent until Gal4 is produced (Brand & Perrimon, 1993). As Gal4 is linked to cell type specificity, it will consequently only activate transcription of UAS controlled genes within the desired cell type. In order to produce transgenic fly lines in which genes are expressed cell specifically, driver lines expressing Gal4 under a cell type-specific promoter are crossed with responder lines expressing the gene of interest under the control of UAS. F1 generations will then contain both constructs, UAS and Gal4 (Brand & Perrimon, 1993).

To refine the activity of Gal4/UAS, the system can be inhibited by introducing Gal80 (Brand & Perrimon, 1993). The regulatory repression of Gal80/Gal4 in yeast has been coined by Douglas and Hawthorne (Douglas & Hawthorne, n.d.). Until date, it has been established that the Gal80 protein unfolds its inhibition by interacting with the transcription site of Gal4 (Gal4 activation domain, Gal4AD) (Hirst et al., 2001; Torchia et al., 1984; Y. Wu et al., 1996). A temperature sensitive variant of Gal80 in yeast has been firstly described by Matsumoto and colleagues (Matsumoto et al., 1978). Following, the temporal and regional gene expression targeting (TARGET) in yeast and *Drosophila* was introduced (McGuire et al., 2003) (Figure 1 B). This system utilizes Gal80ts (a cloned temperature sensitive version of Gal80) under the ubiquitous *tubulin 1α* promoter as a repressor for Gal4. At permissive temperatures of 18°C, Gal80ts unfolds its activity and is thus able to inhibit Gal4 dependent downstream transcription (McGuire et al., 2003). Restrictive temperatures above 29°C no longer allow the effect of Gal80ts. This leads to a temperature dependent control of Gal4 activity and therefore limits the transcription of UAS downstream genes to temporal phases.

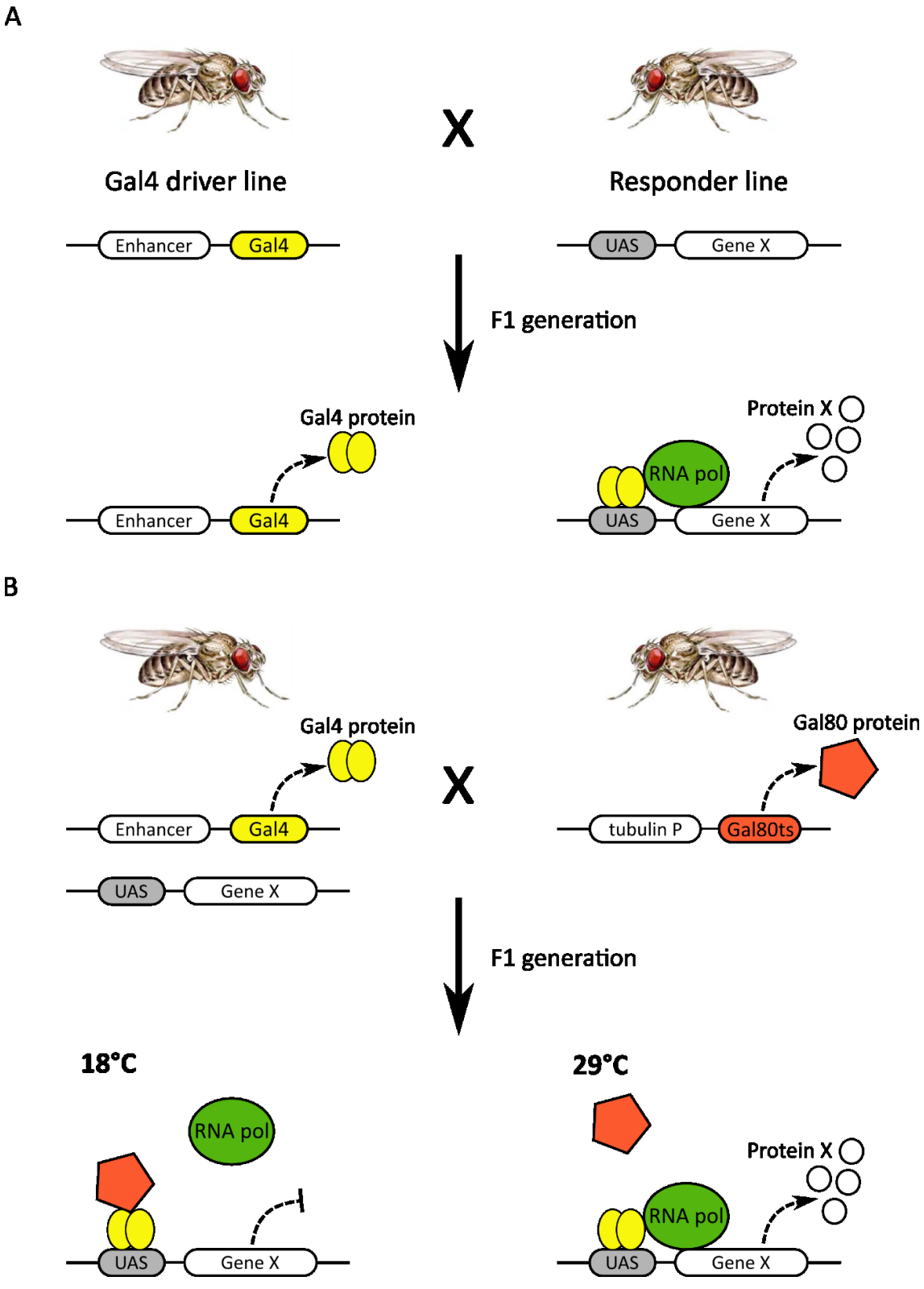

A
Gal4 driver line
X
Responder line
Enhancer
Gal4
UAS
Gene X
F1 generation
Gal4 protein
Protein X
RNA pol
B
Gal4 protein
Gal80 protein
tubulin P
Gal80ts
18°C
29°C

Figure 1: The Gal4/UAS-system

(A): Gal4 driver lines incorporate the Gal4 construct, which is driven by a cell type specific promoter (enhancer). The UAS-target line involves the UAS DNA-construct followed by the gene of interest (Gene X). Two distinct fly lines are crossed to receive a combination of the Gal4 and UAS constructs in their progeny. In progeny that inherits both constructs, Gal4 is expressed cell type-specifically and consequently binds to the UAS (Upstream Activating Sequence) construct. This binding promotes the transcription of the downstream gene of interest. Figure produced on the basis of (Neckameyer & Bhatt, n.d.). **(B):** The Gal4 system may be inhibited with the usage of Gal80. In progeny, which comprise both, Gal4 and Gal80, this protein physically interacts with Gal4, allowing the binding to UAS but prohibiting its transcriptional activity. The temperature sensitive variant of Gal80 is Gal80ts, which enables the use of temperature controlling the activity of a specific gene. At 18°C Gal80ts is active, inhibiting Gal4 activity. Above temperatures of 29°C, Gal80ts can no longer bind Gal4, therefore enabling the transcription of the gene of interest. Figure produced on the basis of (Neckameyer & Bhatt, n.d.).

1.1.1.2 Gal4 technique for real-time and clonal expression (G-TRACE)

To trace specific cell lineages, the Gal4 technique for real-time and clonal expression (G-TRACE) can be utilized (Figure 2). This technique shows real time Gal4 expression but also highlights cell lineages deriving from Gal4 expressing cells (Evans et al., 2009). In order to mark cells via G-TRACE, the method combines two distinct tools: Gal4/UAS (see 1.1.1.1 The Gal4/UAS-System) and the FLP-FRT system. The latter system relies on the principle of recombination which is mediated via the recombinase Flippase (FLP), a protein evolving of the yeast plasmid 2µ circle (Broach et al., 1982). FLP acts site-specifically and binds to DNA sequences referred to as FRT (Flippase recognition target) (Schweizer, 2003). Whenever FLP is expressed, it initiates the cleavage of FRT surrounded sequences via recombination (Broach et al., 1982). This sets the opportunity to control gene expression as genes or sequences of interest can be linked with FRT sites (Schweizer, 2003). In the case of G-TRACE, Gal4 is linked to a cell type specific promoter/enhancer (Evans et al., 2009). The fluorescent marker RFP (red fluorescent protein) is utilized for real time expression whereas lineage based expression is depicted with EGFP (enhanced green fluorescent protein).

RFP and FLP are utilized as UAS - transgenes, consolidating their dependence on Gal4 as RFP and FLP can only be expressed in cells where Gal4 is present. For G-TRACE, FLP targets FRT sites which flank stop codons separating the nEGFP (nuclear EGFP) sequence and the ubiquitous Ubiquitin-p63E (Ubi-p63E) promoter. Once FLP is expressed, FRT - flanked stop codons are excised resulting in the expression of EGFP. In this system, Gal4 is needed for the initiation of lineage reporter expression but not for its maintenance as the permanent removal of the Stop codon leads to an inheritance of EGFP expression in the next generations (Evans et al., 2009).

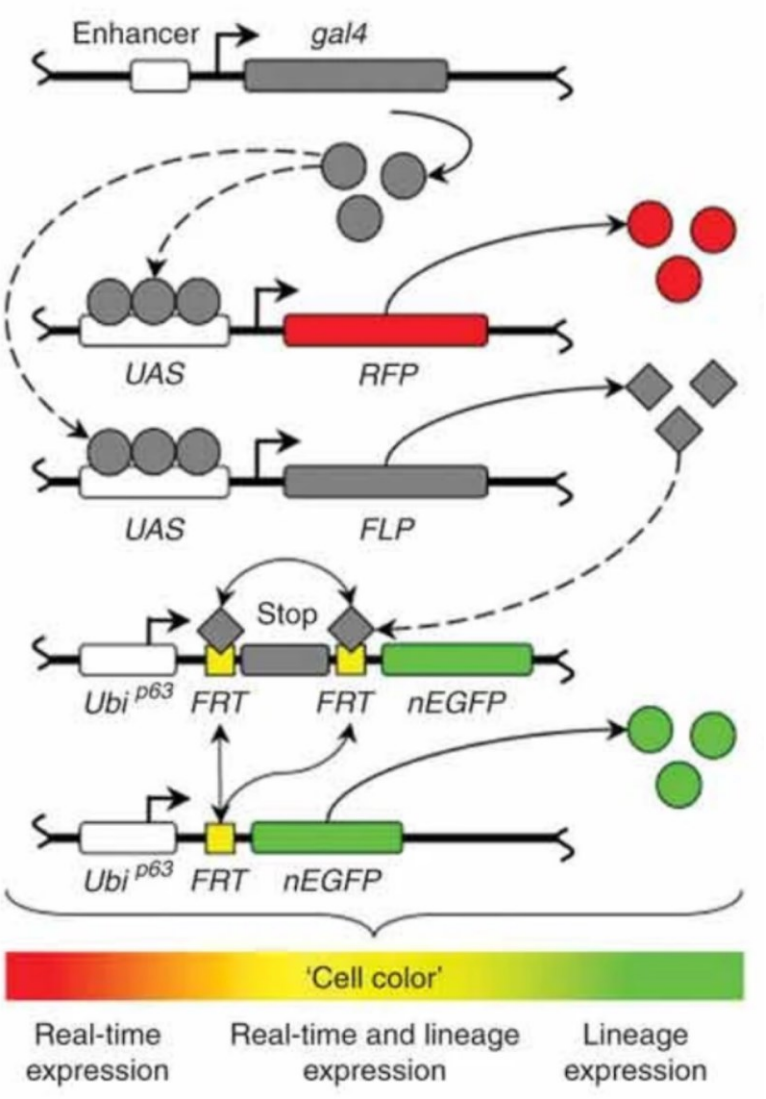


Figure 2: The G-TRACE system

The Gal4 technique for real-time and clonal expression (G-TRACE) is used to trace specific cell lineages in desired tissues while simultaneously marking live expression of cells (Evans et al., 2009). G-TRACE relies on the Gal4/UAS and FLP-FRT system (Evans et al., 2009). Gal4 is expressed under a cell type specific promoter and binds to UAS (Upstream activating sequence). As Gal4 is expressed cell type specifically, UAS downstream genes are only transcribed in those cells with Gal4 production. The Gal4 protein in G-TRACE binds to two distinct UAS-transgenes: UAS-RFP (red fluorescence protein) and UAS-FLP (Flippase). The recombinase FLP targets specific DNA sequences known as FRT (Flippase recognition target). The expression of RFP marks cells, which depict live expression of Gal4 and marks those cells whose lineage is traced. Once FLP is active, it cuts out FRT flanked stop codons upstream of the EGFP (enhanced green fluorescence protein) gene via recombination, which results in the expression of EGFP. This fluorescence marker depicts cell lineages deriving from Gal4 expressing cells. Gal4 is hereby necessary for the initial induction of EGFP expression, however offspring cells express GFP irrespectively of Gal4 expression. Reprinted with permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Methods, G-TRACE: rapid Gal4-based cell lineage analysis in *Drosophila*, Cory J Evans et al., 2009.

1.1.1.3 Ribonucleic Acid Interference

Ribonucleic acid interference, shortly known as RNAi, is a mechanism used by several organisms in order to implement gene silencing on the level of mRNA (Agrawal et al., 2003). It is a natural process which not only regulates gene activity but also acts as a cellular defense mechanism against invading viruses (Agrawal et al., 2003). Since this process was discovered, it was rapidly recognized as a powerful tool to study gene function and revolutionize our understanding of biology (D. H. Kim & Rossi, 2008; M. Rao & Sockanathan, 2005). Broadly speaking, RNAi has become scientifically crucial and with respect to human disease modeling, novel RNAi therapeutics might leverage an alternative in the medication of diseases (Bumcrot et al., 2006). In *Drosophila,* the expression of RNAi constructs can be achieved by linking it to the Gal4 dependent UAS promoter. With the usage of the Gal4/UAS - system, RNAi transgenes can be expressed cell type-specifically (Figure 1 B).

RNAi is initiated by the cleavage of long double stranded RNA (dsRNA) into smaller fragments exerted by the endoribonuclease enzyme dicer (Bernstein et al., 2001; Bumcrot et al., 2006) (Figure 3 A). Dicer in general produces two types of fragments: small interfering RNA (siRNA) which derive from dsRNA and miRNA which are the product of precursor miRNA procession (Lam et al., 2015). Both are considered to be non-coding RNA but differ with respect to their target precision: whereas siRNA are completely complementary to their targets, microRNAs display only partial base pairing (Lam et al., 2015). This results in only one target mRNA for siRNA compared to multiple targets of endogenous sequences for microRNAs (Lam et al., 2015). The production of siRNA and microRNA allows the RNAi machinery to build the RNA-induced silencing complex (RISC) (D. H. Kim & Rossi, 2008). Core representatives of RISC are Argonaute protein family members (D. H. Kim & Rossi, 2008) which are responsible for the cleavage of the loaded siRNA/microRNA, binding only one strand to RISC (guide strand) and releasing the other (Bumcrot et al., 2006; D. H. Kim & Rossi, 2008). The guide strand then directs RISC to the target mRNA by complementary base pairing and leads to its degradation via cleavage (Bumcrot et al., 2006).

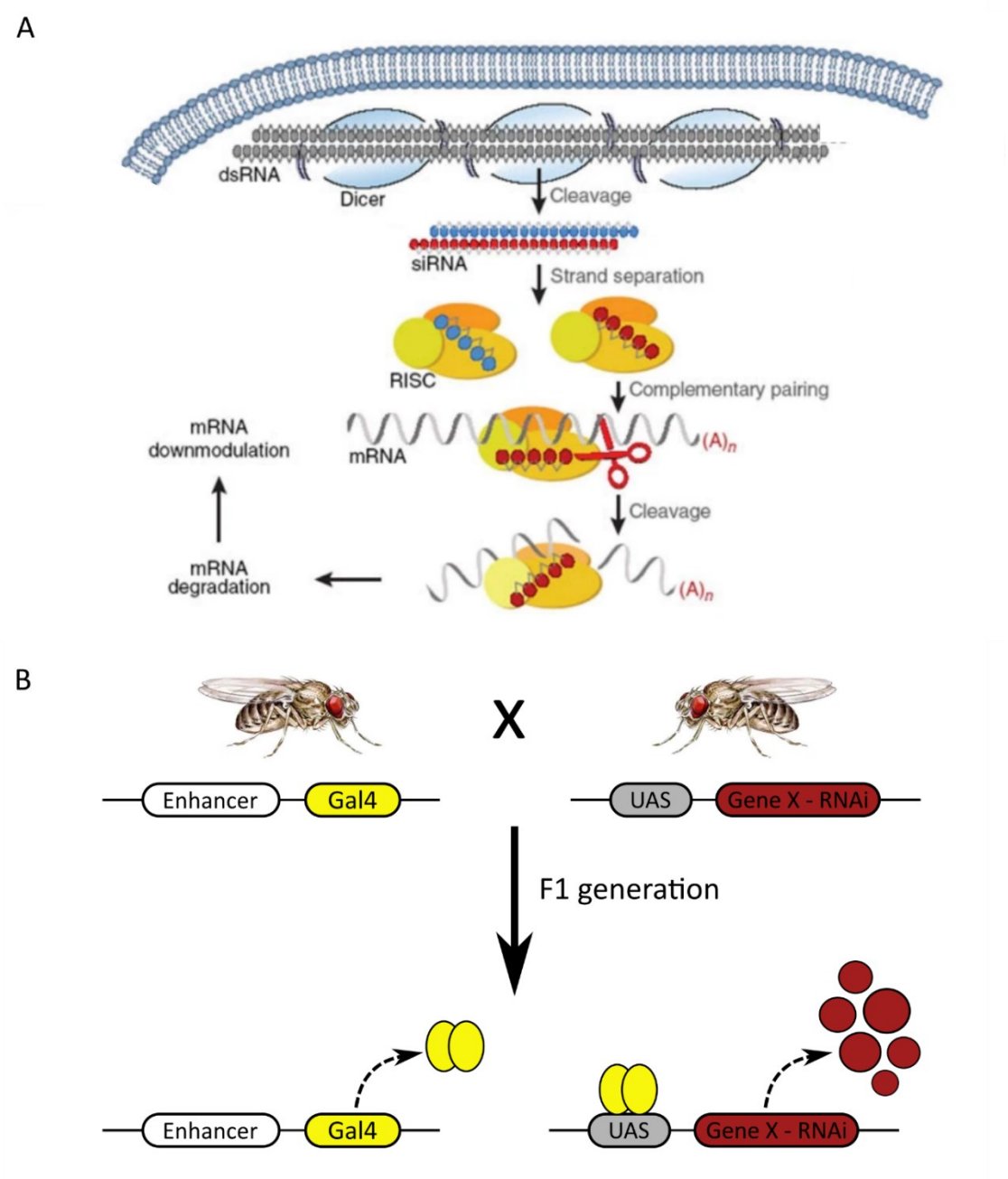


Figure 3: Ribonucleic Acid Interference and its usage in *Drosophila*

(A): Overview of the Ribonucleic Acid Interference (RNAi) process. Long double-stranded RNAs (dsRNA) are transcribed and processed in the nucleus and exported into the cytoplasm. With the help of Dicer, an endoribonuclease enzyme, dsRNAs are cut into shorter fragments producing small interfering RNAs (siRNA). RISC (RNA induced silencing complex) catalyzes the cleavage of the double stranded siRNA, binds the guide strand and releases the other one. The guide strand directs RISC to a complementary target mRNA. The complementary base pairing allows the cleavage, mRNA degradation and consequently its downregulation. Modified and reprinted with permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Chemical

Biology, RNAi therapeutics: a potential new class of pharmaceutical drugs, David Bumcrot et al., 2006. **(B):** Paradigm of RNAi incorporation in *Drosophila* genetics. With the use of the Gal4/UAS-system, RNAi constructs can easily be introduced for tissue specific expression. Figure produced on the basis of (Kelly et al., 2017).

1.2 Stem cells

Stem cells are an organism's master cells, giving rise to all sorts of tissues and cells. They are known to be undifferentiated cell types with the ability to self-renew and produce one or more differentiated cells as descendants (Shen et al., 2004). In general, there are two distinctive types of stem cells: embryonic stem cells (ESC) and adult (or tissue-specific) stem cells (ASC) (Figure 4) (Shen et al., 2004). ESCs derive from the inner cell mass of blastocysts and are pluripotent thus enabling them to generate cells of any lineage and germ layer (Shen et al., 2004; Zakrzewski et al., 2019). ASCs however are more limited in their production range and are linked to tissue specificity by producing cells of a particular lineage (Shen et al., 2004; Zakrzewski et al., 2019). They give rise to daughter cells referred to as progenitor cells which often proliferate transiently and reside at an intermediate state of differentiation with high commitment for further differentiation (Shen et al., 2004). Current paradigms suggest that with each differentiation stage the range of plasticity becomes more restricted, describing the genesis of cells in a hierarchical, unidirectional differentiation fashion (Theise & Krause, 2002). Introduced by Schofield in 1978, many stem cells depend on their adjacent niche environment for proper viability (Schofield, 1978). Niches are defined as compartments, which comprise extracellular matrix components, cells, signals and factors, vascular and nervous systems which all together preserve stem cell activity (Pérez et al., 2018). It has been proposed that niches are distinguished in two structural types: stromal niches, with permanent stromal niche cells contacting the supporting stem cell but generally being independent of stem cell presence, and epithelial niches, devoid of fixed specialized niche cells but contacting moving or developing cells (Morrison & Spradling, 2008). Most *Drosophila* niches are found to function by utilizing two fundamental processes: adhesive interaction and asymmetric signaling (Losick et al., 2011) (Figure 5). Occupancy within a niche is primarily determined by adhesive interactions between stem cells and their niche cells. While niche adhesiveness

is much less developed on the outside of the niche, daughter cells have the possibility to leave the niche environment. Fate determination is regulated by asymmetric signaling, leading to repression of stem cell differentiation within the niche and facilitating stem cell adherence. Outside of the niche, differentiation is being promoted and the level of adhesiveness is reduced (Losick et al., 2011).

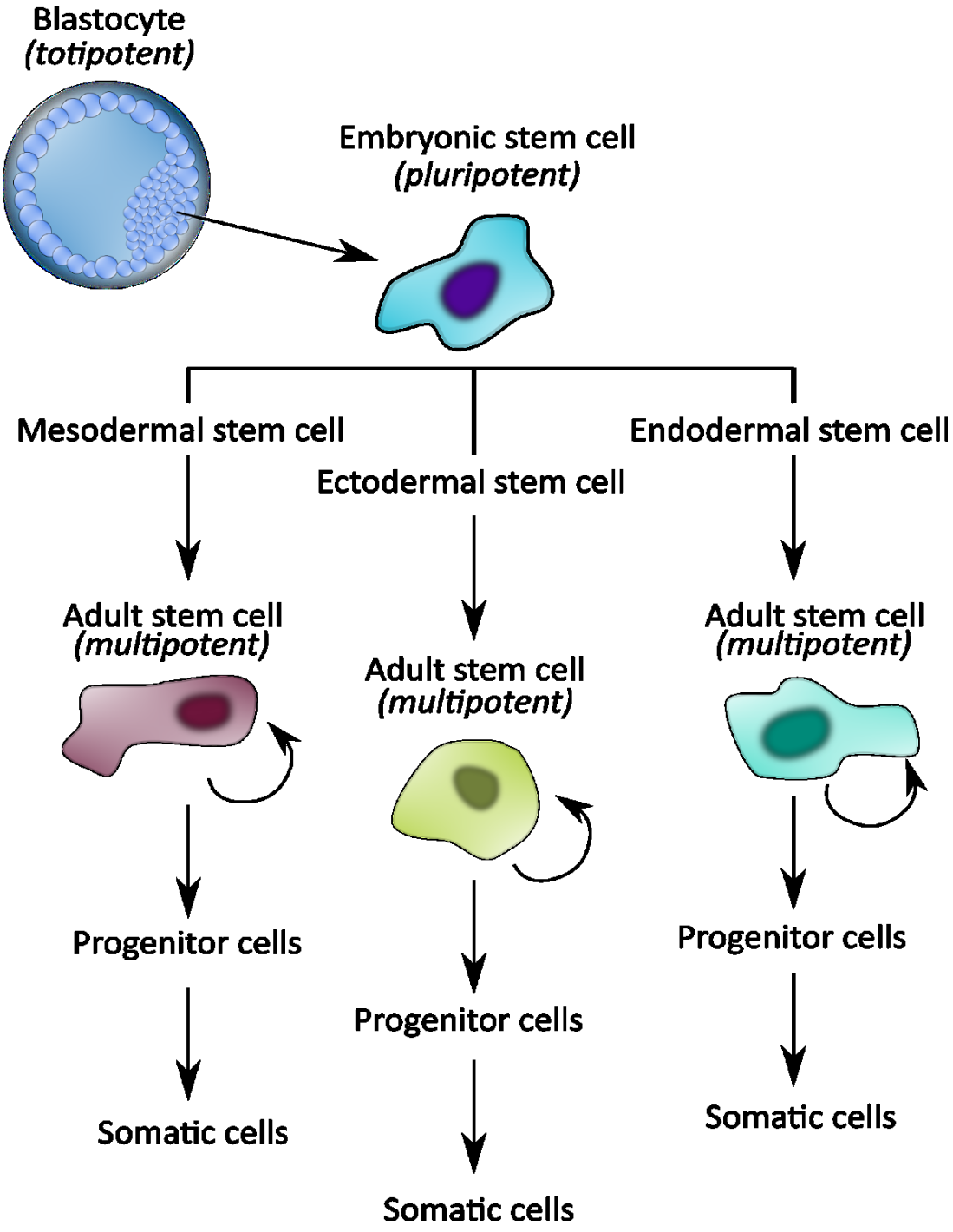


Figure 4: Stem cell categories and differentiation paradigm

There are different types of stem cell categories with respect to their potency levels. Blastocysts are considered to depict totipotent cells, which have the ability to convert to any cell and tissue type. Pluripotent cells, such as embryonic stem cells (ESC), possess the capacity to convert into a variety of differentiated cells. The following differentiation stages for ESCs involve the classification into ectodermal, mesodermal and endodermal stem cells. Adult stem cells (ASC) arise from either ectodermal, mesodermal or endodermal stem cells and are more limited and linked to tissue-specificity. In order to maintain their tissue homeostasis, they possess the ability to self-renew and produce daughter cells, which undergo several differentiation stages until reaching their final cell fate. Figure produced on the basis of (Shen et al., 2004) and (Zakrzewski et al., 2019).

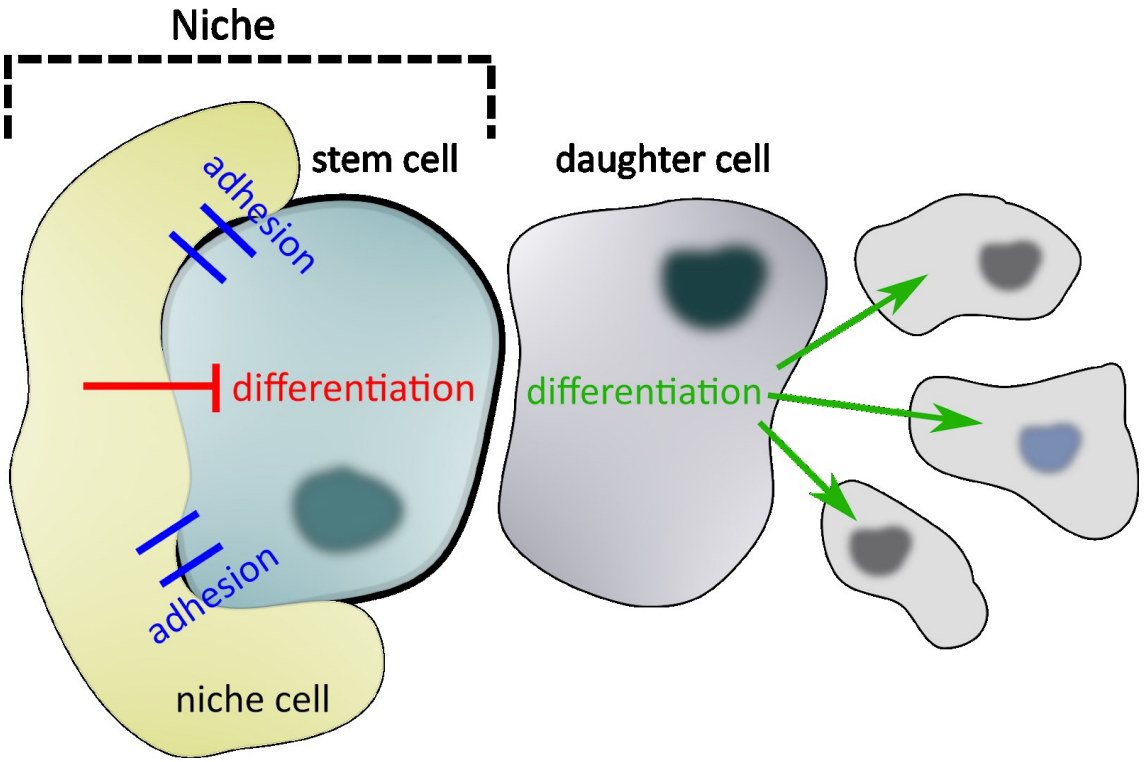


Figure 5: Niche functions for stem cell maintenance

Stem cells depend on their microenvironments, referred to as niches. Most niches promote stem cell function by two distinct principles: adhesive interaction and asymmetric signaling. While adhesiveness is elevated inside the niche, differentiation is repressed. Outside of the niche differentiation is promoted and therefore adhesiveness is reduced. Figure produced on the basis of (Losick et al., 2011).

Stem cells and their niches do not act autonomously but are responsive to environmental stimuli (Drummond-Barbosa, 2008; Losick et al., 2011; Pérez et al., 2018). Broadly speaking, these stimuli comprise nutritional perturbations, stress, infection and age which may alter stem cell activity (Losick et al., 2011; Pérez et al., 2018). In fact, stem cells are as well found to adapt their behavior upon tissue damage or stem cell loss as they might compensate the injury by migrating to the wound or replacing lost neighboring stem cells (Drummond-Barbosa, 2008; Ito et al., 2007; Magavi et al., 2000). Generally, the research on stem cells, niches and their behavior yield a platform for clinical utility as stem cell treatments such as transplantation are hallmarks of regenerative medicine.

1.3 Transdifferentiation

Transdifferentiation questions the paradigm of a cell's fate and lineage, which has been well believed for decades as of yet (Ladewig et al., 2013). It is defined as the process of conversion of a differentiated cell type to another distinctly differentiated cell type (Goro Eguchi & Kodama, 1993; Tosh & Slack, 2002). The term was coined by Selman and Kafatos to elucidate the observation of cells in the labial gland of the silk moth which undergo conversion during metamorphosis without dividing (Selman & Kafatos, 1974). Since then, several studies reported the evidence of transdifferentiating cells in animal models (Bachem et al., 1993; G. Eguchi & Okada, 1973; M. Sambasiva Rao et al., 1986). Interesting examples hereby are the well documented observations of hepatocytes, specialized cells of the liver tissue, emerging in damaged exocrine pancreatic tissue (Dabeva et al., 1997; M. S. Rao et al., 1988). Furthermore, bone marrow cells have been shown to populate tissues of other than mesenchymal origin, although these demonstrations have been considered controversial (Heike & Nakahata, 2005; Petersen et al., 1999; Shen et al., 2004). It has been proposed that the mechanism of transdifferentiation may occur in two versions: either the ancestor cell switches directly to the differentiated offspring cell or the process involves an intermediate stage via dedifferentiation and division (Shen et al., 2004) (Figure 6). Most presumably, the phenomenon relies on genetic regulations such as the alteration of a critical transcription factor which might promote the developmental commitment (Shen et al., 2004). Alternatively, environmental stimuli could as well

lead to the induction of underlying factors for transdifferentiation (Shen et al., 2004). Since transdifferentiation depicts a lucrative and innovative opportunity for clinical application, the question arises how this might be implemented. Cell-based therapy in the sense of artificially producing cells and tissues would constitute a breakthrough with respect to tissue engineering and provides a promising alternative to transplantation in the field of regenerative medicine (Shen et al., 2004). Further research and wise practice of our knowledge on developmental biology could hence enable novel therapeutic strategies and medical procedures.

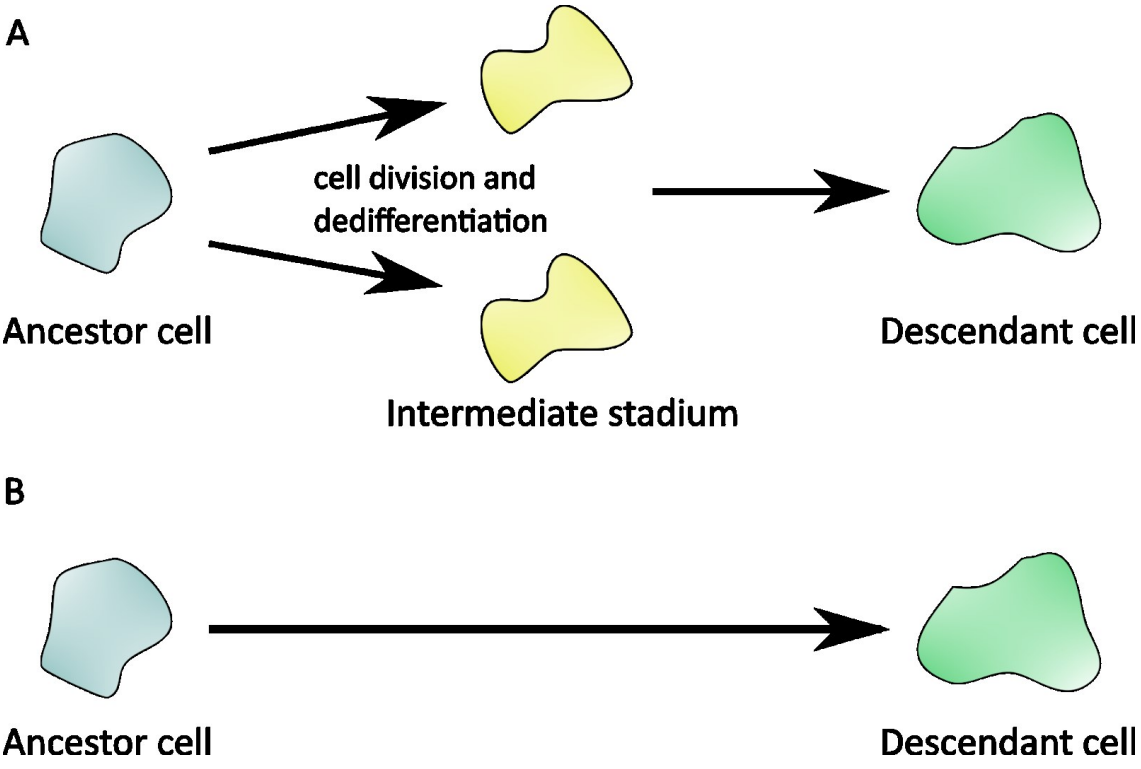


Figure 6: Transdifferentiation modes

There are two considered principles for transdifferentiation. **(A):** Transdifferentiation of an ancestor cell involves an intermediate state, in which the cell divides and dedifferentiates. Given a separate phenotype, the intermediate stadium neither comprises features of the ancestor nor the descendant cell. **(B):** The ancestor cell directly switches to another differentiated cell type. Figure produced on the basis of (Shen et al., 2004).

1.4 Single-cell RNA Sequencing Analysis

Most biological systems incorporate a highly complex structure of diversely arranged cells. These are the fundamental building blocks of every tissue, organ and organism (Macosko et al., 2015). Consequently, the identification of cell types and their roles in crucial mechanisms are essential to study their behavior within maintaining tissue and organ function. For the most part, our knowledge on these systems is limited and therefore our understanding of molecular processes is incomplete (Macosko et al., 2015).

Hence, recent RNA sequencing technologies target cell-type specific transcriptomes to give insight to their genetic picture for further detailed understanding (X. Li & Wang, 2021). Classical approaches of bulk RNA sequencing (bulkRNA-Seq) have been improved and adapted to desired scientific needs, leading to the development of single cell RNA sequencing (scRNA-Seq) (X. Li & Wang, 2021). Bulk RNA-Seq identifies averaged gene expression across a studied sample (G. Chen et al., 2019; X. Li & Wang, 2021). The generalization of gene expression across all cells, however, obscures the transcriptomic signature of rare cell types (X. Li & Wang, 2021). ScRNA-Seq has made it possible to comprehend transcriptomes on the single cell level (Zhang et al., 2021). In contrast to previous methods, scRNA-Seq is able to target more heterogeneous tissues and addresses each cell separately thus depicting a fascinating springboard for (oncological) research (Zhang et al., 2021). This is achieved by the innovative strategy of mapping each collected mRNA to its cell of origin (Macosko et al., 2015). As individual cell expression profiles are in focus, differences and similarities among cells are emphasized which consequently may lead to the identification of (rare) cell types (G. Chen et al., 2019; Kashima et al., 2020; X. Li & Wang, 2021; Zhang et al., 2021).

To date there are several scRNA-Seq protocols developed, all evolving from the first method introduced by Tang and colleagues (Tang et al., 2009). Developed protocols may differ in distinct steps of the scRNA-Seq workflow. Two major approaches of cell isolation are commonly used: droplet- or well-plate-based techniques. Droplet-based approaches (e.g. DropSeq, 10x Chromium) cover tens of thousands of cells at once, which allows us to address high amounts of cells with less effort (Baran-Gale et al., 2018). Hereby, single cell suspensions are

infused in a microcapillary system (Macosko et al., 2015) (Figure 7). With the usage of an oil suspension, cells pass the system to become separately captured into droplets/beads. These droplets are unique and captured in a tube for further procedure (Macosko et al., 2015). Plate-based techniques (e.g. SMART-Seq) portray its benefits in the possibility of amplifying RNA constructs of small quantity tissues, however the throughput is rather low as a small amount of cells is processed (Baran-Gale et al., 2018). Plate-based are often combined with prior cell sorting. This can be achieved via manual pipetting or fluorescence activated cell sorting (FACS) (Baran-Gale et al., 2018; Zhang et al., 2021) (Figure 7). For FACS, targeted cells are fluorescently labeled and are selected within the cell suspension via laser detection (Hu et al., 2016). Selected cells are then captured in individual wells of a plate for further procedure. In order to be compatible with sequencing platforms, scRNA-Seq methods generally need to incorporate cDNA production (Natarajan et al., 2019). Therefore, most of them rely on similar biological principles such as template switching and reverse transcription (Baran-Gale et al., 2018). In the following, a generalized standard workflow for scRNA-Seq is described (Figure 8): cell lysis constitutes the first step of practice. The lysis solution contains primers that are added to the 3' end of the mRNA (Macosko et al., 2015; Picelli et al., 2014). Most methods use oligo (dT) sequences which are complementary to the 3' poly-A tail of the released mRNA (Zhang et al., 2021). In addition, primers hold PCR handles for later PCR amplification (Macosko et al., 2015). Several protocols incorporate primers with sequences for cell barcodes and UMIs (unique molecular identifier) (Hashimshony et al., 2016; Hochgerner et al., 2017; Macosko et al., 2015). In this regard, primers of each cell share the same cell barcode but UMIs differ, which enables the mapping of each mRNA to its cell of origin (Macosko et al., 2015). The addition of cell barcodes ensures that samples can be pooled together without confusion (Zhang et al., 2021). Two properties of the Moloney Murine Leukemia Virus (MMLV) are used for further procedure: reverse transcription and template switching (Picelli, 2017). The MMLV uses the primed mRNA to synthesize the first strand of the resulting cDNA. At the cDNA's 3' end, MMLV adds non templated nucleotides, which allows the docking of TSO (template switching oligo) (Zhang et al., 2021). TSO incorporates complementary base pairing to the cytosine tail of the cDNA and enables the MMLV to switch templates and synthesize the second cDNA

strand (Picelli et al., 2014). The TSO can be connected to the same artificial PCR sequence as the initially used primers in order to use only one PCR primer for the subsequent amplification (Picelli, 2017; Picelli et al., 2014). The resulting cDNA is used as a template for PCR amplification and for further sequencing and analysis ultimately generating large RNA libraries with single cell transcriptomes that are mapped to its gene and cell of origin (Macosko et al., 2015).

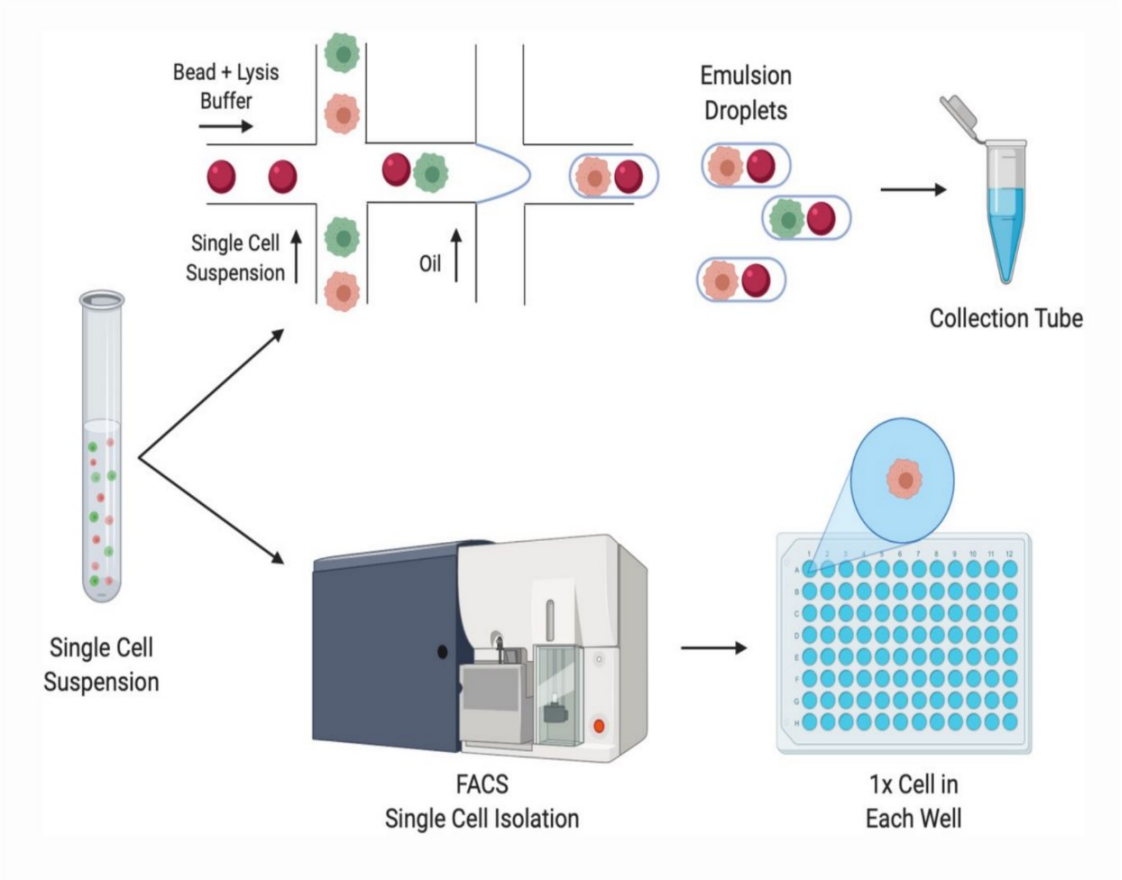


Figure 7: Procedure of droplet and plate-based scRNA-Seq technqiues

Desired tissues are processed into isolated cells, forming a single cell suspension. There are different approaches to sorting single cell suspensions: droplet based techniques or fluorescence activated cell (FAC) sorting. For droplet techniques, single cells are infused into a microcapillary system with oil suspension in order to capture individual cells in beads. These beads are lastly collected in a tube. FAC sorting relies on the detection of fluorescence properties of each cell and sorts cells by their given electrical charges. Cells are then collected separately in wells of a plate. Figure from (Probst et al., 2020) with permission by CC-BY-NC-ND 4.0 International license (CC-BY-NC-ND 4.0).

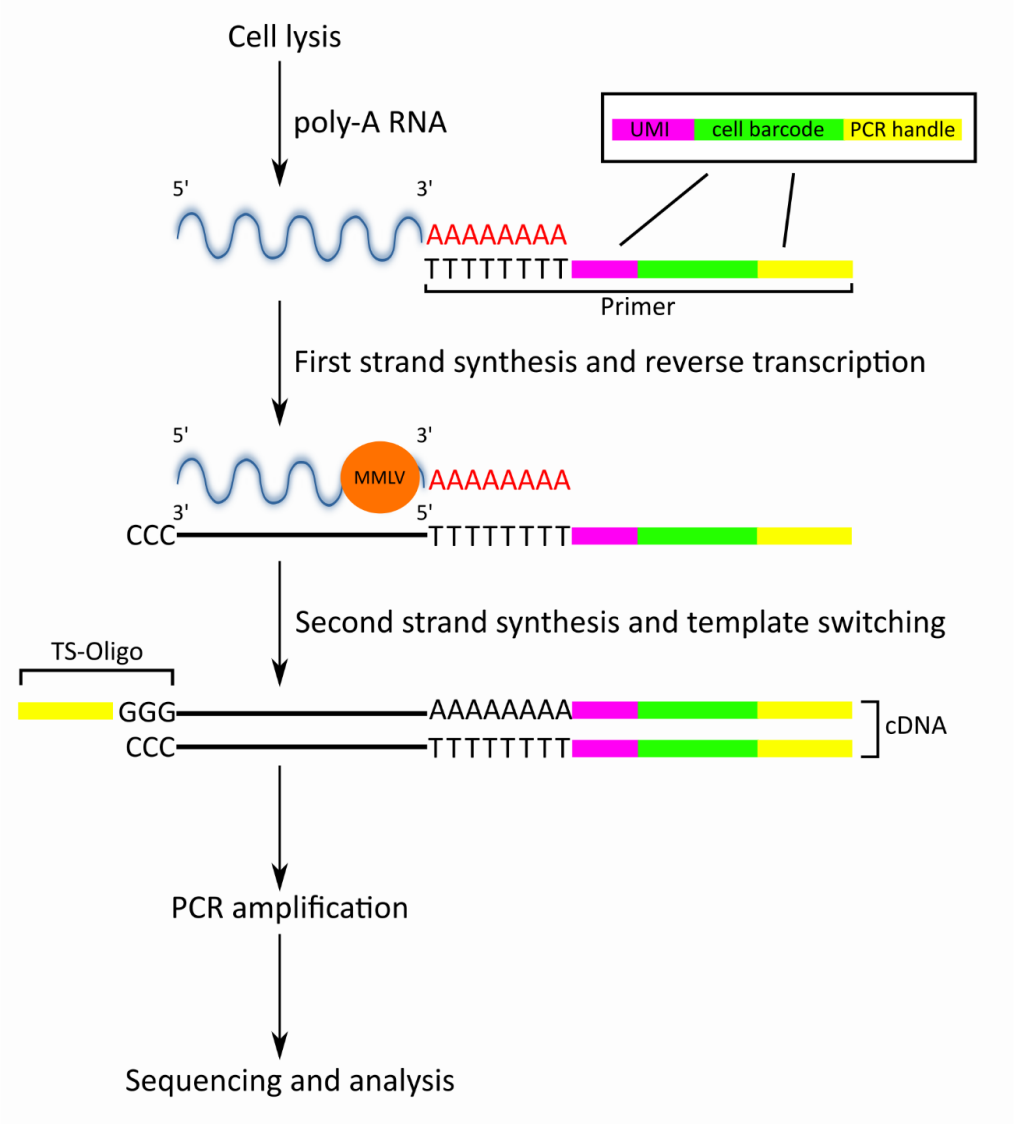


Figure 8: Standardized workflow of scRNA-Seq

Cells are lysed into poly-A RNA. Present primers bind to the poly-A tail of the mRNA via the primer's poly-T tail. These primers incorporate distinct artificial sequences: PCR handles for later amplification (yellow), cell barcodes specific for each cell (green) and unique molecular identifiers (UMI) specific unique sequences for each RNA (magenta). With the help of the Moloney Murine Leukemia Virus (MMLV), the first strand can be synthesized via reverse transcription. MMLV produces a poly-C overhang at the 3' end of the resulting cDNA strand. This enables TS (template switching)-Oligo with a complementary base pairing structure to bind the poly-C construct. These TS-Oligos can be combined with the same PCR handle as the initially used primers so that only one PCR primer is necessary for the following amplification. The binding of TS-Oligo consequently allows MMLV to switch templates and synthesize the second strand, resulting in double-stranded cDNA. These are used as templates for PCR amplification and subsequently sequenced and analyzed.

There are published versions as well as ambitious projects of mapping whole organisms with the usage of scRNA-Seq technologies. These include the Human Cell Atlas (HCA) for the human body as well as the Fly Cell Atlas (FCA) acting as the equivalent for the adult *Drosophila* model organism. Generally, the FCA constitutes the first cohesive dataset incorporating over half a million cells encoding for over 250 distinctive cell types of nearly all tissues of the *Drosophila* organism (H. Li et al., 2022) (Figure 9). For this study, I addressed one of the datasets published in the FCA namely a single nuclei dataset, created with the desire to give particular attention to the research on escort cells (cell type considered further below). Despite that the HCA is only partially completed and still in progress, both projects aim for the representation of a platform to act as a valuable source for research at cellular resolution. These ventures allow us to study cell-cell behavior in detail and affect several aspects of biology and medicine, since the understanding of physiological processes and disease genesis may be addressed (Regev et al., 2017). To point out in the interest of recent events, the HCA even at its early stages contributed to the comprehension of the mechanisms of the SARS-Covid-19 disease (Regev et al., 2017).

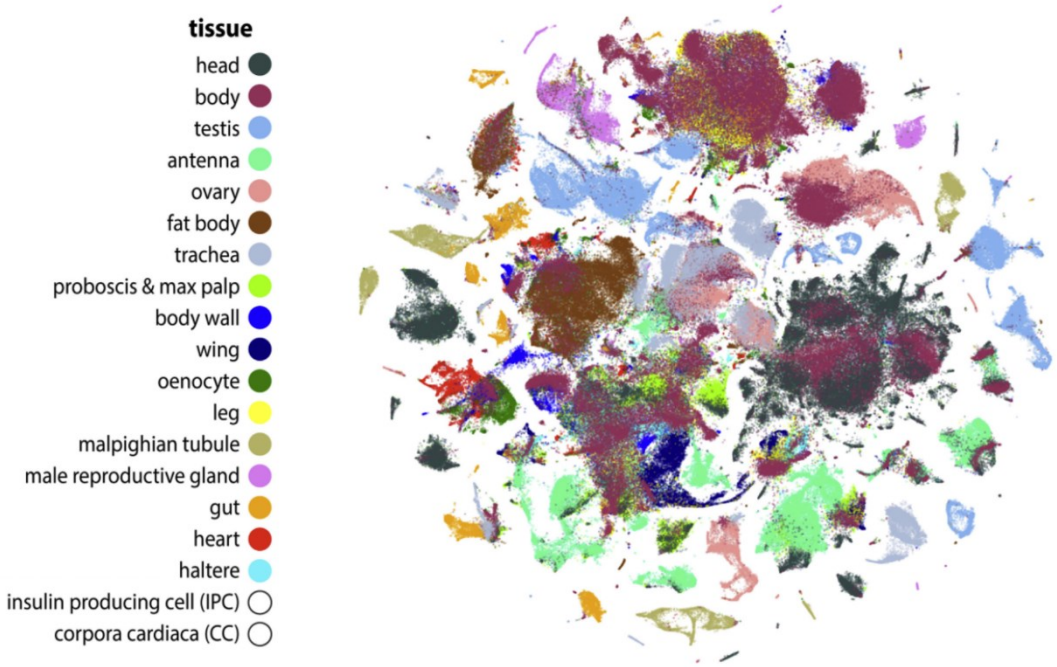


Figure 9: The Fly Cell Atlas

The Fly Cell Atlas (FCA) incorporates a vast amount of tissues and cells of the *Drosophila melanogaster* organism. It implies datasets of single cell sequenced tissues, finalized for further analyses. This map shows the included tissues distinguished by colors. From (H. Li et al., 2022) reprinted with permission from AAAS.

1.5 Anatomy of the female reproductive system of *Drosophila melanogaster*

Within female *Drosophila*, ovaries are located on the anterior end of the abdomen. The reproductive system is composed of paired ovaries, each consisting of about 16 strands of developing follicles referred to as ovarioles (X. Wu et al., 2008) (Figure 10). The ovaries are coated by a peritoneal sheath, holding the ovarioles together and consisting of a fine network of muscle fibers (Bloch Qazi et al., 2003). The ovarioles comprise two distinct morphological structures: the germarium, the most anterior structure of the ovariole, and the vitellarium, containing the developing and increasing egg chambers (Hinnant et al., 2020). Moving posteriorly, each ovariole forms a pedicel at its posterior end (Bloch Qazi et al., 2003). All pedicels unite to build a calyx which joins a lateral oviduct (Bloch Qazi et al., 2003). Subsequently, the two lateral oviducts end into the common oviduct and finally join the uterus. Spermathecae and spermathecal glands (also called parovaria or accessory glands) encircle the anterior tip of the uterus, while the back of the uterus is accompanied by the seminal receptacle (Bloch Qazi et al., 2003). The latter structure is used for storing the majority of the received sperms (Bloch Qazi et al., 2003; Pitnick et al., 1999). The remaining sperm is then stored in spermathecae, capsules connected to the uterus (Bangham et al., 2003; Bloch Qazi et al., 2003). Sperm storage allows the production of progeny for a duration of up to several weeks (Wolfner, 2011). Hereby, secretory glands are crucial for sperm viability, keeping them alive and motile by producing fluids and playing important roles for egg movement throughout the entire reproductive tract (Schnakenberg et al., 2011). Mature oocytes leave the ovaries to enter the uterus for fertilization (Wolfner, 2011). The vagina, located at the posterior end of the uterus, comprises the gonopore, which ultimately serves to discharge the fertilized eggs (Bloch Qazi et al., 2003).

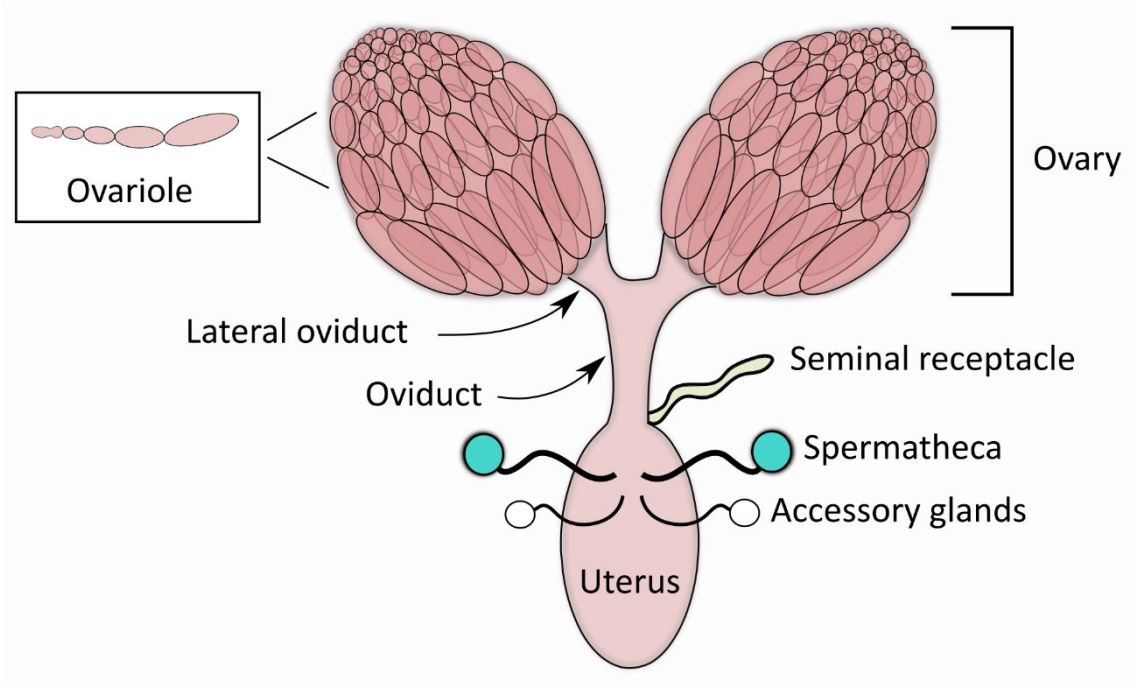


Figure 10: Female reproductive tract in *Drosophila melanogaster*

Anterior view of the female reproductive tract. Female *Drosophila* possess two ovaries covered by a peritoneal sheath holding them together. Each ovary consists of about 16 strands of developing ovarioles. The germarium is the most anteriorly located structure and produces the germline cysts which differentiate into mature oocytes. The posterior ends of each ovariole form pedicels, which unite to a calyx to join the lateral oviduct. These open into the common oviduct, which leads to the uterus. The uterus is accompanied by spermathecae, used for sperm storage, and accessory glands, which secrete fluids for proper sperm viability. The seminal receptacle, storing most of the received sperm, is visible from the posterior face of the uterus. Figure produced on the basis of (Bloch Qazi et al., 2003).

1.5.1 The *Drosophila* germarium

The germarium is located at the anterior tip of each ovariole, classified into four regions (J. M. McLaughlin & Bratu, 2015) (Figure 11 A). Oogenesis starts at the far anterior site of each germarium (Bastock & St Johnston, 2008). The anterior-most Region 1 is represented by two to three GSCs residing in its specialized niche, formed by terminal filament cells (TF) and cap cells (CC) (J. M. McLaughlin & Bratu, 2015). The process of oogenesis involves a timespan of approximately 7 days (Bastock & St Johnston, 2008). In this time, germline cysts pass through 14 morphologically distinct stages - starting with budding and finalizing with

mature eggs (Bastock & St Johnston, 2008). GSCs divide asymmetrically upon adulthood to self-renew and produce daughter cells for further differentiation (Bastock & St Johnston, 2008). Daughter cells are referred to as cystoblasts and undergo 4 circuits of incomplete mitosis to form a cyst of 16 interconnected cells (Bastock & St Johnston, 2008; Fuller & Spradling, 2007). The connections across the cells are known to be cytoplasmic bridges called ring canals (Bastock & St Johnston, 2008). Only one germ cell is selected for further differentiation into the oocyte, while the remaining cells become polyploid nurse cells, supporting the development and supplying the oocyte with nutrients (Bastock & St Johnston, 2008). The early stages of developing progeny of GSCs are accompanied by a population of cells, referred to as escort cells (formerly inner germarial sheath cells/IGS) (J. M. McLaughlin & Bratu, 2015). As the developing cyst exits the escort cell posterior boundary at region 2a/2b, it becomes encapsulated by a layer of pre follicle cells (pFCs) (Rust & Nystul, 2020). These derive from follicle stem cells (FSC), which reside in their specialized niches at the posterior side of the 2a/2b border (J. M. McLaughlin & Bratu, 2015). Followed by a series of cell-cell signaling induced by the germline cyst itself, pFCs initiate further differentiation (Roth & Lynch, 2009). They divide progressively to differentiate into polar cells, characterizing the anterior and posterior poles of the developing follicle, stalk cells, connecting the follicles and main body follicle cells (MBFC), forming a thin layer of cells ensheathing the majority of the follicles outer surface (Rust & Nystul, 2020). This process concludes the posterior end of the germarium and marks the separation process of the newly developed egg chamber which then buds from the germarium. Early stages (1 - 6) of egg chamber development comprise increase in egg chamber size as well as the initiation of follicle cell differentiation and polarization (Roth & Lynch, 2009). Mid stages 7 to 10 continue to polarize, future embryonic axes are specified, the follicle elongates and vitellogenesis occurs through the secretion of eggshell components (J. M. McLaughlin & Bratu, 2015; Roth & Lynch, 2009). Final stages 11 to 14 are characterized by the increase of oocyte volume, nurse cell “dumping” and apoptosis which ultimately leads to the finalized mature egg (Hudson & Cooley, 2014) (Figure 11 B).

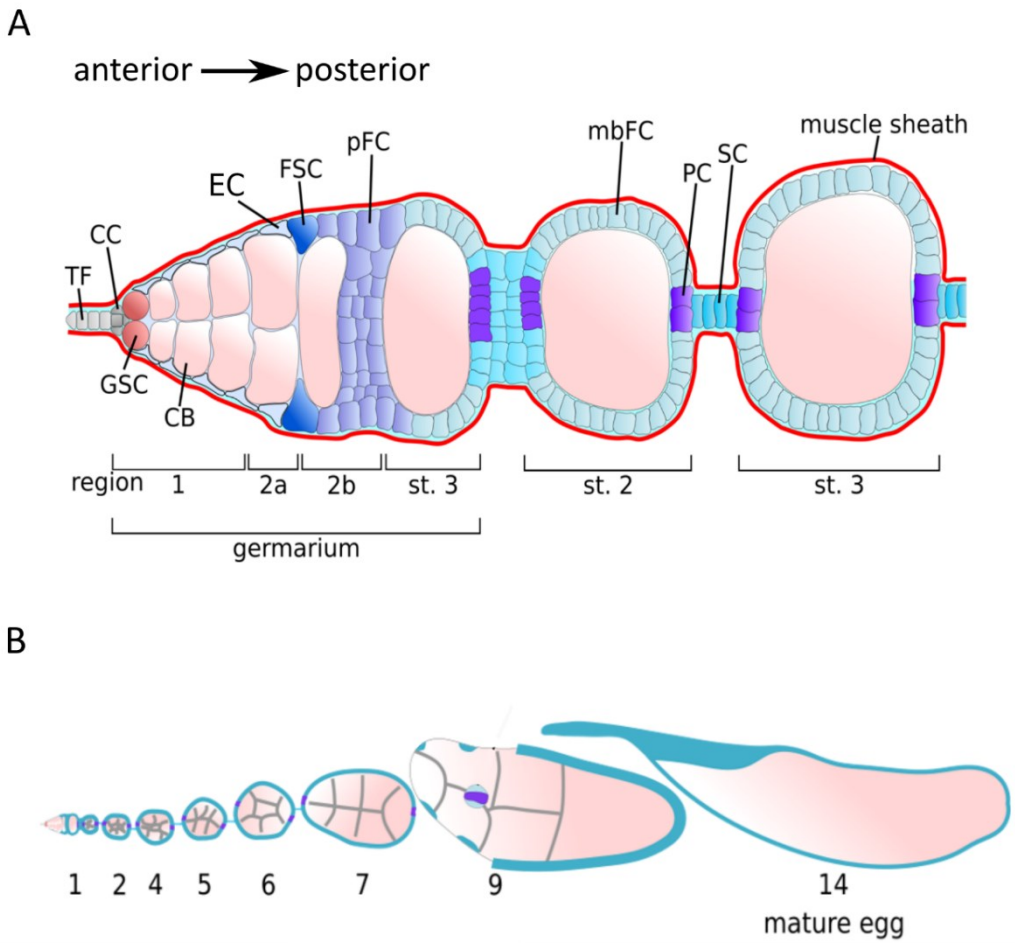


Figure 11: The *Drosophila* Germarium

(A): Graph depicting the germarium and its cell types in an anterior to posterior order. The germarium and the developing follicles are ensheathed by a muscle sheath. The germarium is limited by stage 1, more specific region 1 up to region 3. All following stages comprise individual developing follicles. The most anterior site of the germarium comprises germline stem cells (GSC) and their niche, built by terminal filament (TF) and cap cells (CC). GSCs produce cystoblasts (CB) which differentiate further. These stages are promoted and supported by escort cells (EC), which range from region 1 up to the border of region 2a/2b and are classified into anterior (aEC), central (cEC) and poster ECs (pEC). Posterior to the region 2a/2b border, follicle stem cells (FSC) reside. They produce daughter cells, referred to as prefollicle cells (pFC). Later stages involve budded follicles with main body follicle cells (mbFC), stalk cells (SC) and polar cells (PC). Modified figure by courtesy of Katja Rust. **(B):** Graphical representation of an ovariole containing the 14 stages of oogenesis, resulting in the mature egg at stage 14. Modified figure from (Rust et al., 2020) with permission by the Creative Commons Attribution 4.0 International License.

1.5.2 Escort cells

Escort cells (EC) are considered niche cells and are located in Regions 1 and 2a of the germarium, forming cellular extensions into the center with various lengths and participate in the support of GSC regulation and differentiation processes (Kirilly et al., 2011; J. M. McLaughlin & Bratu, 2015; Schulz et al., 2002). Hereby, the length and motility of EC protrusions are position dependent, suggesting functionally different EC classes (Banisch et al., 2017). Proper EC functioning is necessary for oogenesis, since perturbations within EC signaling lead to insufficient extension building which ultimately results in the failure of GSC differentiation (Banisch et al., 2017; Schulz et al., 2002). Additionally, ECs contribute to the maintenance of FSCs by being an integral component of the FSC niche and providing relevant signaling pathways (further discussed below).
The number of EC subtypes is a controversial topic. Recent scRNA-Seq studies on cells of the adult *Drosophila* ovary have all identified EC subpopulations but disagree with respect to the exact number of EC identities (Rust et al., 2020; Shi et al., 2021; Slaidina et al., 2021; Tu et al., 2021). These are introduced in the following. The identification of distinct subtypes of ECs arose by comparing gene expression patterns among EC cell clusters in scRNA-Seq datasets. Distinct marker genes were calculated for each hypothetical subpopulation. By using enhancer and protein trap lines, cluster localisation within the germarium was visualized, validating the suggestion of several EC subtypes. Certain studies identified the presence of four distinct EC types (IGS 1-4) (Tu et al., 2021) (Figure 12 C). By comparing the extent of gene expressions of the marker genes for each subpopulation throughout the germarium, IGS 1-4 are claimed to be arranged in an anterior to posterior order. In fact, they also demonstrate overlapping gene expression, suggesting the subpopulations to be adjacent to each other (Tu et al., 2021). IGS 1 is assumed to contact GSCs while IGS 2-4 are claimed to support the differentiating GSC progeny (Tu et al., 2021). However, other studies classified the EC domain in only two sectors: anteriorly located ECs (aEC) near region 1 and posterior ECs (pECs) residing next to the FSC site at the 2a/2b border (Shi et al., 2021) (Figure 12 D). aECs are considered to contact cap cells and possibly influence GSCs, whereas pECs are thought to be situated at an intermediate state among aECs and FSCs. This is reflected in gene expression

patterns for pECs, as they overlap with aECs and FSCs (Shi et al., 2021). In contrast, Rust *et al.* and Slaidina *et al.* categorized three EC subtypes by identifying gradients of gene expression throughout EC identities (Rust et al., 2020; Slaidina et al., 2021) (Figure 12 A and B). Investigating the expression of cluster specific markers identified a setup of ECs in an anterior to posterior order: starting in region 1 with the most anterior located aECs, up to the middle of the germarium containing mostly central ECs (cEC) (Rust et al., 2020). pECs reside at the posterior border of region 2a/2b which delineates the EC from the FC compartment, including FSCs (Rust et al., 2020). Given their anatomical position, EC subtypes are assumed to support different states of differentiation of developing GC progeny by creating distinct microenvironments with the secretion of specific sets of molecules (Slaidina et al., 2021).

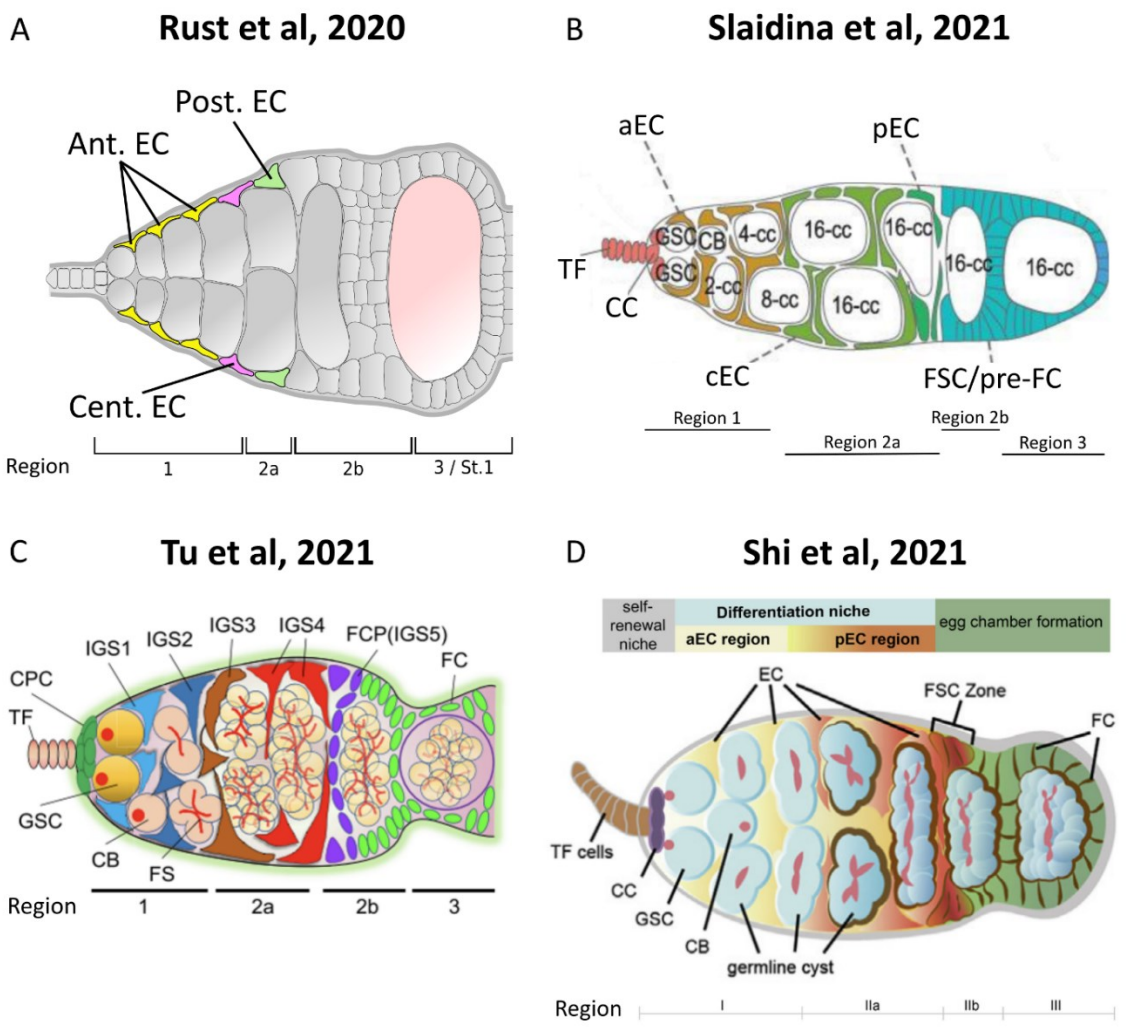


Figure 12: Escort cell identities in the scientific debate

(A): ECs localized in the germarium and classified by their estimated subtypes. According to scRNA-Seq analyses, Rust *et al.* postulates that ECs are divided into three subtypes, ranging from anterior to posterior (Rust et al., 2020). Modified figure from (Rust et al., 2020) with permission by the Creative Commons Attribution 4.0 International License (CC BY 4.0). **(B):** Note that similar results to Rust *et al.* were published by Slaidina *et al.* (Slaidina et al., 2021). By the analysis of scRNA-Seq datasets, Slaidina *et al.* established the existence of three EC subtypes: anterior EC (aEC), central EC (cEC), posterior EC (pEC). Modified figure from (Slaidina et al., 2021) with permission by the Creative Commons License (Attribution-NonCommercial 4.0 International) (CC BY-NC 4.0). **(C):** According to analyses conducted by *Tu et al.* (Tu et al., 2021), calculated IGS identities are visualized in the germarium. IGS 5 is considered to constitute a follicle cell progenitor population (FCP). Modified figure from (Tu et al., 2021) with permission by the Creative Commons Attribution 4.0 International License (CC BY 4.0). **(D):** Depiction of estimated EC compartments according to Shi *et al.* (Shi et al., 2021). Modified and reprinted from Current Biology, Vol. 31, Shi et al., A Progressive Somatic Cell Niche Regulates Germline Cyst Differentiation in the Drosophila Ovary, 18 Pages, 2021, with permission from Elsevier.

1.5.3 Follicle stem cells and niche morphology

FSCs are known to divide with an asymmetric outcome for self-renewal and the production of pFCs which are immediately adjacent to the stem cells (Nystul & Spradling, 2007a; Rust et al., 2020) (Figure 13 A). FSCs were firstly described by Margolis and Spradling (Margolis & Spradling, 1995). They discovered mitotically active cells by heat shock induced site-specific DNA recombination, producing clones with genetically heritable markers and comparing clone patterns after multiple time points. The observations proved the foundation to the assumption of two actively dividing stem cells per germarium by comparing the contribution of each lineage to a follicle (Margolis & Spradling, 1995). Nonetheless, how many FSCs actively reside in the germarium is still under debate as recent publications question this paradigm. Several studies are consistent with the presence of two FSCs which are Fasciclin III (Fas III) positive. Fas III is a surface protein expressed by all follicle cells and thus suggests that FSCs reside in the Fas III area (Fadiga & Nystul, 2019; Kirilly et al., 2005; Margolis & Spradling, 1995;

Nystul & Spradling, 2007b; Rust et al., 2020). In fact, it has been deliberated that the contribution of the FSC lineage may vary and the number of FSCs consequently may fluctuate from two to four (Fadiga & Nystul, 2019). In contrast, there are contradicting perceptions, postulating the existence of about 14-16 FSCs arranged in three anterior to posterior layers (Reilein et al., 2017). Posterior layer FSCs are considered to mainly produce pFCS and to be Fas III positive. Anterior and middle layer FSCs are claimed to give rise to predominantly ECs and are thought to be located anterior to the Fas III boundary, thus being Fas III negative (Reilein et al., 2017). However, FSCs may move from one layer to another and therefore switch in the production of FCs and ECs (Reilein et al., 2017).

The *Drosophila* FSCs provide an attractive model to study epithelial stem cell niches as they have been elusive to detailed understanding (Nystul & Spradling, 2007b; Rust & Nystul, 2020). Whereas well studied niches, such as those of GSCs, are more static and stem cells are maintained by the direct contact to their dedicated niche cells, it has been established that FSCs are thought to reside in highly dynamic niche environments (Margolis & Spradling, 1995; Nystul & Spradling, 2007b). In this context, ECs have been reported to constitute components of the FSC niche as they induce significant signaling pathways and are assumed to form adherens junctions with FSCs (Kim-Yip & Nystul, 2018; Sahai-Hernandez & Nystul, 2013; Song & Xie, 2002). A strategy for FSC substitution has been established by Nystul and Spradling (Nystul & Spradling, 2007b): FSCs possess the ability to replace lost neighboring stem cells by lateral cross-migration of daughter cells (Figure 13 B). In general, one FSC daughter remains in the niche while the other regularly migrates away (Nystul & Spradling, 2007b). As they are able to substitute the stem cell of the counter niche, daughter cells therefore compete for niche occupancy (Nystul & Spradling, 2007b). Unconventionally, FSCs may also be replaced by cell types of another cell lineage (Rust et al., 2020). This interesting phenomenon is considered in the subsequent chapter.

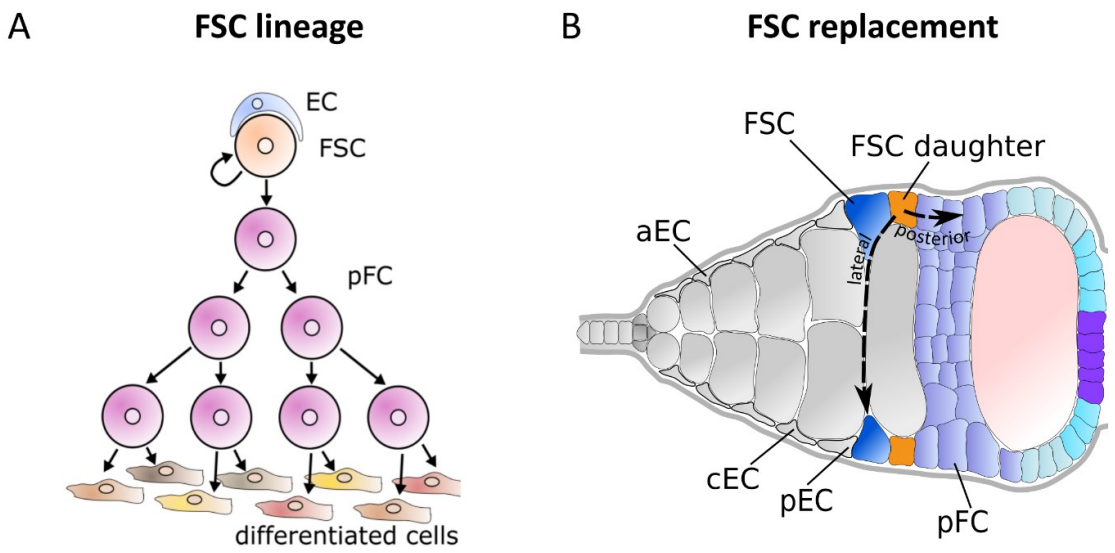


Figure 13: The follicle stem cell lineage and conventional FSC replacement
(A) A graphical representation of the FSC (follicle stem cell) lineage. Escort cells (EC) assume a niche function for FSCs. These produce the follicle epithelium by self-renewal and production of differentiating daughter cells, called prefollicle cells (pFC). pFCs undergo several rounds of division until reaching final differentiation. Figure by courtesy of Katja Rust. **(B)** Illustration of FSC replacement. FSCs divide with an asymmetric outcome, where one FSC daughter (orange) regularly remains fixed and then differentiates posteriorly and the other migrates laterally to the counter niche. The cross-migration of the FSC daughter leads to the competition for niche occupancy and may ultimately result in stem cell replacement. Modified figure by courtesy of Katja Rust.

1.6 Environmental stressors and FSC niche cell plasticity

Stressors are conditions that are defined to provoke stress responses in organisms which involve a cohesive set of behavioral and physiological changes. These include biotic factors, such as food availability and infections, abiotic factors, in the manner of the exposure to toxicants, and psychological/emotional threats (Schulte et al., 2014).

1.6.1 Diet perturbation leads to EC fate conversion

The most common environmental challenge in an animal's life are food limitations and fluctuations (Lee & Jang, 2014). As starvation states are considered to depict severe stress conditions, these may lead to adjusted strategies on survival (Linford et al., 2015). In this context, it was discovered that a distinct subpopulation of ECs can convert into follicle stem cells upon starvation (Rust et al., 2020) (Figure 14 A). For this, EC lineages were traced with the result that EC-derived FSC clones occurred only in starved flies but not in the well fed standard conditions. This phenomenon remains poorly understood but prompted the idea that environmental changes such as nutrient availability could induce niche cell plasticity in order to compensate for FSC deficit (Figure 14 B). As cellular fate conversion events have been reported for other stem cell niches as well (Tetteh et al., 2016; Voog et al., 2014), it is possible that transdifferentiation might display a more general theme of stem cell niches and their recovery strategies on stem cell dysfunction or loss.

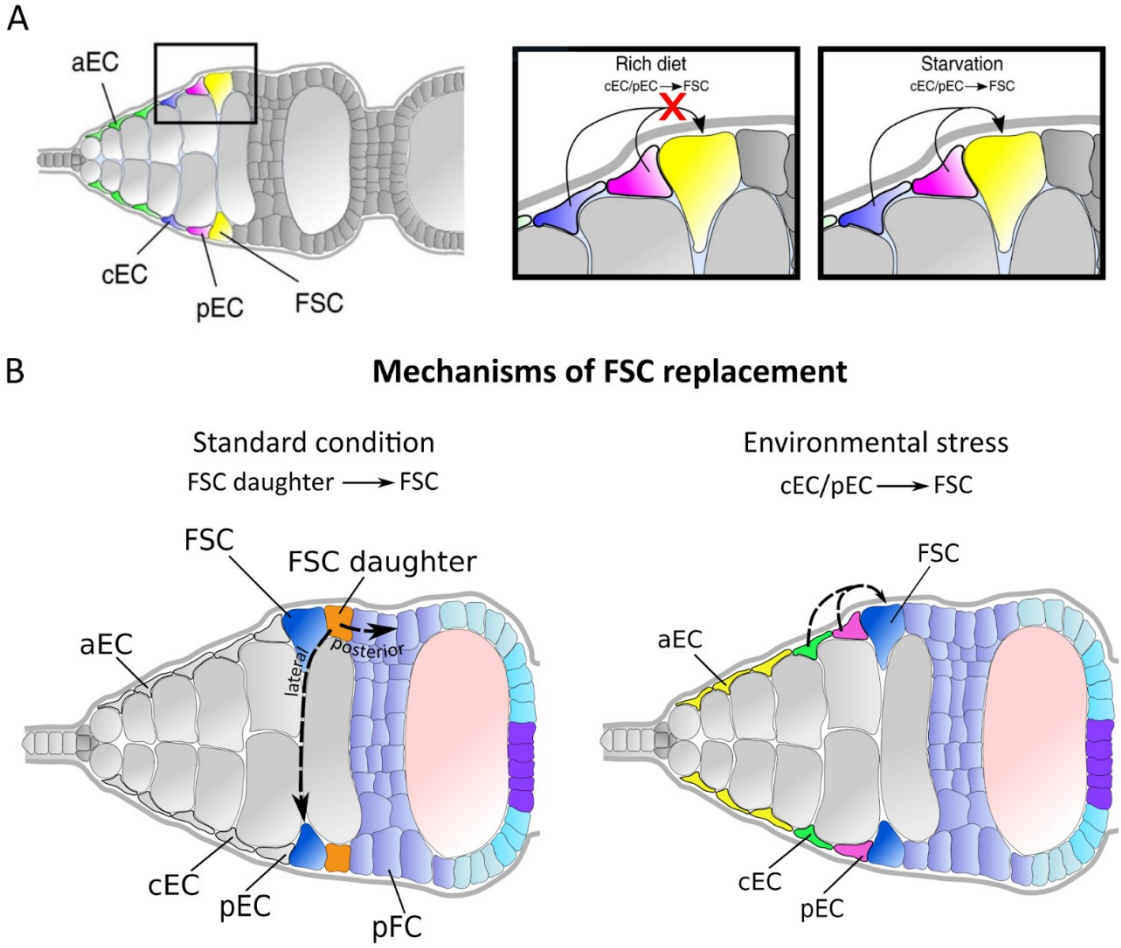


Figure 14: EC fate conversion as mechanism for FSC replacement

(A): Left: Localizations of EC subpopulations in the germarium: anterior EC (green), central EC (blue) and posterior EC (magenta). FSCs (yellow). Right: Central and/or posterior ECs possess the ability to convert to FSCs under starvation, but not under standard rich diet conditions. Figure from (Rust et al., 2020) with permission by the Creative Commons Attribution 4.0 International License (CC BY 4.0). **(B)** Comparison of FSC replacement strategies. Left: FSC daughters are assumed to substitute FSCs by lateral cross-migration under standard conditions. Right: In response to environmental stress, central ECs (green) and posterior ECs (magenta) are supposed to undergo cell fate conversion in order to preserve tissue homeostasis. Modified figures by courtesy of Katja Rust.

1.6.2 Toll signaling

In order to discover which genes control the EC conversion event genetically, Rust *et al.* were able to establish that *Toll* overexpression, among others, leads to a significant increase in the production of FSC clones (Rust et al., 2020). In flies, the Toll pathway is responsible for embryonic development and immune activities (Valanne et al., 2011). The Toll pathway receptor Toll belongs to the nine recorded Toll-like receptors (TLR) encoded in the *Drosophila* genome. For immune regulation processes, the pathway is induced by gram positive bacteria or fungi, leading to an activation of cellular response and the production of antimicrobial peptides (AMP). In *Drosophila* embryonic development, the pathway can be induced independent of bacteria and fungi and assumes a crucial role in dorsal-ventral (DV) patterning of the embryo (Valanne et al., 2011). There are ten described human TLRs residing in the plasma membranes of monocytes and natural killer cells assuming a fundamental role in the innate immune system (Siegmund-Schultze, 2007). Since the discovery of human TLRs, the *Drosophila* Toll pathway serves as a model of this evolutionary conserved signaling cascade (Valanne et al., 2011). The research on *Drosophila* as a model could therefore leverage the application of therapeutic strategies which could particularly target TLRs. The *Drosophila* Toll Pathway is activated by the ligand spätzle (Valanne et al., 2011) (Figure 15). The binding to its receptor requires conformational change of spätzle. The identification of gram positive bacteria or fungi by extracellular recognition factors triggers the proteolysis of the inactive precursor version to the processed mode of spätzle. Now the ligand can bind to its receptor, forming a ligand-receptor complex and provoking a series of cascade signaling (Valanne et al., 2011). Via intracellular TIR domains, the Toll receptor binds to the MyD88 protein. Following, the death domains (DD) of the adaptor protein Tube mediate a complex formation of MyD88, Tube and the kinase Pelle. This leads to the phosphorylation of *Drosophila* IκB factor Cactus. Although it is not finally confirmed, it is believed that Pelle is responsible for the induction of Cactus phosphorylation. In non-signal conditions, Cactus is bound to the NF-κB transcription factors Dorsal and/ or Dif repressing their nuclear localisation and activity (Valanne et al., 2011). Phosphorylation of Cactus triggers its degradation and consequently enables the nuclear translocation of Dorsal/Dif which leads to

the transcription of target genes (AMP). With the production of AMPs the Toll pathway is hereby active in humoral immune response as well as it participates in cellular response by phagocytosis and encapsulation of microbes and killing parasites (Valanne et al., 2011). High levels of AMP, however, are a direct stimulant to the AMP-activated protein kinase (AMPK), depicting an intracellular energy and stress sensor (Wang et al., 2012).

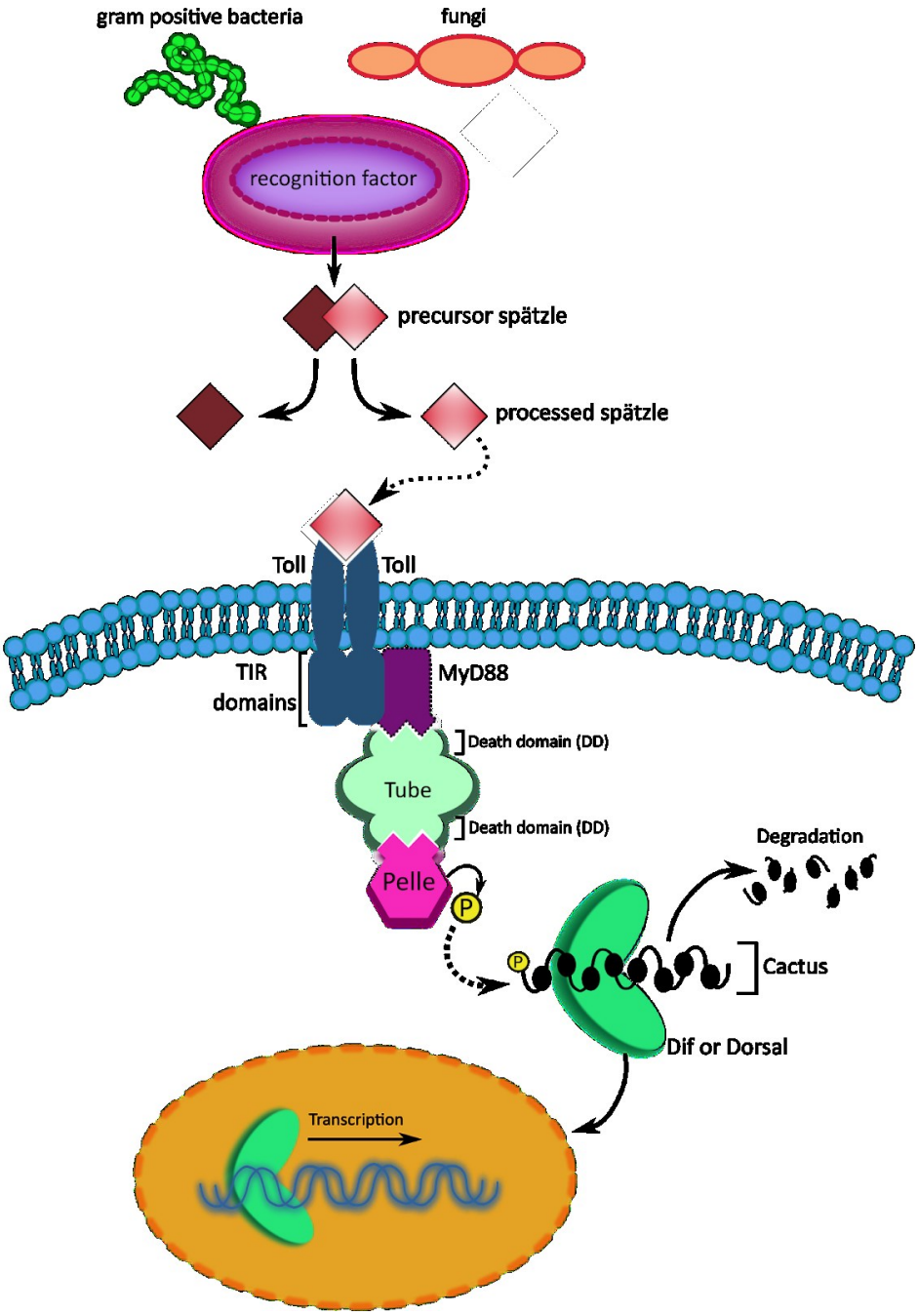


Figure 15: The Toll pathway in *Drosophila melanogaster*

Spätzle represents the special ligand of the Toll receptor. The Toll pathway may be activated via two principles: immune response or in embryonic development. For immune response, recognition factors detect gram positive bacteria and fungi and therefore promote the proteolysis of the precursor version of spätzle. This results in conformational change of spätzle, which enables the processed spätzle version to bind the Toll receptor. The binding triggers the downstream activation of the Toll signaling. The receptor subsequently binds the MyD88 protein via its TIR domain and promotes the further formation of a MyD88, Tube and Pelle complex via death domains (DD). In the following, Cactus becomes phosphorylated and degrades its binding to Dif/Dorsal in non-signaling

conditions. Cactus hereby acts as a negative regulator for Dif/Dorsal functioning. When Cactus inhibition is abolished, Dif/Dorsal are translocated into the nucleus, triggering the transcription of distinct target genes. Figure produced on the basis of (Lindsay & Wasserman, 2014) and (Valanne et al., 2011).

1.6.3 Bisphenol A

Bisphenol A (BPA) is a monomer which is used extensively in the chemical industry given by the fact that the production of BPA rose to 5,5 million metric tons in 2011 (Rochester, 2013). BPA is used as base material for the production of polycarbonate, which then is utilized for the production of many consumer products such as plastics, PVC, thermal receipts, sunglasses – nearly every type of plastic utensil (Rochester, 2013; Wazir & Mokbel, 2019). However, there are several more application fields for BPA: for instance, a major section is used for the manufacture of epoxy resins, which find their practice in the internal coating of canned food and beverages (Wazir & Mokbel, 2019). Moreover, the dental industry incorporates certain variants of BPA in the usage of dental materials such as fillings and sealings (Fleisch et al., 2010). Bis-GMA (bisphenol A-glycidyl methacrylate) and Bis-DMA (bisphenol A-dimethacrylate) are hereby the dominant forms of BPA (Fleisch et al., 2010). Additionally, BPA is found in objects not normally associated with chemicals: contact lenses, children toys and healthcare products (Konieczna et al., 2015; Mustieles et al., 2020; Wazir & Mokbel, 2019). Therefore, the chemical is a widely used substance and exposure to it is nearly inevitable (Wazir & Mokbel, 2019). BPA may be absorbed via different ways: oral, inhalation or transdermal (Konieczna et al., 2015; Mustieles et al., 2020; Rochester, 2013). The most common modality is displayed with the subjection of contaminated products. Whenever BPA is not fully polymerized or exposed to high temperatures causing depolymerization, BPA may leak (Almeida et al., 2018). Consequences of constant exposure may lead to severe pathologies, which firstly may be attributed to the phenolic structure of BPA. BPA hereby displays its character as an endocrine disrupting chemical (EDC), as it is shown to bind the estrogen and endocrine receptors (ER) (Konieczna et al., 2015; Mustieles et al., 2020; Rochester, 2013). Although its binding affinity is considered to be at least 10.000 fold lower than that of estradiol, several studies have reported that BPA still stimulates cellular responses at low doses

(Rochester, 2013; Vandenberg et al., 2007). In this context, it has been established that BPA affects several endpoints of fertility (Konieczna et al., 2015; Mustieles et al., 2020; Rochester, 2013). Its lipophilic structure enables the chemical to cross cell membranes unrestricted, including placental and blood-brain barriers (Balakrishnan et al., 2010; Nishikawa et al., 2010; Sun et al., 2002). Behavioral changes and impacts on neurodevelopment may be observed in *Drosophila* (Nguyen et al., 2021). Likewise, BPA does not only seem to affect hormone and neuronal processes, but is also connected to metabolic disruption. Several investigations in model organisms have reported that BPA exposure leads to the emergence of metabolic dysfunction (Marmugi et al., 2012; Nunez et al., 2001; Williams et al., 2014). Since studies on BPA arose in recent years and unfolded the more problematic nature of BPA, distinct agencies have adapted their criteria for BPA usage: according to the European Food Safety Authority (EFSA), the tolerable daily intake (TDI) of BPA was reduced and its toxicity was re-evaluated. Moreover, the European Commission had issued a law in 2011 to ban the usage of BPA as additive or excipient in baby bottles.

1.7 Scope of this study

Stem cell research on vertebrates is a highly complex endeavor. The complexity of tissue morphologies, the lack of distinct cell markers and reproducibility challenges due to low abundance, impede the usage and research on stem cells *in vivo*. In contrast, the *Drosophila* germarium provides a relatively simple microenvironment to study the behavior of stem cells in their niches (Bastock & St Johnston, 2008). The objective of the present study is to examine environmental impacts on stem cells, niches and their strategies on stem cell replacement by exemplifying the *Drosophila* follicle stem cell (FSC) environment. The ability of niche cell fate conversion under specific conditions in *Drosophila* has been verified in different studies (Rust et al., 2020; Voog et al., 2014). Postulating that the transdifferentiation of escort cells (Ecs), niche cells of the follicle stem cells in the *Drosophila* germarium, is a generalized mechanism of stem cell replacement upon environmental perturbations (Figure 14 B), distinct trigger conditions for this phenomenon ought to be found. Thus, this study involves a candidate screen of different environmental conditions and surveys the rate of EC conversion events in germaria of exposed flies. For the analysis of EC lineages, established fly lines should be validated. Since starvation and the overexpression of *Toll* have been found to affect EC conversion (Rust et al., 2020), the screen includes dietary factors, *Toll* knockdown, mating restraint and the exposure to toxicants. For each condition three replicates with five flies and approximately 450 germaria should be analyzed.

Recent scRNA-Seq techniques present themselves as one of the most auxiliary tools in biomedical research. Given the debate of the number of EC subtypes (Rust et al., 2020; Shi et al., 2021; Slaidina et al., 2021; Tu et al., 2021), this study seeks to identify EC subpopulations via separate analysis of the FCA EC scRNA-Seq dataset. For better resolution, the dataset should be integrated in published datasets on ECs. Consequently identified subpopulations and the expression of subtype specific marker genes should be validated with reporter fly lines. The purpose of identifying EC subtypes and subtype specific marker genes via downstream analysis could enhance further EC research and facilitate investigations on elementary differences among the EC population. This should as well provide benefits for tools and methods of EC tracing.

2. Materials

2.1 Chemicals and solutions

Chemicals were purchased from Sigma-Aldrich and ThermoFisher Scientific.

Substance	Procedure
4% PFA (0,5L)	preheat 300 ml of ddH_2O (temperature needs to be below 60°C) add 20 g of PFA with constant stirring, add 1M NaOH until PFA is dissolved add 50 ml of 10x PBS using 1M HCl, set the pH to 7,0 - 7,2 fill up to 500 ml with ddH_2O
10x PBS (1L)	dissolve NaCl (80 g), Na_2HPO_4 (14,2 g), KH_2PO_4 (2,72 g) and KCl (2 g) in ddH_2O adjust the pH to 7,4 using HCl autoclave the solution store at room temperature
PBST	add 0,2% Triton X-100 to PBS
Block	add 0,5% BSA (Bovine Serum Albumin) to PBST
Fluoromount-G	mount ovaries in Fluoromount-G after staining (see 3.3) store the ovaries in the dark and at 7°C
BPA	weigh the desired amount and grind BPA beads (see 2.4) add and dilute the powder into the fly food or wet yeast

2.2 Antibodies

For the initial steps of immunofluorescence staining, listed primary antibodies were used in this study. DSHB = Developmental Studies Hybridoma Bank.

Antibody	Host	Dilution	Source
Anti-GFP	Rabit	1:1000	Cell Signaling #2956
Anti-RFP	Rat	1:1000	ChromoTek 5F8
Anti-Fas III	Mouse	1:100	DSHB #7G10
Anti-beta Galactosidase	Mouse	1:500	Promega Z378A

For the subsequent immunofluorescence staining procedure, following secondary antibodies were used:

Antibody	Conjugation	Dilution	Source
Goat anti-rabbit	488	1:500	Thermo Fisher Scientific (A-11008)
Goat anti-rat	555	1:1000	Thermo Fisher Scientific (A-21434)
Goat anti-mouse	647	1:1000	Thermo Fisher Scientific (A-21236)
Goat anti-mouse	555	1:1000	Thermo Fisher Scientific (A-21424)

2.3 Fly Stocks

The following fly lines have been used in this study. BL = Bloomington Drosophila Stock Center. VDRC = Vienna Drosophila Resource Center.

Fly stock	Genotype	Description	Reference
Recombinant fly lines			
Gal80^{ts}/ CyO; *fax-Gal4*, G-TRACE/ TM2	P{w[+mC]=tubP-GAL80[ts]}10/CyO;PBac{w[+mC]=IT.GAL4}fax[0122-G4], P{w[+mC]=UAS-RedStinger}6, P{w[+mC]=UAS-FLP.Exel}3, P{w[+mC]=Ubi-p63E(FRT.STOP)Stinger}15 F2/TM2	ubiquitous expression of heat shock sensitive Gal80^{ts} for inhibiting Gal4 driver; EC-specific *fax-Gal4* driver and G-TRACE lineage tracing system	Obtained from Rust Lab Generated from: *tub-Gal80^{ts}* BL#7108 *fax-Gal4* BL#77520 G-TRACE BL#28281
Fly lines for environmental stress conditions			
Gal80^{ts}/ CyO; *fax-Gal4*/ TM6B	P{w[+mC]=tubP-GAL80[ts]}10/CyO;PBac{w[+mC]=IT.GAL4}fax[0122-G4]/TM6B	expression of heat shock sensitive Gal80^{ts} and *fax-Gal4* driver	Obtained from Rust Lab *tub-Gal80^{ts}* BL#7108 *fax-Gal4* BL#77520
G-TRACE	P{w[+mC]=UAS-RedStinger}6, P{w[+mC]=UAS-FLP.Exel}3, P{w[+mC]=Ubi-p63E(FRT.STOP)Stinger}15 F2	G-TRACE lineage tracing system	BL #28281

Toll pathway fly lines			
UAS-dl-RNAi/CyO; G-TRACE/ TM6B	P{y[+t7.7]v[+t1.8]=TRiP.GL00610}attP40/CyO; P{w[+mC]=UAS-RedStinger}6, P{w[+mC]=UAS-FLP.Exel}3, P{w[+mC]=Ubi-p63E(FRT.STOP)Stinger}15F2/TM6B	knockdown of dorsal protein, a component of the Toll pathway, combined with G-TRACE	Obtained from Rust Lab Generated from: *UAS-dl-RNAi* BL#36650 G-TRACE BL#28281
UAS-tl-RNAi/CyO; G-TRACE/ TM6B	P{KK103505}VIE-260B/CyO; P{w[+mC]=UAS-RedStinger}6, P{w[+mC]=UAS-FLP.Exel}3, P{w[+mC]=Ubi-p63E(FRT.STOP)Stinger}15F2/TM6B	knockdown of *Toll* receptor combined with G-TRACE	Obtained from Rust Lab Generated from: *UAS-tl-RNAi* VDRC #100078 G-TRACE BL#28281
GFP marker fly lines			
DOR-GFP	Mi{y[+mDint2]=MIC}DOR[MI06007]/TM3, Sb[1] Ser[1]	expresses GFP tagged *DOR* under the endogenous promoter	BL #42129
Nox-GFP	Mi{y[+mDint2]=MIC}Nox[MI15634]/SM6a	expresses GFP tagged *Nox* under the endogenous promoter	BL #61114

CycA-GFP	TI{GFP[3xP3.cLa]=CRIMIC.GT14}CycA[CR00017]/TM3, Sb[1] Ser[1]	expresses GFP tagged *CycA* under the endogenous promoter	BL #78866
CG42524-GFP	Mi{y[+mDint2]=MIC}CG42524[MI02322]	expresses GFP tagged *CG42524* under the endogenous promoter	BL #34314
CG42524-GFP	Mi{y[+mDint2]=MIC}CG42524[MI02405]	expresses GFP tagged *CG42524* under the endogenous promoter	BL #37321
CG42524-GFP	Mi{y[+mDint2]=MIC}CG42524[MI04404]	expresses GFP tagged *CG42524* under the endogenous promoter	BL #37447
CrebA-GFP	Mi{y[+mDint2]=MIC}CrebA[MI06441]/TM3, Sb[1] Ser[1]	expresses GFP tagged *CrebA* under the endogenous promoter	BL #44167
CG9220-GFP	Mi{y[+mDint2]=MIC}Chsy[MI01619]/FM7h	expresses GFP tagged *CG9220* under the endogenous promoter	BL #33131

Gal4 marker fly lines			
Irc-Gal4	Mi{GFP[E.3xP3]=ET1}Irc[MB11278]	expresses Gal4 under the *Irc* promoter	BL #29191
Nox-Gal4	TI{GFP[3xP3.cLa]=CRIMIC.TG4.0}Nox[CR00693-TG4.0]/SM6a	expresses Gal4 under the *Nox* promoter	BL #78988
DOR-Gal4	Mi{GFP[E.3xP3]=ET1}DOR[MB01323]	expresses Gal4 under the *DOR* promoter	BL #23307
CG9449-Gal4	TI{GFP[3xP3.cLa]=CRIMIC.TG4.0}CG9449[CR01691-TG4.0]	expresses Gal4 under the *CG9449* promoter	BL #91274
CG42524-Gal4	TI{GFP[3xP3.cLa]=CRIMIC.TG4.2}CG42524[CR01445-TG4.2]/SM6a	expresses Gal4 under the *CG42524* promoter	BL #86393
LacZ marker fly lines			
CrebA-lacZ	P{ry[+t7.2]=PZ}CrebA[03576] ry[506]/TM3, P{ry[+t7.2]=ftz-lacC}SC1, ry[RK] Sb[1] Ser[1]	expresses lacZ under the *CrebA* promoter	BL #10183
Reporter fly line			
CD8RFP/CyO; *fax-GFP*/TM6B	P{w[+mC]=UAS-mCD8.mRFP.LG}10b/CyO; P{w[+mC]=PTT-GC}fax[YC0102]/TM6B	expresses membrane bound RFP under the *UAS* promoter and GFP-tagged *fax* under its endogenous promoter	Obtained from Rust Lab Generated from: *UAS-CD8RFP* BL #27399 *fax-GFP* BL #50870

2.4 Recipes for fly work

BPA: molecular weight 228, 29 g/Mol

mM of BPA	BPA per liter	BPA per 200 grams food	BPA per 1 gram food	Final amount
BPA solid fly food				
20	4,5658 g	0,9136 g	0,004568 g	200 g (0,9136 g BPA)
1	0,22829 g	0,045658 g	0,00022829 g	200 g (0,045658 g BPA)
0,1	0,022829 g	0,0045658 g	0,00002283 g	200 g (0,0045658 g BPA)
BPA diluted wet yeast				
20	4,5658 g	0,9136 g	0,004568 g	145 g (0,66236 g BPA)
1	0,22829 g	0,045658 g	0,00022829 g	126 g (0,02876454 g BPA)
0,1	0,022829 g	0,0045658 g	0,00002283 g	136 g (0,00310 g BPA)

Standard fly food medium:

- Dilute Agar Agar (310 g) and cornmeal (2850 g) in 31 L water
- Add 4 L of water diluted with soy flour (390 g) and brewer's yeast (675 g)
- Add malt extract (1800 g) and treacle (1400 g)
- Boil for 10 minutes with constant stirring
- Cool down to 60°C with constant stirring
- Add propionic acid (200 ml) and nipagin solution (10%, 600 ml)

Wet yeast for high fat diet test condtion:

- Weigh 17 g of active dry yeast
- Add 11 ml of liquid coconut oil
- Add 25 ml of water and mix all components until equally blended

Sucrose solution for protein starvation test condition:

- Dilute 13,693 g of sucrose in water until equally blended
- Fill up to 200 ml water

2.5 Microscope and Imaging System

Confocal microscope: Leica TCS SP8 with an HC PL APO CS2 63×/1.4 oil objective
Stereoscopic microscope: Zeiss Stemi 508 with Schott EasyLED lighting

2.6 Applications

FIJI: Schindelin *et al*., 2012
Inkscape: Inkscape Project, Free Software Foundation Inc.
RStudio: RStudio Team (2022), RStudio: Integrated Development Environment for R
LAS-X Leica: Leica Microsystems CMS GmbH
Word / Excel / Powerpoint: Microsoft Office Package 2013

3. Methods

3.1 Bioinformatic Analysis

3.1.1 Analysis of a single cell RNA-sequencing dataset

All single cell sequencing analyses were performed in RStudio 1.4.1717 using Seurat v3 (Stuart et al., 2019). I followed the standard workflow of the Seurat algorithm (Seurat Vignette “Seurat - Guided Clustering Tutorial”; Satija et al., 2015; Stuart et al., 2019). Pre-processing steps comprised the filtration of low quality cells based on QC metrics such as mitochondrial expression (percent.mt), the number of genes (nFeature_RNA) and the number of UMIs (nCount_RNA) using the subset() command. For the help of filtration, QC metrics were visualized with the usage of VlnPlot() and FeatureScatter(). After filtration, the data was normalized by default settings using NormalizeData() and variable features across the dataset were computed via the FindVariableFeatures() function. For principle component analysis (PCA = dimensional reduction technique), ScaleData() and RunPCA() commands were utilized. PCA results were visualized with print(), VizDimLoadings(), DimPlot() and DimHeatmap(). To identify and select significant PCs included in further analysis, JackStraw(), ScoreJackstraw(), JackStrawPlot() and ElbowPlot() were used. Cell clustering was performed with the aid of FindNeighbors(), FindClusters() and adjusting the resolution factor to achieve appropriate clustering. Cluster visualization was implemented with RunUMAP() and, if required, clusters were merged with WhichCells() and SetIdent(). Cluster specific markers were identified by running a Wilcoxon test with the default settings of the FindAllMarkers() command. Top 10 markers were reported with %>%() of the dplyr package (Hadley Wickham, Romain François, Lionel Henry, 2022). For the illustration of marker expression, DotPlot(), FeaturePlot() and DoHeatmap() were utilized. In order to annotate cell identities, CellSelector(), SetIdent(), append(), subset() commands were used and depicted with FeaturePlot() and TSNEPlot(). Reference mapping of a reference and a query dataset was performed using FindTransferAnchors() and MapQuery().

3.1.2 Dataset integration

For dataset integration, RStudio 1.4.1717 and the standard guideline for scRNA-Seq integration by Seurat v4 (Hao et al., 2021) was used (Seurat Vignette “Introduction to scRNA-seq integration”; Stuart et al., 2019). Prior to integration, individual datasets were separately analyzed according to the workflow described in 4.1.1. Subsequently, desired datasets were merged together to a single object by using the merge() command. With SplitObject() the merged file was split into a dataset list of individual Seurat objects. Therefore, normalization via NormalizeData() and the computation of variable features with FindVariableFeatures() was performed on each dataset independently. Variable features across all datasets were identified with SelectIntegrationFeatures() in order to define connecting anchors between the datasets with FindIntegrationAnchors(). Following, a single integrated data assay was created with IntegrateData() for a conjoint downstream analysis of all cells. The modified und unprocessed versions were specified by DefaultAssay(). According to standard guidelines for dimensional reduction, default commands via ScaleData() and RunPCA() were performed. Cell clustering was implemented by the functions FindNeighbors() and FindClusters(). The final resolution factor was adjusted until well defined clusters were obtained. The individual datasets were assigned to their corresponding sources by means of append(), WhichCells() and subset(). Cluster visualization was illustrated with DimPlot() and UMAPPlot() and cluster specific markers were calculated with FindAllMarkers(). The top 10 markers for each cluster were specified with the dplyr package function %>%() and depicted with DotPlot() and FeaturePlot(). The subset() command was utilized to extract specific cells out of the merged object. Certain clusters were merged together by WhichCells() and SetIdent(). Specific clusters were highlighted with cells.highlight = WhichCells().

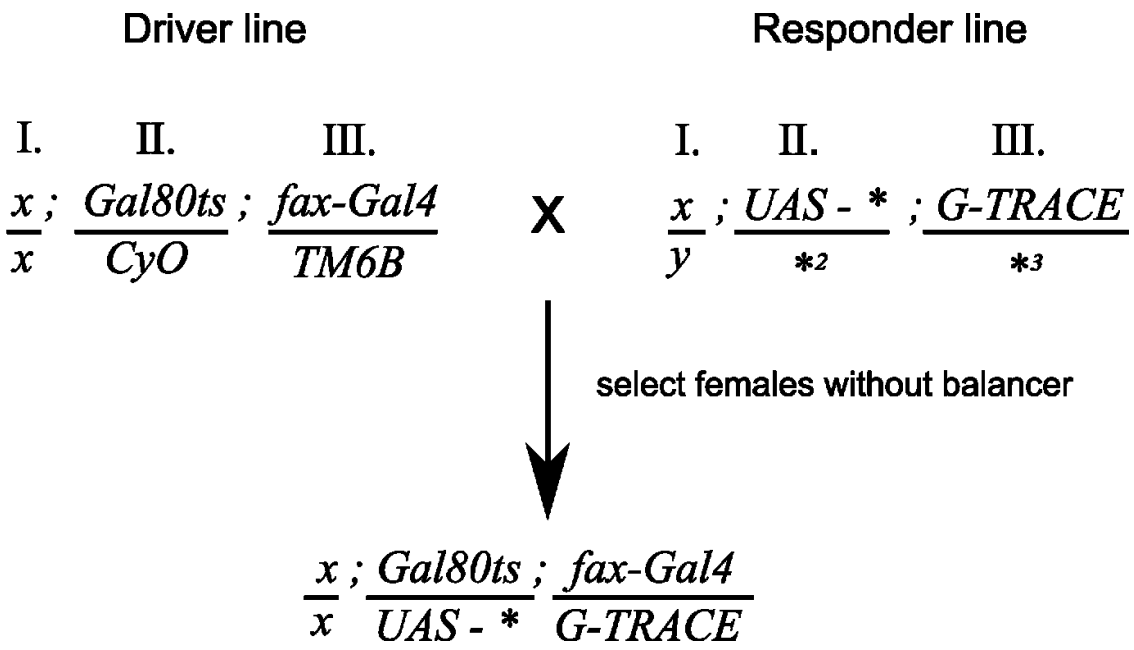
3.2 Fly husbandry

For all experiments comprising Gal80ts, flies were bred according to the breeding scheme (Figure 16) and reared at 18°C to assure the inactivity of the G-TRACE system during the development of offspring flies. Gal80ts controls were conducted by feeding wet yeast to newly hatched flies for a duration of 3 days at

restrictive temperatures of 18°C until dissection. For all fly lines containing G-TRACE, the adult F1 generations were shifted to 29°C to induce G-TRACE post eclosion. Only female flies without the characteristics of CyO, TM6B or TM2 were dissected and stained. For G-TRACE controls, examination after 7 or 14 days was implemented, feeding flies wet yeast for the desired time interval until dissection. All other genotypes were reared under standard lab conditions at 25°C and fed wet yeast daily for 3 consecutive days prior to dissection. Flies of the desired genotype (Gal80ts; *fax-Gal4*, G-TRACE) underwent a set of various stress conditions within time intervals of 14 days:

Condition	**Procedure**
Mating restraint	Virgin female offspring flies were collected, maintained separately from male flies and fed wet yeast daily for 14 days until dissection
Water starvation	7 days of rich diet (feeding wet yeast daily) followed by 24 hours of starvation (shifting flies to empty vials containing a KimWipe soaked with water) and then transferring flies back towards a rich diet for 6 more days until dissection
Protein starvation	7 days of rich diet followed by 3 days of protein starvation (shifting flies to empty vials containing a KimWipe soaked with 200mM sucrose solution) and then transferring flies back towards a rich diet for 4 more days until dissection
High fat diet	Flies were kept in empty vials containing a KimWipe soaked with water and fed wet yeast (blended with 20% of liquid coconut oil) daily for 14 days until dissection

Bisphenol-A (BPA) exposure	Flies underwent BPA exposure for 14 consecutive days and were exposed to food (solid fly food and wet yeast) containing different concentrations of BPA (0,1mM, 1mM, 20mM). For the concentrations given, I used values from known literature regarding BPA exposure to *Drosophila* as a guideline (Begum et al., 2020; Kaur et al., 2015; Nguyen et al., 2021). Vials were previously prepared with solid food containing either 0,1mM, 1mM or 20mM of Bisphenol-A. Additionally, flies were fed daily with wet yeast containing the same diluted concentrations of Bisphenol-A.



Symbols	Meaning
X	X chromosome
Y	Y chromosome
Gal80ts	Temperature sensitive Gal80
fax-Gal4	Gal4 construct under *fax* promoter
UAS - *	Gal4 dependant transgene under the control of *UAS* promoter
*2	Either balancer or wild type gene
*3	Either G-TRACE or wild type gene
CyO	Balancer chromosome
TM6B	Balancer chromosome

Figure 16: Breeding scheme of fly lines used in this study

Presentation of the breeding procedure for experiments conducted in this study. Symbols are described in the table.

3.3 Immunofluorescence staining

Flies were dissected in PBS at room temperature (RT) within a timespan of maximum 1 hour. Ovaries were kept on ice whenever the dissection procedure involved more time. The ovaries were fixed for 15 minutes with 4% PFA at RT while swaying. The fixative was removed and the ovaries were washed 3 times with PBS and afterwards blocked for 1 hour in blocking solution on a tube rocker at RT. Once the block was removed, the primary antibodies were diluted in 0,5 ml blocking solution per sample and the incubation was performed overnight at 4°C on a tube rocker. The next day, ovaries were rinsed with 0,5 - 1ml 3 times for 15 minutes using blocking solution and leaving the ovaries on the tube rocker at RT. The secondary antibodies were diluted in 0,5 ml blocking solution per sample and incubated for 4 hours on the tube rocker at RT. To prevent exposure to light, the ovaries were covered during the incubation procedure. As soon as the incubation process ended, the ovaries were washed 3 times á 10 minutes with Block and twice with PBS for 10 minutes at RT. After removing the washing solution, ovaries were mounted in DAPI Fluoromount-G (Thermo Fisher Scientific, OBO10200) and analyzed with confocal microscopy.

3.4 Image Analysis

All images were acquired with a Leica Sp8 confocal microscope. The settings for recording images such as laser intensity, width of wavelengths and resolution were kept constant across screening of different experiments. Each image was acquired by using the 63x objective and clones were identified by taking Fas III positive cells as reference. Any image processing such as brightness, contrast, cropping or rotation was performed via FIJI and figures were created with Inkscape.

3.5 Statistics and Reproducibility

All bioinformatic analysis of data was performed in RStudio. The generation of graphs was prepared in Microsoft Excel (Microsoft office package 2013). I performed at least three independent biological and technical replicates and conducted paired controls for each experimental condition. A two-sided student's T-test was used for the calculation of p-values.

4. Results

4.1 Single cell sequencing analysis

4.1.1 Cluster specific markers defined by the analysis of EC enriched FCA dataset

The Fly Cell Atlas (FCA) represents a dataset, incorporating a vast majority of tissues of *Drosophila melanogaster* and exhibits approximately 250 already annotated cell types (H. Li et al., 2022). As an incorporation of the FCA, the FCA EC dataset was created for the purpose of gaining detailed insight in the EC nature by enriching mainly cells of the EC identities in the dataset. For this, the EC specific Gal4 driver *Wnt4-Gal4* was crossed to the reporter *UAS-unc84:: GFP* in order to mark ECs within the tissue. Ovaries expressing unc84::GFP were dissected and older developing stages were removed. In order to enrich the dataset with ECs, GFP positive nuclei were sorted with fluorescence activated cell sorting (FACS) and sequenced via SMART-Seq (H. Li et al., 2022). I performed downstream analysis on the raw FCA EC dataset (see 3.1.1 Analysis of a single cell RNA-sequencing dataset) following the workflow of the Seurat algorithm (Seurat Vignette “Seurat - Guided Clustering Tutorial”; Satija et al., 2015; Stuart et al., 2019). The aim hereby was to identify potential EC subtypes and detect subtype specific marker genes, which consequently may be used for investigating EC subtypes in the germarium. Starting off with a total amount of 900 cells incorporated in the dataset, the raw data was filtered by features depicting low quality cells: the height of mitochondrial gene expression (high mitochondrial contamination is correlated with dying cells), the quantity of detected genes and total reads per cell. After these filtering steps, I obtained a file containing 713 high quality cells. In the following, I computed features, which were variable across all cells in the dataset. This serves to identify significant principal components, depicting the basis for cell clustering. Cell clusters in this context represent spatially contiguous cell groups, which share transcriptional similarities. In the following, the visualization of cell clustering is depicted using UMAP (Uniform manifold approximation and projection). In the UMAP each cell is represented as a point. This algorithm works as a dimensional reduction system and for two

dimensional visualization, cells with similar characteristics are depicted closer to each other thus cells further afar of each other are more likely to be different. Cell clustering sets the opportunity to detect subtypes within a cell population and UMAP contributes to its visualization. For the FCA EC dataset, I sought to generate well defined cell clusters, where clusters are clearly separated from each other and each cluster would portray a distinct cell type. According to prior studies (Rust et al., 2020; Shi et al., 2021; Slaidina et al., 2021; Tu et al., 2021), I expected the generation of at least three distinct EC clusters. I adjusted the resolution factor for clustering and determined the final factor of 0,5, concluding five ultimate cell clusters (Figure 17 A). I assumed that these clusters portrayed EC subpopulations and calculated cluster specific marker genes. These marker genes are calculated via differential expression and define a particular cluster by comparing it to all other clusters. Hence, good quality marker genes show high expression rates with respect to the expression strength and the amount of cells expressing the gene in the particular cluster. The visualization of the expression rate of the respective marker genes was performed by using FeaturePlots and DotPlots. FeaturePlots map the expression of marker genes on the UMAP of the clustered version of the dataset. The expression of each marker gene is represented by a color gradient, with blue indicating high expression and gray indicating low expression. Optimally, cluster-specific marker genes on the FeaturePlot should show an expression limited to the respective cluster. DotPlots show graphs incorporating cluster identity on the y axis and marker genes on the x axis. The visualization of expression is depicted in the DotPlot by two parameters: the size of the dot indicates the amount of cells in the cluster that express the particular marker gene and the color of the dot shows the expression strength (blue = high expression, gray = low expression). Strong cluster specific marker genes demonstrate large and dark blue dots and their expression is mainly limited to their respective clusters. Markers for the five clusters of the FCA EC dataset were visualized via DotPlot incorporating the top 10 of calculated genes (Figure 17 B). Marker genes for cluster 3 and 4 meet the criteria for good quality markers as cluster specificity and strong expressions (big and blue dot) are given. However, the identification of good marker genes for cluster 0, 1 and 2 failed: for most of the markers of cluster 0 only 25% of cells express the particular gene (small dot); marker genes for clusters 1 and 2, on the other hand,

show a higher expression in relation to the cells expressing the particular gene (big dots), but their expression is not limited to their respective cluster.

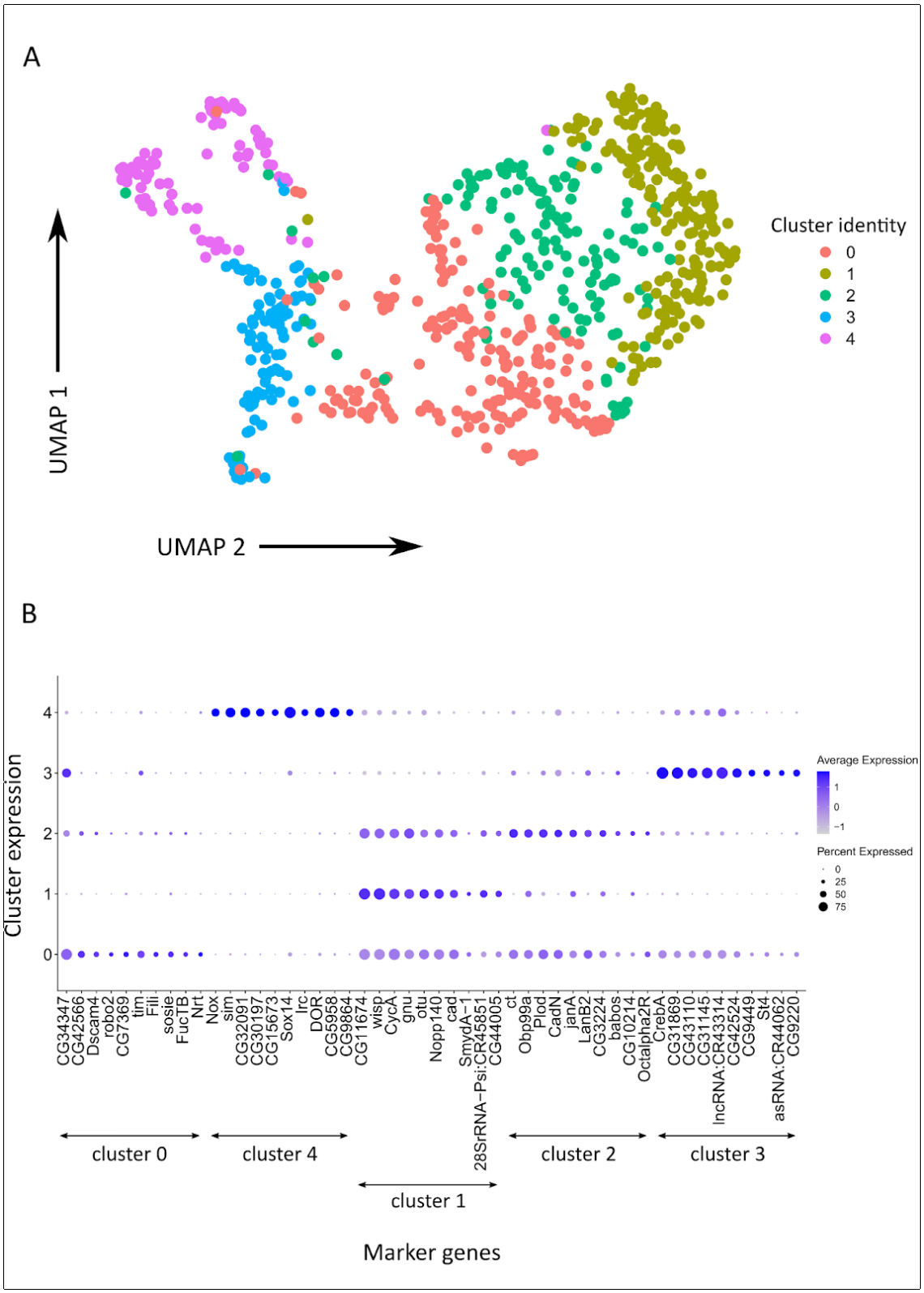


Figure 17: Initial clustering of the FCA EC dataset produces five cell clusters
(A) UMAP dimensionality reduction of FCA EC clustering. Each point in the UMAP represents a cell. Similarities of the cell transcriptomes are represented in the UMAP by smaller distances between the cells. The more different the cells the further the distance between them. This results in the formation of certain clusters, which are distinguished from each other by different colors. Initial clustering for the FCA EC dataset resulted in five clusters (0, 1, 2, 3, 4). **(B)** DotPlot visualization of top 10 calculated cluster specific markers. The size of the dot indicates the percentage of cells in a cluster that express the particular gene. The color indicates the average expression strength (blue = high expression, gray = low expression). Calculated markers for clusters 3 and 4 show high expression strength and hardly any expression in others from their own cluster. Note that marker genes calculated for clusters 0,1 and 2 in comparison to cluster 4 and 3 depict non stable expression patterns, as expressions are either equally present among other clusters or depict average expression rates of rather low levels.

I therefore reasoned that prior clustering needed to be adjusted in order to generate transcriptionally distinct clusters for the entire dataset. Since clusters 0, 1, and 2 coexpressed marker genes (Figure 17 B), these clusters were merged together into one single subset, creating cluster 5 and leaving the dataset with a definite arrangement of three clusters (Figure 18 A). For cluster 5 the process of marker calculation was repeated. The top 10 marker genes of the three final clusters are visualized using DotPlot (Figure 18 B). Noticeably, the recalculated marker genes of cluster 5 now demonstrate high expression rates (big and blue dots) and cluster specificity. For further analysis, marker genes of each cluster were sorted by evaluating their expressions in DotPlots and FeaturePlots. Cluster specific marker genes with the most specific expression patterns were selected: *Irc*, *Nox* and *DOR* representing cluster 4, *CrebA*, *CG9449*, *CG9220* and *CG42524* for cluster 3 and *gnu*, *mtrm* and *stet* for cluster 5 (Figure 19). Table 1 gives an overview of the genetic function of the selected marker genes of each cluster (Table 1).

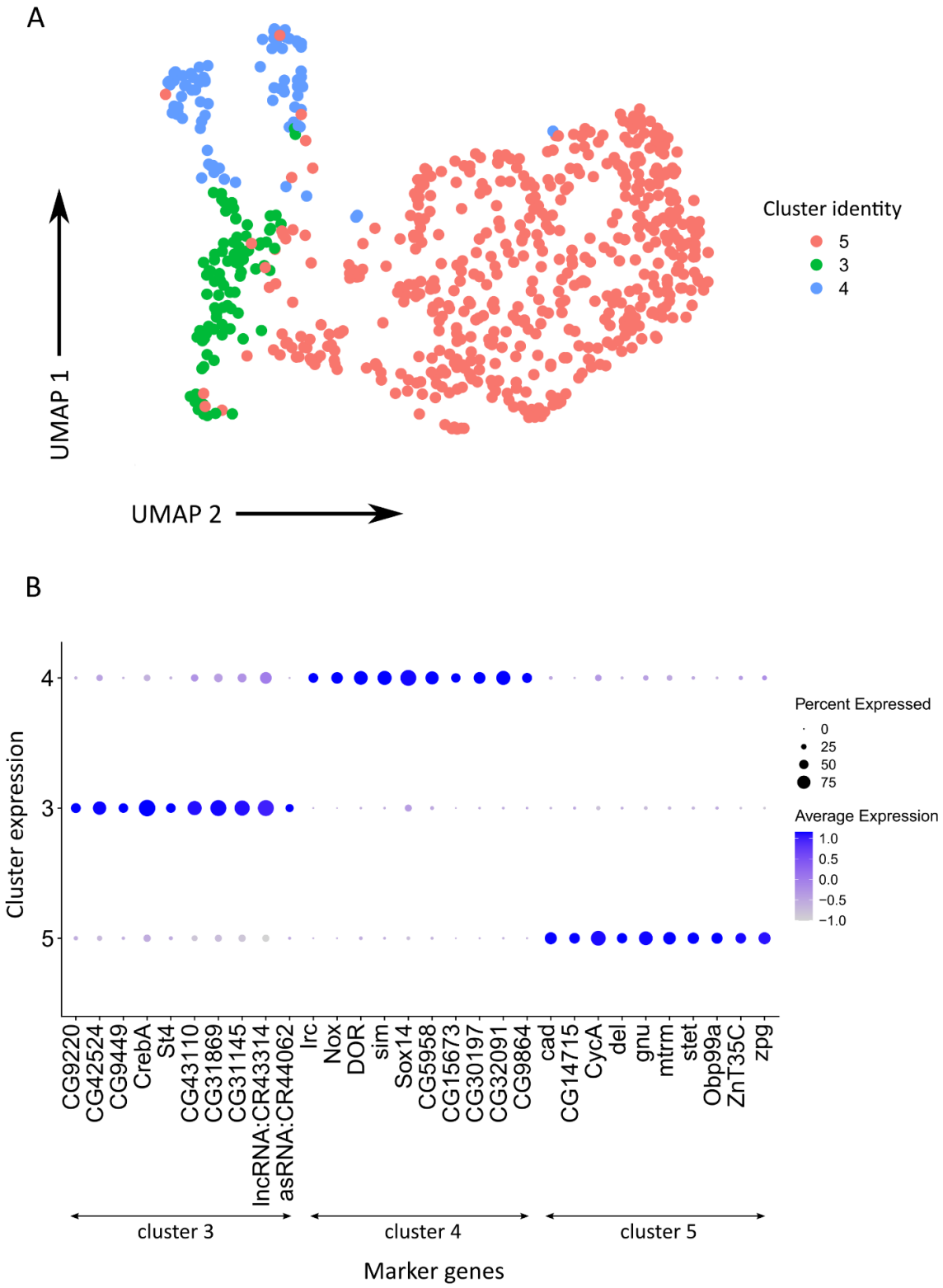

A
UMAP 1
UMAP 2
Cluster identity
5
3
4
B
Cluster expression
4
3
5
Percent Expressed
0
25
50
75
Average Expression
1.0
0.5
0.0
−0.5
−1.0
CG9220
CG42524
CG9449
CrebA
St4
CG43110
CG31869
CG31145
lncRNA:CR43314
asRNA:CR44062
Irc
Nox
DOR
sim
Sox14
CG5958
CG15673
CG30197
CG32091
CG9864
cad
CG14715
CycA
del
gnu
mtrm
stet
Obp99a
ZnT35C
zpg
cluster 3
cluster 4
cluster 5
Marker genes

Figure 18: Optimization of FCA EC clustering yields 3 final clusters

(A) Due to overlapping marker gene expression, prior clusters 0,1 and 2 were merged together to a single cell cluster, referred to as cluster 5. Cluster 5 is shown in red. **(B)** DotPlot visualization of top 10 calculated marker genes for each final cluster. Expression strength is depicted by a color gradient (blue for high expression; gray for low expression). The representation of the amount of cells in a cluster expressing the particular gene (in percentages) is given by the size of the dot. The x axis contains the calculated marker genes while the y axis reflects the expression in the respective clusters. All clusters show cluster specific expression of the marker genes. Note that marker genes of cluster 5 deliver a more cluster limited outcome, compared to prior acquired genes of the individual clusters 0, 1 and 2 (Figure 17 B).

Table 1: Overview of the genetic background of selected markers of the FCA EC dataset

Genetic information for each selected marker gene of the respective FCA EC cluster as summarized on FlyBase.

Marker gene	Cluster	Genetic background
Irc	4	response to oxidative stress
Nox	4	positive regulator for calcium-mediated signaling
DOR	4	regulates autophagosome formation and protein degradation via ecdysone signaling
CrebA	3	transcription factor
CG9449	3	phosphatase activity
CG9220	3	chondroitin sulfate biosynthetic process
CG42524	3	undescribed
gnu	5	protein kinase activator activity
mtrm	5	regulation of meiotic cell cycle
stet	5	activator to ligands of the product of Egfr

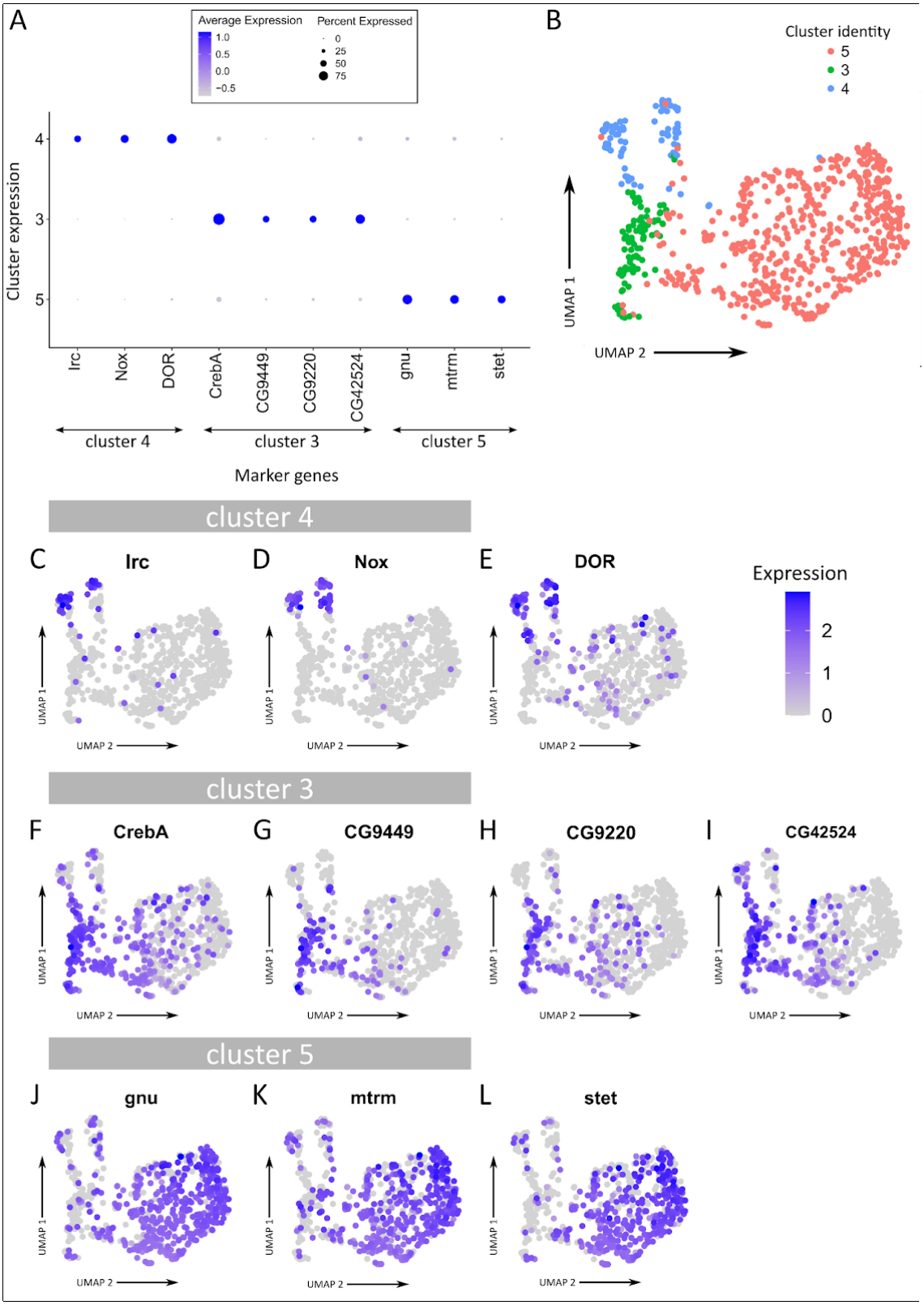

A
Average Expression
1.0
0.5
0.0
−0.5
Percent Expressed
0
25
50
75
Cluster expression
4
3
5
Irc
Nox
DOR
CrebA
CG9449
CG9220
CG42524
gnu
mtrm
stet
cluster 4
cluster 3
cluster 5
Marker genes
B
Cluster identity
5
3
4
UMAP 1
UMAP 2
cluster 4
C
Irc
D
Nox
E
DOR
Expression
2
1
0
cluster 3
F
CrebA
G
CG9449
H
CG9220
I
CG42524
cluster 5
J
gnu
K
mtrm
L
stet

Figure 19: Selection of specific marker genes for each cluster of the FCA EC dataset by DotPlots and FeaturePlots

(A) Visualization of the expression of selected cluster specific markers in the respective clusters of the FCA EC dataset using DotPlot. The expression rate of cells expressing the gene is represented by the size of the dot. The expression strength is indicated by the color of the dot (blue = high expression, gray = low expression). **(B)** UMAP representation of final FCA EC clustering with 3 definite clusters. UMAP plots the cells in 2D in dependence of their transcriptome similarity. Each dot in the UMAP represents a cell. Cells with genetic similarities are shown closer together on the UMAP, resulting in the formation of clusters (highlighted by different colors). **(C-L)** FeaturePlots of the selected cluster specific markers of the three final clusters in the FCA EC dataset. FeaturePlots plot the respective marker on the UMAP of the dataset. Each dot in the FeaturePlot represents a cell. The expression strength of the respective marker is depicted by a color gradient (blue = high expression, gray = low expression).

With the aspiration to assign each cluster of the FCA EC dataset a cell identity, I compared the outcomes of the FCA EC analysis to published scRNA-Seq studies on ECs. First, I made use of the published marker genes for EC subtypes of Tu *et al.* (Tu et al., 2021), Shi *et al.* (Shi et al., 2021) and Rust *et al.* (Rust et al., 2020) (Figure 20 A, B and C). I surveyed the respective marker genes in the FCA EC dataset. This expression is visualized via DotPlot (Figure 20 D). Interestingly, most of the marker genes displayed little expressions in all of the clusters of the FCA EC dataset. However, several published marker genes of Rust and Shi *et al.* did show cluster specific expression in the FCA EC dataset: *pcs* (Shi *et al.*, pEC) and *Fer2LCH* (Rust *et al.*, cEC) for cluster 4, *Pdk1* (Rust *et al.*, aEC), *dally* (Rust *et al.*, pEC), *corto* (Rust *et al.*, pEC) and *mamo* (Rust *et al.*, pEC) for cluster 3 and *Idgf6* (Rust *et al.*, aEC), *bic* (Rust *et al.*, pEC), *msk* (Rust *et al.*, pEC) and *CG15093* (Shi *et al.*, aEC) mainly for cluster 5 (Figure 20 D). I plotted these marker genes on the FCA EC dataset via FeaturePlot with the expectation to use the expression to identify cluster identities for the FCA EC clusters. Notably, the published marker genes for EC subtypes were not confined to clusters within the FCA EC dataset (Figure 21 C-L). Hence, this analysis did not yield in the identification of cluster identity of the FCA EC dataset. Sequencing depth of single cell datasets is relatively low and may have obstructed this analysis. Thus, I

tested whether marker genes of the FCA EC dataset were expressed in EC subtypes in published datasets: *Irc*, *Nox, DOR* (cluster 4 of FCA EC)*, CrebA*, *CG9449*, *CG9220, CG42524* (cluster 3 of FCA EC)*, gnu*, *mtrm* and *stet* (cluster 5 of FCA EC) were plotted in the datasets of Shi *et al.*, Tu *et al.* and Rust *et al.* (Figure 22). Rust *et al.* provided an already clustered version of its data incorporating cell type annotations of each cluster (Rust et al., 2020). Note that cluster identities were not published for Tu and Shi *et al.* (Shi et al., 2021; Tu et al., 2021). In order to compare the FCA EC dataset with Tu *et al.* and Shi *et al.*, I performed analysis of the raw datasets using the standard workflow of Seurat (Seurat Vignette "Seurat - Guided Clustering Tutorial"; Satija et al., 2015; Stuart et al., 2019, see 3.1.1 Analysis of a single cell RNA-sequencing dataset). Further, I mapped published marker genes of each study in their own dataset for the identification of cluster identities. Since Tu *et al.* established the existence of four EC subtypes (Tu et al., 2021), I sought to reproduce this in the dataset. Marker gene expression (EC subtype markers see Figure 20 C) for Tu *et al.* however delivered the identification of three EC identities (aEC, cEC and pEC) and one FC compartment. Shi *et al.* divides the EC population in two compartments (aEC and pEC) (Shi et al., 2021). Marker overlap in the dataset (EC subtype markers see Figure 20 A) results in the annotation of three clusters: aEC, pEC and FC. Subsequently, the expression of selected markers of the FCA EC clusters was explored in the annotated datasets of Tu, Shi and Rust *et al.* using DotPlot (Figure 22). Indicated by the DotPlots, marker genes of the clusters of the FCA EC dataset did not show defined expression patterns in the EC subtype clusters of Tu, Shi and Rust *et al.* (Figure 22). All marker genes except for *Irc* (cEC in Rust *et al.*) show little expression in the EC identities of the Rust *et al.* dataset. However, *Irc* is not expressed in the dataset of Shi *et al.* but is detected in the FC compartment in Tu *et al*. This trend is applicable to any marker gene in the DotPlots. In addition, cluster 5 markers were surprisingly highly expressed in germ cell populations of Rust *et al.* (Figure 22 C). As these findings are not indicative for the final annotation of cell identities in the FCA EC clusters, it rather raises the question whether the FCA EC dataset contains cells from identities other than ECs.

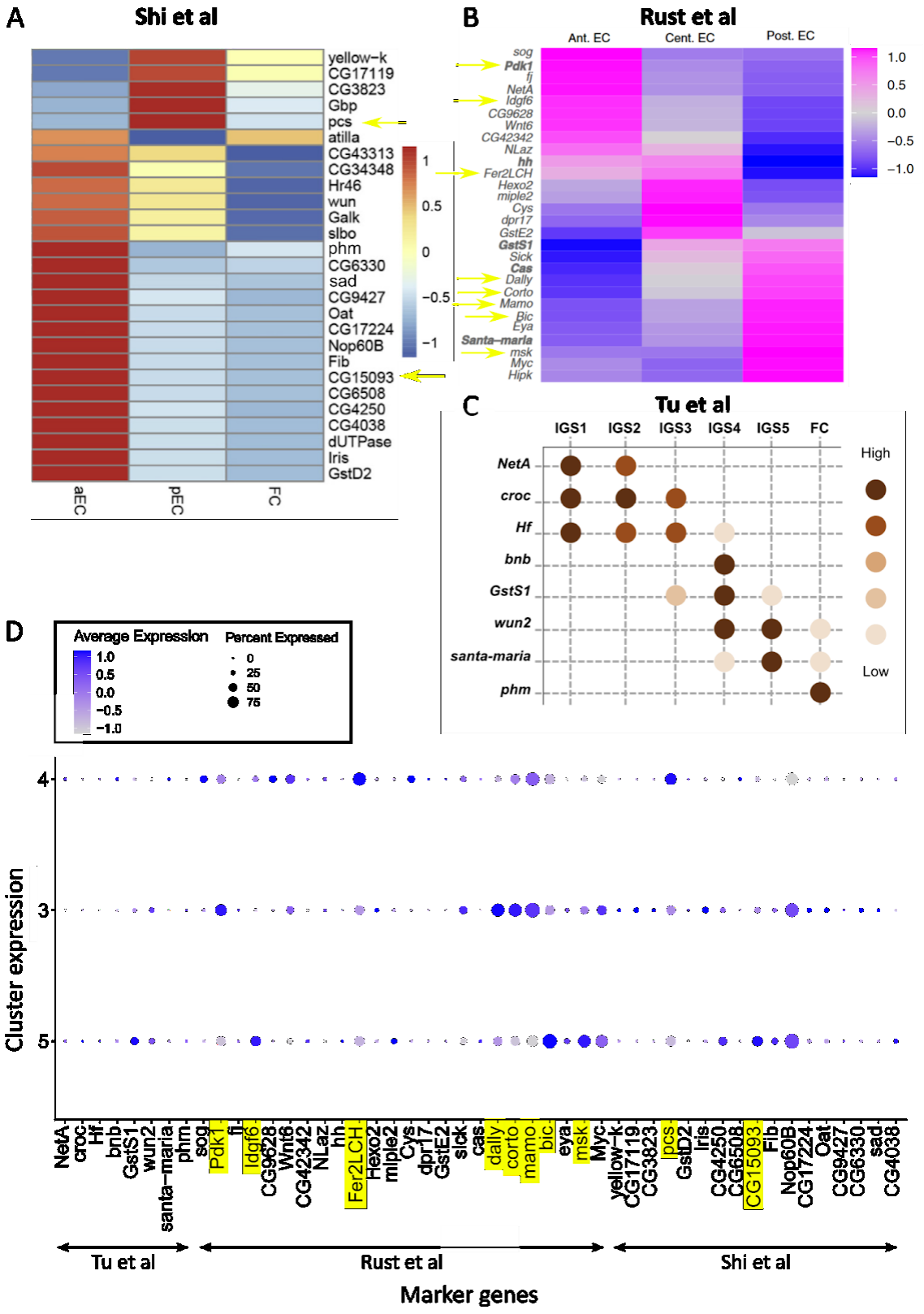
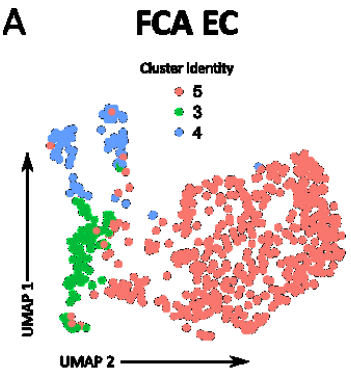

A
Shi et al
yellow−k
CG17119
CG3823
Gbp
pcs
atilla
CG43313
CG34348
Hr46
wun
Galk
slbo
phm
CG6330
sad
CG9427
Oat
CG17224
Nop60B
Fib
CG15093
CG6508
CG4250
CG4038
dUTPase
Iris
GstD2
aEC
pEC
FC
1
0.5
0
−0.5
−1
B
Rust et al
Ant. EC
Cent. EC
Post. EC
sog
Pdk1
fj
NetA
Idgf6
CG9628
Wnt6
CG42342
NLaz
hh
Fer2LCH
Hexo2
mlple2
Cys
dpr17
GstE2
GstS1
Sick
Cas
Dally
Corto
Mamo
Bic
Eya
Santa−maria
msk
Myc
Hipk
1.0
0.5
0.0
−0.5
−1.0
C
Tu et al
IGS1
IGS2
IGS3
IGS4
IGS5
FC
NetA
croc
Hf
bnb
GstS1
wun2
santa-maria
phm
High
Low
D
Average Expression
1.0
0.5
0.0
−0.5
−1.0
Percent Expressed
0
25
50
75
Cluster expression
4
3
5
NetA
croc
Hf
bnb
GstS1
wun2
santa−maria
phm
sog
Pdk1
fj
Idgf6
CG9628
Wnt6
CG42342
NLaz
hh
Fer2LCH
Hexo2
mlple2
Cys
dpr17
GstE2
sick
cas
dally
corto
mamo
bic
eya
msk
Myc
yellow−k
CG17119
CG3823
pcs
GstD2
Iris
CG4250
CG6508
CG15093
Fib
Nop60B
CG17224
Oat
CG9427
CG6330
sad
CG4038
Tu et al
Rust et al
Shi et al
Marker genes

Figure 20: Published EC marker genes by Rust, Shi and Tu *et al.* do not contribute to the assignment of cluster identities of the FCA EC dataset

(A) Published marker genes of Shi *et al.* (Shi et al., 2021) for anterior and posterior ECs (aEC/pEC). The expression strength is depicted by the color (red = high expression, blue= low expression). Yellow arrows point at marker genes which were selected via DotPlot. Modified and reprinted from Current Biology, Vol. 31, Shi et al., A Progressive Somatic Cell Niche Regulates Germline Cyst Differentiation in the Drosophila Ovary, 18 Pages, 2021, with permission from Elsevier. **(B)** EC marker genes of Rust *et al.* (Rust et al., 2020) for anterior, central and posterior ECs. Magenta highlighted bars indicate a high level of expression. Yellow arrows point at selected marker genes. Modified figure from (Rust et al., 2020) with permission by the Creative Commons Attribution 4.0 International License (CC BY 4.0). **(C)** Marker genes released by Tu *et al.* (Tu et al., 2021) for four EC identities (IGS1-4) and the FC compartment (IGS5+FC). Note that some marker genes are assigned to different cell identities in each of the three studies mentioned above: *phm* in Shi *et al.* for anterior ECs and in Tu *et al.* for FCs; *santa-maria* in Rust *et al.* for posterior ECs and in Tu *et al.* for FCs. Figure from (Tu et al., 2021) with permission by the Creative Commons Attribution 4.0 International License (CC BY 4.0). **(D)** Depiction of the expression of published EC marker genes of Rust, Shi and Tu *et al.* in the FCA EC dataset using DotPlot. The x axis contains the respective marker genes of each study. Same marker genes were listed in only one of the three studies. The y axis represents the cluster identities of the FCA EC dataset. Expression in the DotPlot is visualized by two parameters: the size of the dot indicates the percentage of cells in the cluster that express the particular gene; the color of the dot refers to the expression strength - blue for high expression levels and gray for low expression levels. Yellow highlighted marker genes show cluster specificity for FCA EC clusters: high expressions for cluster 4 are identifiable for *pcs* (Shi *et al.*, pEC) and *Fer2LCH* (Rust *et al.*, cEC); cluster 3 specific expression is represented by *Pdk1* (Rust *et al*., aEC), *dally* (Rust *et al*., pEC), *corto* (Rust *et al*., pEC), *mamo* (Rust *et al*., pEC); *Idgf6* (Rust *et al*., aEC), *bic* (Rust *et al*., pEC), *msk* (Rust *et al*., pEC) and *CG15093* (Shi *et al*., aEC) are mostly limited to the expression in cluster 5.

A



B

Marker	Source	DotPlot FCA EC
pcs	Shi et al, 2021	cluster 4
CG15093	Shi et al, 2021	cluster 5
Pdk1	Rust et al, 2020	cluster 3
mamo	Rust et al, 2020	cluster 3
dally	Rust et al, 2020	cluster 3
corto	Rust et al, 2020	cluster 3
Fer2LCH	Rust et al, 2020	cluster 4
bic	Rust et al, 2020	cluster 5
msk	Rust et al, 2020	cluster 5
Idgf6	Rust et al, 2020	cluster 5

Expression in cluster 3

Expression in cluster 4

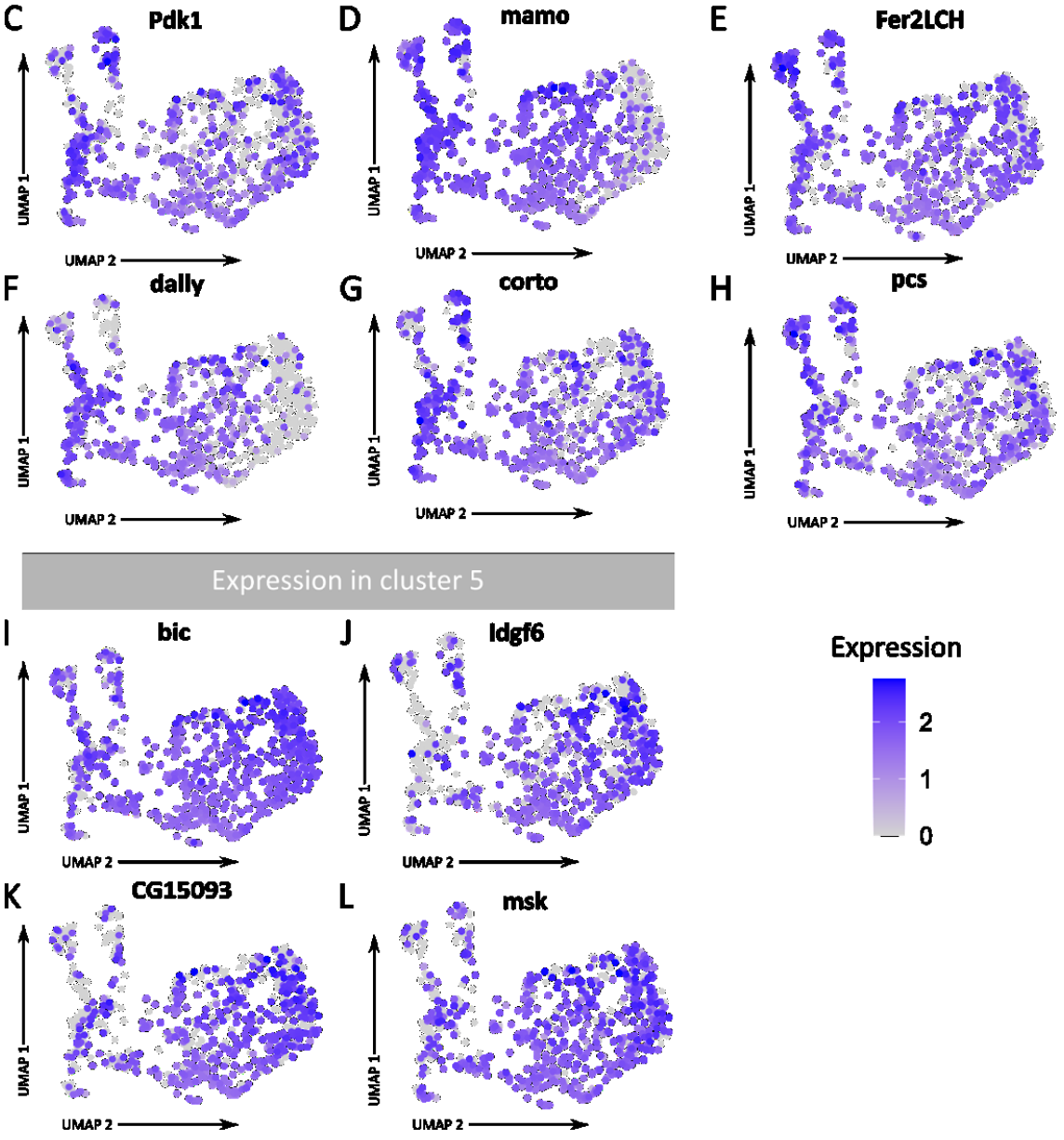


Figure 21: FeaturePlot of selected published marker genes of Rust, Shi and Tu *et al.* depict no cluster specificity for FCA EC clusters

(A) UMAP of the final clustered version of the FCA EC dataset. In the UMAP each dot depicts a cell of the dataset. Cells with similarities in transcriptomes are depicted closer to each other and are thus summarized in cell clusters highlighted by different colors. **(B)** Table of selected marker genes of Rust, Shi and Tu *et al.* with respect to their expressions in the DotPlot of the FCA EC dataset. **(C-L)** FeaturePlots of selected marker genes of Rust, Shi and Tu *et al.* mapped on the UMAP of the FCA EC dataset. Marker gene expression is depicted with a color gradient (blue = high expression; gray = low expression). Note that selected marker genes show no cluster specificity. All marker genes tend to be present in all FCA EC clusters.

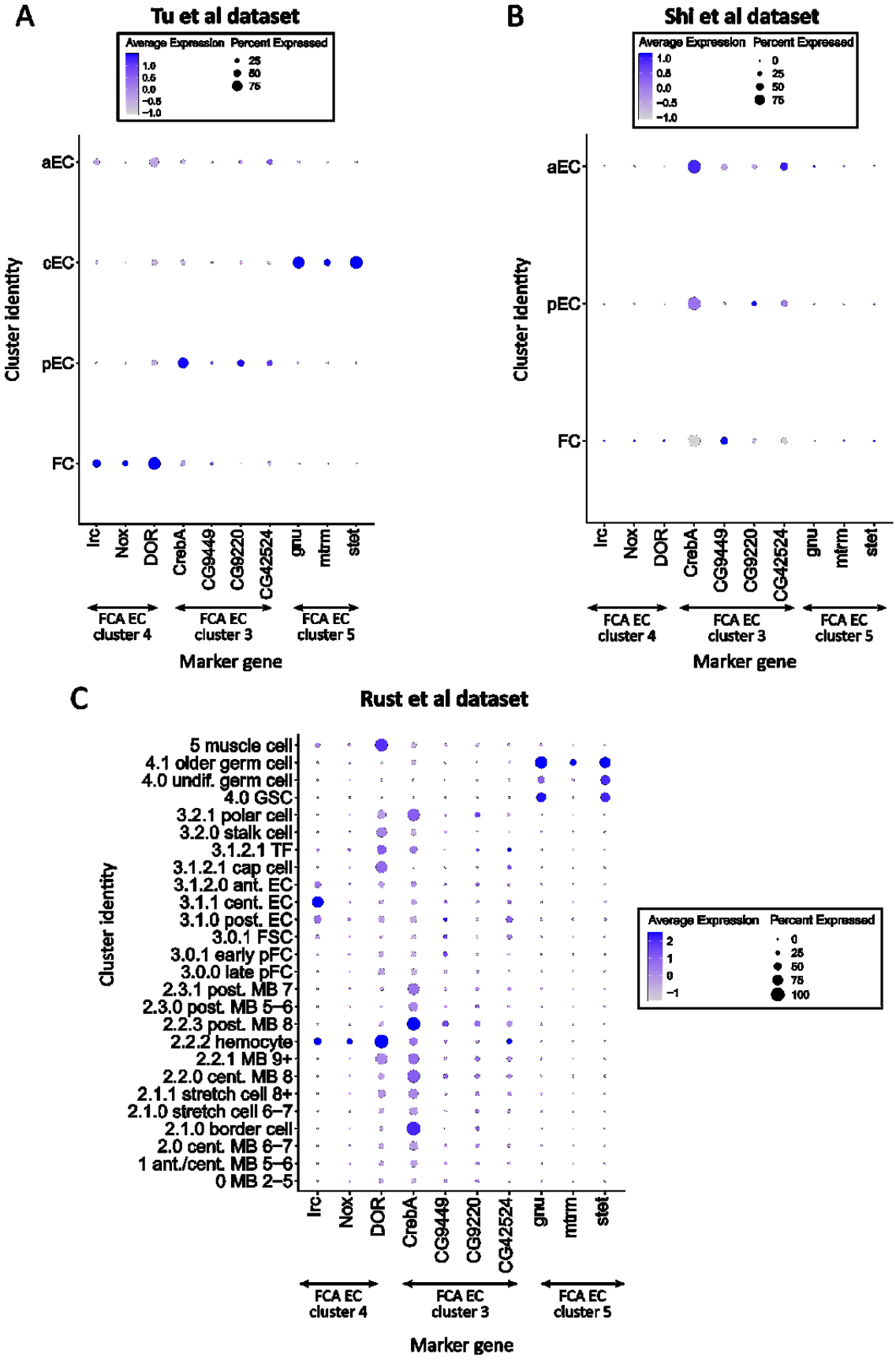

A
Tu et al dataset
Average Expression
Percent Expressed
Cluster identity
aEC
cEC
pEC
FC
Irc
Nox
DOR
CrebA
CG9449
CG9220
CG42524
gnu
mtrm
stet
FCA EC cluster 4
FCA EC cluster 3
FCA EC cluster 5
Marker gene
B
Shi et al dataset
C
Rust et al dataset
5 muscle cell
4.1 older germ cell
4.0 undif. germ cell
4.0 GSC
3.2.1 polar cell
3.2.0 stalk cell
3.1.2.1 TF
3.1.2.1 cap cell
3.1.2.0 ant. EC
3.1.1 cent. EC
3.1.0 post. EC
3.0.1 FSC
3.0.1 early pFC
3.0.0 late pFC
2.3.1 post. MB 7
2.3.0 post. MB 5–6
2.2.3 post. MB 8
2.2.2 hemocyte
2.2.1 MB 9+
2.2.0 cent. MB 8
2.1.1 stretch cell 8+
2.1.0 stretch cell 6–7
2.1.0 border cell
2.0 cent. MB 6–7
1 ant./cent. MB 5–6
0 MB 2–5

Figure 22: Plotting selected cluster specific markers of the FCA EC dataset in Rust, Shi and Tu *et al.* datasets does not lead to identification of cluster identities of the FCA EC dataset

(A-C) DotPlots of selected FCA EC marker genes in the Tu *et al.* (2021), Shi *et al.* (2021) and Rust *et al.* (2020) datasets. Tu and Shi *et al.* datasets did not incorporate finalized clustering and cluster annotations. After running analysis on the raw datasets, published marker genes of Tu *et al.* and Shi *et al.* were mapped in its respective dataset in order to obtain EC and FC clusters. The expression of the FCA EC marker genes is represented in the DotPlots by the size and color of the dots: the size refers to the amount of cells in the cluster that express the gene whereas the expression strength is depicted in the color (blue = high expression, gray = low expression). Note that FCA EC marker genes are displayed in different clusters of the respective datasets: in Rust *et al.* all FCA EC marker genes show little expression in the EC compartments with the exception of *Irc*, whereas *Irc* is expressed in Tu *et al.* in the FC compartment and in Shi *et al.* there is hardly any expression for *Irc*.

4.1.2 Verification of FCA EC data and calculated markers

4.1.2.1 Data integration of EC datasets

High resolutions of datasets (i.e. the amount of incorporated cells) are beneficial for analysis and to accurately define cell coherences. This can be achieved by single-cell data integration in that separate datasets are merged together for a conjoint analysis (Stuart et al., 2019). For the further purpose of studying ECs in the *Drosophila* germarium in detail, I performed data integration of all published EC datasets. For this, I followed the standard workflow of Seurat v4 (Seurat Vignette “Introduction to scRNA-seq integration”; Stuart et al., 2019). Hereby, I sought to treat a greater input of cells in order to define subtypes within the EC population. I included several EC datasets of which I performed separate downstream analysis: FCA EC (H. Li et al., 2022), Tu *et al.* (initially three independent datasets, Tu et al., 2021) and Shi *et al.* (Shi et al., 2021). Furthermore, EC subsets of Rust *et al.* (Rust et al., 2020) (three datasets), Slaidina *et al.* (Slaidina et al., 2021) (five datasets) and Jevitt *et al.* (Jevitt et al., 2020) were incorporated. As data integration requires a sufficient amount of cells in order to apply the default settings, I therefore performed the data integration

process with a total amount of 11 individual datasets from 6 sources, excluding certain datasets with a fewer number of 102 cells (Table 2). This generated a file with a total amount of 8545 cells.

Table 2: EC datasets and associated sources for the single-cell data integration

Data integration was performed using the Seurat workflow v4 (detailed description available at SatijaLab). Yellow marked datasets were excluded, as the amount of cells was insufficient for the integration procedure.

Source	**Dataset**	**Cells**
Shi *et al.*, 2021	Shi	2345
Tu *et al.*, 2021	Tu aEC	585
	Tu mEC	1351
	Tu pEC	2496
Jevitt *et al.*, 2020	Jevitt	250
Slaidina *et al.*, 2020	Slaidina 1	158
	Slaidina 2	31
	Slaidina 4	83
	Slaidina 5	102
	Slaidina 6	134
Li *et al.*, 2022	FCA EC	713
Rust *et al.*, 2020	Rust 1	13
	Rust 2	141
	Rust 3	270

After running the process of dataset integration (see 3.1.2 Dataset integration), I identified six clusters using a resolution factor of 0,08 (Figure 23 A). Unexpectedly, the visualization of the clustering via UMAP exhibited cluster 1 to overlap with several other clusters, in particular cluster 2 (Figure 23 B). In order to understand how the datasets contribute to the clusters, I split the UMAP by the dataset source (Figure 23 C). Noticeably, mostly the cells of the Shi and Tu *et al.* dataset contribute to the overlapping cluster 1 and 2 respectively.

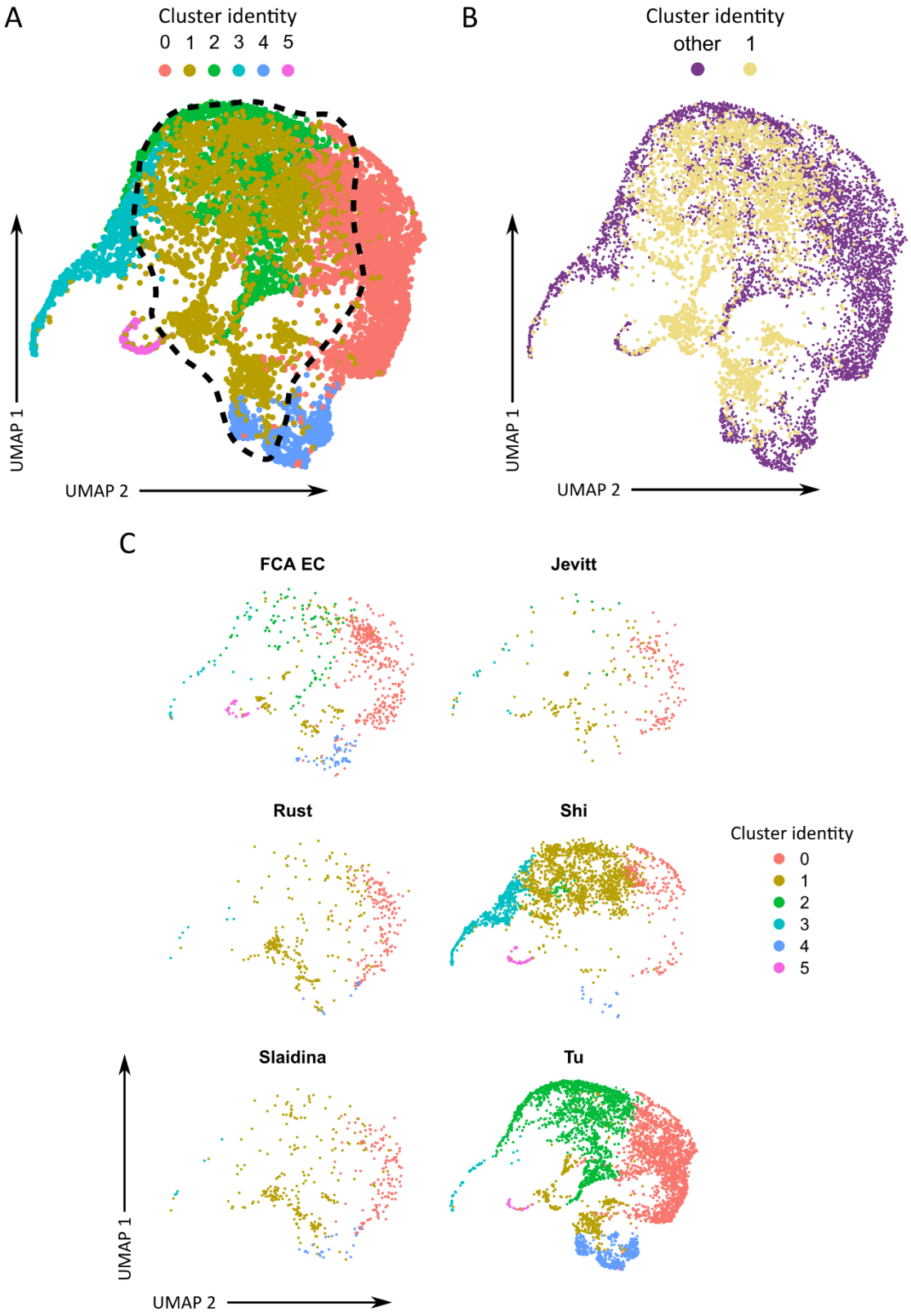

A
Cluster identity
0 1 2 3 4 5
UMAP 1
UMAP 2
B
Cluster identity
other 1
UMAP 1
UMAP 2
C
FCA EC
Jevitt
Rust
Shi
Slaidina
Tu
Cluster identity
0
1
2
3
4
5
UMAP 1
UMAP 2

Figure 23: Integration of published EC datasets reveals an overlap of clusters
(A) UMAP depicting the final clustering with six ultimate clusters of the integration of published EC datasets. In the UMAP each dot represents a cell. Based on transcriptome similarities, cells are portrayed closer to each other resulting in cell clusters. The black dashed circle outlines cluster 1 which overlaps with cluster 2. **(B)** UMAP of the merged object highlighting cluster 1 and depicting all other clusters in the same color. **(C)** UMAP split by its dataset source indicates that the Tu and Shi *et al.* datasets contribute the most to the overlapping clusters 1 and 2.

For further analysis, marker genes were calculated for each cluster of the merged object. The top 4 of cluster specific marker genes are presented via DotPlot (Figure 24 B). Marker genes for clusters 0, 3, 4 and 5 displayed distinct expression patterns as cluster specificity and expression strength (big and blue dots) are given. Interestingly, I was not able to define marker genes with a consistent expression throughout cluster 1. For cluster 2 a similar trend was observable as calculated marker genes do not limit their expression to cluster 2 only (Figure 24 B). In an attempt to further analyze these clusters, I created subsets only containing cells of either cluster 1 or 2 and extracted them out of the merged object (Figure 24 C-D). The isolated subsets were once again clustered. The subset of former cluster 2 resulted in a generation of six clusters under a resolution factor of 0,15. Due to overlapping gene expression patterns, I merged two subclusters together, resulting in five final clusters for the subset of cluster 2 (Figure 24 C). Hereby, I was able to generate adequate marker genes which were visualized using DotPlot (Figure 24 E). These markers depicted mostly the anticipated cluster specific expression. In contrast to this, it was not manageable to adjust the resolution factor for cluster 1 and acquire well defined clusters. An example for unsuccessful clustering is shown in Figure 24 D with a resolution factor of 0,001. This resulted in the lack of specific marker gene calculation. These results prompted the conclusion that the integration was impossible due to large batch effects, which were particularly noticeable among the cells of cluster 1. Since I discovered that cells of the Tu and Shi *et al.* datasets contributed to cluster 1 the most, these datasets are presumably not integrable due to low quality characteristics (e.g. doublets). Thus, the data integration was not involved in the further process of this study.

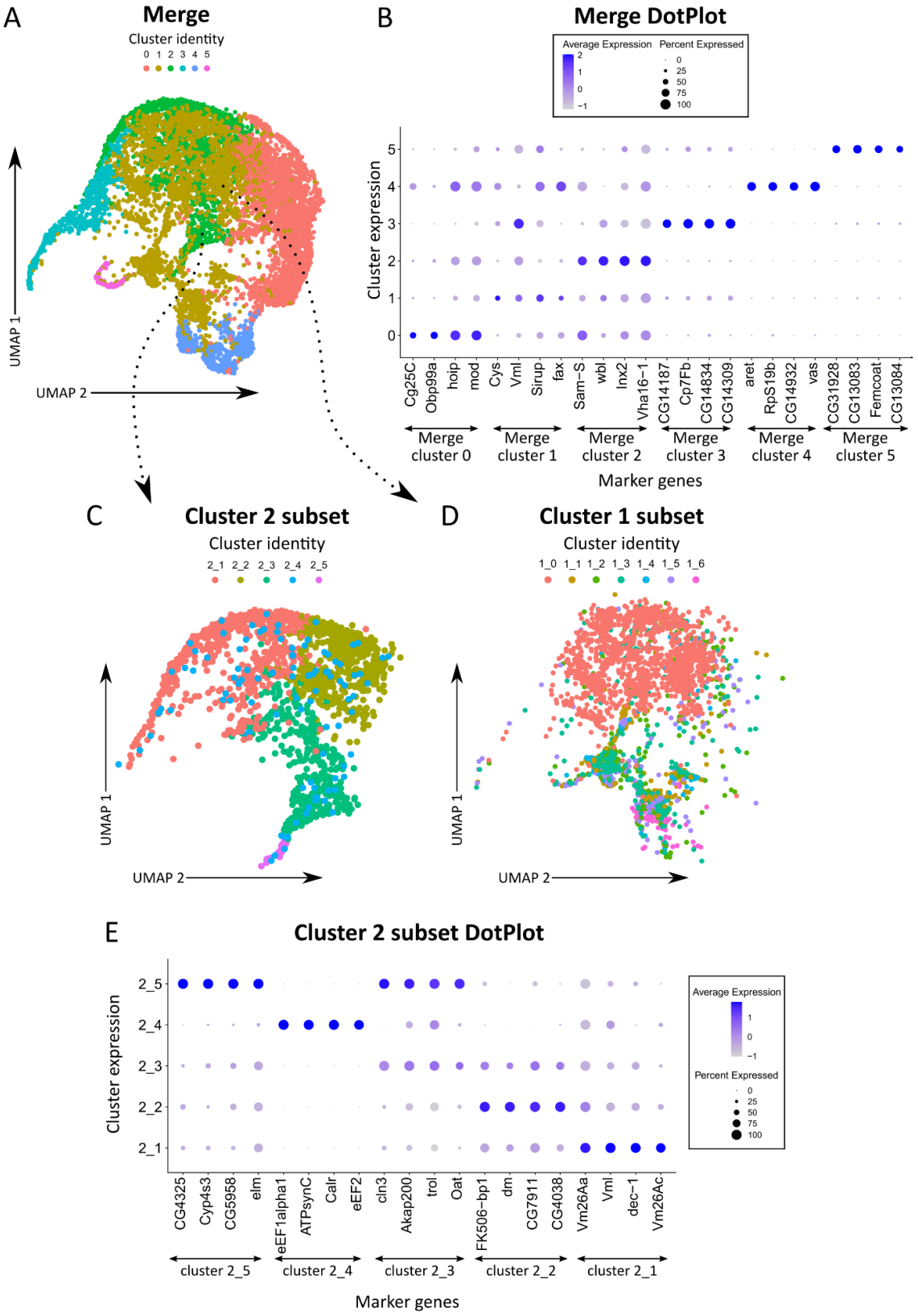

A
Merge
Cluster identity
0 1 2 3 4 5
UMAP 1
UMAP 2
B
Merge DotPlot
Average Expression
Percent Expressed
Cluster expression
Cg25C
Obp99a
hoip
mod
Cys
Vml
Sirup
fax
Sam−S
wbl
lnx2
Vha16−1
CG14187
Cp7Fb
CG14834
CG14309
aret
RpS19b
CG14932
vas
CG31928
CG13083
Femcoat
CG13084
Merge cluster 0
Merge cluster 1
Merge cluster 2
Merge cluster 3
Merge cluster 4
Merge cluster 5
Marker genes
C
Cluster 2 subset
Cluster identity
2_1 2_2 2_3 2_4 2_5
UMAP 1
UMAP 2
D
Cluster 1 subset
Cluster identity
1_0 1_1 1_2 1_3 1_4 1_5 1_6
UMAP 1
UMAP 2
E
Cluster 2 subset DotPlot
Cluster expression
CG4325
Cyp4s3
CG5958
elm
eEF1alpha1
ATPsynC
Calr
eEF2
cln3
Akap200
trol
Oat
FK506−bp1
dm
CG7911
CG4038
Vm26Aa
Vml
dec−1
Vm26Ac
cluster 2_5
cluster 2_4
cluster 2_3
cluster 2_2
cluster 2_1
Marker genes
Average Expression
Percent Expressed

Figure 24: Cluster specific markers are identified for all but cluster 1 of the data integration of published EC datasets

(A) UMAP with six distinct clusters of the merged EC object. Each cell is depicted as a dot in the UMAP. Clusters are acquired based on the transcriptomic similarities of the cells. Similar cells are shown closer to each other and result in cluster cell compounds highlighted in different colors. Dashed arrows point at UMAPs of subsets of cluster 1 and 2. **(B)** DotPlot of the top 4 calculated marker genes for each of the clusters of the merged EC object. Expression strength is represented by the color of the dot (blue = high expression, gray = low expression) while the dot size indicates the amount of cells in the cluster that express the particular gene. Note that marker genes for cluster 0, 3, 4 and 5 show expressions limited to its respective cluster while marker genes for cluster 1 and 2 distribute their expression over several clusters. **(C)** In order to analyze the cells of cluster 2 separately, a subset of the merged EC object containing only cells of cluster 2 was performed. UMAP showing the clustering of the subset of cluster 2 with the formation of five distinct clusters. Note that cells of the cluster 2_4 do not form a coherent cluster. **(D)** Analogous to (C), cells of cluster 1 of the merged EC object were extracted to perform a separate analysis. Clustering via Seurat did not result in the formation of contiguous cell clusters. **(E)** DotPlot of calculated marker genes for the subset of cluster 2. Despite 2_5 and 2_3 which show overlapping marker gene expression, calculated marker genes exhibit distinct cluster specificity.

4.1.2.2 No EC marker genes identified by the validation of FCA EC markers

In order to investigate whether new EC specific genes were calculated via the analysis of the FCA EC dataset, I sought to investigate the expression pattern of the calculated markers within the *Drosophila* germarium. In the following, I focussed on selected marker genes of the FCA EC clusters 3 and 4. Marker genes for FCA EC cluster 5 were excluded from further proceedings as the overlap of *gnu*, *mtrm* and *stet* (cluster 5 FCA EC) with the Rust *et al.* dataset depicted high expressions in germ cells and no expression in the EC compartment, suggesting that the FCA EC dataset contains a population of germ cells (Figure 22 C).

I used fly lines with either GFP protein traps or expression of Gal4/lacZ under the marker gene promoter. For immunofluorescence staining of protein trap lines, I used Fas III to identify follicle cells and GFP for the expression of GFP tagged

marker genes. Gal4 driver lines were crossed with reporter fly lines of the genotype *UAS-CD8::RFP*/CyO; *fax::GFP*/TM6B. Follicle cells were stained for Fas III. Since *fax* (*failed axon connections*) is required for membrane function in ECs, fax::GFP marks the somatic EC membranes. CD8 represents a transmembrane protein and the expression of RFP via *UAS-CD8::RFP* consequently marks the cell surfaces of the desired gene. LacZ coded fly lines were stained for lacZ. After staining procedures, I observed the expression by using confocal microscopy. I was not able to detect any fluorescence activity of *Irc-Gal4*, *Nox-Gal4*, *CG9449-Gal4*, *CG42524-Gal4*, *DOR-Gal4*, *CrebA-lacZ*, *CG9220-GFP*, *Irc-GFP*, *Nox-GFP*, *CrebA-GFP*, *CG42524-GFP* (Figure 25, Figure 26). In contrast to this, *DOR-GFP* was detected in stalk cells, which were identified by their particular morphology (Figure 26 D).

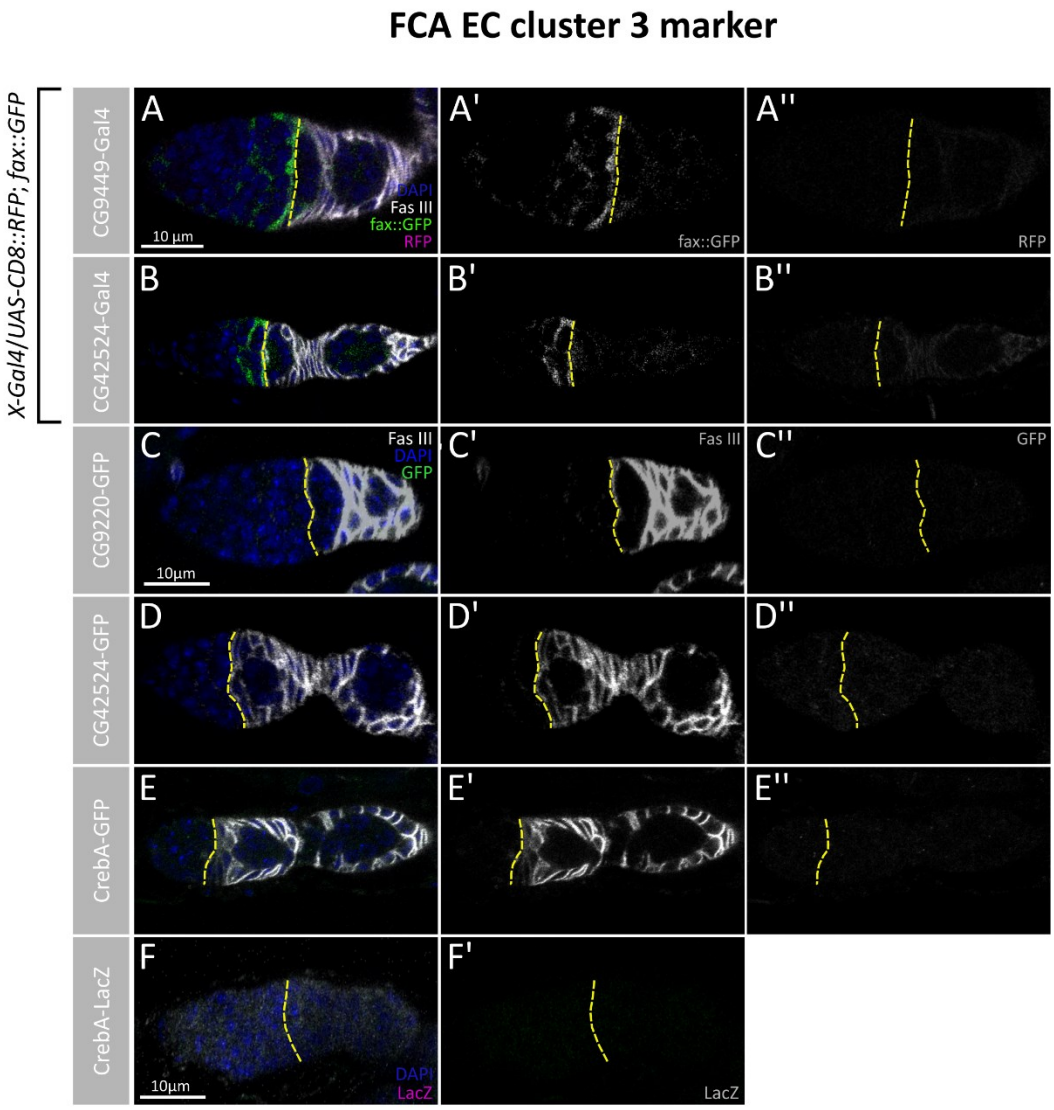


Figure 25: Selected FCA EC cluster 3 markers show no measurable fluorescence expression

(A-B) Germaria of flies expressing Gal4 under the marker gene promoter (*X*), crossed with the *UAS-CD8::RFP*; *fax::GFP*, which marks ECs, and stained for DAPI (blue), RFP (magenta), GFP (green) and the follicle cell marker Fas III (white). Yellow dashed lines represent the Region 2a/2b border, defining the mark between posterior EC compartments and the FC section, including FSCs. **(A'-B')** GFP channel in white. *fax::GFP* is matching the membranes of ECs and reside anterior to the Region2a/2b border (green, A-B, white, A'-B') **(A''-B'')** RFP channel in white. RFP activity should show the live expression of the respective marker gene, but was not detected for both marker gene candidates (A-B; A''-B''). **(C-E)** Germaria of flies expressing the GFP tagged marker gene under the endogenous promoter, stained for DAPI (blue), Fas III (white) and GFP (green). On the Region 2a/2b border, the demarcation of ECs and FCs is marked by the yellow dashed line. **(C'-E')** Fas III channel in white. **(C''-E'')** GFP channel in white. No recognizable expression was detected for all GFP tagged marker genes. **(F)** Germaria of flies expressing lacZ under the *CrebA* promoter, stained for lacZ (magenta) and DAPI (blue). **(F')** LacZ channel in white. The yellow dashed line marks the boundary between ECs and FCs. For *CrebA-lacZ* no expression was measurable.

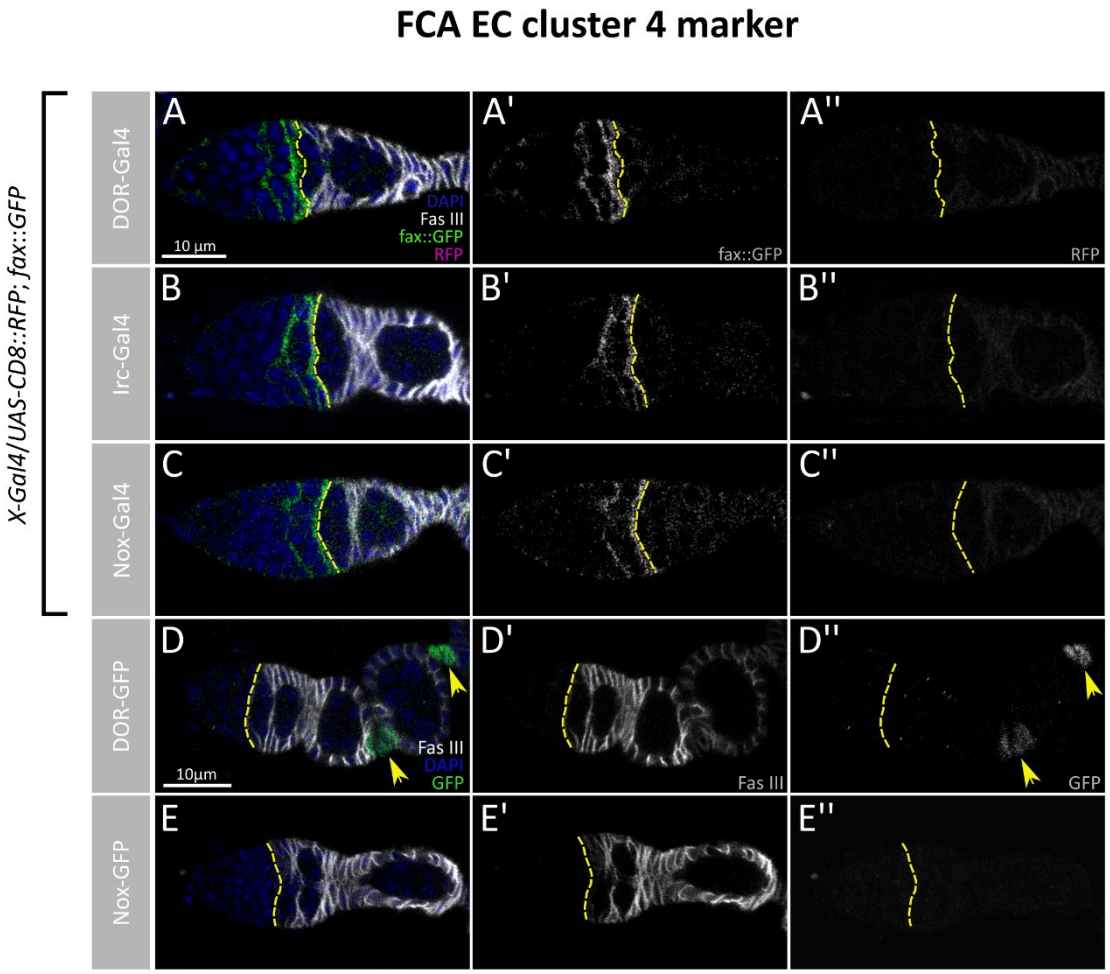


Figure 26: FCA EC cluster 4 marker *DOR* is expressed in stalk cells

(A-C) Germaria of flies with the genotype *X-Gal4*/*UAS-CD8::RFP*; *fax*::*GFP* stained for DAPI (blue), Fas III (white), GFP (green) and RFP (magenta). X in the genotype stands as a placeholder for the respective marker gene. *fax::GFP* marks EC membranes (green). Fas III constitutes a follicle cell marker. Yellow dashed lines indicate the delineation of ECs and FCs, including FSCs, at the Region 2a/2b border. **(A'-C')** GFP channel in white. Membranes of ECs are marked by *fax::GFP* and are located anterior to the FC compartment (white). **(A''-C'')** RFP channel in white. Gal4 is expressed under the marker gene promoter, therefore RFP activity should mark the expression of the respective marker gene. For all marker gene Gal4 fly lines no expression was detected. **(D-E)** Germaria of flies of GFP marker gene trap lines stained for DAPI (blue), GFP (green) and Fas III (white). At the Region 2a/2b border the EC compartment is delimited from the FC section represented by the yellow dashed lines. *DOR-GFP* is expressed in the area of stalk cells identified by their distinct stalk-like morphology (D, yellow arrowheads). **(D'-E')** Fas III channel in white. **(D''-E'')** GFP channel in white. The GFP expression is pointed out by yellow arrowheads (D'').

To confirm the observed expression of *DOR*, I mapped *DOR* in Rust *et al.* and the FCA ovary dataset: both show *DOR* activity in stalk cells but no significant expression in the EC compartments (Figure 27 E, Figure 28 B). Comparing *DOR* to several cell type specific marker genes via DotPlot, an obvious difference of validated EC subtype markers such as *wun2* (cEC, pEC) and *ptc* (aEC, cEC,pEC) in comparison to *DOR* is emphasized: while *wun2* and *ptc* limit their expression mostly to the EC identities, *DOR* shows little expression in ECs (Figure 27 G). As *DOR* is also active in polar, stalk, cap and terminal filament cells, the direct comparison to strong markers for those cells (*CG46339* = stalk and terminal filament cells, *upd1* = polar cell*, en* = cap and terminal filament cell*)* demonstrate that *DOR* depicts no clear cell specificity, since the expression strength for polar, stalk, cap and terminal filament cells are almost equally high (Figure 27 G).

Cell type prediction via Seurat v4 Reference Mapping (Seurat Vignette "Multimodal reference mapping"; Hao et al., 2021) uses a reference dataset in order to explore cell identities of a query dataset. In the course of the workflow, anchors (transcriptomic similarities) between the reference and the query dataset are computed. These are used for the transfer of the reference (e.g. cell type labels) to the query data, which is finally visualized via UMAP. Conclusive cell type prediction of the FCA EC cells using assigned cell type annotations of the Rust *et al.* dataset, reveal largely cell identities other than ECs (Figure 27 F). These observations together with previous results of the scRNA-Seq analysis unravel that the FCA EC dataset is mostly not composed of EC identities.

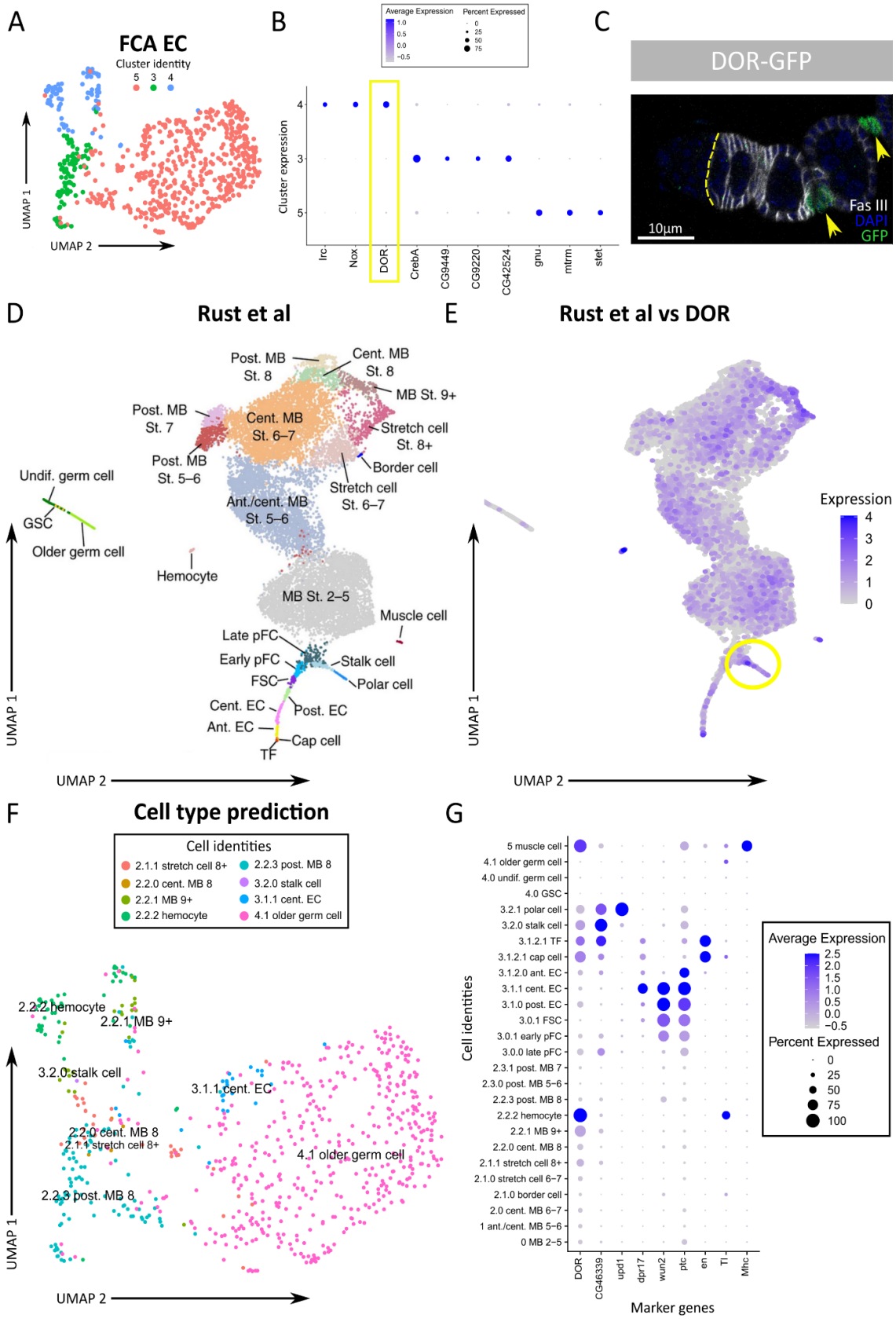
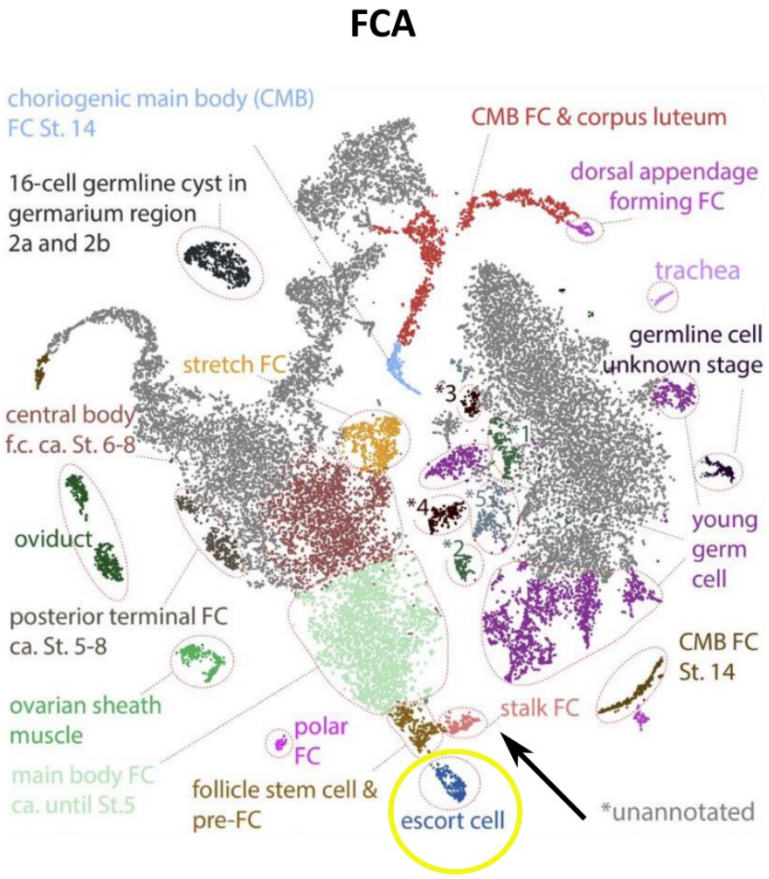

A
FCA EC
Cluster identity
5 3 4
UMAP 1
UMAP 2
B
Average Expression
Percent Expressed
Cluster expression
Irc
Nox
DOR
CrebA
CG9449
CG9220
CG42524
gnu
mtrm
stet
C
DOR-GFP
Fas III
DAPI
GFP
10µm
D
Rust et al
Post. MB St. 8
Cent. MB St. 8
MB St. 9+
Post. MB St. 7
Cent. MB St. 6–7
Stretch cell St. 8+
Post. MB St. 5–6
Border cell
Stretch cell St. 6–7
Undif. germ cell
GSC
Older germ cell
Ant./cent. MB St. 5–6
Hemocyte
MB St. 2–5
Muscle cell
Late pFC
Early pFC
Stalk cell
FSC
Polar cell
Cent. EC
Post. EC
Ant. EC
Cap cell
TF
UMAP 1
UMAP 2
E
Rust et al vs DOR
Expression
UMAP 1
UMAP 2
F
Cell type prediction
Cell identities
2.1.1 stretch cell 8+
2.2.0 cent. MB 8
2.2.1 MB 9+
2.2.2 hemocyte
2.2.3 post. MB 8
3.2.0 stalk cell
3.1.1 cent. EC
4.1 older germ cell
2.2.2 hemocyte
2.2.1 MB 9+
3.2.0 stalk cell
3.1.1 cent. EC
2.2.0 cent. MB 8
2.1.1 stretch cell 8+
4.1 older germ cell
2.2.3 post. MB 8
UMAP 1
UMAP 2
G
5 muscle cell
4.1 older germ cell
4.0 undif. germ cell
4.0 GSC
3.2.1 polar cell
3.2.0 stalk cell
3.1.2.1 TF
3.1.2.1 cap cell
3.1.2.0 ant. EC
3.1.1 cent. EC
3.1.0 post. EC
3.0.1 FSC
3.0.1 early pFC
3.0.0 late pFC
2.3.1 post. MB 7
2.3.0 post. MB 5–6
2.2.3 post. MB 8
2.2.2 hemocyte
2.2.1 MB 9+
2.2.0 cent. MB 8
2.1.1 stretch cell 8+
2.1.0 stretch cell 6–7
2.1.0 border cell
2.0 cent. MB 6–7
1 ant./cent. MB 5–6
0 MB 2–5
Cell identities
DOR
CG46339
upd1
dpr17
wun2
ptc
en
Tl
Mhc
Marker genes
Average Expression
Percent Expressed

Figure 27: The FCA EC dataset largely does not contain ECs

(A) UMAP plot of the final clustered version of the FCA EC dataset with a generation of three clusters (3, 4 and 5). In the UMAP each cell is represented by a dot. The more similar the transcriptomes of cells, the closer the cells are represented to each other in the UMAP. Cell associations of contiguous cells are then defined as clusters. **(B)** DotPlot of selected marker genes for each generated cluster of the FCA EC dataset. The average expression strength is represented by the color of the dot (blue = high expression; gray = low expression) whereas the size of the dot indicates the amount of cells in a cluster which express the particular gene. The yellow box highlights the expression of *DOR* which is expressed exclusively in cluster 4. **(C)** Germarium of a fly expressing GFP tagged *DOR* under the endogenous promoter stained for DAPI (blue), Fas III (white) and GFP (green). Fas III depicts a follicle cell marker. The yellow dashed line marks the boundary between the EC compartment and the FC section with FSCs at the Region 2a/2b border. Yellow arrowheads point at *DOR* expression in stalk cells which are identified by their typical morphology (green). **(D)** UMAP of the clustered version of the Rust *et al.* dataset containing 26 clusters with assigned cluster identities. Adapted from (Rust et al., 2020). **(E)** FeaturePlot of the Rust *et al.* dataset featuring the FCA EC cluster 4 marker gene *DOR*. Expression for *DOR* is displayed in several clusters of the Rust *et al.* dataset. The yellow circle points out the expression of *DOR* in stalk cells. **(F)** Cell type prediction of the cells of the FCA EC dataset using Rust *et al.'s* assigned cell identities and visualized via UMAPPlot. Cell type prediction reveals largely other cell identities than ECs. Stalk cell identities are predicted in cluster 4 of the FCA EC dataset. **(G)** DotPlot comparison of *DOR* (FCA EC cluster 4 marker) and several validated marker genes for different cell types using the *Rust et al.'s* cluster annotations: *CG46339* (stalk cell marker), *upd1* (polar cell marker), *dpr17* (cEC marker), *wun2* (cEC and pEC marker), *ptc* (aEC, cEC and pEC marker), *en* (cap and terminal filament cell marker), *Tl* (hemocyte marker), *Mhc* (muscle cell marker). *DOR* shows sparse expression in the individual EC identities, whereas it is highly expressed in hemocytes, muscle cells, polar cells, cap cells and stalk cells. Note that marker genes specific for ECs such as *wun2*, *ptc* and *dpr17* display a different expression scheme, limiting their expression mostly to the EC section.

A


FCA
choriogenic main body (CMB)
FC St. 14
CMB FC & corpus luteum
dorsal appendage
forming FC
16-cell germline cyst in
germarium region
2a and 2b
trachea
germline cell
unknown stage
stretch FC
central body
f.c. ca. St. 6-8
*3
1
*4
5
*2
young
germ
cell
oviduct
posterior terminal FC
ca. St. 5-8
CMB FC
St. 14
ovarian sheath
muscle
stalk FC
polar
FC
main body FC
ca. until St.5
follicle stem cell &
pre-FC
escort cell
*unannotated

B

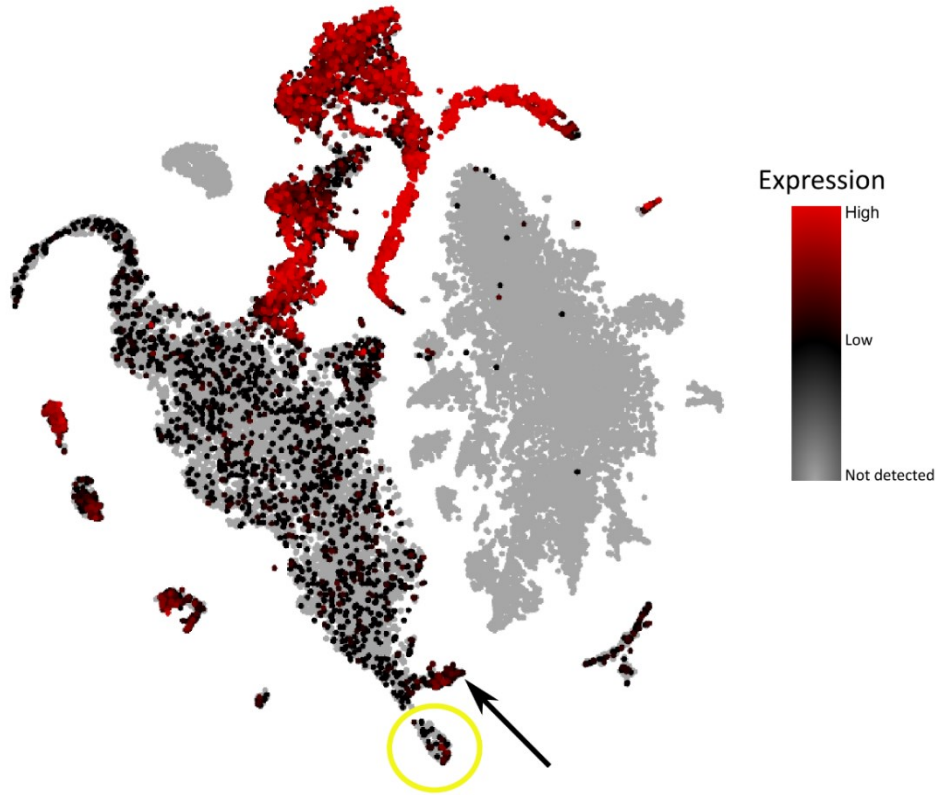

FCA vs DOR
Expression
High
Low
Not detected

Figure 28: *DOR* is expressed in main body follicle and stalk cells in the FCA ovary data

(A) T-SNE (t-distributed stochastic neighbor embedding) visualization of assigned cell identities of the FCA data regarding cells of the *Drosophila melanogaster* ovary. In the T-SNE, cells with similar transcriptomes are visualized closer to each other resulting in the formation of cell clusters. The yellow circle points out the cluster of escort cells, whereas the black arrow points at the group of stalk cells. Modified figure from (H. Li et al., 2022) reprinted with permission from AAAS. **(B)** *DOR* in the FCA ovary data exhibits high expression in later stages of oogenesis (different stages of main body follicle cells) and in the area of stalk cells (black arrow) while low expression is detected in the cluster of escort cells (yellow circle).

4.2 Escort cell lineage tracing

4.2.1 Experimental design

For lineage tracing EC descendants, fly lines of the genotype Gal80ts; *fax-Gal4*, G-TRACE were used. Since *fax* is expressed in all ECs, Gal4 production via *fax-Gal4* is EC specific and G-TRACE is thus only activated in Gal4 expressing ECs. *tub-Gal80ts* was used to control Gal4 dependent expressions due to temperature sensitivity. Gal80ts activity was assured by rearing flies at restrictive temperatures of 18°C. Upon eclosion, flies were shifted to a permissive temperature of 29°C, inhibiting the activity of Gal80ts. Hence, G-TRACE activity was limited to the adulthood of flies. Utilizing this system, it was possible to mark clones deriving from Gal4 expressing ECs. Direct and live expression of Gal4 activity was visualized with the expression of the red fluorescent protein (RFP), whereas the offspring lineages of RFP positive cells were marked with the green fluorescent protein (GFP). GFP expression is inherited to the next generation of cells within the lineage independent of Gal4 (see 1.1.1.2 Gal4 technique for real-time and clonal expression (G-TRACE)). Since *fax-Gal4* is sporadically expressed throughout the EC population and ECs typically do not contribute to the production of other cell types under standard conditions (Rust et al., 2020), I expected that RFP and GFP expressions were limited to the EC section from Region 1 to the border of Region 2a/2b. Hence, no fluorescence was presumed in Fas III positive cells under standard well fed conditions.

In this study, I screened for the expression of GFP and RFP within germaria of flies undergoing different test conditions. In general, there were variable patterns for GFP and RFP expression detectable (Figure 29). As anticipated for GFP monitoring, the following expression patterns were obtained: unmarked germaria, characterized by the appearance of RFP positive ECs but no expression of GFP (Figure 29 B); GFP positive EC clones, with GFP expression in ECs anterior to the Region 2a/2b border (Figure 29 A); GFP positive FC clones, represented by GFP expression in Fas III positive cells posterior to the Region 2a/2b, which are usually accompanied by GFP marked ECs anterior to the Region 2a/2b border (Figure 29 C). RFP was observable in two distinct expression patterns: RFP positive ECs, where RFP expression was limited to the EC compartment (anterior to the Region 2a/2b border) (Figure 29 A) and, unexpectedly, the expression of RFP in Fas III positive cells (RFP positive FCs, posterior to the Region 2a/2b border) (Figure 29 D). The occurrence of RFP positive FCs is addressed further below.

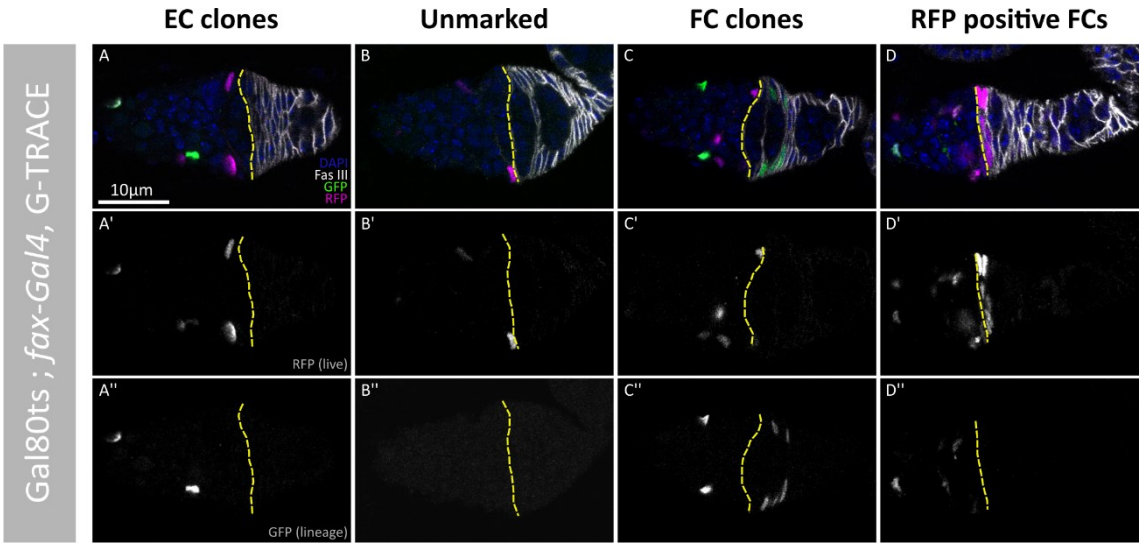


Figure 29: Expression patterns of RFP and GFP clones for G-TRACE EC lineage tracing

(A-D) Obtained expression patterns via the usage of G-TRACE in germaria from flies of the genotype Gal80ts; *fax-Gal4*/G-TRACE stained for DAPI (blue), Fas III (white), GFP (green) and RFP (magenta). **(A'-D')** RFP channels in white. **(A''-D'')** GFP channels in white. Yellow dashed lines depict the Region 2a/2b border, defining the compartments of ECs and FCs which include FSCs. I observed four different expression patterns: **(A-A'')** Germaria with EC clones: ECs are marked anterior to the Region 2a/2b border with RFP, depicting Gal4 live expression, and GFP, representing EC clones deriving from Gal4 live expression. This expression pattern was expected for RFP in general and for

GFP in standard well fed conditions. **(B-B'')** Unmarked germaria: RFP marked ECs residing anterior to the Region 2a/2b showing no expression of GFP. **(C-C'')** Germaria with FC clones: GFP expression in Fas III positive cells posterior to the Region 2a/2b border are considered to depict clones deriving from the EC compartment anterior to the Region 2a/2b border. This observation represents the EC conversion events. In addition, FC clones are usually accompanied by RFP and GFP marked ECs anterior to the Region 2a/2b border. **(D-D'')** RFP positive FCs: RFP positive cells are detected in the Fas III positive area of FCs.

For all controls in this study, fly lines of the genotypes Gal80ts; *fax-Gal4*, G-TRACE were crossed and bred at 18°C. First, to confirm Gal80ts activity, progeny at 18°C was examined after three days. Gal80ts activity was proven as no GFP or RFP expression was detected in all germaria (Figure 30 D). Second, proper function of Gal4 and downstream activation of G-TRACE was tested by the observations at two different time points. Postulating that complete G-TRACE induction requires a distinct time and may take up to two weeks, flies were examined after 7 and 14 days: after 7 days the gradual induction of G-TRACE is portrayed by 52% of germaria remaining unmarked, 48% depicting GFP positive EC clones and no FC clones (n=134 germaria isolated from 7 flies, one repeat) (Figure 30 A) while for RFP expression 92% of germaria were positive for RFP in ECs and only 8% in FCs (n=134 germaria) (Figure 30 C). Observations of germaria after 14 days (Figure 30 E) are consistent with the assumption of G-TRACE induction: only 5% of all germaria were unmarked, 94% demonstrated GFP positive EC clones and in only 1% of all cases FC clones emerged (n=188 germaria isolated from 8 flies, one repeat) (Figure 30 A). RFP expression was increased to the benefit of RFP positive ECs with 94% whereas the number of RFP positive FCs decreased to 6% (n=188 germaria, 8 flies) (Figure 30 C).
Paired controls for environmental test conditions were all conducted analogously to above described procedures, including rich diets at 29°C for 14 days. As expected, GFP positive FC clones remained at low levels for all control conditions (FC clones < 0,5%, n=2745 germaria, each control three repeats) (Figure 30 B). Taking all 14 days well fed controls into account and comparing them to one against another, no significance was observed with respect to FC clone emergence in germaria ($p > 0.05$) (Table 3). This depicts the independence of the

respective controls conducted on different days, suggesting the authentication of the control series. Since these results depict a stable clonal marking system due to reasonable expression rates, I decided to utilize Gal80ts; *fax-Gal4*, G-TRACE as fly line for experiments conducted in this study.

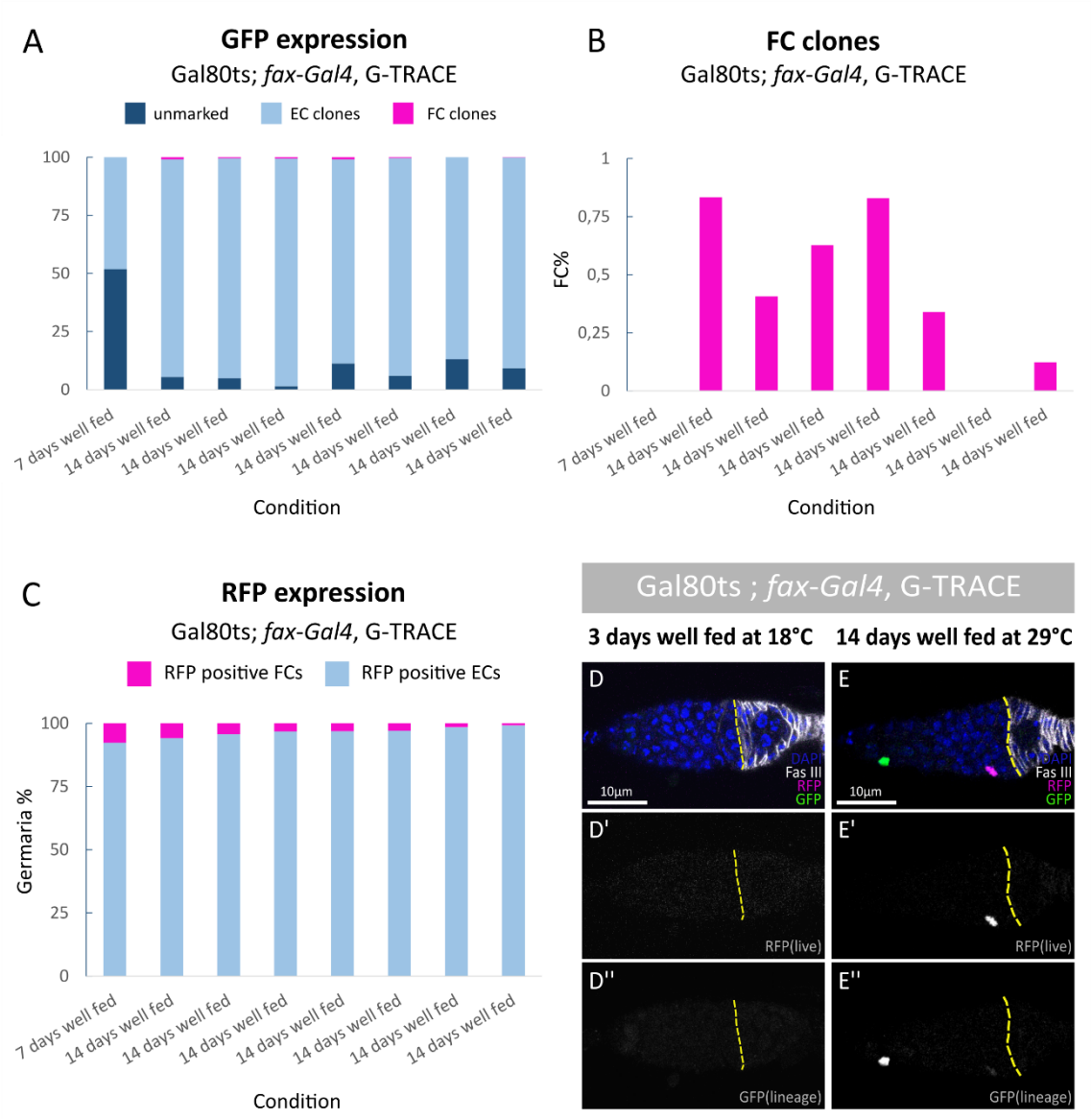


Figure 30: Control conditions provide low ratios of FC clones

Statistics of control conditions for flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE. For **(A-C)** from left to right: 7 days well fed n=134 germaria, one repeat, 14 days well fed n=188 germaria, one repreat. All subsequent controls correspond to the paired controls for the environmental screen with n=460 germaria (3 repeats), n=443 germaria (3 repeats), n=474 germaria (3 repeats), n=459 germaria (3 repeats), n=446 germaria (3 repeats) and n=463 germaria (3 repeats). **(A)** Bar plot of well fed controls versus percentages of unmarked germaria: ECs only expressing RFP but no GFP (dark blue),

germaria with EC clones: GFP marked ECs anterior to Region 2a/2b border (blue) and germaria with FC clones: GFP positive FC clones posterior to Region 2a/2b border (magenta). Note the high rate of unmarked germaria at the 7 days well fed control and the reduction of the unmarked rate after the 14 days well fed control. This demonstrates that G-TRACE induction involves a distinct time span. **(B)** Percentages of FC clone production among all control conditions. The maximum rate was approximately 0,75%, depicting the overall low emergence of FC clones in controls. **(C)** Bar plot of well fed controls versus percentages of germaria with RFP positive ECs: RFP detectable in ECs anterior to the Region 2a/2b border (blue) and RFP positive FCs: RFP present in Fas III positive cells posterior to the Region 2a/2b border (magenta). **(D-D'')** Control for proper Gal80ts function of germaria with the genotype Gal80ts; *fax-Gal4*, G-TRACE stained for DAPI (blue), Fas III (white), RFP (magenta) and GFP (green) after 3 days on a rich diet at 18°C - no fluorescence activity was observable indicating the proper Gal80ts function. **(E-E'')** Control for Gal4 and G-TRACE induction of germaria with the genotype Gal80ts; *fax-Gal4*, G-TRACE stained for DAPI (blue), Fas III (white), RFP (magenta) and GFP (green) after 14 days well fed control at 29°C. GFP and RFP marked ECs are present, verifying the proper function of Gal4/G-TRACE.

Table 3: P-value comparison of well fed control conditions in this study

Control referes to all 14 days well fed conditions.

	Control	Control	Control	Control	Control	Control	Control
Control	-	0,74793534	0,85490788	0,8349765	0,89787143	0,98809758	0,83629982
Control	0,74793534	-	0,87611099	0,58993077	0,65671244	0,75926392	0,95552209
Control	0,85490788	0,87611099	-	0,68249445	0,75306296	0,86726315	0,9463674
Control	0,8349765	0,58993077	0,68249445	-	0,94196635	0,82301206	0,70469214
Control	0,89787143	0,65671244	0,75306296	0,94196635	-	0,8862699	0,75677007
Control	0,98809758	0,75926392	0,86726315	0,82301206	0,8862699	-	0,84582427
Control	0,83629982	0,95552209	0,9463674	0,70469214	0,75677007	0,84582427	-

4.2.2 Validation of recombinant fly lines

In order to reduce the crossing steps for EC conversion experiments, I sought to identify a functional recombinant with the genotype Gal80ts; *fax-Gal4*, G-TRACE. The recombinant fly lines tested in the following are referred to as #1, #3 and #5 and were initially intended for the implementation of further experimental conditions. To confirm that each component of the flie's genotype was properly working, I firstly tested the activity of Gal80ts. For this, flies were reared and maintained at 18°C and examined three days post hatching. I found that in 100% of all cases, no GFP or RFP expression was detectable, suggesting the accurate function of Gal80ts. Upon eclosion at 18°C, F1 generations of each line were shifted to 29°C to investigate Gal4 and the induction of G-TRACE. In order to validate their correct function, flies were hereby reviewed after 7 days first: for #1 36% of all germaria remained unmarked while GFP positive EC clones were present in 58% and in roughly 6% of all cases FC clones appeared (n=135 germaria isolated from 12 flies; p=0,0183, one repeat) (Figure 31 A). Surprisingly, even higher expression rates for RFP positive FCs were detected with 13% while the remaining germaria expressed RFP positive EC clones (87%, n=135 germaria, 12 flies) (Figure 31 B). #3 delivered 29% of unmarked germaria, 70% of GFP positive EC clones and only 1% of FC clones (n=133 germaria isolated from 8 flies; p=0,713, one repeat) (Figure 31 A). The highest score of RFP positive FCs with 31% was observed in #3 (in contrast to 69% RFP positive EC clones, n=133 germaria, 8 flies) (Figure 31 B). #5 displayed similar trends with 27% unmarked germaria, 71% GFP marked EC clones and the emergence of FC clones in 2% of all cases (n=144 germaria isolated from 8 flies; p=0,441, one repeat) (Figure 31 A). RFP positive FCs rose to 21% and left 79% of all germaria with RFP positive EC clones (Figure 31 B).

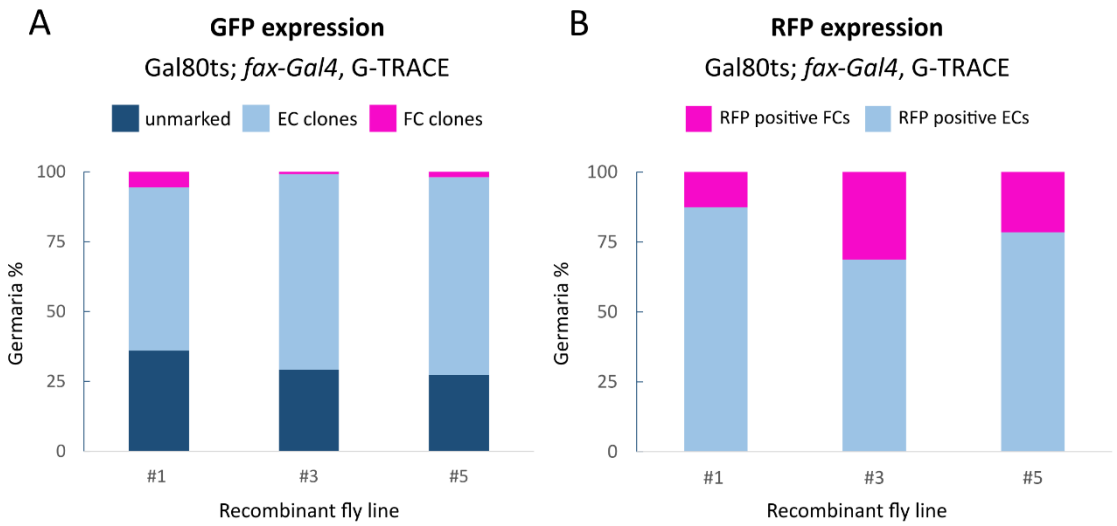


Figure 31: High ratios of FC clones presented after 7 days rich diet of recombinants

Barplots showing GFP **(A)** and RFP **(B)** expression in germaria of three independent recombinant fly lines (#1, #3, #5) with the genotype Gal80ts, *fax-Gal4*, G-TRACE fed a rich diet for 7 days at 29°C. #1 (n=135 germaria, 12 flies, p=0,0183, one repeat), #3 (n=133 germaria, 8 flies; p=0,713, one repeat), #5 (n=144 germaria isolated from 8 flies; p=0,441, one repeat). **(A):** A high amount of germaria did not contain GFP positive cells (dark blue). The majority of germaria contained GFP positive ECs (blue). A surprisingly high amount of germaria presented with GFP positive FC clones (magenta). (**B):** Many germaria contained RFP positive FCs identified by Fas III expression (magenta). Remaining germaria showed RFP expression in ECs (blue).

Postulating that after 14 days more cells had converted to a GFP positive state and therefore the number of unmarked germaria might be reduced, germaria of flies were examined after a time span of 14 days. Consistent with this, I observed the complete reduction of unmarked germaria for #5, whereas #3 reduced the amount to 10% and #1 to 9,5% (Figure 32 A). However, GFP and RFP expression were altered unexpectedly (Figure 32 A-B): #1 showed a surprisingly high amount of RFP positive FCs (24% in n=120 germaria isolated from 10 flies, p=0,734, one repeat) while no FC clones were detected; #3 presented FC clones in 2% of all cases (p=0,538), whereas RFP positive FCs were present in 10% of germaria (n=95 isolated from 10 flies, one repeat); #5 portrayed a significant increase in FC clones with 7% in total (n=130 germaria isolated from 11 flies; p=0,011, one

repeat) and even higher quantities of RFP positive FCs (26%, n=130 germaria isolated from 11 flies).

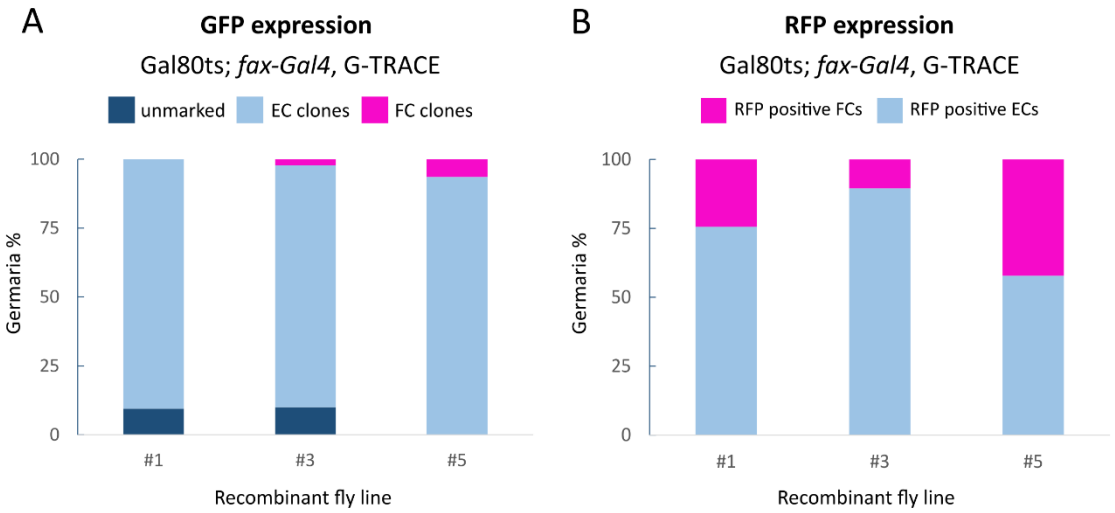


Figure 32: Increase in FC clones and RFP positive FCs after 14 days rich diet of recombinants

Graphs of 14 days rich diet at 29°C for three recombinant fly lines with the genotype Gal80ts, *fax-Gal4*, G-TRACE (#1, #3 and #5) split for GFP expression **(A)** and RFP expression **(B)**. #1 (n=120 germaria, 10 flies, p=0,734, one repeat), #3 (n=95, 10 flies, p=0,538, one repeat) and #5 (n=130 germaria, 11 flies, p=0,011, one repeat). **(A):** The ratio of unmarked germaria was reduced for all recombinants (dark blue). The amount of germaria with FC clones increased for #3 and #5, while no FC clones were detected in germaria of #1 (magenta). **(B):** Germaria with RFP positive ECs are depicted in blue. Note the high amount of germaria with RFP positive FCs in all three recombinant fly lines (magenta).

Since GFP and RFP rates were particularly high with respect to their expression in FCs, I reasoned that the examined flies potentially became homozygous for *fax-Gal4*, G-TRACE. This could result in a double fold amount of Gal4 present which would consequently enable an increased activation of GFP and RFP expression. To confirm this postulation, flies were again richly fed for 14 consecutive days but priorly sorted by the presence of the TM2 balancer. TM2 positive flies represented the heterozygous whereas TM2 negative flies depicted the homozygous section. I anticipated the heterozygous flies to display reduced GFP and RFP expression in FCs as there is less Gal4 present. Figure 33 shows

the GFP and RFP expression after 14 days of standard rich diet conditions for each sample, isolated by the property of homozygous or heterozygous genotypes. I observed alterations in all three samples respectively for GFP and RFP expression: #1 depicted an obvious decrease in FC clones from 36% in homozygous flies (n=118 germaria isolated from 6 flies, one repeat) to 3% in heterozygous flies (n=141 germaria isolated from 6 flies, p<0.001, one repeat). Analogously, a reduction in RFP positive FCs in homozygous flies (6% in n=118 germaria from 6 flies) compared to heterozygous flies (1% in n=141 germaria from 6 flies) was observed (Figure 33). #3 delivered no significant differences in FC clones with 13% in homozygous flies (n=101 germaria isolated from 6 flies, one repeat) and 12% in heterozygous flies (n=90 germaria isolated from 6 flies, p=0,798, one repeat). Strikingly, the number of unmarked germaria and RFP positive FCs was reduced for heterozygous flies of #3, from 14% unmarked in the homozygous section (n=101 germaria) to 2% unmarked in heterozygous flies (n=90 germaria) and 48% RFP positive FCs (n=101 germaria) in homozygous flies to 14% (n=90 germaria) in heterozygous flies (Figure 33). For #5 I detected an increase in FC clones with 4% in heterozygous flies (n=122 germaria isolated from 6 flies, p=0,332, one repeat) compared to 2% in homozygous flies (n=89 germaria isolated from 6 flies, one repeat) whereas unmarked germaria were completely reduced from 11% in homozygous fies (n=89 germaria) to 0% within heterozygous flies (n=122 germaria). RFP positive FCs were minimized by 36% in homozygous flies (n=89 germaria) to 4% in heterozygous flies (n=122 germaria) (Figure 33). These outcomes confirmed that heterozygosity had significant effects on the expression rate of GFP and RFP in the respective recombinant fly line. In order to investigate FC clone induction triggered by distinct conditions, the clonal marking system needs to function reliably. Due to time constraints and the lack of reliability in GFP and RFP expression patterns of the investigated recombinant fly lines, these were not used for further implementation in this study. However, novel populations of heterozygous flies of #1, #3 and #5 were bred for future potential usage for Rust Lab, whereas particularly #1 tends to be the best suited candidate for future experiments.

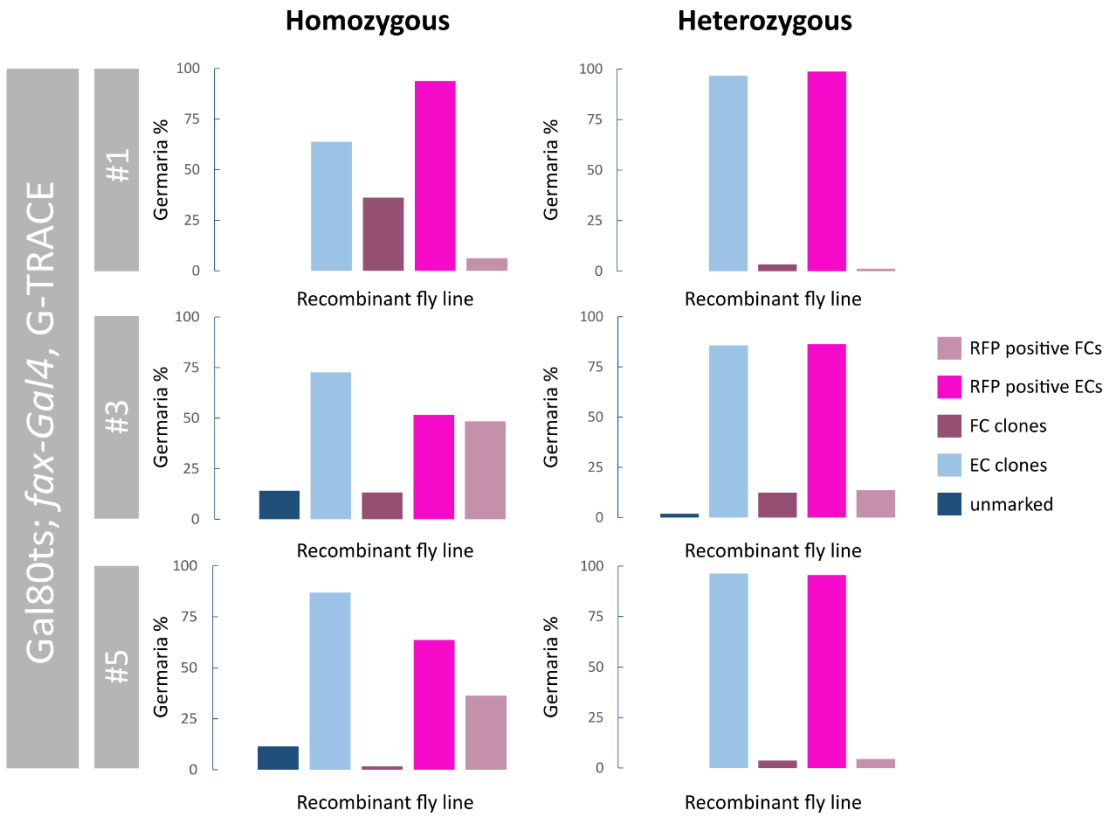


Figure 33: Heterozygosity for *fax-Gal4*, G-TRACE reduces FC clone emergence in recombinant fly lines

Bar plots for three independent recombinant fly lines (#1, #3, #5) with the genotype Gal80ts; *fax-Gal4*, G-TRACE passing 14 days rich diet at 29°C. Each fly line was sorted for homo- or heterozygosity for the feature of *fax-Gal4*, G-TRACE. Homozygous samples are presented with high ratios of germaria with FC clones (purple) and RFP positive FCs (rose). Note that the heterozygous control group reduced the amount of germaria with FC clones and RFP positive FCs for all three fly lines remarkably. #1 homozygous (n=118 germaria, 6 flies, one repeat), #1 heterozygous (n=141 germaria, 6 flies, $p<0.001$, one repeat); #3 homozygous (n=101 germaria, 6 flies, one repeat), #3 heterozygous (n=90 germaria, 6 flies, p=0,798, one repeat); #5 homozygous (n=89 germaria, 6 flies, one repeat), #5 heterozygous (n=122 germaria, 6 flies, p=0,332, one repeat).

4.2.3 Candidate screen of environmental stressors reveals identification of potential initiators on EC conversion

To investigate potential factors triggering EC conversion to FSCs, I conducted a series of experiments exposing flies to specific environmental stress conditions. Chosen conditions were established to have impacts on a flies' or human existence (Table 4).

Table 4: Test conditions and their influences on the *Drosophila* organism

Test condition	Impacts
Absence of mating partners	decreases longevity of female *D. melanogaster* (Tracey Chapman et al., 1995; Lung et al., n.d.-a; Pitnick & García-Gonzá Lez, 2002), mating regulates ovulation and egg-laying (Lung et al., n.d.-b) and the cell number in the flie's midgut (Reiff et al., 2015; White et al., 2021).
Starvation	nutritional stress; established trigger for EC conversion (Rust et al., 2020)
Protein starvation	nutritional stress; established trigger for EC conversion (Rust et al., 2020)
High fat diet	causes metabolic disorders, reduction in fertility and interferes with insulin-signaling (S. Liao et al., 2021; Nayak & Mishra, 2021)
BPA	causes fertility issues in humans (Konieczna et al., 2015); affects behavior, neuronal development and lipid metabolism in *D. melanogaster* (Nguyen et al., 2021; Williams et al., 2014)

Fundamental assumptions to this practice are based on the fact that conversion may be triggered by stress which leads to the complete loss or loss of function of FSCs (1.6.1 Diet perturbation leads to EC fate conversion). For this, fly lines with the genotype Gal80ts; *fax-Gal4*, G-TRACE were utilized. Flies were hereby set up to undergo the distinct test condition within 14 days paired with a well fed control. In this context, I sought to identify trigger conditions which lead to an increase of germaria with FC clones. For this, germaria with the emergence of FC clones were quantified. Additionally, the average number of flies with FC clones was also calculated for the respective stress condition.

4.2.3.1 Diet perturbations

I tested three types of dietary perturbations: protein starvation, in which flies are starved on sucrose solutions for three consecutive days, water starvation, where flies are only exposed to water for 24 hours and high fat diet, feeding flies a high fat nutrition (20% coconut oil) for a total of 14 days. For protein starvation the following outcomes were determined: in n=457 germaria isolated from 20 flies (three repeats), 4% revealed FC clones while 93% showed EC clones and 2% were unmarked, in contrast to the paired well fed control with 0,8% FC clones, 88% EC clones and 11,2% unmarked germaria (n=474 germaria, 16 flies, three repeats) (Figure 34 A). For this condition the p-value counts p=0,029, therefore falling under the standard threshold of 0,05, proving a significant difference of test condition versus negative control (Figure 34 B-C). Furthermore, the quantification of flies showing conversion events proves the impact of the respective test condition: 45% (9 out of 20 flies) of all flies tested for protein starvation displayed FC clones in contrast to 18,75% (3 out of 16 flies) in the control (Figure 34 D). Exposing flies to high fat food leads to FSC conversion in 1% of all cases and 95% EC clones (n=503 germaria, 14 flies, p=0,59, three repeats) in contrast to the paired control portraying only one tenth of FC clones (0,1%, n=463, 15 flies, three repeats) and 90% EC clones (Figure 34 A-B). Although the results of the high fat nutritional regimen did not produce significant outcomes according to its p-value (p=0,593) (Figure 34 B-C), I found that 35,7% (5 out of 14 flies) of all flies examined were positive for FC clones. This stands out from the paired control, as roughly 5 times more flies reveal the presence of FC clones (35,7% compared

to 6,7%, 1 out of 15 flies) (Figure 34 D). Since this approaches the average FC clone emergence for flies of the protein starvation condition (45%), high fat exposure should be considered in follow up studies on FSC conversion. For water starved flies, only in 1% of all cases FC clones were detected (n=554 germaria of 17 flies, p=0,67, three repeats) while 8% of germaria remained unmarked and 91% produced EC clones (n=554 germaria) (Figure 34 A-B). Paired well fed controls exhibit 0,4% of germaria with FC clones, 94,7% EC clones and 4,9% of unmarked germaria (n=460 germaria of 16 flies, three repeats) (Figure 34 A-B). Previous studies established starvation to be a significant inducer of FSC conversion (Rust et al., 2020). Regardless that the results for water starvation in this study could not clearly reproduce prior evidences with respect to the significance level, a trend for the increase of germaria with FC clones was noticeable: water starvation doubles the amount of flies with FC clones in contrast to its paired well fed control (starved: 23,5%, 4 out of 17 flies; control: 12,5%, 2 out of 16 flies) (Figure 34 D) while all water starved conditions together elevate the amount of flies with FC clones at an almost 3 times higher frequency compared to the sum of all well fed controls (all starved: 30%, 12 out of 40 flies; all control: 11%, 11 out of 101 flies).

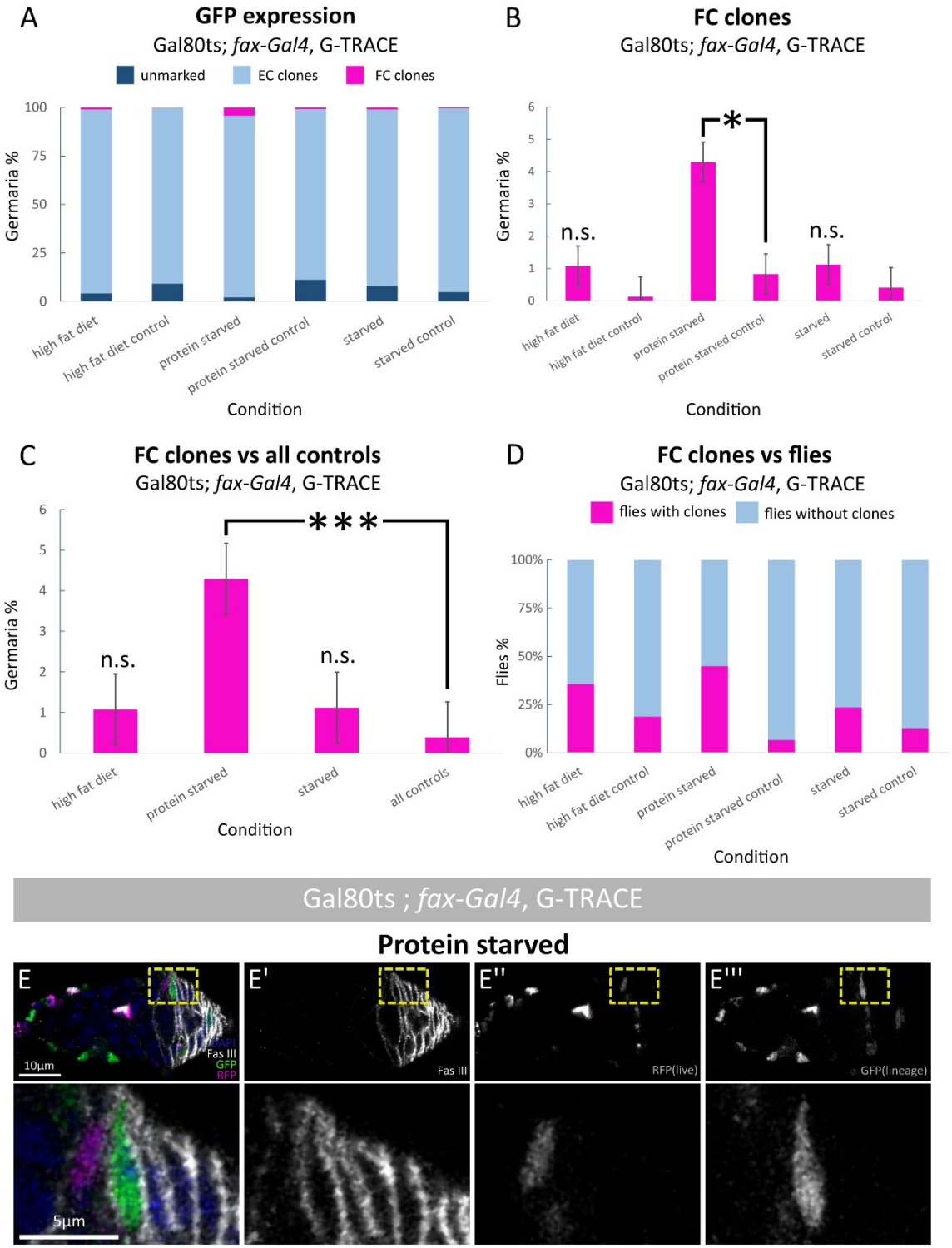


Figure 34: Protein starvation reveals the highest production of FC clones among all diet stressors

(A): Bar plot showing GFP expression in germaria of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE undergoing three different environmental stress conditions (high fat diet, protein starvation and starvation) for 14 days at 29°C. GFP expression is split for three categories: unmarked (dark blue): germaria with no measurable GFP expression; EC clones (light blue): GFP expression in ECs anterior to the Region 2a/2b border; FC

clones (magenta): GFP expression in the Fas III positive area. Each test condition is paired with a control group of flies fed standard fly food for 14 days at 29°C. FC clone emergence remained at low rates for control conditions while protein starvation in particular elevated the amount of germaria with FC clones. Protein starvation (n=457 germaria, 20 flies, 3 repeats) and its paired control (n=474 germaria, 16 flies, 3 repeats); High fat diet (n=503 germaria, 14 flies, 3 repeats) and its paired control (n=463 germaria, 15 flies, 3 repeats); Water starvation (n=554 germaria, 17 flies, 3 repeats) and its paired control (n=460 germaria, 16 flies, 3 repeats). **(B):** Closer outline of germaria with FC clones versus test condition of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE. Error bars depict S.E.M. Asterisk depicts significance (*= p<0,05; ** = p<0,01; *** = p<0,001) whereas not significant values are described with n.s. High fat diet (p=0,593, n.s.) and water starvation (p=0,67, n.s.) increased the detection of germaria positive for FC clones in comparison to their paired controls. Protein starvation elevated the ratio of FC clones notably (p=0,029; *). **(C):** Bar chart for germaria with FC clones versus test conditions of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE. The standard error is indicated by the error bars. The control condition (all controls) for this bar chart incorporates quantifications of all control conditions conducted in this study (n=2325 germaria, 101 flies, 18 repeats). High fat diet (p=0,61942894, n.s.) and water starvation (p=0,37764953, n.s.) increased the amount germaria with FC clones, while protein starvation altered the ratio strikingly (p=0,00065024, ***) in comparison to all controls. **(D):** Comparison of flies with or without clones of the genotype Gal80ts; *fax-Gal4*, G-TRACE. For all stress conditions more flies with FC clones were detected. Protein starvation (45%, 9 out of 20 flies) and its paired control (18,75%, 3 out of 16 flies), water starvation (23,5%, 4 out of 17 flies) and its paired control (12,5%, 2 out of 16 flies), high fat diet (35,7%, 5 out of 14 flies) and its paired control (6,7%, 1 out of 15 flies). **(E-E'''):** A representative image of a follicle cell clone in germaria of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE that underwent protein starvation, stained for DAPI (blue), Fas III (white), RFP (magenta) and GFP (green). **(E'):** Fas III channel in white. **(E''):** RFP channel in white. **(E'''):** GFP channel in white. Insets hint at the appearance of FC clones. These clones are present in the area of Fas III cells, but depict only GFP expression and not RFP, indicating that clones have derived from EC lineages.

4.2.3.2 Mating restraint

For investigating the influence of mating on the conversion events, virgin female *Drosophila* were collected and held separate from male flies. Following results were assessed for the mating restraint: 1% of all germaria (n=479 isolated from 16 flies, p=0,61, three repeats) displayed FC clones while 96 % showed EC clones and 3% of all germaria remained unmarked (Figure 35 A-B). The paired control delivered a detection of 0,4% of germaria with FC clones, 94,7% EC clones and 4,9% unmarked germaria (n=460 germaria isolated from 16 flies, three repeats) (Figure 35 A-B). With respect to the quantification of flies with or without clones, around the virgin section of flies (37,5%, 6 out of 16 flies) more than two times more flies with clones were detected compared to the paired control (12,5%, 2 out of 16 flies) (Figure 35 C).

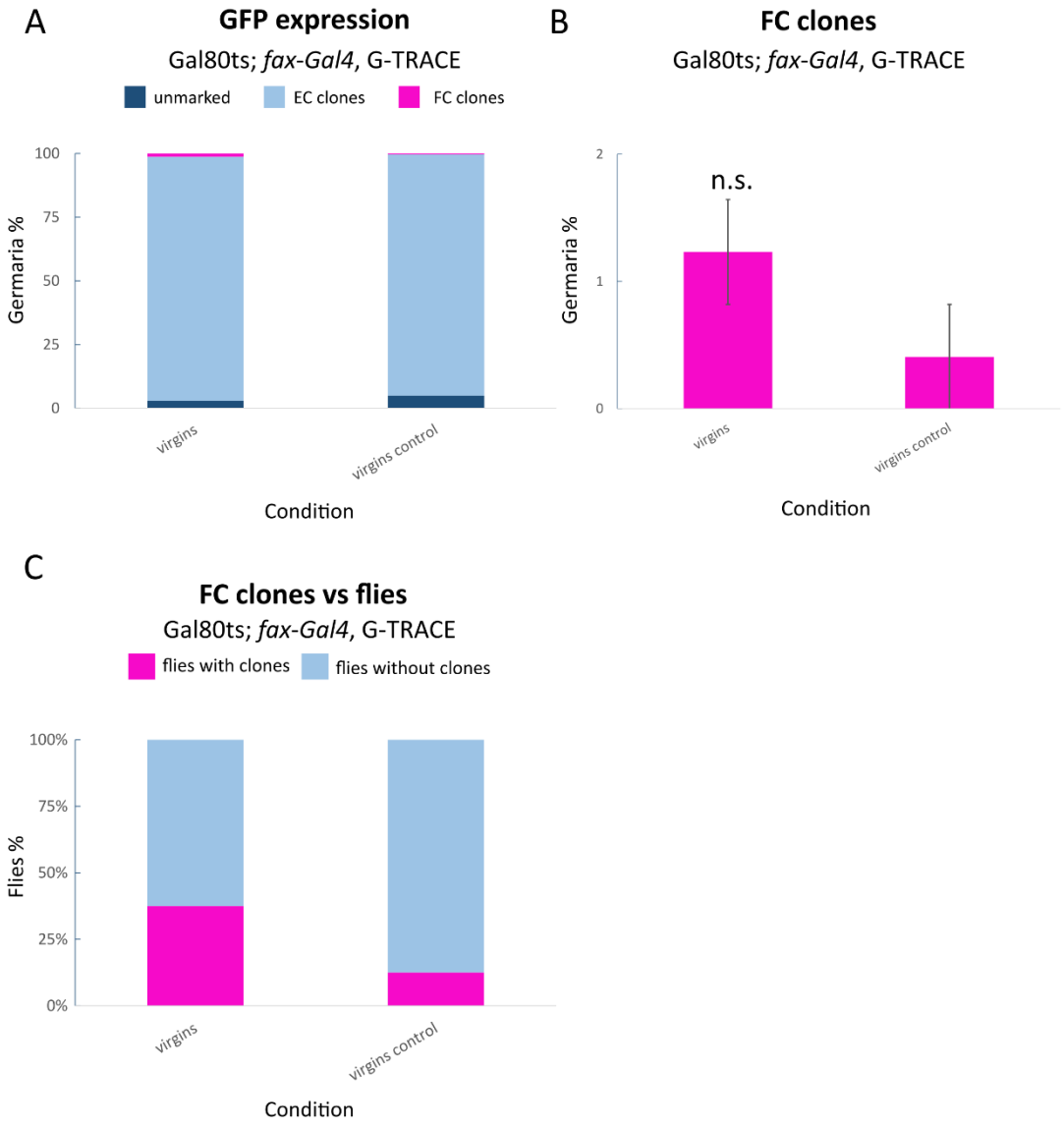


Figure 35: Mating restraint does not produce notable ratios of FC clones

Bar plots showing GFP expression **(A)**, FC clone production **(B)** and the comparison of flies with or without the presence of FC clones **(C)** for germaria of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE under mating restraint and rich diet for 14 days at 29°C. Paired controls without mating restraint were analogously fed a rich diet at 29°C. **(A):** Quantifications of germaria for three distinct GFP expression patterns: unmarked germaria, where no GFP was detected; germaria with EC clones: ECs expressing GFP anterior to the Fas III positive area; germaria with FC clones: GFP expression measurable within the Fas III positive area. The mating restrained section (virgins) (n=479, 16 flies, 3 repeats) slightly increased the amount of germaria with FC clones in comparison to its paired control (n=460 germaria, 16 flies, 3 repeats) **(B):** Virgin test conditions did not raise the amount of FC clones significantly (p=0,61, n.s.). Error bars indicate S.E.M. No significance is depicted with n.s. **(C):** The intersection of flies with FC clones was more than doubled by flies undergoing mating restraint (37,5%, 6 out of 16 flies) in comparison to its control (12,5%, 2 out of 16 flies).

4.2.3.3 BPA exposure

Since BPA is known to affect human fertility and is broadly present in our environment (Konieczna et al., 2015; Mustieles et al., 2020; Rochester, 2013), I examined whether BPA influences EC conversion in *Drosophila*. Therefore, flies were exposed to fly food which contained different concentrations of BPA for 14 consecutive days (0,1mM, 1mM, 20mM). The concentration of 0,1mM of BPA produced in 2% of all cases FC clones, 93% EC clones and 5% unmarked germaria (n=452 germaria isolated from 17 flies, p=0,35, three repeats) (Figure 36 A-B); fly food with a concentration of 1mM BPA led to 3% FC clones, 7% unmarked germaria and 90% EC clones (n=435 germaria isolated from 15 flies, p=0,12, three repeats) (Figure 36 A-B); 20mM of BPA generated 2% of FC clones, 92% EC clones and 6% unmarked germaria (n=464 germaria isolated from 15 flies, p=0,34, three repeats) (Figure 36 A-B). In general, the results of the BPA testings clearly differ from the paired well fed controls by exposing FC clones in only 0,4 % of all germaria, 94,7 % EC clones and 4,9% unmarked germaria (n=460 germaria isolated from 16 flies, three repeats) (Figure 36 A-B). While generated p-values depict non-significance in contrast to their paired control (Figure 36 B), I found that by comparing 1mM of BPA to all controls conducted in

this study (n=2325 germaria, 101 flies, 18 repeats) delineated a significant difference with respect to FC clone production (p=0,046) (Figure 36 C). Additionally, I scored the second highest amount for the quantification of flies with clones within this condition: for the concentration of 1mM BPA the average amount of flies with clones counts 60% (9 out of 15 flies) (Figure 36 D). Compared to the protein starvation condition (see 4.2.3.1 Diet perturbations; 45%, 9 out 20 flies), this implies that BPA portrays an initiator to FSC conversion, pending additional experiments to confirm this assumption. Interestingly, the exposure to 20mM of BPA was expected to produce even more FC clones than the lower concentrations. Instead, it represented rather a decrease with respect to the detection of germaria with FC clones (BPA 0,1mM 2%; BPA 1mM 3%; BPA 20mM 2%). Quantifications for the amount of flies with FC clones (26,7%, 4 out of 15 flies) resulted in a lower intersection than for 0,1mM of BPA (33,33%, 5 out of 15 flies) and 1mM of BPA (60%, 9 out of 15 flies) (Figure 36 D).

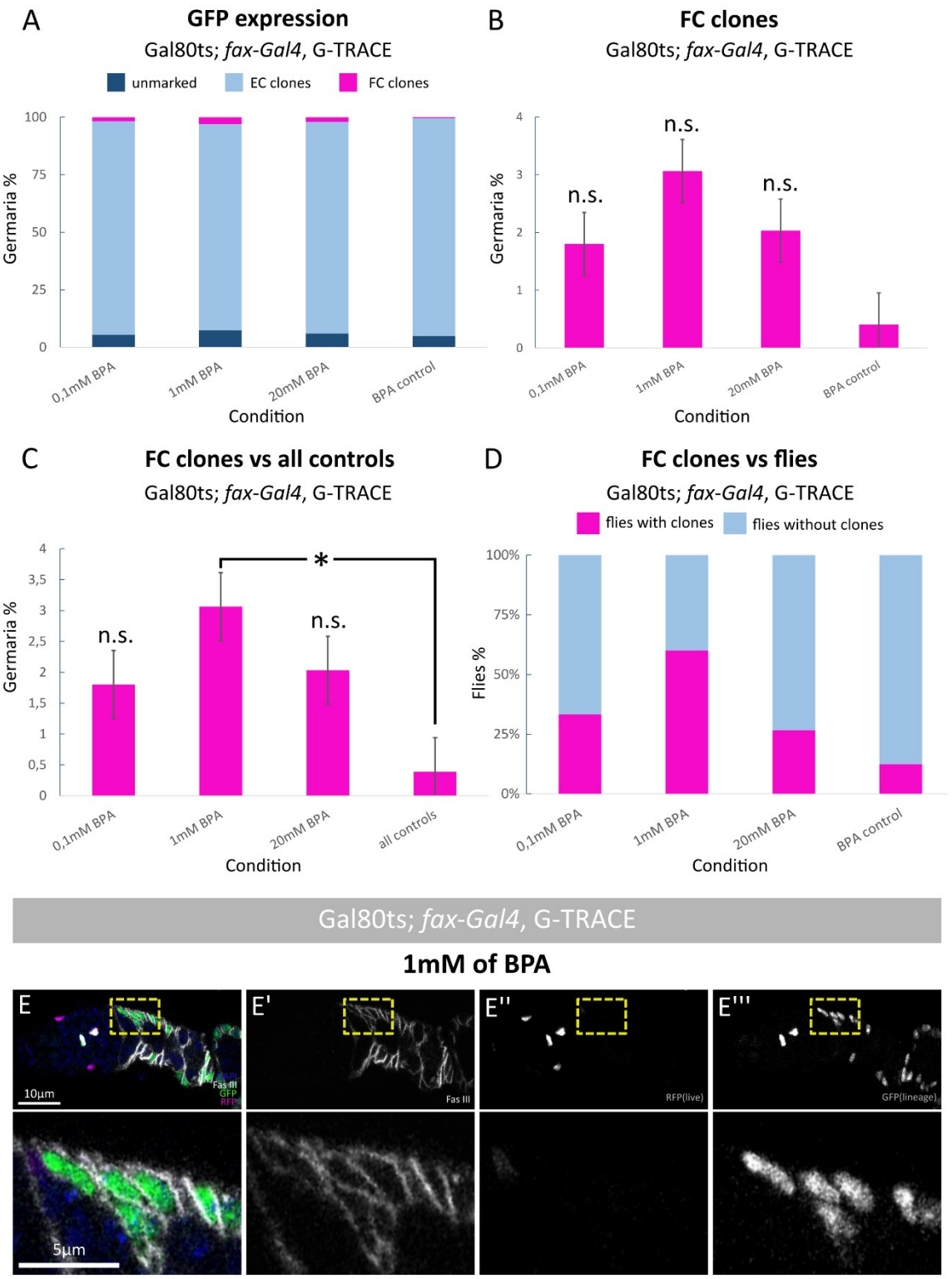


Figure 36: BPA exposure hints to depict a trigger for EC conversion

(A): Quantifications of germaria for three patterns of GFP expression versus three different concentrations of BPA and their paired control. Germaria were extracted from flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE that were exposed to 14 days of BPA contaminated food in different concentrations. Flies for the paired control (BPA control) were fed standard fly food for 14 consecutive days. Germaria were quantified by

features of not showing GFP expression (unmarked, dark blue), GFP expression in EC clones anterior to the Region 2a/2b border (EC clones, light blue) and GFP expression detectable in the Fas III positive follicle cell area (FC clones, magenta). All BPA conditions raised the amount of germaria with FC clones in comparison to the control condition. 0,1mM of BPA (n=452 germaria, 17 flies, 3 repeats), 1mM of BPA (n=435 germaria, 15 flies, 3 repeats), 20mM of BPA (n=464 germaria, 15 flies, 3 repeats), paired control (n=460 germaria, 16 flies, 3 repeats). **(B):** Germaria with FC clones of the genotype Gal80ts; *fax-Gal4*, G-TRACE versus BPA test conditions. Error indicators show the standard error. 0,1mM of BPA (p=0,35, n.s.) and 20mM of BPA (p=0,34, n.s.) delivered a 5 fold increase of FC clone emergence in comparison to the paired control. 1mM of BPA (p=0,12, n.s.) heightened the amount of germaria with FC clones by 7,5 with respect to the paired control. N.s. indicates no significance (p>0,05). **(C):** Bar plot showing the comparison of germaria with FC clones of the genotype Gal80ts; *fax-Gal4*, G-TRACE for BPA exposure and all control conditions executed in this study (all controls, n=2325 germaria, 101 flies, 18 repeats). The S.E.M. is indicated by error bars. Asterisk indicates significance (* = p<0,05; ** = p<0,01; *** = p<0,001). 1mM of BPA (p=0,04647188, *) increases the amount of FC clones significantly in comparison to all controls conducted. **(D):** Line-up of flies with or without clones for BPA and control conditions. The intersection of FC clones in flies is increased for all BPA concentrations. 1mM of BPA demonstrates the highest ratio (60%, 9 out of 15 flies) in comparison to its paired control with 12,5% (2 out of 16 flies). **(E-E'''):** A representative depiction of a follicle cell clone induced by 1mM of BPA in germaria of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE. Germaria were stained for DAPI (blue), Fas III (white), RFP (magenta) and GFP (green). **(E'):** Fas III channel in white. **(E''):** RFP channel in white, showing RFP live expression **(E'''):** GFP channel in white, depicting the GFP lineage. Insets point out the appearance of FC clones present in the Fas III positive area. As FC clones are not RFP positive, these clones must have derived from EC lineages anterior to the Region 2a/2b border.

4.2.4 The Toll signaling pathway in EC conversion

Based on previous studies elucidating that the overexpression of the Toll pathway leads to the emergence of FC clones at a significant higher frequency (Rust et al., 2020), I examined the behavior of EC conversion via knockdown of *Toll* components under standard well fed versus starvation conditions. This experiment served to further investigate whether *Toll* represents a regulator for niche conversion processes. For downregulation of the Toll pathway, two distinct

RNAi fly lines were utilized: *UAS-tl-RNAi* for the knockdown of the *Toll* receptor and *UAS-dl-RNAi* silencing the dorsal protein within the pathway. In order to obtain the desired genotype, I combined the RNAi constructs with G-TRACE, Gal80ts and *fax-Gal4*. A paired control of the standard fly line Gal80ts; *fax-Gal4*, G-TRACE underwent the same conditions of rich diet and starvation procedures. For the inquiry of flies kept under well fed conditions, the following outcomes were assessed: *tl-RNAi* revealed the emergence of germaria with FC clones in 0,6% of all cases, 76,7% EC clones and a high amount of unmarked germaria with 22,7% (n=492 isolated from 18 flies, p=0,75, three repeats) (Figure 37 A-B). The knockdown of the dorsal protein produced only 0,2% FC clones, 87,8% EC clones and 12% unmarked germaria (n=450 germaria isolated from 22 flies, p=0,92, three repeats) (Figure 37 A-B). Interestingly, the paired control did not show any FC clones but 87% EC clones and a high number of unmarked germaria as well (13%, n=446 germaria isolated from 18 flies, three repeats) (Figure 37 A-B). Since the amount of unmarked germaria appeared to be high for all of the rich diet experiments, a separate analysis of calculating the FC clone percentage was conducted by rejecting the number of unmarked germaria from the calculations. I reasoned that considering germaria without GFP positive ECs would lead to an underestimation of the conversion rate, since conversion events can only be tracked when ECs express GFP. However, these calculations did not deliver strikingly different numbers with respect to an increase of the FC clone percentage among the RNAi fly lines (0,1% increase for both conditions). I also compared the RNAi samples with the total amount of all controls conducted in this study (n=2325 germaria, 101 flies, 18 repeats). Expectantly, both RNAi fly lines did not differ significantly from other controls (*dl-RNAi*, p=0,82; *tl-RNAi*, p=0,93) (Figure 37 C). Moreover, *tl-RNAi* flies displayed an average number of flies with clones of 5,5% (1 out of 18 flies), while the ratio for *dl-RNAi* samples resides at 4,5% (1 out of 22 flies) (Figure 37 D).

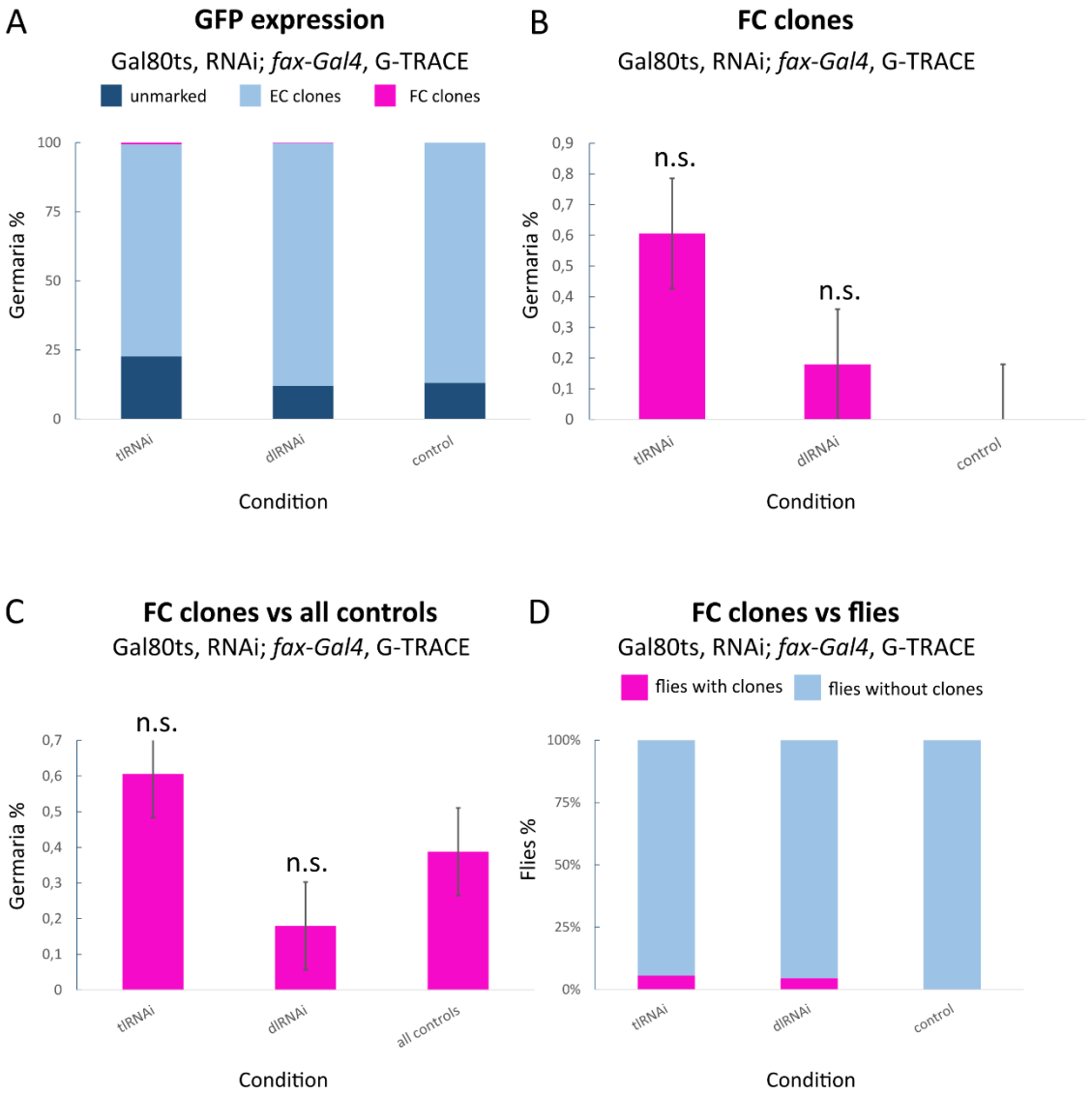


Figure 37: RNAi fly lines deliver overall low percentages of FC clones

(A) Column diagram comparing GFP expression in germaria of flies with the genotype Gal80ts, RNAi; *fax-Gal4*, G-TRACE for *Toll* knockdown fly lines and Gal80ts; *fax-Gal4*, G-TRACE for the paired control. GFP expression is split for three expression patterns: unmarked germaria depict germaria with no measurable GFP expression; EC clones constitute GFP expression in ECs anterior to the Region 2a/2b border; FC clones show GFP expression in the Fas III positive area posterior to the Region 2a/2b border. RNAi samples (*tl-RNAi* n=492 germaria, 18 flies, 3 repeats; *dl-RNAi* n=450 germaria, 22 flies, 3 repeats) and the paired control condition (n=446 germaria, 18 flies, 3 repeats) do not produce notable amounts of FC clones. **(B)** Bar plot of RNAi conditions and their paired control versus percentages of germaria with FC clones. Error bars indicate the standard error. No significant increase in germaria with FC clones was detected for RNAi fly lines (*tl-RNAi* p=0,75; *dl-RNAi* p=0,92) and is depicted with n.s. (no significance, p>0,05). **(C)** Bar plot of RNAi test conditions versus the sum of all controls in this study (n=2325

germaria, 101 flies, 18 repeats). RNAi fly lines do not differ significantly from control conditions with respect to FC clone production (*tl-RNAi* p=0,93; *dl-RNAi* p=0,82). **(D)** Bar plot comparison of flies with or without clones for RNAi fly lines and their paired control. Low percentages were detected for the intersection of flies positive for FC clones (*tl-RNAi* = 5,5%, 1 out of 18 flies; *dl-RNAi* = 4,5%, 1 out of 22 flies, control = 0%, 0 out of 18 flies).

To confirm that perturbations in the Toll pathway may lead to a rescue of starvation-induced EC conversion, *tl-* and *dl-RNAi* flies were water starved for 24 hours. Flies of the genotype Gal80ts; *fax-Gal4*, G-TRACE were starved analogously for the starved control and richly fed for 14 days for the well fed control. Confirming the trend that starvation induces EC conversion, starved flies of the control group produced four times more germaria with FC clones (1,3%, n=554 germaria, 23 flies, p=0,55, three repeats) than well fed flies of the control group (0,3%, n=459 germaria, 16 flies, three repeats) (Figure 38 A-B). Knockdown of *Toll* components produced even higher levels of FC clones in response to starvation: 4,2% for the knockdown of the dorsal protein (n=457 germaria isolated from 15 flies, p=0,07, three repeats) and 1,5% for the *Toll* receptor knockdown (n=482 germaria isolated from 16 flies, p=0,48, three repeats) (Figure 38 A-B). Testing starved RNAi samples against the sum of all controls conducted in this study (n=2325 germaria, 101 flies, 18 repeats), a significant difference with respect to FC clone emergence is displayed for *dl-RNAi* populations (p=0,023) (Figure 38 C). Consistent with this trend, starved *dl-RNAi* samples produced significantly more FC clones in contrast to their well fed *dl-RNAi* control (p=0,042). This trend is reflected in the comparison of flies with or without clones: the amount of flies with clones in knockdown samples, 53,3% for dorsal knockdown (8 out of 15 flies) and 43,75% for *Toll* receptor knockdown (7 out of 16 flies), remained to be higher in comparison to the starved flies of the control group (21,7%, 5 out of 23 flies) (Figure 38 D) and their well fed RNAi controls (*dl-RNAI*, 4,5%; *tl-RNAi*, 5,5%). These outcomes indicate that *Toll* knockdown does not seem to rescue starvation induced EC conversion.

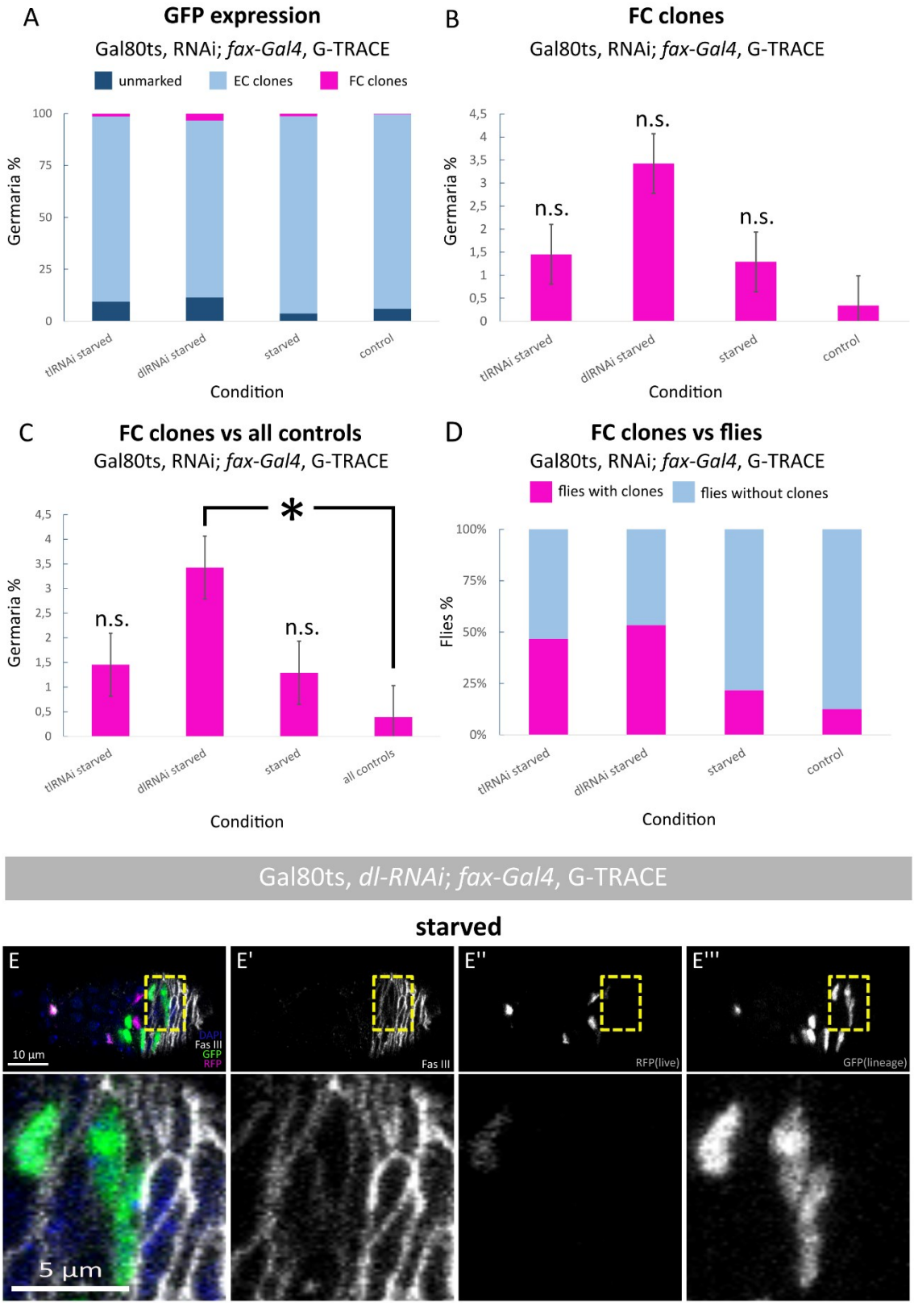


Figure 38: High ratios of FC clones obtained from starved RNAi fly lines

Quantifications of **(A)** GFP expression, **(B)** and **(C)** germaria with FC clones and **(D)** division of flies with or without clones for flies of the genotype Gal80ts, RNAi; *fax-Gal4*, G-TRACE undergoing starvation and Gal80ts; *fax-Gal4*, G-TRACE for the paired control conditions. Error bars in **(B)** and **(C)** indicate S.E.M. Significance is depicted with

asterisks (* = p<0,05; ** = p<0,01; *** = p<0,001) while no significance is described with n.s. (p>0,05). **(A)** Germaria are quantified by three distinct GFP appearances: unmarked (dark blue) for germaria with no detectable GFP expression; EC clones (light blue) for GFP in ECs anterior to the Region 2a/2b border and FC clones (magenta) for GFP expressing cells in the Fas III positive area. All starvation samples increase the amount of germaria with FC clones in comparison to the well fed control. Note that RNAi fly lines particularly show high percentages of FC clones in germaria. *dl-RNAi* starved (n=457 germaria, 15 flies, three repeats), *tl-RNAi* starved (n=482 germaria, 16 flies three repeats), control (well fed) (n=459 germaria, 16 flies, 3 repeats), control (starved) (n=554 germaria, 23 flies, 3 repeats). **(B)** Starved (RNAi) fly lines in comparison to their paired control. The starved control group (p=0,55, n.s.) and *tl-RNAi* starved fly lines (p=0,48, n.s.) double the amount of germaria with FC clones, while starvation of *dl-RNAi* fly lines (0,07, n.s.) lead to a five fold increase. **(C)** Germaria with FC clones in starved (RNAi) fly lines versus the sum of all control conditions of this study (n=2325 germaria, 101 flies, 18 repeats). A notable increase is detected for starvation of *dl-RNAi* fly lines (p=0,02388815, *). **(D)** For both starved RNAi fly lines, FC clones are detected in approximately half of all flies examined (*dl-RNAi* starved 53,3%, 8 out of 15 flies; *tl-RNAi* starved 43,75%, 7 out of 16 flies; control (well fed) 12,5%, 2 out of 16 flies; control (starved) 21,7%, 5 out of 23 flies). **(E)** Follicle cell clone representation in germaria of starved flies of the genotype Gal80ts, *dl-RNAi*; *fax-Gal4*, G-TRACE stained for DAPI (blue), Fas III (white), GFP (green) and RFP (magenta). **(E'-E''')** Fas III, RFP and GFP channels are depicted in white. FC clones appear in the Fas III positive area posterior to the Region 2a/2b border. These clones are GFP positive but RFP negative, which indicates the derivation of EC lineages.

4.2.5 Increase of RFP positive FCs in stress conditions

As *fax-Gal4* activity was priorly described to be limited to the EC compartment (Rust et al., 2020), I expected to only detect GFP in FCs but no RFP fluorescence. However, I observed RFP expression exceeding the EC compartment into the Fas III positive area of FCs in the absence of GFP positive FC clones. These observations were determined for all control as well as stressor conditions, however, the amount of RFP positive FCs was particularly increased for stress conditions in comparison to well fed controls. Figure 39 shows the rate of RFP positive FCs among the different test conditions: the most prominent example is hereby portrayed by the protein starvation test condition, which demonstrates a fivefold increase of RFP positive FCs, with 15% (n=457 germaria, 20 flies,

p<0,0001, three repeats) compared to its well fed control with only 3% (n=474 germaria, 16 flies, three repeats). Flies exposed to a high fat diet contained germaria with RFP positive FCs in 7% (n=503 germaria, 14 flies, p=0,002, three repeats) of all cases, while its paired control revealed the lowest levels of all conditions taken together (0,8%, n=503 germaria, 15 flies, three repeats). RFP activity in RNAi fly lines however was obviously narrowed: starved flies in this context produced 6% of germaria with RFP positive FCs (n=554 germaria, 23 flies, p=0,077, three repeats), followed by a more than eightfold lower percentage for *tl-RNAi* starved flies (0,7%, n=482 germaria, 16 flies, p=0,275, three repeats) and more than fourfold lower percentage for *dl-RNAi* starved flies (1,38%, n=457 germaria, 15 flies, p=0,447, three repeats). Paired well fed controls in both RNAi conditions, well fed and starved, produced each an amount of about 3% of germaria with RFP positive FCs (well fed circuit: n=446 germaria, 18 flies, three repeats; starved circuit: n=459 germaria, 16 flies, three repeats). Both well fed RNAi fly lines showed germaria with RFP positive FCs only in 0,7% of all cases (*dl-RNAi*: n=450, 22 flies, p=0,665, three repeats; *tl-RNAi*: n=492, 18 flies, p=0,664, three repeats). While the amount of GFP positive clones for virgin female flies were at rather low levels (see 4.2.3.2 Mating restraint), 9% of germaria contained RFP positive FCs (n=479 germaria, 16 flies, p=0,010, three repeats). Almost equally high numbers were detected for 1mM of BPA (10%, n=435 germaria, 15 flies, p=0,0052, three repeats) and 20mM of BPA (8%, n=464 germaria, 15 flies, p=0,052, three repeats); 0,1mM of BPA displayed 12% of germaria with RFP positive FCs (n=452 germaria, 17 flies, p=0,0002, three repeats) whereas starved samples delivered a total amount of 6% (n=482 germaria, 17 flies, p=0,46, three repeats). The paired well fed control for mating restraint, BPA and starvation investigations displayed RFP positive FCs in 4% of all germaria (n=460 germaria, 16 flies, three repeats).

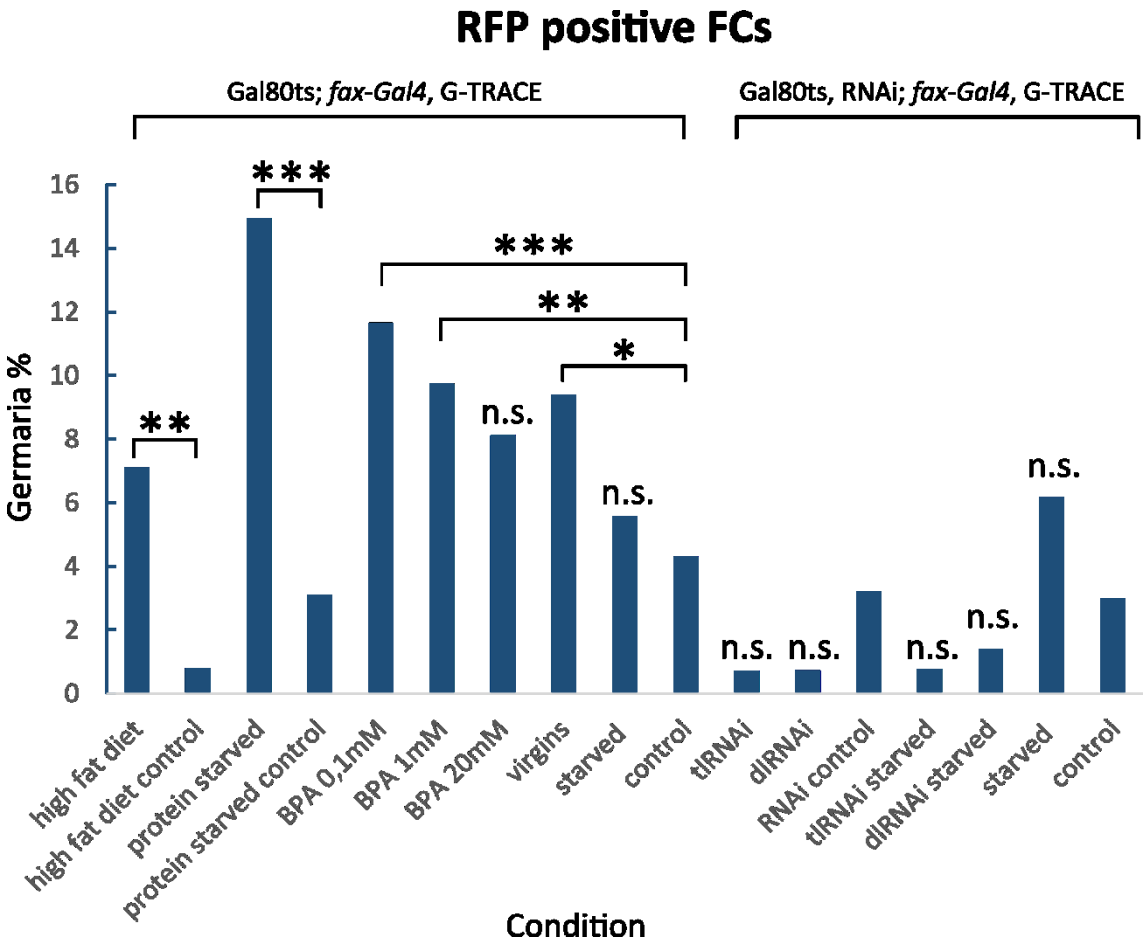


Figure 39: RFP positive FCs are distinctly increased in stress conditions

Quantification of RFP positive FCs in germaria of flies with the genotype Gal80ts; *fax-Gal4*, G-TRACE and Gal80ts, RNAi; *fax-Gal4*, G-TRACE of all conditions in this study. Significance is depicted with asterisks (* = p<0,05; ** = p<0,01; *** = p<0,001) and non significance is illustrated with n.s. (p>0,05). Significance values (p) are calculated for each condition in comparison to their respective paired control. Number (n) of analyzed germaria for each condition from left to right: high fat diet (n=503, p=0,002, 14 flies, 3 repeats), high fat diet control (n=463, 15 flies, 3 repeats), protein starvation (n=457, 20 flies, p<0,0001, 3 repeats), protein starvation control (n=474, 16 flies, 3 repeats), BPA 0,1mM (n=452, 17 flies, p=0,0002, 3 repeats), BPA 1mM (n=435, 15 flies, p=0,0052, 3 repeats), BPA 20mM (n=464, 15 flies, p=0,052, 3 repeats), virgins/mating restraint (n=479, 16 flies, p=0,010, 3 repeats), starved (diet perturbation) (n=482, 17 flies, p=0,46, 3 repeats), control for BPA, virgins and starvation (n=460, 16 flies, 3 repeats), *tl-RNAi* (n=492, 18 flies, p=0,664, 3 repeats), *dl-RNAi* (n=450, 22 flies, p=0,665, 3 repeats), RNAi control (n=446, 18 flies, 3 repeats), *tl-RNAi* starved (n=482, 16 flies, p=0,275, 3 repeats), *dl-RNAi* starved (n=457, 15 flies, p=0,447, 3 repeats), starved (starved control in RNAi starvation circuit) (n=554, 23 flies, p=0,077, 3 repeats), RNAi starved control (n=459, 16 flies, 3 repeats). Note that RFP positive FCs are particularly elevated for all stress conditions in comparison to their paired controls.

5. Discussion

Understanding the nature of stem cells depicts fundamental baselines for abundant biological and biomedical research. Ever since it was discovered that stem cells could be applied as therapeutic mediums, the hope arose to cure or at least successfully treat diseases thought to be irremediable. For this, not only are the crucial key events of stem cell maintenance decisive per se, but also their behavior towards their niches and responding to environmental conditions. In this context, the FSC lineage offers distinct benefits, as the ovarian tissue may easily be extracted for analysis. Moreover, FSCs depict the most proliferative lineage in *Drosophila* and several genetic pathways are conserved between the insect and humans. Here I have investigated how niche cells behave upon exposure towards different environmental conditions. Furthermore, I examined whether novel marker genes for ECs, which function as niche cells for FSCs, could be calculated by analyzing the scRNA-Seq FCA EC dataset. I provide evidence that niche cells of the FSC lineage in *Drosophila melanogaster*, have the potential to convert to FSCs upon nutritional perturbations.

5.1 *Fax-Gal4* driver reveals technical difficulties in screen design for monitoring EC lineages

In this study I made use of a Gal4 driver line induced via the promoter for the *failed axon connections (fax)* gene, which encodes a membrane protein expressed throughout the whole EC population. This Gal4 driver was used to induce the G-TRACE lineage tracing system. The *fax-Gal4* was previously identified to be weak and only sparsely labeled ECs (Rust et al., 2020). Consistent with this, I established similar conclusions: the *fax-Gal4* driver set constitutes a moderate driver, which revealed technical difficulties and are discussed in the following.

5.1.1 Driver strength and Gal4 status

To categorize driver strength for *fax-Gal4* in this context, driver intensity depends on distinct parameters: by using G-TRACE, high activity of driver are necessary to provoke GFP than RFP expression (see 1.1.1.2 Gal4 technique for real-time and

clonal expression (G-TRACE)); strong drivers therefore result in high amounts of Gal4, thus overcoming a threshold effect and being able to mark more GFP positive cells (Evans et al., 2009). This was reflected in my data, as the experimental set up of the direct comparison between homozygous and heterozygous genotypes of flies delivered evident results (4.2.2 Validation of recombinant fly lines). Interestingly, the surprisingly high frequency of GFP as well as RFP expression among FCs in well fed conditions for the recombinant genotypes, was reduced by the control group of heterozygous flies (Figure 33). This validates the assumption that homozygosity involved a presumably two fold higher amount of driver and Gal4 production which consequently produced more RFP and GFP marked cells. Since Gal4 is strongly elevated in homozygosity for *fax-Gal4*, G-TRACE, I reasoned that Gal4 output and consequently driver strength resides at much lower levels for the standard heterozygous conditions. In the heterozygous context, I observed a rather random marking pattern of ECs for all conditions, as expected, and little to no expression in the FC lineage. In general, more cells per germarium were positive for RFP than GFP and unmarked germaria were observed throughout. The assessed marking scheme proves the marking variability when using *fax-Gal4* to drive G-TRACE, which is consistent with prior reports of marking patterns for weaker drivers (Evans et al., 2009).

5.1.2 Cell specificity of the *fax-Gal4* driver

Desired characteristics of driver sets monitoring ECs in the germarium, comprise not only driver strength but also cell specific expression. *Fax-Gal4* is only active within the EC population, as RFP expression was detected among ECs and GFP positive cells were mostly limited to the EC compartment (Rust et al., 2020). Contrary to this, I detected RFP positive FCs, both in well fed as well as in nutritional and knockdown test conditions (see 4.2.5 Increase of RFP positive FCs in stress conditions). Interestingly, the amount of RFP expression was distinctly elevated in stress conditions versus their controls (Figure 39). Two conditions may result in RFP positive FCs. First, RFP detections among FCs could possibly be attributed to the expression of *fax* within FCs. By plotting *fax* in scRNA-Seq datasets containing FCs as well as ECs, *fax* is as well active around FSCs although more sparsely than in ECs (Rust et al., 2020). Hence, low levels of *fax*

promoter activity may explain the observations of RFP expression in FSCs. Such a *fax* expression in FCs could elicit the emergence of GFP positive FC clones, thus making it difficult to monitor EC conversion. However, it is established that *fax* is insulin-signaling mediated and therefore down regulated via poor dietary (Hsu & Drummond-Barbosa, 2017; Su et al., 2018). Since I detected relatively high rates of RFP expression around starvation assays (see Figure 39), it is thus not likely that regular *fax* expression in FCs is responsible for RFP and subsequent nascent GFP marked cells. Moreover, Rust *et al.* could show that GFP positive FC clones also appeared at 18°C when Gal4 is repressed by Gal80, confuting the activity of *fax-Gal4* in FCs (Supplementary Information in Rust et al., 2020). Why then is it possible that RFP expression is evidently increased in the course of different test conditions? Similar observations are assessed by the usage of another EC driver set (data not shown), also displaying RFP expression in FCs in combination with GFP marked FC clones. Second, it is reasonable that the detected RFP was inherited to FCs in recent EC conversion events resulting from GFP unmarked ECs. This explanation is in agreement with more frequent RFP positive FCs in stress conditions increasing EC conversion. Otherwise, fluorescence activity is associated with a certain half-life, leaving the possibility open that during the time point of observation, detected cells did not degrade all of the RFP yet (Evans et al., 2009).

5.1.3 Statistics

Despite the technical difficulties of the screen design, this study produced significant outcomes for certain test conditions, implying that *fax-Gal4* in combination with the G-TRACE lineage system can be employed to monitor EC conversion events. Owing to the tools, the conversion event can only be backtraced with low frequencies. As this study was seeking to examine a broad set of environmental conditions, it therefore incorporated a sample size of n=150 germaria with a confidence level of 85%. In conclusion, subsequent follow ups should examine trigger events detected through this screen in detail. Hereby, the application of larger sample sizes should be considered in order to exclude “jackpot events”. These are understood as the detection of rare phenomena which appear to have a greater impact in a distinct sample size. For a confidence

level of at least 95%, an evaluation of n=240 germaria should be incorporated in future follow ups (according to Cochran's formula for finite population sizes; Glen, n.d.).

5.2 Identification of escort cell specific tools by single cell RNA-sequencing

The knowledge that the method may be improved by the usage of a different driver set, led me to explore novel EC specific marker genes. For this, I attempted to identify genes specific for each EC subtype by analyzing a single nuclei RNA-sequencing dataset containing EC transcriptomes (H. Li et al., 2022). In addition, I integrated the dataset with other published EC single cell RNA-sequencing datasets (Jevitt et al., 2020; Rust et al., 2020; Shi et al., 2021; Slaidina et al., 2021; Tu et al., 2021). Notably, technical difficulties occurred during the integration procedure: required cell clustering was not feasible (see 4.1.2.1 Data integration of EC datasets, Figure 24). This inhibited further desired calculations for EC subtype specific marker genes. Further, the FCA EC dataset was supposed to contain mostly ECs, as the cells were marked with the EC-specific *Wnt4*>GFP, FAC-sorted and sequenced (H. Li et al., 2022). However, marker genes of cell types identified in the FCA EC dataset were either found to be expressed in stalk cells or no expression in ECs could be detected (see 4.1.2.2 No EC marker genes identified by the validation of FCA EC markers). Moreover, cell identity prediction using a confirmed dataset with all ovarian cell types (Rust et al., 2020) revealed that the FCA EC dataset was not enriched for ECs (Figure 27 F). Consistent with these findings, the research on the genetic background of selected marker genes of the FCA EC dataset (see Table 1) reveal gene functions which are attributable to cell populations other than ECs: *mtrm* and *stet* are expected to be associated with GSCs, *Nox* is anticipated in muscle cells and *DOR* is expected in later stages of oogenesis. This ultimately validated the conclusion that the FCA EC dataset does not contain ECs but cells of other ovarian origin for the most part. In agreement with cell sorting efficiency limitations of FACS (X. Liao et al., 2016), technical troubles in FAC-sorting may have caused the FCA EC dataset not containing ECs widely. Otherwise, the selection of an unwanted cell type for FAC-sorting explains the capture of cell identities other than ECs alternatively. In summary, another

attempt may be executed for the integration of all EC datasets, as this analysis would offer higher resolutions with respect to the incorporated cell amount.

5.3 Fas III status of FSCs

Many studies have established that FSCs are Fas III positive (Fadiga & Nystul, 2019; Kirilly et al., 2005; Margolis & Spradling, 1995; Nystul & Spradling, 2007b). Yet, a recent study has challenged this notion: Reilein et al (Reilein et al., 2017) propose the existence of 16 FSCs present in three rings, with a posterior ring of Fas III positive cells and two immediately anterior located rings containing Fas III negative cells (see 1.5.3 Follicle stem cells and niche morphology). These anterior rings contain cells that are considered ECs in this study. Stem cells are defined as those cells that are capable of producing differentiating daughter cells under normal conditions. Evaluating the experimental setup, it is notable that GFP positive FC clones appear only with a low frequency among all well fed control conditions (see 4.2.1 Experimental design, Figure 30). With the usage of *fax-Gal4*, I acquired fluorescence marked cells immediately anterior to the Fas III positive border on a regular basis. When deliberating that cells adjacent to this border could be FSCs, GFP positive FC clones would have been expected regularly in all control conditions. Instead, these FC clones in well fed conditions are more likely to stem from the fact that ECs convert to FSCs sporadically or that *fax* may potentially be expressed in FSCs on rare occasions as well (see 5.1.2 Cell specificity of the *fax-Gal4* driver). Moreover, this study provides evidence that clone production is influenced by a set of different stressors, strongly arguing for the presence of conversion events. Although the field debates about the Fas III status of FSCs, my data classes with the hypothesis that FSCs are Fas III positive. I therefore conclude that the posterior compartment at the 2a/2b border comprises ECs, which possess the capability to convert to FSCs under certain circumstances.

5.4 Metabolic regimen displays a critical role in EC conversion

Drosophila melanogaster has been widely used for nutritional research. Growing interest arose for medicine, as phenotypic expressions of diet-induced human diseases may be mimicked (Lüersen et al., 2019). Distinct parameters such as life span and fertility may be attributed to nutritional modifications (Staats et al., 2018). Therefore, a flie's metabolic fitness is well known to be linked to its diet (Tennessen et al., 2014). There are several approaches in interfering with the metabolic state by changing the intake of food, which ranges from total starvation to the exposure to high sugars. The present study incorporated different diet variables in order to unravel their effects on EC conversion events. Evaluating the outcomes for the diet assays, I conclude that strong stressors are necessary to induce cell fate switching. These events possibly rely on drastic impacts on FSCs, such as loss or disruption of their function triggered by different conditions (Figure 40).

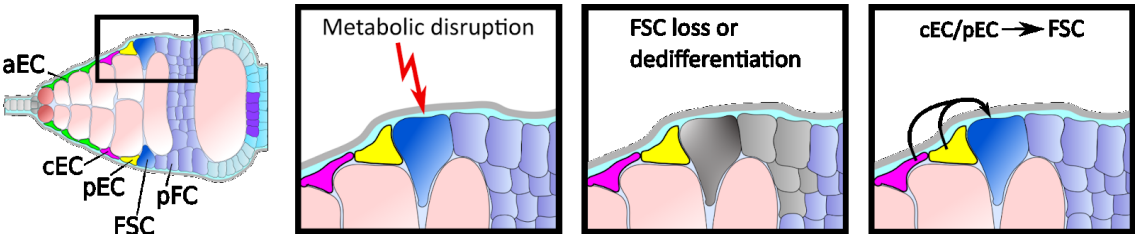


Figure 40: Metabolism as elicitor for EC conversion

Paradigm of EC conversion upon metabolic disruption. External circumstances, which impair proper metabolism, might lead to a loss or loss of function of FSCs. In order to maintain a functional epithelium, cECs/pECs undergo cell fate switching and convert to FSCs to preserve tissue homeostasis. Figure produced by courtesy of Katja Rust.

5.4.1 Diet induced cell fate switching

Nutritional deprivation causes EC conversion (Rust et al., 2020). In this project, several starvation assays were applied. In contrast to previously published results, water starvation did not trigger significant levels of EC conversion. This fact may be owed to the above discussed low activity of the *fax-Gal4* tool (see 5.1.1 Driver strength and Gal4 status). Nevertheless, I still propose starvation to induce EC conversion relevantly, as RNAi starvation assays (considered further below)

delivered obvious outcomes (see 4.2.4 The Toll signaling pathway in EC conversion, Figure 38). Moreover, protein deprivation by starving flies on sucrose solution gathered significant statistics of triggering EC conversion (4.2.3.1 Diet perturbations, Figure 34). These findings delineate consistency with prior studies establishing protein starvation as a trigger for EC conversion (Rust et al., 2020). Therefore, I reason that protein deprivation as a modification of starvation promotes niche plasticity. Previous studies have shown that protein starvation regulates the evolutionary conserved PI3K and TOR pathways in *Drosophila* (L. Li et al., 2010; Tettweiler et al., 2005). Both pathways are crucial to cell and organismal growth as they control the initiation of translation for many proteins. Either may be activated independent of each other, however remaining linked by activation of TOR via downstream effectors of PI3K namely the Rheb protein. It was previously established that an overexpression of *Rheb* leads to a significant increase in EC conversion in the *Drosophila* ovary (Rust et al., 2020). Moreover, recent studies found that in response to manipulating certain effectors of the TOR pathway cell fate switching is promoted in *Dictyostelium* (Jaiswal & Kimmel, 2019). In fact, TOR has been proven to particularly contribute to the regulation of FSC proliferation, whereas pFCs seem to be unaffected by TOR (Rust & Nystul, 2020). Taken together these circumstances suggest a causal coherence between the TOR pathway and the EC conversion event, since perturbations within TOR due to external influences may lead to a disruption of FSCs. As TOR and PI3K are directly activated by the insulin pathway (Hopkins et al., 2020b), TOR is as well connected with high fat diet induced phenotypes within *Drosophila* (Birse et al., 2010). The crucial importance of TOR hereby is underlined by demonstrating that high fat diet induced obesity and cardiac dysfunctioning in *Drosophila* may be rescued by the inhibition of TOR (Birse et al., 2010). Furthermore, a high fat diet is established to affect various metabolic variables, including the decrease in fecundity, elevation in triacylglycerol (TAG) levels and the promotion of insulin-resistance (S. Liao et al., 2021; Nayak & Mishra, 2021). For this study, I found that a high fat diet displayed an increase in EC conversion, although in a non-significant manner (see 4.2.3.1 Diet perturbations). The statistical outcome with respect to its significance may be due to the tool design (see 5.1.1 Driver strength and Gal4 status). Hence, future experiments are further necessary to validate an induction of EC conversion via high fat diet. Yet, since a notable trend for clone production was

still detectable, I reason that a high fat diet might constitute a catalyst for inducing EC conversion. Based on the above described circumstances, this might be accomplished by the influence of niche plasticity in the FSC lineage via the activity of the TOR pathway, similar to the effects of protein starvation.

5.4.2 Mating status does not interact with conversion

The *Drosophila* metabolism is not only directly influenced by nutrient intake, but also indirectly through the mating status of females. Copulation is associated with a negative effect on female longevity (Tracey Chapman et al., 1995; Lung et al., n.d.-a; Pitnick & García-Gonzá Lez, 2002). Proteins of the male ejaculate are reported to contribute to the males reproductive success, by provoking an increase in egg-laying and ovulation (T. Chapman et al., 2001; P. S. Chen et al., 1988; Lung et al., n.d.-a; Xue, 2000). Additionally, mating increases the number of cells in the midgut, therefore extending its size (Reiff et al., 2015; White et al., 2021). This is considered essential, as the opportunity of increased food intake may facilitate reproduction, which is energetically rather costly (Reiff et al., 2015). However, the unmated females did not display increased levels of EC conversion (see 5.4.2 Mating status does not interact with conversion, Figure 35). I conclude that the influence of the mating status is unlikely to affect the metabolism severely enough to induce EC conversion.

5.4.3 BPA influences EC conversion

Bisphenol A (BPA hereinafter) has received attention as an endocrine disrupting chemical (EDC) (see 1.6.3 Bisphenol A; Konieczna et al., 2015; Mustieles et al., 2020; Rochester, 2013). Increased BPA exposure is shown to be involved in several impacts on an organism's functioning: human pathogenesis hereby involve male and female infertility, hormone dependent tumors and the emergence of metabolic disorders (Konieczna et al., 2015). BPA exposure to *Drosophila melanogaster* is reported to affect behavior, neuronal development and the flie's lipid metabolism (Nguyen et al., 2021; Williams et al., 2014). I found that subjecting flies to 1mM BPA induces EC conversion. Interestingly, the rate of EC conversion did not rise proportionally with the increase of concentration (see 4.2.3.3 BPA exposure, Figure 36). This phenomenon was as well observed in

another study testing BPA in the *Drosophila* organism, eliciting non-linear dose responses (Weiner et al., 2014). This may stem from the fact that endocrine disrupting chemicals produce non-monotonic dose response curves (NMDRC) like natural hormones (Vandenberg, 2014; Vandenberg et al., 2012). Since endocrine disrupting chemicals share hormone-like characteristics, the mechanisms of non-monotonic dose response curves are typically related to interactions of hormones with their receptors and occur with a frequency of about 20% of all BPA experiments (Vandenberg, 2014). In agreement of strong stressors driving EC conversion, BPA hereby potentially displays severe impacts on the *Drosophila* organism. This may include direct effects on FSCs or impacting FSCs indirectly via affecting regulatory cycles, such as metabolic disruption or targeting neuronal structures. As BPA was shown to reduce the proliferative activity of stem cells and their progeny in different animal models (K. Kim et al., 2007; Tiwari et al., 2015), this may be applicable to a broader set of stem cells. Therefore, FSC fitness might be directly targeted by downregulation of proliferation via BPA intake. On the other hand, BPA is structurally similar to estradiol, binds to ER (estrogen and endocrine receptors) in vertebrates and seems to disrupt fertility via estrogenic activity in the hypothalamus in humans (Konieczna et al., 2015). Although no binding assay has confirmed the capacity of BPA binding dERR, the *Drosophila* ortholog of estrogen related receptors, yet (Nguyen et al., 2021), it is still considerable that BPA might bind to dERR. Supportive for this consideration is the fact that BPA experiments within this study produced non-monotonic dose response curves, which indicate typical hormone activity and interaction of hormone and receptor, respectively. dERR activity shows crucial importance in metabolism, since it was established that dERR mutant larvae displayed severe effects, including reductions in ATP volume, downregulation of genes involved in carbohydrate metabolism and therefore high levels of circulating sugar (Tennessen et al., 2011). Future studies could assess this hypothesis by evaluating the outcomes of dERR knockdown or overactivation in adult flies with respect to the induction of EC conversion. Furthermore, the assumption of BPA interacting with dERR could be addressed by testing if the BPA provoked phenotype can be rescued by knockdown of dERR and vice versa. Analogous to diet induced conversion (see 5.4.1 Diet induced cell fate switching), BPA is shown to interfere with the organism's insulin-signaling, by reducing the amount

of insulin-like peptides (ILs) (Williams et al., 2014). This prompts the idea that the insulin pathway may be involved in the event of EC conversion, since several conditions are associated with it. Overexpression of a dominant negative allele of the insulin receptor *InR*^{K1409A} or *foxo* did not induce EC conversion (Rust et al., 2020). This hints at the more important role of PI3K and TOR for the conversion event, as the insulin receptor and *foxo* display regulating functions within the pathway: *foxo* is a negative regulator for the TOR pathway, whereas the insulin receptor *InR*^{K1409A} downstream stimulates the activity of PI3K (Hay, 2011; Hopkins et al., 2020a). Raising flies on BPA food is found to leave them more susceptible towards starvation and it is considered that BPA might have impacts on feeding behavior (Williams et al., 2014). Therefore, BPA may alternatively be related to starvation like states, as nutrient intake is potentially disturbed. This might ultimately lead to higher sensitivity towards hunger and result in an elevation of stress for the flie's organism. In conclusion, there are different eventualities of how cell fate switching in the FSC lineage might be promoted via BPA exposure. Regardless, BPA is considered to evoke strong responses on cell level.

5.5 Knockdown of *Toll* does not prevent starvation induced EC conversion

Overactivation of the Toll pathway was shown to induce EC conversion (Rust et al., 2020). This may be due to high levels of AMP produced by the activation of *Toll*. AMP activated kinase (AMPK) is stimulated by the increase in AMP, functioning as an intracellular stress sensor and regulator to energy homeostasis (Wang et al., 2012). Since it is shown that nutritional depletion leads to an activation of AMPK initiated by the production of AMP, it is considerable that an overactivation of *Toll* leads to similar events. Since *Toll* overexpression fuels EC conversion, the pathway is sufficient to induce in the process. Here, I investigated whether the Toll pathway is also necessary for EC conversion. For this, I compared well fed and starved samples incorporating *Toll* RNAi knockdowns (see 4.2.4 The Toll signaling pathway in EC conversion). If necessary, I expected the clone production around RNAi samples between starved and well fed assays to not vary significantly, as the knockdown of Toll pathway would not allow starvation

induced conversion events. Nonetheless, RNAi fly lines exposed to starvation revealed equal or higher levels of EC conversion compared to starved flies of the control group (Figure 38). I therefore reason that the knockdown of Toll signaling components does not rescue starvation triggered EC conversion. In conclusion, I postulate that the Toll pathway is sufficient to facilitate EC conversion, but is not necessary. On the other hand, it may be argued that the utilized RNAi fly lines did not result in a sufficient knockdown of the targeted RNA. In order to exclude these doubts, qPCR or in situ hybridization could be applied to validate these RNAi fly lines. However, the same RNAi constructs were successfully utilized in other studies (C. N. McLaughlin et al., 2016; Yang & Hultmark, 2017), implying their duly knockdown capacity for this study. Alternatively, as *fax-Gal4* itself does not display significant driver strength, but was used to drive the RNAi; G-TRACE construct in this design, it may be considered that this potentially led to a lower induction of the knockdown. Here, the usage of another stronger driver could overcome these technical difficulties.

5.6 Perspectives

Subsequent experiments could include discovered trigger conditions of this environmental screen by focussing on these stress conditions in particular. This could as well comprise assays targeting the TOR pathway. In this context, an optimized tool set could be beneficial. The usage of recently described escort cell specific drivers, such as *dpr17* and *wun2*, might deliver even more reliable results. Potential approaches for future live imaging of the conversion event could be executed by the ablation of FSCs. Furthermore, following analyses on scRNA-Seq could unravel additional suitable marker genes for each escort cell subtype, which could facilitate the retracement of such phenomena. In addition, scRNA-Seq of FAC-sorted starved and well fed germaria may be applied in order to trace which genes are up- or downregulated via nutritional deprivation and identify candidates regulating niche cell conversion.

6. Summary

Proper tissue homeostasis is often regulated and maintained by stem cells, which undergo self-renewal to furthermore produce differentiating daughter cells. Tissues are exposed to a vast amount of environmental influences including stress conditions which may cause harm to the stem cell preservation. In order to maintain homeostasis, cell lineages are subject to different tasks which also include strategies on stem cell replacement. An advantageous model organism for the study of stem cells *in vivo* is the insect organism *Drosophila melanogaster*. In particular, the *Drosophila* follicle stem cell lineage is a model example of stem cell based epithelial cell lineages due to the conservation of tissue organization and its easy accessibility. The lineage is located in the ovaries of female flies and contributes to crucial functions within oogenesis as it produces the epithelial coating of the developing oocyte. Recent studies demonstrated that cells independent of the follicle stem cell lineage in *Drosophila* may convert to stem cells upon environmental perturbation. When flies are exposed to food deprivation, specialized niche cells, known as escort cells, are equipped with the ability to convert to follicle stem cells. In the absence of food deprivation, overactivation of Toll signaling, a pathway responsible for embryonic development and immune activities in flies, can induce escort cell conversion. The conversion of differentiated cells to another differentiated cell type is referred to as transdifferentiation and in the context of escort cell conversion, this might depict a mechanism of stem cell recovery. Since starvation has been reported to induce escort cell conversion in *Drosophila*, this study seeks to identify more environmental stressors which could trigger these conversion events. Therefore, the follicle stem cell lineage in *Drosophila* has been used and flies were exposed to a number of environmental factors including water and protein starvation, exposure to the Bisphenol A chemical, mating restraint and high fat diet. In addition, with the usage of ribonucleic acid interference fly lines, Toll pathway knockdown flies were exposed to starvation in order to investigate whether the pathway controls the conversion event. To visualize the conversion, fly lines were utilized in which escort cell lineages could be traced in the tissue. This was achieved via an escort cell specific promoter in combination with a clonal marking system.

The environmental screen revealed that protein starvation and the exposure to Bisphenol A induces escort cell conversion, whereas several other conditions indicate the same trend. Additionally, in starved *Toll* knockdown flies, escort cell conversion increased compared to its control. I therefore reasoned that the Toll pathway is not necessary for escort cell conversion.

Guided by the discoveries of different escort cell subtypes, escort cells were furthermore examined specifically within this study. With the aspiration to identify these subpopulations and associated marker genes which could hence facilitate the application of novel tools for escort cell subtypes, an escort cell specific single cell sequencing dataset was analyzed and compared to published escort cell datasets. The analysis of the escort cell specific dataset showed that calculated marker genes were not present in the escort cell compartments, which led me to the conclusion that the dataset was mostly composed of cells other than escort cell origin.

Despite methodological difficulties such as driver strength or statistical sample size, this study produced significant results regarding the effect of environmental factors on escort cell conversion. I concluded that the conversion event is highly dependent on the exposure to distinct environmental conditions. Considering results from this and other studies, metabolism and the evolutionary conserved Target of Rapamycin pathway seem to be involved in this process. In this context, I postulate that strong environmental stressors lead to follicle stem cell loss or dedifferentiation as a consequence of metabolic disruption. Hence, cell fate switching of escort cells would depict a recovery strategy on stem cell loss. This study argues for the capability of escort cells to convert to follicle stem cells and therefore maintain tissue homeostasis of the *Drosophila* follicle epithelium. This underlines the versatile response of niches to their environment and provides us with a more detailed insight in niche and stem cell networks.

7. Zusammenfassung

Gewebehomöostase wird häufig von Stammzellen reguliert und aufrechterhalten, welche sich selbst erneuern und differenzierte Tochterzellen produzieren. Gewebe sind einer Vielzahl von Umwelteinflüssen ausgesetzt, darunter auch Stressbedingungen, die der Erhaltung der Stammzellen schaden können. Um die Homöostase aufrechtzuerhalten, unterliegen Zelllinien verschiedenen Aufgaben, zu denen auch Strategien zum Ersatz von Stammzellen gehören. Ein vorteilhafter Modellorganismus für die Untersuchung von Stammzellen *in vivo* ist der Insekten-Organismus *Drosophila melanogaster*. Insbesondere die follikuläre Stammzelllinie in *Drosophila* ist aufgrund der Konservierung der Gewebeorganisation und der leichten Zugänglichkeit ein Modellbeispiel für stammzell-basierte epitheliale Zelllinien. Diese befindet sich in den Eierstöcken der weiblichen Fliegen und trägt zu entscheidenden Funktionen innerhalb der Oogenese bei, da sie die epitheliale Hülle der sich entwickelnden Eizelle bildet. Kürzlich durchgeführte Studien haben gezeigt, dass Zellen, die unabhängig von der follikulären Stammzelllinie in *Drosophila* sind, sich bei Umweltstörungen in Stammzellen umwandeln können. Setzt man Fliegen Nahrungsentzug aus, sind spezialisierte Nischenzellen, sogenannte Escort Zellen, in der Lage, sich in follikuläre Stammzellen umzuwandeln. Ohne Nahrungsentzug kann eine Überaktivierung des Toll-Signalwegs, der für die Embryonalentwicklung und die Immunaktivität der Fliegen verantwortlich ist, die Escort-Zell-Umwandlung auslösen. Die Umwandlung einer differenzierten Zelle in einen anderen differenzierten Zelltyp wird als Transdifferenzierung bezeichnet, und im Zusammenhang der Umwandlung von Escort Zellen könnte dies einen Mechanismus des Stammzellersatzes darstellen. Da berichtet wurde, dass Hungern die Umwandlung von Escort Zellen in *Drosophila* auslöst, sollen in dieser Studie weitere Umweltstressoren identifiziert werden, die diese Umwandlungsvorgänge begünstigen. Demnach wurde die follikuläre Stammzelllinie von *Drosophila* verwendet und die Fliegen einer Reihe von Umweltfaktoren ausgesetzt: Hungern und Proteinmangel, Exposition gegenüber der Chemikalie Bisphenol A, Paarungshemmung und fettreiche Ernährung. Mit Hilfe von Ribonukleinsäure-Interferenz Fliegenlinien, wurden Fliegen, bei denen der Toll-Signalweg ausgeschaltet war, Hunger ausgesetzt, um zu untersuchen,

ob der Signalweg die Umwandlung kontrolliert. Um die Umwandlung sichtbar zu machen, wurden Fliegenlinien verwendet, bei denen die Escort-Zelllinien im Gewebe zurückverfolgt werden konnten. Dies wurde durch einen Escort Zell spezifischen Promotor in Kombination mit einem klonalen Markierungssystem erreicht. Das Umwelt-Screening ergab, dass Proteinmangel und die Exposition gegenüber Bisphenol A die Umwandlung von Escort Zellen auslösen, während mehrere andere Bedingungen den gleichen Trend indizieren. Darüber hinaus war die Umwandlung von Escort Zellen in ausgehungerten Fliegen, bei denen *Toll* ausgeschaltet wurde, im Vergleich zur Kontrolle erhöht. Dies führte zu der Annahme, dass der Toll-Signalweg für die Umwandlung von Escort Zellen nicht notwendig ist.

Aufgrund der Entdeckungen verschiedener Escort Zell-Subtypen wurden in dieser Studie Escort Zellen speziell untersucht. Um die Anwendung neuer Methoden für Escort Zell-Subtypen zu ermöglichen, sollten Subpopulationen und zugehörige Markergene mittels Analyse eines Escort Zell spezifischen Einzelzell-Sequenzierungs Datensatzes identifiziert werden und mit bereits veröffentlichten Escort Zell Datensätzen verglichen werden. Die Analyse des Escort Zell spezifischen Datensatzes zeigte, dass berechnete Markergene in den Escort Zell Kompartimenten nicht vorhanden waren, welches letztlich zu der Schlussfolgerung führte, dass der Datensatz größtenteils aus anderen Zellidentitäten bestand.

Trotz methodischer Schwierigkeiten, wie Treiberstärke oder Stichprobengröße, lieferte diese Studie signifikante Ergebnisse hinsichtlich der Auswirkungen von Umweltfaktoren auf die Konversion von Escort Zellen. Ich schlussfolgere folglich, dass die Konversion in hohem Maße von der Exposition gegenüber bestimmten Umweltbedingungen abhängig ist. In Anbetracht der Ergebnisse dieser und anderer Studien, scheinen der Stoffwechsel und der evolutionär konservierte Target of Rapamycin-Signalweg an diesem Prozess beteiligt zu sein. In diesem Zusammenhang postuliere ich, dass starke Umweltstressoren zum Verlust oder zur Dedifferenzierung von follikulären Stammzellen als Folge einer Störung des Stoffwechsels führen. Die Umwandlung des Zellschicksals der Escort Zellen stelle daher eine Strategie zur Kompensation des Stammzell-Verlustes dar. Diese Studie spricht für die Fähigkeit von Escort Zellen, sich in follikuläre Stammzellen umzuwandeln und so die Gewebehomöostase des *Drosophila*

Follikelepithels aufrechtzuerhalten. Dies unterstreicht die vielseitige Reaktion von Nischen auf ihre Umgebung und verschafft uns einen detaillierteren Einblick in Nischen- und Stammzellnetzwerke.

8. Bibliography

Agrawal, N., Dasaradhi, P. V. N., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., & Mukherjee, S. K. (2003). RNA Interference: Biology, Mechanism, and Applications. *Microbiology and Molecular Biology Reviews*, *67*(4), 657. https://doi.org/10.1128/MMBR.67.4.657-685.2003

Almeida, S., Raposo, A., Almeida-González, M., & Carrascosa, C. (2018). Bisphenol A: Food Exposure and Impact on Human Health. *Comprehensive Reviews in Food Science and Food Safety*, *17*(6), 1503–1517. https://doi.org/10.1111/1541-4337.12388

Bachem, M. G., Sell, K. M., Melchior, R., Kropf, J., Eller, T., & Gressner, A. M. (1993). Tumor necrosis factor alpha (TNFα) and transforming growth factor β1 (TGFβ1) stimulate fibronectin synthesis and the transdifferentiation of fat-storing cells in the rat liver into myofibroblasts. *Virchows Archiv B Cell Pathology Including Molecular Pathology*, *63*(1), 123–130. https://doi.org/10.1007/BF02899251

Balakrishnan, B., Henare, K., Thorstensen, E. B., Ponnampalam, A. P., & Mitchell, M. D. (2010). Transfer of bisphenol A across the human placenta. *American Journal of Obstetrics and Gynecology*, *202*(4), 393.e1-393.e7. https://doi.org/10.1016/J.AJOG.2010.01.025

Bangham, J., Chapman, T., Smith, H. K., & Partridge, L. (2003). *Influence of female reproductive anatomy on the outcome of sperm competition in Drosophila melanogaster*. https://doi.org/10.1098/rspb.2002.2237

Banisch, T. U., Maimon, I., Dadosh, T., & Gilboa, L. (2017). Escort cells generate a dynamic compartment for germline stem cell differentiation via combined Stat and Erk signalling. *Development (Cambridge)*, *144*(11), 1937–1947. https://doi.org/10.1242/DEV.143727/VIDEO-11

Baran-Gale, J., Chandra, T., & Kirschner, K. (2018). Experimental design for single-cell RNA sequencing. *Briefings in Functional Genomics*, *17*(4), 233–239. https://doi.org/10.1093/BFGP/ELX035

Bastock, R., & St Johnston, D. (2008). Drosophila oogenesis. *Current Biology*, *18*(23), R1082–R1087. https://doi.org/10.1016/J.CUB.2008.09.011

Begum, M., Paul, P., Das, D., & Ghosh, S. (2020). Endocrine-disrupting plasticizer Bisphenol A (BPA) exposure causes change in behavioral attributes in Drosophila melanogaster. *Toxicology and Environmental Health Sciences*, *12*(3), 237–246. https://doi.org/10.1007/S13530-020-00052-8

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature 2001 409:6818*, *409*(6818), 363–366. https://doi.org/10.1038/35053110

Birse, R. T., Choi, J., Reardon, K., Rodriguez, J., Graham, S., Diop, S., Ocorr, K., Bodmer, R., & Oldham, S. (2010). High-fat-diet-induced obesity and heart dysfunction are regulated by the TOR pathway in Drosophila. *Cell Metabolism*, *12*(5), 533–544. https://doi.org/10.1016/J.CMET.2010.09.014

Bloch Qazi, M. C., Heifetz, Y., & Wolfner, M. F. (2003). The developments

between gametogenesis and fertilization: ovulation and female sperm storage in drosophila melanogaster. *Developmental Biology*, *256*(2), 195–211. https://doi.org/10.1016/S0012-1606(02)00125-2

Brand, A. H., & Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, *118*(2), 401–415. https://doi.org/10.1242/DEV.118.2.401

Broach, J. R., Guarascio, V. R., & Jayaram, M. (1982). Recombination within the yeast plasmid 2μ circle is site-specific. *Cell*, *29*(1), 227–234. https://doi.org/10.1016/0092-8674(82)90107-6

Bumcrot, D., Manoharan, M., Koteliansky, V., & Sah, D. W. Y. (2006). RNAi therapeutics: a potential new class of pharmaceutical drugs. *Nature Chemical Biology 2006 2:12*, *2*(12), 711–719. https://doi.org/10.1038/nchembio839

Chapman, T., Herndon, L. A., Heifetz, Y., Partridge, L., & Wolfner, M. F. (2001). The Acp26Aa seminal fluid protein is a modulator of early egg hatchability in Drosophila melanogaster. *Proceedings. Biological Sciences*, *268*(1477), 1647–1654. https://doi.org/10.1098/RSPB.2001.1684

Chapman, Tracey, Liddle, L. F., Kalb, J. M., Wolfner, M. F., & Partridge, L. (1995). Cost of mating in Drosophila melanogaster females is mediated by male accessory gland products. *Nature*, *373*(6511), 241–244. https://doi.org/10.1038/373241a0

Chen, G., Ning, B., & Shi, T. (2019). Single-cell RNA-seq technologies and related computational data analysis. *Frontiers in Genetics*, *10*(APR), 317. https://doi.org/10.3389/FGENE.2019.00317/BIBTEX

Chen, P. S., Stumm-Zollinger, E., Aigaki, T., Balmer, J., Bienz, M., & Böhlen, P. (1988). A male accessory gland peptide that regulates reproductive behavior of female D. melanogaster. *Cell*, *54*(3), 291–298. https://doi.org/10.1016/0092-8674(88)90192-4

Dabeva, M. D., Hwang, S. G., Vasa, S. R. G., Hurston, E., Novikoff, P. M., Hixson, D. C., Gupta, S., & Shafritz, D. A. (1997). Differentiation of pancreatic epithelial progenitor cells into hepatocytes following transplantation into rat liver. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(14), 7356–7361. https://doi.org/10.1073/PNAS.94.14.7356/ASSET/22165096-964C-4967-8012-AC24513D63CA/ASSETS/GRAPHIC/PQ1471493004.JPEG

Douglas, H. C., & Hawthorne, D. C. (n.d.). *REGULATION OF GENES CONTROLLING SYNTHESIS OF THE GALACTOSE PATHWAY ENZYMES IN YEAST1*.

Drummond-Barbosa, D. (2008). Stem Cells, Their Niches and the Systemic Environment: An Aging Network. *Genetics*, *180*(4), 1787. https://doi.org/10.1534/GENETICS.108.098244

Eguchi, G., & Okada, T. S. (1973). Differentiation of Lens Tissue from the Progeny of Chick Retinal Pigment Cells Cultured In Vitro: A Demonstration of a Switch of Cell Types in Clonal Cell Culture. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(5), 1495. https://doi.org/10.1073/PNAS.70.5.1495

Eguchi, Goro, & Kodama, R. (1993). Transdifferentiation. *Current Opinion in Cell Biology*, *5*(6), 1023–1028. https://doi.org/10.1016/0955-0674(93)90087-7

Evans, C. J., Olson, J. M., Ngo, K. T., Kim, E., Lee, N. E., Kuoy, E., Patananan, A. N., Sitz, D., Tran, P. T., 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 2009 6:8*, *6*(8), 603–605. https://doi.org/10.1038/nmeth.1356

Fadiga, J., & Nystul, T. G. (2019). The follicle epithelium in the Drosophila ovary is maintained by a small number of stem cells. *ELife*, *8*. https://doi.org/10.7554/ELIFE.49050

Fleisch, A. F., Sheffield, P. E., Chinn, C., Edelstein, B. L., & Landrigan, P. J. (2010). Bisphenol A and Related Compounds in Dental Materials. *Pediatrics*, *126*(4), 760. https://doi.org/10.1542/PEDS.2009-2693

Fuller, M. T., & Spradling, A. C. (2007). Male and female Drosophila germline stem cells: Two versions of immortality. *Science*, *316*(5823), 402–404. https://doi.org/10.1126/SCIENCE.1140861

Glen, S. (n.d.). *Sample Size in Statistics (How to Find it): Excel, Cochran's Formula, General Tips"*. https://www.statisticshowto.com/probability-and-statistics/find-sample-size/

Hadley Wickham, Romain François, Lionel Henry, K. M. (2022). *dplyr: A Grammar of Data Manipulation*.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell*, *184*(13), 3573-3587.e29. https://doi.org/10.1016/J.CELL.2021.04.048/ATTACHMENT/1990279C-028D-47D4-A5D3-C89B04E14795/MMC3.XLSX

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K. J., Rozenblatt-Rosen, O., Dor, Y., Regev, A., & Yanai, I. (2016). CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biology*, *17*(1), 1–7. https://doi.org/10.1186/S13059-016-0938-8/FIGURES/4

Hay, N. (2011). Interplay between FOXO, TOR, and Akt. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1813*(11), 1965–1970. https://doi.org/10.1016/J.BBAMCR.2011.03.013

Heike, T., & Nakahata, T. (2005). Stem cell plasticity in hematopoietic system. *Ensho Saisei*, *25*(2), 90–101. https://doi.org/10.2492/JSIR.25.90

Hinnant, T. D., Merkle, J. A., & Ables, E. T. (2020). Coordinating Proliferation, Polarity, and Cell Fate in the Drosophila Female Germline. *Frontiers in Cell and Developmental Biology*, *8*, 19. https://doi.org/10.3389/FCELL.2020.00019/BIBTEX

Hirst, M., Ho, C., Sabourin, L., Rudnicki, M., Penn, L., & Sadowski, I. (2001). A two-hybrid system for transactivator bait proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(15),

8726–8731. https://doi.org/10.1073/PNAS.141413598/ASSET/207184D3-DF33-4422-A85F-C15F6B4778FF/ASSETS/GRAPHIC/PQ1414135005.JPEG

Hochgerner, H., Lönnerberg, P., Hodge, R., Mikes, J., Heskol, A., Hubschle, H., Lin, P., Picelli, S., La Manno, G., Ratz, M., Dunne, J., Husain, S., Lein, E., Srinivasan, M., Zeisel, A., & Linnarsson, S. (2017). STRT-seq-2i: dual-index 5' single cell and nucleus RNA-seq on an addressable microwell array. *Scientific Reports 2017 7:1*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-16546-4

Hopkins, B. D., Goncalves, M. D., & Cantley, L. C. (2020a). Insulin–PI3K signalling: an evolutionarily insulated metabolic driver of cancer. *Nature Reviews Endocrinology 2020 16:5*, *16*(5), 276–283. https://doi.org/10.1038/s41574-020-0329-9

Hopkins, B. D., Goncalves, M. D., & Cantley, L. C. (2020b). Insulin–PI3K signalling: an evolutionarily insulated metabolic driver of cancer. *Nature Reviews Endocrinology*, *16*(5), 276–283. https://doi.org/10.1038/S41574-020-0329-9

Hsu, H. J., & Drummond-Barbosa, D. (2017). A visual screen for diet-regulated proteins in the Drosophila ovary using GFP protein trap lines. *Gene Expression Patterns*, *23–24*, 13–21. https://doi.org/10.1016/J.GEP.2017.01.001

Hu, P., Zhang, W., Xin, H., & Deng, G. (2016). Single cell isolation and analysis. *Frontiers in Cell and Developmental Biology*, *4*(OCT), 116. https://doi.org/10.3389/FCELL.2016.00116/BIBTEX

Hudson, A. M., & Cooley, L. (2014). Methods for studying oogenesis. *Methods (San Diego, Calif.)*, *68*(1), 207. https://doi.org/10.1016/J.YMETH.2014.01.005

Ito, M., Yang, Z., Andl, T., Cui, C., Kim, N., Millar, S. E., & Cotsarelis, G. (2007). Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding. *Nature*, *447*(7142), 316–320. https://doi.org/10.1038/NATURE05766

Jaiswal, P., & Kimmel, A. R. (2019). mTORC1/AMPK responses define a core gene set for developmental cell fate switching. *BMC Biology*, *17*(1), 58. https://doi.org/10.1186/S12915-019-0673-1

Jennings, B. H. (2011). Drosophila – a versatile model in biology & medicine. *Materials Today*, *14*(5), 190–195. https://doi.org/10.1016/S1369-7021(11)70113-4

Jevitt, A., Chatterjee, D., Xie, G., Wang, X. F., Otwell, T., Huang, Y. C., & Deng, W. M. (2020). A single-cell atlas of adult Drosophila ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLOS Biology*, *18*(4), e3000538. https://doi.org/10.1371/JOURNAL.PBIO.3000538

Kashima, Y., Sakamoto, Y., Kaneko, K., Seki, M., Suzuki, Y., & Suzuki, A. (2020). Single-cell sequencing techniques from individual to multiomics analyses. *Experimental & Molecular Medicine 2020 52:9*, *52*(9), 1419–1427. https://doi.org/10.1038/s12276-020-00499-2

Kaur, K., Simon, A. F., Chauhan, V., & Chauhan, A. (2015). Effect of bisphenol A on Drosophila melanogaster behavior – A new model for the studies on neurodevelopmental disorders. *Behavioural Brain Research*, *284*, 77–84. https://doi.org/10.1016/J.BBR.2015.02.001

Kelly, S. M., Elchert, A., & Kahl, M. (2017). Dissection and immunofluorescent staining of mushroom body and photoreceptor neurons in adult Drosophila melanogaster brains. *Journal of Visualized Experiments*, *2017*(129). https://doi.org/10.3791/56174

Kim-Yip, R. P., & Nystul, T. G. (2018). Wingless promotes EGFR signaling in follicle stem cells to maintain self-renewal. *Development (Cambridge, England)*, *145*(23). https://doi.org/10.1242/DEV.168716

Kim, D. H., & Rossi, J. J. (2008). RNAi mechanisms and applications. *BioTechniques*, *44*(5), 613. https://doi.org/10.2144/000112792

Kim, K., Son, T. G., Kim, S. J., Kim, H. S., Kim, T. S., Han, S. Y., & Lee, J. (2007). Suppressive effects of bisphenol A on the proliferation of neural progenitor cells. *Journal of Toxicology and Environmental Health. Part A*, *70*(15–16), 1288–1295. https://doi.org/10.1080/15287390701434216

Kirilly, D., Spana, E. P., Perrimon, N., Padgett, R. W., & Xie, T. (2005). BMP signaling is required for controlling somatic stem cell self-renewal in the Drosophila ovary. *Developmental Cell*, *9*(5), 651–662. https://doi.org/10.1016/J.DEVCEL.2005.09.013/ATTACHMENT/F454887A-162A-471B-8217-80D6DFDE16B6/MMC1.PDF

Kirilly, D., Wang, S., & Xie, T. (2011). Self-maintained escort cells form a germline stem cell differentiation niche. *Development*, *138*(23), 5087–5097. https://doi.org/10.1242/DEV.067850

Konieczna, A., Rutkowska, A., & Rachoń, D. (2015). Health risk of exposure to Bisphenol A (BPA). *Roczniki Panstwowego Zakladu Higieny*, *66*(1), 5–11.

Ladewig, J., Koch, P., & Brüstle, O. (2013). Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. *Nature Reviews Molecular Cell Biology 2013 14:4*, *14*(4), 225–236. https://doi.org/10.1038/nrm3543

Lam, J. K. W., Chow, M. Y. T., Zhang, Y., & Leung, S. W. S. (2015). siRNA Versus miRNA as Therapeutics for Gene Silencing. *Molecular Therapy - Nucleic Acids*, *4*(9), e252. https://doi.org/10.1038/MTNA.2015.23

Lee, K. P., & Jang, T. (2014). Exploring the nutritional basis of starvation resistance in Drosophila melanogaster. *Functional Ecology*, *28*(5), 1144–1155. https://doi.org/10.1111/1365-2435.12247

Li, H., Janssens, J., de Waegeneer, M., Kolluru, S. S., Davie, K., Gardeux, V., Saelens, W., David, F. P. A., Brbić, M., Spanier, K., Leskovec, J., McLaughlin, C. N., Xie, Q., Jones, R. C., Brueckner, K., Shim, J., Tattikota, S. G., Schnorrer, F., Rust, K., … Aerts, S. (2022). Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science*, *375*(6584). https://doi.org/10.1126/SCIENCE.ABK2432/SUPPL_FILE/SCIENCE.ABK2432_MDAR_REPRODUCIBILITY_CHECKLIST.PDF

Li, L., Edgar, B. A., & Grewal, S. S. (2010). Nutritional control of gene expression

in Drosophila larvae via TOR, Myc and a novel cis-regulatory element. *BMC Cell Biology*, *11*. https://doi.org/10.1186/1471-2121-11-7

Li, X., & Wang, C. Y. (2021). From bulk, single-cell to spatial RNA sequencing. *International Journal of Oral Science 2021 13:1*, *13*(1), 1–6. https://doi.org/10.1038/s41368-021-00146-0

Liao, S., Amcoff, M., & Nässel, D. R. (2021). Impact of high-fat diet on lifespan, metabolism, fecundity and behavioral senescence in Drosophila. *Insect Biochemistry and Molecular Biology*, *133*. https://doi.org/10.1016/J.IBMB.2020.103495

Liao, X., Makris, M., & Luo, X. M. (2016). Fluorescence-activated Cell Sorting for Purification of Plasmacytoid Dendritic Cells from the Mouse Bone Marrow. *Journal of Visualized Experiments : JoVE*, *2016*(117), 54641. https://doi.org/10.3791/54641

Lindsay, S. A., & Wasserman, S. A. (2014). Conventional and non-conventional Drosophila Toll signaling. *Developmental & Comparative Immunology*, *42*(1), 16–24. https://doi.org/10.1016/J.DCI.2013.04.011

Linford, N. J., Ro, J., Chung, B. Y., & Pletcher, S. D. (2015). Gustatory and metabolic perception of nutrient stress in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2587–2592. https://doi.org/10.1073/PNAS.1401501112/SUPPL_FILE/PNAS.201401501SI.PDF

Losick, V. P., Morris, L. X., Fox, D. T., & Spradling, A. (2011). Drosophila Stem Cell Niches: A Decade of Discovery Suggests a Unified View of Stem Cell Regulation. *Developmental Cell*, *21*(1), 159–171. https://doi.org/10.1016/J.DEVCEL.2011.06.018

Lüersen, K., Röder, T., & Rimbach, G. (2019). Drosophila melanogaster in nutrition research—the importance of standardizing experimental diets. *Genes & Nutrition*, *14*(1). https://doi.org/10.1186/S12263-019-0627-9

Lung, O., Tram, U., Finnerty, C. M., Eipper-Mains, M. A., Kalb, J. M., & Wolfner, M. F. (n.d.-a). *The Drosophila melanogaster Seminal Fluid Protein Acp62F Is a Protease Inhibitor That Is Toxic Upon Ectopic Expression*. Retrieved April 3, 2022, from www.doe-mbi.ucla.edu/people/frsvr/frsvr.html;

Lung, O., Tram, U., Finnerty, C. M., Eipper-Mains, M. A., Kalb, J. M., & Wolfner, M. F. (n.d.-b). *The Drosophila melanogaster Seminal Fluid Protein Acp62F Is a Protease Inhibitor That Is Toxic Upon Ectopic Expression*. Retrieved June 20, 2022, from www.doe-mbi.ucla.edu/people/frsvr/frsvr.html;

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., 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. https://doi.org/10.1016/J.CELL.2015.05.002/ATTACHMENT/4A61AA26-CCB0-48AC-9D09-485C9571EDBD/MMC6.ZIP

Magavi, S. S., Leavitt, B. R., & Macklis, J. D. (2000). Induction of neurogenesis

in the neocortex of adult mice. *Nature*, *405*(6789), 951–955. https://doi.org/10.1038/35016083

Margolis, J., & Spradling, A. (1995). Identification and behavior of epithelial stem cells in the Drosophila ovary. *Development*, *121*(11), 3797–3807. https://doi.org/10.1242/DEV.121.11.3797

Marmugi, A., Ducheix, S., Lasserre, F., Polizzi, A., Paris, A., Priymenko, N., Bertrand-Michel, J., Pineau, T., Guillou, H., Martin, P. G. P., & Mselli-Lakhal, L. (2012). Low doses of bisphenol A induce gene expression related to lipid synthesis and trigger triglyceride accumulation in adult mouse liver. *Hepatology (Baltimore, Md.)*, *55*(2), 395–407. https://doi.org/10.1002/HEP.24685

Matsumoto, K., Toh-e, A., & Oshima, Y. (1978). Genetic control of galactokinase synthesis in Saccharomyces cerevisiae: evidence for constitutive expression of the positive regulatory gene gal4. *Journal of Bacteriology*, *134*(2), 446–457. https://doi.org/10.1128/JB.134.2.446-457.1978

McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K., & Davis, R. L. (2003). Spatiotemporal Rescue of Memory Dysfunction in Drosophila. *Science*, *302*(5651), 1765–1768. https://doi.org/10.1126/SCIENCE.1089035

McLaughlin, C. N., Nechipurenko, I. V., Liu, N., & Broihier, H. T. (2016). A Toll receptor–FoxO pathway represses Pavarotti/MKLP1 to promote microtubule dynamics in motoneurons. *The Journal of Cell Biology*, *214*(4), 459. https://doi.org/10.1083/JCB.201601014

McLaughlin, J. M., & Bratu, D. P. (2015). Drosophila melanogaster oogenesis: An overview. *Methods in Molecular Biology*, *1328*. https://doi.org/10.1007/978-1-4939-2851-4_1

Mirzoyan, Z., Sollazzo, M., Allocca, M., Valenza, A. M., Grifoni, D., & Bellosta, P. (2019). Drosophila melanogaster: A model organism to study cancer. *Frontiers in Genetics*, *10*, 51. https://doi.org/10.3389/FGENE.2019.00051/BIBTEX

Morrison, S. J., & Spradling, A. C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell*, *132*(4), 598. https://doi.org/10.1016/J.CELL.2008.01.038

Mustieles, V., D'Cruz, S. C., Couderq, S., Rodríguez-Carrillo, A., Fini, J. B., Hofer, T., Steffensen, I. L., Dirven, H., Barouki, R., Olea, N., Fernández, M. F., & David, A. (2020). Bisphenol A and its analogues: A comprehensive review to identify and prioritize effect biomarkers for human biomonitoring. *Environment International*, *144*, 105811. https://doi.org/10.1016/J.ENVINT.2020.105811

Natarajan, K. N., Miao, Z., Jiang, M., Huang, X., Zhou, H., Xie, J., Wang, C., Qin, S., Zhao, Z., Wu, L., Yang, N., Li, B., Hou, Y., Liu, S., & Teichmann, S. A. (2019). Comparative analysis of sequencing technologies for single-cell transcriptomics. *Genome Biology*, *20*(1), 1–8. https://doi.org/10.1186/S13059-019-1676-5/FIGURES/2

Nayak, N., & Mishra, M. (2021). High fat diet induced abnormalities in metabolism, growth, behavior, and circadian clock in Drosophila

melanogaster. *Life Sciences*, *281*. https://doi.org/10.1016/J.LFS.2021.119758

Neckameyer, W. S., & Bhatt, P. (n.d.). *Protocols to Study Behavior in Drosophila*. https://doi.org/10.1007/978-1-4939-6371-3_19

Nguyen, U., Tinsley, B., Sen, Y., Stein, J., Palacios, Y., Ceballos, A., Welch, C., Nzenkue, K., Penn, A., Murphy, L., Leodones, K., Casiquin, J., Ivory, I., Ghenta, K., Danziger, K., Widman, E., Newman, J., Triplehorn, M., Hindi, Z., & Mulligan, K. (2021). Exposure to bisphenol A differentially impacts neurodevelopment and behavior in Drosophila melanogaster from distinct genetic backgrounds. *Neurotoxicology*, *82*, 146–157. https://doi.org/10.1016/j.neuro.2020.12.007

Nishikawa, M., Iwano, H., Yanagisawa, R., Koike, N., Inoue, H., & Yokota, H. (2010). Placental Transfer of Conjugated Bisphenol A and Subsequent Reactivation in the Rat Fetus. *Environmental Health Perspectives*, *118*(9), 1196. https://doi.org/10.1289/EHP.0901575

Nunez, A. A., Kannan, K., Giesy, J. P., Fang, J., & Clemens, L. G. (2001). Effects of Bisphenol A on energy balance and accumulation in brown adipose tissue in rats. *Chemosphere*, *42*(8), 917–922. https://doi.org/10.1016/S0045-6535(00)00196-X

Nystul, T., & Spradling, A. (2007a). An epithelial niche in the Drosophila ovary undergoes long-range stem cell replacement. *Cell Stem Cell*, *1*(3), 277–285. https://doi.org/10.1016/J.STEM.2007.07.009

Nystul, T., & Spradling, A. (2007b). An Epithelial Niche in the Drosophila Ovary Undergoes Long-Range Stem Cell Replacement. *Cell Stem Cell*, *1*(3), 277–285. https://doi.org/10.1016/J.STEM.2007.07.009

Pérez, L. M., De Lucas, B., & Gálvez, B. G. (2018). Unhealthy Stem Cells: When Health Conditions Upset Stem Cell Properties. *Cellular Physiology and Biochemistry*, *46*(5), 1999–2016. https://doi.org/10.1159/000489440

Petersen, B. E., Bowen, W. C., Patrene, K. D., Mars, W. M., Sullivan, A. K., Murase, N., Boggs, S. S., Greenberger, J. S., & Goff, J. P. (1999). Bone marrow as a potential source of hepatic oval cells. *Science*, *284*(5417), 1168–1170. https://doi.org/10.1126/SCIENCE.284.5417.1168

Picelli, S. (2017). Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biology*, *14*(5), 637. https://doi.org/10.1080/15476286.2016.1201618

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols 2013 9:1*, *9*(1), 171–181. https://doi.org/10.1038/nprot.2014.006

Pitnick, S., & García-Gonzá Lez, F. (2002). *Harm to females increases with male body size in Drosophila melanogaster*. https://doi.org/10.1098/rspb.2002.2090

Pitnick, S., Markow, T., & Spicer, G. S. (1999). EVOLUTION OF MULTIPLE KINDS OF FEMALE SPERM-STORAGE ORGANS IN DROSOPHILA. *Evolution*, *6*, 1804–1822.

Probst, V., Pacheco, F., Nielsen, F. C., & Bagger, F. O. (2020). Benchmarking full-length transcript single cell RNA sequencing protocols. *BioRxiv*, 2020.07.29.225201. https://doi.org/10.1101/2020.07.29.225201

Rao, M. S., Dwivedi, R. S., Subborao, V., Usman, M. I., Scarpelli, D. G., Nemali, M. R., Yeldandi, A., Thangada, S., Kumar, S., & Reddy, J. K. (1988). Almost total conversion of pancreas to liver in the adult rat: A reliable model to study transdifferentiation. *Biochemical and Biophysical Research Communications*, *156*(1), 131–136. https://doi.org/10.1016/S0006-291X(88)80814-3

Rao, M. Sambasiva, Scarpelli, D. G., & Reddy, J. K. (1986). Transdifferentiated hepatocytes in rat pancreas. *Current Topics in Developmental Biology*, *20*(C), 63–78. https://doi.org/10.1016/S0070-2153(08)60654-7

Rao, M., & Sockanathan, S. (2005). Molecular mechanisms of RNAi: Implications for development and disease. *Birth Defects Research Part C: Embryo Today: Reviews*, *75*(1), 28–42. https://doi.org/10.1002/BDRC.20030

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., … Yosef, N. (2017). The human cell atlas. *ELife*, *6*. https://doi.org/10.7554/ELIFE.27041

Reiff, T., Jacobson, J., Cognigni, P., Antonello, Z., Ballesta, E., Tan, K. J., Yew, J. Y., Dominguez, M., & Miguel-Aliaga, I. (2015). Endocrine remodelling of the adult intestine sustains reproduction in Drosophila. *ELife*, *4*(JULY 2015). https://doi.org/10.7554/ELIFE.06930

Reilein, A., Melamed, D., Park, K. S., Berg, A., Cimetta, E., Tandon, N., Vunjak-Novakovic, G., Finkelstein, S., & Kalderon, D. (2017). Alternative direct stem cell derivatives defined by stem cell location and graded Wnt signaling. *Nature Cell Biology*, *19*(5), 433. https://doi.org/10.1038/NCB3505

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., & Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. *Genome Research*, *11*(6), 1114–1125. https://doi.org/10.1101/GR.169101

Rochester, J. R. (2013). Bisphenol A and human health: A review of the literature. *Reproductive Toxicology*, *42*, 132–155. https://doi.org/10.1016/J.REPROTOX.2013.08.008

Roth, S., & Lynch, J. A. (2009). Symmetry Breaking During Drosophila Oogenesis. *Cold Spring Harbor Perspectives in Biology*, *1*(2), a001891. https://doi.org/10.1101/CSHPERSPECT.A001891

Rust, K., Byrnes, L. E., Yu, K. S., Park, J. S., Sneddon, J. B., Tward, A. D., & Nystul, T. G. (2020). A single-cell atlas and lineage analysis of the adult Drosophila ovary. *Nature Communications 2020 11:1*, *11*(1), 1–17. https://doi.org/10.1038/s41467-020-19361-0

Rust, K., & Nystul, T. (2020). Signal transduction in the early Drosophila follicle stem cell lineage. *Current Opinion in Insect Science*, *37*, 39–48. https://doi.org/10.1016/J.COIS.2019.11.005

Sahai-Hernandez, P., & Nystul, T. G. (2013). A dynamic population of stromal

cells contributes to the follicle stem cell niche in the Drosophila ovary. *Development (Cambridge, England)*, *140*(22), 4490–4498. https://doi.org/10.1242/DEV.098558

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology 2015 33:5*, *33*(5), 495–502. https://doi.org/10.1038/nbt.3192

Schnakenberg, S. L., Matias, W. R., & Siegal, M. L. (2011). Sperm-Storage Defects and Live Birth in Drosophila Females Lacking Spermathecal Secretory Cells. *PLOS Biology*, *9*(11), e1001192. https://doi.org/10.1371/JOURNAL.PBIO.1001192

Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells*, *4*(1–2), 7–25. https://europepmc.org/article/med/747780

Schulte, P. M., Davies, S. A., Dow, J. A. T., & Lukowiak, K. (2014). What is environmental stress? Insights from fish living in a variable environment. *Journal of Experimental Biology*, *217*(1), 23–34. https://doi.org/10.1242/JEB.089722

Schulz, C., Wood, C. G., Jones, D. L., Tazuke, S. I., & Fuller, M. T. (2002). Signaling from germ cells mediated by the rhomboid homolog stet organizes encapsulation by somatic support cells. *Development*, *129*(19), 4523–4534. https://doi.org/10.1242/DEV.129.19.4523

Schweizer, H. P. (2003). Applications of the Saccharomyces cerevisiae Flp-FRT System in Bacterial Genetics. *J Mol Microbiol Biotechnol*, *5*, 67–77. https://doi.org/10.1159/000069976

Selman, K., & Kafatos, F. C. (1974). Transdifferentiation in the labial gland of silk moths: is DNA required for cellular metamorphosis? *Cell Differentiation*, *3*(2), 81–94. https://doi.org/10.1016/0045-6039(74)90030-X

Shen, C. N., Burke, Z. D., & Tosh, D. (2004). Transdifferentiation, Metaplasia and Tissue Regeneration. *Organogenesis*, *1*(2), 36. https://doi.org/10.4161/ORG.1.2.1409

Shi, J., Jin, Z., Yu, Y., Zhang, Y., Yang, F., Huang, H., Cai, T., & Xi, R. (2021). A Progressive Somatic Cell Niche Regulates Germline Cyst Differentiation in the Drosophila Ovary. *Current Biology*, *31*(4), 840-852.e5. https://doi.org/10.1016/J.CUB.2020.11.053

Siegmund-Schultze, N. (2007). Toll-like-Rezeptoren: Neue Zielstruktur f�r immunstimulierende Medikamente. *Dtsch Arztebl International*, *104*(16), A-1072-. https://www.aerzteblatt.de/int/article.asp?id=55316

Slaidina, M., Gupta, S., Banisch, T. U., & Lehmann, R. (2021). A single-cell atlas reveals unanticipated cell type complexity in Drosophila ovaries. *Genome Research*, *31*(10), 1938–1951. https://doi.org/10.1101/GR.274340.120/-/DC1

Song, X., & Xie, T. (2002). DE-cadherin-mediated cell adhesion is essential for maintaining somatic stem cells in the Drosophila ovary. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(23), 14813–14818. https://doi.org/10.1073/PNAS.232389399/ASSET/742B50AE-6B4C-4772-

87A3-A1AFF4701A4A/ASSETS/GRAPHIC/PQ2323893004.JPEG

Staats, S., Lüersen, K., Wagner, A. E., & Rimbach, G. (2018). Drosophila melanogaster as a Versatile Model Organism in Food and Nutrition Research. *Journal of Agricultural and Food Chemistry*, *66*(15), 3737–3753. https://doi.org/10.1021/ACS.JAFC.7B05900/ASSET/IMAGES/ACS.JAFC.7B05900.SOCIAL.JPEG_V03

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, *177*(7), 1888-1902.e21. https://doi.org/10.1016/J.CELL.2019.05.031/ATTACHMENT/2F8B9EBE-54E6-43EB-9EF2-949B6BDA8BA2/MMC3.PDF

Su, Y. H., Rastegri, E., Kao, S. H., Lai, C. M., Lin, K. Y., Liao, H. Y., Wang, M. H., & Hsu, H. J. (2018). Diet regulates membrane extension and survival of niche escort cells for germline homeostasis via insulin signaling. *Development (Cambridge)*, *145*(7). https://doi.org/10.1242/DEV.159186/264561/AM/DIET-REGULATES-MEMBRANE-EXTENSION-AND-SURVIVAL-OF

Sun, Y., Nakashima, M. N., Takahashi, M., Kuroda, N., & Nakashima, K. (2002). Determination of bisphenol A in rat brain by microdialysis and column switching high-performance liquid chromatography with fluorescence detection. *Biomedical Chromatography : BMC*, *16*(5), 319–326. https://doi.org/10.1002/BMC.161

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods 2009 6:5*, *6*(5), 377–382. https://doi.org/10.1038/nmeth.1315

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

Tennessen, J. M., Barry, W. E., Cox, J., & Thummel, C. S. (2014). Methods for studying metabolism in Drosophila. *Methods (San Diego, Calif.)*, *68*(1), 105. https://doi.org/10.1016/J.YMETH.2014.02.034

Tetteh, P. W., Basak, O., Farin, H. F., Wiebrands, K., Kretzschmar, K., Begthel, H., Van Den Born, M., Korving, J., De Sauvage, F., Van Es, J. H., Van Oudenaarden, A., & Clevers, H. (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell*, *18*(2), 203–213. https://doi.org/10.1016/J.STEM.2016.01.001

Tettweiler, G., Miron, M., Jenkins, M., Sonenberg, N., & Lasko, P. F. (2005). Starvation and oxidative stress resistance in Drosophila are mediated through the eIF4E-binding protein, d4E-BP. *Genes & Development*, *19*(16), 1840. https://doi.org/10.1101/GAD.1311805

Theise, N. D., & Krause, D. S. (2002). Toward a new paradigm of cell plasticity. *Leukemia*, *16*(4), 542–548. https://doi.org/10.1038/sj.leu.2402445

Tiwari, S. K., Agarwal, S., Seth, B., Yadav, A., Ray, R. S., Mishra, V. N., &

Chaturvedi, R. K. (2015). Inhibitory Effects of Bisphenol-A on Neural Stem Cells Proliferation and Differentiation in the Rat Brain Are Dependent on Wnt/β-Catenin Pathway. *Molecular Neurobiology*, *52*(3), 1735–1757. https://doi.org/10.1007/S12035-014-8940-1

Tolwinski, N. S. (2017). Introduction: Drosophila—A Model System for Developmental Biology. *Journal of Developmental Biology*, *5*(3). https://doi.org/10.3390/JDB5030009

Torchia, T. E., Hamilton, R. W., Cano, C. L., & Hopper, J. E. (1984). Disruption of regulatory gene GAL80 in Saccharomyces cerevisiae: effects on carbon-controlled regulation of the galactose/melibiose pathway genes. *Molecular and Cellular Biology*, *4*(8), 1521. https://doi.org/10.1128/MCB.4.8.1521

Tosh, D., & Slack, J. M. W. (2002). How cells change their phenotype. *Nature Reviews Molecular Cell Biology 2002 3:3*, *3*(3), 187–194. https://doi.org/10.1038/nrm761

Tu, R., Duan, B., Perera, A., Haug, J., Xie, T., Song, X., Chen, S., Scott, A., Hall, K., Blanck, J., Degraffenreid, D., & Li, H. (2021). Article Multiple Niche Compartments Orchestrate Stepwise Germline Stem Cell Progeny Differentiation Multiple Niche Compartments Orchestrate Stepwise Germline Stem Cell Progeny Differentiation. *Current Biology*, *31*, 827–839. https://doi.org/10.1016/j.cub.2020.12.024

Ugur, B., Chen, K., & Bellen, H. J. (2016). Drosophila tools and assays for the study of human diseases. *Disease Models & Mechanisms*, *9*(3), 235–244. https://doi.org/10.1242/DMM.023762

Valanne, S., Wang, J.-H., & Rämet, M. (2011). The Drosophila Toll Signaling Pathway. *The Journal of Immunology*, *186*(2), 649–656. https://doi.org/10.4049/JIMMUNOL.1002302

Vandenberg, L. N. (2014). NON-MONOTONIC DOSE RESPONSES IN STUDIES OF ENDOCRINE DIS-RUPTING CHEMICALS: BISPHENOL A AS A CASE STUDY. *Formerly Nonlinearity in Biology*, *12*, 259–276. https://doi.org/10.2203/dose-response.13-020.Vandenberg

Vandenberg, L. N., Colborn, T., Hayes, T. B., Heindel, J. J., Jacobs, D. R., Lee, D. H., Shioda, T., Soto, A. M., vom Saal, F. S., Welshons, W. V., Zoeller, R. T., & Myers, J. P. (2012). Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocrine Reviews*, *33*(3), 378. https://doi.org/10.1210/ER.2011-1050

Vandenberg, L. N., Hauser, R., Marcus, M., Olea, N., & Welshons, W. V. (2007). Human exposure to bisphenol A (BPA). *Reproductive Toxicology*, *24*(2), 139–177. https://doi.org/10.1016/J.REPROTOX.2007.07.010

Voog, J., Sandall, S. L., Hime, G. R., Resende, L. P. F., Loza-Coll, M., Aslanian, A., Yates, J. R., Hunter, T., Fuller, M. T., & Jones, D. L. (2014). Escargot restricts niche cell to stem cell conversion in the Drosophila testis. *Cell Reports*, *7*(3), 722. https://doi.org/10.1016/J.CELREP.2014.04.025

Wang, S., Song, P., & Zou, M. H. (2012). AMP-activated protein kinase, stress responses and cardiovascular diseases. *Clinical Science (London, England : 1979)*, *122*(12), 555. https://doi.org/10.1042/CS20110625

Wazir, U., & Mokbel, K. (2019). Bisphenol A: A Concise Review of Literature and

a Discussion of Health and Regulatory Implications. *In Vivo*, *33*(5), 1421. https://doi.org/10.21873/INVIVO.11619

Weiner, A. K., Ramirez, A., Zintel, T., Rose, R. W., Wolff, E., Parker, A. L., Bennett, K., Johndreau, K., Rachfalski, C., Zhou, J., & Smith, S. T. (2014). Bisphenol A affects larval growth and advances the onset of metamorphosis in Drosophila melanogaster. *Ecotoxicology and Environmental Safety*, *101*(1), 7–13. https://doi.org/10.1016/J.ECOENV.2013.12.008

White, M. A., Bonfini, A., Wolfner, M. F., & Buchon, N. (2021). Drosophila melanogaster sex peptide regulates mated female midgut morphology and physiology. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(1). https://doi.org/10.1073/PNAS.2018112118/SUPPL_FILE/PNAS.2018112118.SD03.XLSX

Williams, M. J., Wang, Y., Klockars, A., Monica Lind, P., Fredriksson, R., & Schiöth, H. B. (2014). Exposure to Bisphenol A Affects Lipid Metabolism in Drosophila melanogaster. *Basic & Clinical Pharmacology & Toxicology*, *114*(5), 414–420. https://doi.org/10.1111/BCPT.12170

Wolfner, M. F. (2011). Precious Essences: Female Secretions Promote Sperm Storage in Drosophila. *PLoS Biology*, *9*(11). https://doi.org/10.1371/JOURNAL.PBIO.1001191

Wu, X., Tanwar, P. S., & Raftery, L. A. (2008). Drosophila follicle cells: morphogenesis in an eggshell. *Seminars in Cell & Developmental Biology*, *19*(3), 271. https://doi.org/10.1016/J.SEMCDB.2008.01.004

Wu, Y., Reece, R. J., & Ptashne, M. (1996). Quantitation of putative activator-target affinities predicts transcriptional activating potentials. *EMBO Journal*, *15*(15), 3951–3963. https://doi.org/10.1002/J.1460-2075.1996.TB00769.X

Xue, L. (2000). Drosophila female sexual behavior induced by sterile males showing copulation complementation. *Proceedings of the National Academy of Sciences*, *97*(7), 3272–3275. https://doi.org/10.1073/PNAS.060018897

Yang, H., & Hultmark, D. (2017). Drosophila muscles regulate the immune response against wasp infection via carbohydrate metabolism. *Scientific Reports 2017 7:1*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-15940-2

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: Past, present, and future. *Stem Cell Research and Therapy*, *10*(1), 1–22. https://doi.org/10.1186/S13287-019-1165-5/FIGURES/8

Zhang, Y., Wang, D., Peng, M., Tang, L., Ouyang, J., Xiong, F., Guo, C., Tang, Y., Zhou, Y., Liao, Q., Wu, X., Wang, H., Yu, J., Li, Y., Li, X., Li, G., Zeng, Z., Tan, Y., & Xiong, W. (2021). Single-cell RNA sequencing in cancer research. *Journal of Experimental & Clinical Cancer Research 2021 40:1*, *40*(1), 1–17. https://doi.org/10.1186/S13046-021-01874-1

9. Appendix

R-scripts

Analysis of a single cell RNA-sequencing dataset

```
#analysis of scRNA-Seq datasets
#X stands for respective dataset (FCA EC, Tu, Shi)
getwd()
setwd()
library(Seurat)
library(SeuratObject)
library(ggplot2)
library(SeuratData)
library(patchwork)

X <- readRDS(file.choose())

X[["percent.mt"]] <- PercentageFeatureSet(X, pattern = "MT")
VlnPlot(X, features = c("nFeature_RNA", "nCount_RNA", "percent.mt"),ncol=3)
plot1 <- FeatureScatter(X, feature1 = "nCount_RNA", feature2 = "percent.mt")
plot2 <- FeatureScatter(X, feature1 = "nCount_RNA", feature2 =
"nFeature_RNA")
plot1 + plot2
X<- subset(X, subset = nFeature_RNA > & nFeature_RNA < & nCount_RNA >
& nCount_RNA < & percent.mt <)
VlnPlot(X, features = c("nFeature_RNA", "nCount_RNA", "percent.mt"),ncol=3)
X <- NormalizeData(X, normalization.method = "LogNormalize", scale.factor =
10000)
X <- FindVariableFeatures(X, selection.method = "vst", nfeatures = 3000)
top10 <- head(VariableFeatures(X), 10)
plot1 <- VariableFeaturePlot(X)
plot2 <- LabelPoints(plot = plot1, points = top10, repel = TRUE)
plot1 + plot2
all.genes <- rownames(X)
X <- ScaleData(X, features = all.genes)
X <- RunPCA(X, features = VariableFeatures(object = X))
print(X[["pca"]], dims = 1:5, nfeatures = 5)
VizDimLoadings(X, dims = 1:2, reduction = "pca")
DimPlot(X, reduction = "pca")
DimHeatmap(X, dims = 1:20, cells = 500, balanced = TRUE)
X <- JackStraw(X, num.replicate = 100)
X <- ScoreJackStraw(X, dims = 1:20)
JackStrawPlot(X, dims = 1:20)
ElbowPlot(X)
X <- FindNeighbors(X, dims = )
X <- FindClusters(X, resolution = )
```

```
X <- RunUMAP(X, dims = )
UMAPPlot(X)
X.markers <- FindAllMarkers(X, only.pos = TRUE, min.pct = 0.25,
logfc.threshold = 0.25)
top10_X <- X.markers %>% group_by(cluster) %>% top_n(n = 10, wt =
avg_log2FC)
View(top10_X)
write.csv(top10_X, 'top10_X.csv')
DotPlot(X, features=c('') ) + theme(axis.text.y = element_text(size = 20),
axis.text.x = element_text(angle = 90, size = 20,hjust= 1,vjust=0.2))
FeaturePlot(X, '')

#cluster merge
WhichCells(X, idents = c(''))
X<- SetIdent(X, cells = WhichCells(X, idents = c('')), '')

saveRDS(X, 'X.rds')

#cluster annotation
FeaturePlot(X, features = '', order = T)
X <- CellSelector(FeaturePlot(X, features = '', order = T))
DimPlot(X, cells.highlight = c(''), pt.size = .5, reduction = "") +
scale_color_manual(labels = c(''), values = c("magenta", "#009E73"))
X <- RunTSNE(X)
UMAPPlot(X)
X <- SetIdent(X, cells = WhichCells(X, idents = c('), '')
X <- subset(X, cells = WhichCells(X, idents = ''))
y <- CellSelector(UMAPPlot(X))
y <- append(y, WhichCells(X, idents = ''))

#reference mapping of a query and reference dataset
query <- readRDS(file.choose())
reference <- readRDS(file.choose())
anchors <- FindTransferAnchors(reference = reference, query = query, dims = ,
reference.reduction = "pca")
predictions <- TransferData(anchorset = anchors, refdata = , dims = )
UMAPPlot(query, group.by = "predicted.id", pt.size = 1, label = T)
```

Data integration:

```
library(Seurat)
library(SeuratData)
library(patchwork)
library(SeuratObject)
library(dplyr)
getwd()
setwd()

#X stands for respective dataset (Rust, Jevitt, Slaidina, Tu, Shi)

X<- readRDS(file.choose())

X@meta.data[, "dataset"] <- "X"
X <- RenameCells(X, add.cell.id = "X")

X@meta.data[, "source"] <- "X"
X <- append(WhichCells(list$X), WhichCells(list$X)
X <- subset(combined, cells = X)
X@meta.data[, "source"] <- "X"
combined2 <- merge(x = X, y = c(X),  merge.data = T, project = "")
UMAPPlot (combined2, split.by = "source", ncol = 2, pt.size = 2)

#check for cell amount in datasets; dont include cells with less than 102 cells
all <- merge(x = , y = ,  merge.data = T, project = "")
all <- NormalizeData(all)
DefaultAssay(all) <- "RNA"

all.list <- SplitObject(all, split.by = "dataset")
all.list <- lapply(X = all.list, FUN = function(x) {
  x <- NormalizeData(x)
  x <- FindVariableFeatures(x, selection.method = "vst", nfeatures = 2000)})

features <- SelectIntegrationFeatures(object.list = all.list)
anchors <- FindIntegrationAnchors(object.list = all.list, anchor.features =
features)
combined <- IntegrateData(anchorset = anchors)
DefaultAssay(combined) <- "integrated"
combined <- ScaleData(combined, verbose = FALSE)
combined <- RunPCA(combined, npcs = 30, verbose = FALSE)
combined <- RunUMAP(combined, reduction = "pca", dims = 1:30)
combined <- FindNeighbors(combined, reduction = "pca", dims = 1:30)
combined <- FindClusters(combined, resolution = )
p1 <- DimPlot(combined, reduction = "umap", group.by = "dataset")
p2 <- DimPlot(combined, reduction = "umap", label = TRUE, repel = TRUE)
p1 + p2
DimPlot(combined, reduction = "umap", group.by = "dataset")
```

```
saveRDS()

combined.markers <- FindAllMarkers(combined, only.pos = TRUE, min.pct =
0.25, logfc.threshold = 0.25)
top10 <- combined.markers %>% group_by(cluster) %>% top_n(n = 10, wt =
avg_log2FC)
View(top10)
DotPlot(combined, features = '')
plot1 <- UMAPPlot(combined, pt.size = 3)
plot2 <- FeaturePlot(combined, '', pt.size = 3, order = T)
plot1 + plot2
DimPlot(combined, cells.highlight = WhichCells(combined, idents = ""), pt.size =
1) +
  scale_color_manual(labels = c("other", ""), values = c("magenta", "#009E74"))


#subset of cluster 2 of combined, repeat for cluster 1 of combined
subset_2 <- subset(combined, cells = WhichCells(combined, idents = '2'))
UMAPPlot(subset_2)
subset_2 <-NormalizeData(subset_2)
subset_2 <-FindVariableFeatures(subset_2)
subset_2 <- ScaleData(subset_2, verbose = FALSE)
subset_2 <- RunPCA(subset_2, npcs = 30, verbose = FALSE)
subset_2 <- FindNeighbors(subset_2, reduction = "pca", dims = 1:30)
subset_2 <- FindClusters(subset_2, resolution = )
DimPlot(subset_2, reduction = "umap")
WhichCells(subset_2, idents = c('3', '1'))
subset2_merge <- SetIdent(subset_2, cells = WhichCells(subset_2, idents =
c('3', '1')), '8')
subset2_merge.markers <- FindAllMarkers(subset2_merge, only.pos = TRUE,
min.pct = 0.25, logfc.threshold = 0.25)
top10 <- subset2_merge.markers %>% group_by(cluster) %>% top_n(n = 10,
wt = avg_log2FC)
View(top10)
plot1 <- UMAPPlot(subset2_merge, pt.size = 3)
plot2 <- FeaturePlot(object = subset2_merge, features = '', pt.size = 3, order =
T, reduction = 'umap')
plot1 + plot2
saveRDS(subset2_merge, '.rds')
```

T-test calculation:

```
all <- read.csv(file.choose(), sep = ";")
t <- pairwise.t.test(all$FC.per, all$Condition, p.adjust.method = "none")

setwd()

write.csv(t$p.value, 'all_p_values_by_condition.csv')

all <- read.csv(file.choose(), sep = ";")
t <- pairwise.t.test(all$FC.per, all$Condition_date, p.adjust.method = "none")

write.csv(t$p.value, 'all_p_values_by_condition_by_date.csv')

all <- read.csv(file.choose(), sep = ";")
t <- pairwise.t.test(all$RFP.FC, all$Condition, p.adjust.method = "none")

write.csv(t$p.value, 'all_p_values_by_condition_RFP.csv')

all <- read.csv(file.choose(), sep = ";")
t <- pairwise.t.test(all$RFP.FC, all$Condition, p.adjust.method = "none")

write.csv(t$p.value, 'all_p_values_by_condition_by_date_RFP.csv')
```

EC lineage tracing FC clones (p-values calculated by condition and date)

	14 days cross 21.6	**14 days vial 1 -TM2 23.6**	**14 days vial 1 +TM2 23.6**	**14 days vial 12.6**	**14 days vial 3 -TM2 23.6**	**14 days vial 3 +TM2 23.6**	**14 days vial 3 2.6**
14 days vial 1 -TM2 23.6	6,47E-35	NA	NA	NA	NA	NA	NA
14 days vial 1 +TM2 23.6	0,33434638	8,70E-28	NA	NA	NA	NA	NA
14 days vial 12.6	0,7342247	8,23E-39	0,18659664	NA	NA	NA	NA
14 days vial 3 -TM2 23.6	3,15E-06	2,27E-15	0,00048751	2,38E-07	NA	NA	NA
14 days vial 3 +TM2 23.6	1,10E-05	3,54E-16	0,00120749	9,89E-07	0,79836409	NA	NA
14 days vial 3 2.6	0,53822352	3,87E-35	0,65658164	0,31129781	1,53E-05	5,19E-05	NA
14 days vial 5 -TM2 23.6	0,92081683	5,53E-31	0,41780069	0,67755605	1,90E-05	5,60E-05	0,64445291
14 days vial 5 +TM2 23.6	0,25626989	3,59E-27	0,8735165	0,13411339	0,00086358	0,00206522	0,53359104
14 days vial 5 2.6	0,01159579	4,74E-29	0,19683696	0,00231607	0,0070907	0,01616157	0,04325026
3 days cross 10.6	0,74032164	1,40E-37	0,19563656	1	4,07E-07	1,61E-06	0,32436433
7 days cross 14.6	0,75567102	1,77E-34	0,22018015	1	1,54E-06	5,33E-06	0,35813274
7 days vial 1 26.5	0,0349291	1,43E-30	0,36396824	0,00913238	0,00213647	0,00536406	0,11450735
7 days vial 3 26.5	0,9456544	4,95E-33	0,38220277	0,69036234	7,95E-06	2,55E-05	0,60243492
7 days vial 5 26.5	0,63447835	4,15E-33	0,59915455	0,40068319	2,28E-05	7,17E-05	0,90890415

	14 days cross 21.6	14 days vial 1 -TM2 23.6	14 days vial 1 +TM2 23.6	14 days vial 12.6	14 days vial 3 -TM2 23.6	14 days vial 3 +TM2 23.6	14 days vial 3 2.6
BPA 0,1mM 15.7	0,5784 1013	1,31E-40	0,550776 31	0,316 58964	1,57E-06	6,71E-06	0,8927 7039
BPA 1mM 15.7	0,2828 1556	1,67E-37	0,915688 18	0,122 37867	2,06E-05	7,43E-05	0,6619 812
BPA 20 mM 15.7	0,5559 6215	2,31E-39	0,585036 22	0,305 1426	2,82E-06	1,14E-05	0,9332 8891
Control 15.7	0,8458 2427	5,49E-43	0,206082 79	0,848 80118	6,39E-08	3,15E-07	0,3509 136
Control 16.7	0,8362 9982	4,92E-43	0,202168 87	0,859 06457	6,03E-08	2,98E-07	0,3442 0864
Control 26.7	0,9555 2209	5,18E-42	0,298989 65	0,645 9596	2,08E-07	9,70E-07	0,5064 7464
Control 3.7	0,9463 674	2,30E-44	0,237923 62	0,731 4783	5,27E-08	2,74E-07	0,4086 1834
Control 5.7	0,7046 9214	9,95E-45	0,148044 47	1	1,67E-08	8,98E-08	0,2510 2131
Control 6.8	0,7567 7007	7,54E-43	0,174087 47	0,950 29121	4,71E-08	2,33E-07	0,2951 9466
dIRNAi starved 16.7	0,2138 2037	7,35E-37	0,961900 59	0,084 38462	3,97E-05	0,000137 6	0,5359 9955
dIRNAi 5.7	0,7513 483	4,59E-46	0,157057 45	0,936 82857	1,14E-08	6,46E-08	0,2678 9487
High fat 6.8	0,8876 1589	1,92E-40	0,347098 77	0,589 07982	5,13E-07	2,23E-06	0,5796 2992
Starved 15.7	0,8843 9352	3,39E-42	0,333672 54	0,575 17091	2,52E-07	1,18E-06	0,5644 701
Starved 16.7	0,8018 8279	3,56E-44	0,361325 36	0,485 86977	1,51E-07	7,67E-07	0,6178 0067
Sucrose 26.7	0,0710 4591	3,64E-37	0,613778 93	0,018 19457	0,000133 93	0,000448 67	0,2307 0158
tIRNAi starved 16.7	0,7126 5007	8,44E-41	0,449474 66	0,426 68244	8,57E-07	3,73E-06	0,7423 4397
tIRNAi 5.7	0,8934 0433	9,03E-44	0,220083 49	0,791 80098	5,52E-08	2,81E-07	0,3764 9519
Virgins 15.7	0,8278 2994	2,20E-41	0,372090 52	0,526 73174	4,36E-07	1,96E-06	0,6236 1937

	14 days vial 5 -TM2 23.6	14 days vial 5 +TM2 23.6	14 days vial 5 2.6	3 days cross 10.6	7 days cross 14.6	7 days vial 1 26.5	7 days vial 3 26.5
14 days vial 5 2.6	0,027 27111	0,267009	NA	NA	NA	NA	NA
3 days cross 10.6	0,683 71334	0,1421350 3	0,0030 5337	NA	NA	NA	NA
7 days cross 14.6	0,699 53545	0,1642997 1	0,0058 5569	1	NA	NA	NA
7 days vial 1 26.5	0,067 73618	0,4672601 9	0,6480 0615	0,0112 4777	0,0183 8803	NA	NA
7 days vial 3 26.5	0,973 60752	0,2989057 8	0,0185 554	0,6968 6687	0,7134 1163	0,0505 8691	NA
7 days vial 5 26.5	0,733 25768	0,4866221 4	0,0436 7373	0,4119 8509	0,4411 2663	0,1094 5705	0,6956 6951
BPA 0,1mM 15.7	0,697 46293	0,4295347	0,0155 9172	0,3329 2773	0,3740 8569	0,0549 3651	0,6512 5276
BPA 1mM 15.7	0,388 3825	0,7671328 1	0,0755 3287	0,1346 6897	0,1681 3085	0,1974 6497	0,3419 3466
BPA 20 mM 15.7	0,672 57532	0,4616214 8	0,0209 4299	0,3207 3766	0,3603 4259	0,0684 1655	0,6268 4436
Control 15.7	0,773 33135	0,1454533 5	0,0013 6758	0,8536 6288	0,8653 3065	0,0066 8854	0,7920 5284
Control 16.7	0,764 91248	0,1424422 9	0,0013 0612	0,8636 0323	0,8744 9295	0,0064 2513	0,7830 9865
Control 26.7	0,950 80571	0,2186361 2	0,0033 9698	0,6567 226	0,6827 9248	0,0147 4619	0,9804 0892
Control 3.7	0,860 46131	0,1684874 6	0,0014 1262	0,7405 5623	0,7621 6187	0,0073 3443	0,8851 7599
Control 5.7	0,648 76592	0,1008747	0,0005 1599	1	1	0,0029 5041	0,6594 1698
Control 6.8	0,695 18292	0,1214414	0,0010 1463	0,9518 6772	0,9556 597	0,0050 4785	0,7089 3207
dIRNAi starved 16.7	0,309 65284	0,8866812 2	0,1115 6563	0,0946 8714	0,1236 1979	0,2703 2037	0,2660 6282
dIRNAi 5.7	0,688 72639	0,1065724 9	0,0004 3694	0,9391 0685	0,9444 6721	0,0027 0613	0,7019 0803
High fat 6.8	0,985 36637	0,2592193 9	0,0058 833	0,6006 4207	0,6289 8262	0,0228 2263	0,9528 7583
Starved 15.7	0,985 45602	0,2460846 6	0,0041 3755	0,5879 3787	0,6189 2696	0,0177 6802	0,9519 4777
Starved 16.7	0,914 3439	0,2657183 1	0,0035 4628	0,5019 6879	0,5408 2804	0,0167 8294	0,8753 5112
Sucrose 26.7	0,131 52601	0,7586247 7	0,2627 5183	0,0226 5171	0,0371 3456	0,5472 0173	0,1010 1516
tIRNAi starved 16.7	0,825 08071	0,3431084 2	0,0096 4225	0,4418 6776	0,4794 7699	0,0360 6983	0,7840 1685

	14 days vial 5 -TM2 23.6	14 days vial 5 +TM2 23.6	14 days vial 5 2.6	3 days cross 10.6	7 days cross 14.6	7 days vial 1 26.5	7 days vial 3 26.5
tIRNAi 5.7	0,814 47197	0,1554106 9	0,0013 4796	0,7987 057	0,8151 9193	0,0068 3762	0,8359 923
Virgins 15.7	0,932 59229	0,278263	0,0059 0392	0,5402 5914	0,5733 5501	0,0237 2811	0,8965 9079

	7 days vial 5 26.5	BPA 0,1mM 15.7	BPA 1mM 15.7	BPA 20 mM 15.7	Control 15.7	Control 16.7	Control 26.7
BPA 0,1mM 15.7	0,9989 7045	NA	NA	NA	NA	NA	NA
BPA 1mM 15.7	0,5950 0152	0,5122 3185	NA	NA	NA	NA	NA
BPA 20 mM 15.7	0,9633 782	0,9559 9321	0,560 35886	NA	NA	NA	NA
Control 15.7	0,4573 1979	0,3547 4501	0,123 23942	0,34 1490 03	NA	NA	NA
Control 16.7	0,4499 875	0,3469 3866	0,119 72441	0,33 4105 77	0,988097 58	NA	NA
Control 26.7	0,6218 6322	0,5387 6501	0,214 6539	0,51 5634 41	0,759263 92	0,747935 34	NA
Control 3.7	0,5251 3818	0,4194 5091	0,144 81157	0,40 2621 41	0,867263 15	0,854907 88	0,87611 099
Control 5.7	0,3482 1612	0,2380 5482	0,071 37318	0,23 1221 24	0,823012 06	0,834976 5	0,58993 077
Control 6.8	0,3939 7396	0,2915 1378	0,097 46703	0,28 1577 33	0,886269 9	0,897871 43	0,65671 244
dIRNAi starved 16.7	0,4833 0121	0,3872 3206	0,838 9502	0,43 2230 26	0,080667 39	0,078166 89	0,14797 975
dIRNAi 5.7	0,3722 2311	0,2533 6426	0,073 05206	0,24 6110 47	0,887151 29	0,899844 61	0,63719 637
High fat 6.8	0,6927 8929	0,6264 9096	0,272 71282	0,59 9420 7	0,688277 81	0,677701 81	0,91625 057

	7 days vial 5 26.5	BPA 0,1mM 15.7	BPA 1mM 15.7	BPA 20 mM 15.7	Control 15.7	Control 16.7	Control 26.7
Starved 15.7	0,68248451	0,60800182	0,24958822	0,5810849 4	0,67399379	0,66297416	0,91269254
Starved 16.7	0,74269926	0,67229757	0,26846224	0,6408270 5	0,56535321	0,55445319	0,80864169
Sucrose 26.7	0,21526959	0,11666681	0,40245293	0,1447291 9	0,01250291	0,01196501	0,02942926
tIRNAI starved 16.7	0,85658449	0,82100325	0,38701737	0,7843428 5	0,49077807	0,48144705	0,70189252
tIRNAi 5.7	0,48811006	0,3828898	0,13188166	0,3681271 3	0,93675899	0,92455761	0,81343478
Virgins 15.7	0,74023455	0,67906818	0,29534038	0,6488555 05	0,6139438	0,60350755	0,84299927

	Control 3.7	Control 5.7	Control 6.8	dIRNAi starved 16.7	dIRNAi 5.7	High fat 6.8	Starved 15.7
Control 5.7	0,6824944 5	NA	NA	NA	NA	NA	NA
Control 6.8	0,7530629 6	0,9419663 5	NA	NA	NA	NA	NA
dIRNAi starved 16.7	0,0940444 6	0,0440082 8	0,0629714 6	NA	NA	NA	NA
dIRNAi 5.7	0,7396288 1	0,9242292 9	0,9886183 1	0,0441434 6	NA	NA	NA
High fat 6.8	0,7944856 1	0,5302771 1	0,5937781 9	0,1949625 1	0,57155056	NA	NA
Starved 15.7	0,7838752 1	0,5090122 2	0,5763813 9	0,1737468 7	0,5497952	0,999351 15	NA
Starved 16.7	0,6680272 6	0,4016322 2	0,4723216 6	0,1836375 6	0,43316809	0,905230 05	0,8988641
Sucrose 26.7	0,0134102 8	0,0049327 8	0,0092751 1	0,5351396 6	0,00425557	0,046933 68	0,03580481
tIRNAI starved 16.7	0,5759522 5	0,3511045 5	0,4117841 5	0,2840841 6	0,37713239	0,791278 69	0,78032138

	Control 3.7	Control 5.7	Control 6.8	dIRNAi starved 16.7	dIRNAi 5.7	High fat 6.8	Starved 15.7
tIRNAi 5.7	0,92936986	0,75478856	0,82197307	0,08579152	0,81616575	0,737225 17	0,72433938
Virgins 15.7	0,71537149	0,45772836	0,52269483	0,21039178	0,4935825	0,93 1316 78	0,92717677

	Starved 16.7	Sucrose 26.7	tIRNAI starved 16.7	tIRNAi 5.7
Sucrose 26.7	0,03299883	NA	NA	NA
tIRNAI starved 16.7	0,86222168	0,07557358	NA	NA
tIRNAi 5.7	0,61136041	0,01259961	0,52885428	NA
Virgins 15.7	0,97839415	0,04878194	0,85343828	0,6601108

EC lineage tracing FC clones (p-values calculated by condition)

	14 days cross	14 days vial 1	14 days vial 1 -TM2	14 days vial 1 +TM2	14 days vial 3	14 days vial 3 -TM2	14 days vial 3 +TM2
14 days vial 1	0,73254858	NA	NA	NA	NA	NA	NA
14 days vial 1 -TM2	2,34E-35	2,64E-39	NA	NA	NA	NA	NA
14 days vial 1 +TM2	0,33118647	0,18372411	3,91E-28	NA	NA	NA	NA
14 days vial 3	0,5355623	0,3081355	1,39E-35	0,65447571	NA	NA	NA
14 days vial 3 -TM2	2,71E-06	1,98E-07	1,49E-15	0,0004474	1,35E-05	NA	NA
14 days vial 3 +TM2	9,58E-06	8,39E-07	2,27E-16	0,00112063	4,63E-05	0,79707075	NA
14 days vial 5	0,01106079	0,00216667	2,05E-29	0,19391666	0,04189685	0,00672386	0,0154768
14 days vial 5 -TM2	0,92029935	0,67556196	2,25E-31	0,41475427	0,64228393	1,67E-05	5,01E-05
14 days vial 5 +TM2	0,2531647	0,13157982	1,65E-27	0,8726942	0,53091118	0,00079816	0,0019293

	14 days cross	14 days vial 1	14 days vial 1 -TM2	14 days vial 1 +TM2	14 days vial 3	14 days vial 3 -TM2	14 days vial 3 +TM2
3 days cross	0,7386 8095	1	4,66E-38	0,1927215 9	0,3212 0165	3,41E-07	1,37E-06
7 days cross	0,7541 2049	1	6,50E-35	0,2171695 8	0,3549 8944	1,31E-06	4,63E-06
7 days vial 1	0,0337 5263	0,0086 8623	5,91E-31	0,3608310 5	0,1121 4972	0,0019 9668	0,0050694 7
7 days vial 3	0,9452 9862	0,6884 3818	1,90E-33	0,3790896 1	0,6000 5752	6,93E-06	2,26E-05
7 days vial 5	0,6322 5844	0,3976 0154	1,58E-33	0,5967615 9	0,9083 0946	2,01E-05	6,42E-05
BPA 0,1 mM	0,5759 2125	0,3134 2662	3,98E-41	0,5481669	0,8920 7157	1,34E-06	5,84E-06
BPA 1mM	0,2796 7093	0,1199 466	5,56E-38	0,9151374 3	0,6599 0373	1,82E-05	6,66E-05
BPA 20 mM	0,5533 7471	0,3019 821	7,27E-40	0,5825774 9	0,9328 5251	2,43E-06	9,97E-06
Control	0,8305 2154	0,8011 5076	5,16E-53	0,1512964 2	0,2614 5055	7,43E-10	5,72E-09
dIRNAi	0,7497 7224	0,9364 1523	1,17E-46	0,1543545 4	0,2647 695	9,12E-09	5,30E-08
dIRNAi starved	0,2108 3195	0,0823 6738	2,50E-37	0,9616509 6	0,5333 2935	3,53E-05	0,0001243 1
High fat	0,8868 8393	0,5866 3972	5,86E-41	0,3439460 6	0,5771 4656	4,32E-07	1,91E-06
starved	0,8237 2016	0,4824 6091	3,49E-48	0,3164832 3	0,5570 4716	2,64E-08	1,60E-07
sucrose	0,0692 1348	0,0174 4807	1,23E-37	0,6114561 9	0,2276 5813	0,0001 2095	0,0004113 4
tIRNAi	0,8927 096	0,7904 6754	2,48E-44	0,2170732 5	0,3733 7373	4,52E-08	2,35E-07
tIRNAI starved	0,7108 5037	0,4236 5646	2,54E-41	0,4465075 4	0,7407 1508	7,26E-07	3,22E-06
virgins	0,8267 1888	0,5240 2477	6,51E-42	0,3689631 4	0,6213 4502	3,66E-07	1,68E-06

	14 days vial 5	14 days vial 5 -TM2	14 days vial 5 +TM2	3 days cross	7 days cross	7 days vial 1	7 days vial 3
14 days vial 5 -TM2	0,0262 7678	NA	NA	NA	NA	NA	NA
14 days vial 5 +TM2	0,2638 85	0,3292 7706	NA	NA	NA	NA	NA
3 days cross	0,0028 6604	0,6817 5272	0,1395381	NA	NA	NA	NA

	14 days vial 5	14 days vial 5 -TM2	14 days vial 5 +TM2	3 days cross	7 days cross	7 days vial 1	7 days vial 3
7 days cross	0,0055 3996	0,6976 621	0,1615505 3	1	NA	NA	NA
7 days vial 1	0,6458 5551	0,0659 5354	0,4643450 5	0,0107 2493	0,0176 3579	NA	NA
7 days vial 3	0,0177 9822	0,9734 3453	0,2957482 1	0,6949 7868	0,7116 1624	0,0490 9216	NA
7 days vial 5	0,0423 1181	0,7315 7595	0,4837692 1	0,4089 2607	0,4381 3692	0,1071 5009	0,6937 7468
BPA 0,1 mM	0,0149 2475	0,6955 7805	0,4265155 9	0,3297 6727	0,3709 6094	0,0533 6354	0,6491 1898
BPA 1mM	0,0736 3554	0,3852 7909	0,7656504 3	0,1321 3085	0,1653 5848	0,1945 4191	0,3387 7855
BPA 20 mM	0,0201 1704	0,6705 5436	0,4586893 1	0,3175 7462	0,3572 0151	0,0666 2353	0,6245 8607
Control	8,01E-05	0,7522 5498	0,0980827	0,8103 1632	0,8308 3684	0,0008 3494	0,7703 0186
dIRNAi	0,0004 0045	0,6867 9322	0,1042955 3	0,9387 0836	0,9441 0368	0,0025 3638	0,7000 4791
dIRNAi starved	0,1092 3723	0,3064 9091	0,8859432 5	0,0925 426	0,1211 7647	0,2671 9137	0,2629 4032
High fat	0,0055 664	0,9852 7042	0,2561086 4	0,5982 5616	0,6267 3503	0,0219 447	0,9525 672
starved	0,0013 3634	0,9412 9405	0,2252917 2	0,5008 9119	0,5444 2411	0,0082 6644	0,9013 842
sucrose	0,2596 3481	0,1290 1389	0,7570917 4	0,0217 7844	0,0359 0927	0,5445 7738	0,0987 9831
tIRNAi	0,0012 5268	0,8132 7784	0,1527187 6	0,7974 1446	0,8140 0226	0,0064 8104	0,8349 3233
tIRNAi starved	0,0091 7715	0,8239 525	0,3399530 2	0,4388 8	0,4766 0044	0,0348 6792	0,7826 361
virgins	0,0055 8614	0,9321 5137	0,2751234 8	0,5376 0619	0,5708 4345	0,0228 2576	0,8959 1659

	7 days vial 5	BPA 0,1mM	BPA 1mM	BPA 20 mM	Control	dIRNAi	dIRNAi starved
BPA 0,1 mM	0,9989 637	NA	NA	NA	NA	NA	NA
BPA 1mM	0,5925 8899	0,5094 6959	NA	NA	NA	NA	NA
BPA 20 mM	0,9631 3824	0,9557 0495	0,557 7902 7	NA	NA	NA	NA
Control	0,3869 2215	0,2248 1921	0,046 4718 8	0,2220 8565	NA	NA	NA

	7 days vial 5	BPA 0,1mM	BPA 1mM	BPA 20 mM	Control	dlRNAi	dlRNAi starved
dlRNAi	0,36909589	0,25026487	0,0711902 5	0,24302687	0,82154034	NA	NA
dlRNAi starved	0,48043721	0,38412678	0,8379088	0,42921777	0,02388815	0,04277216	NA
High fat	0,69087851	0,62423091	0,2695804 4	0,59702899	0,61942894	0,56903097	0,19205056
starved	0,69278148	0,59584001	0,2012178 9	0,56721548	0,37764953	0,41173587	0,12751599
sucrose	0,21227597	0,11428823	0,3993746 6	0,14211276	0,00065024	0,00401062	0,532466
tlRNAi	0,48526214	0,37977768	0,1293665 6	0,36499478	0,93364445	0,81498212	0,08375615
tlRNAl starved	0,85565441	0,81984965	0,3839117 4	0,78296407	0,37433467	0,37401182	0,28093822
virgins	0,73859335	0,67708228	0,2921850 1	0,64670881	0,5193769	0,4907533	0,20741605

	High fat	starved	sucrose	tlRNAi	tlRNAl starved
starved	0,94036126	NA	NA	NA	NA
sucrose	0,04550775	0,01400261	NA	NA	NA
tlRNAi	0,73556646	0,61296954	0,01203018	NA	NA
tlRNAl starved	0,78994206	0,80179021	0,07367568	0,52615564	NA
virgins	0,93086755	0,97703255	0,04732082	0,65802344	0,85248827

EC lineage tracig RFP positive FSCs (p-values calculated by condition and date):

	BPA 0,1mM 15.7	BPA, 1mM 15.7	BPA, 20mM 15.7	Control 15.7	Control 16.7	Control 26.7	Control 5.7
BPA, 1mM 15.7	0,3908665	NA	NA	NA	NA	NA	NA
BPA, 20mM 15.7	0,09070903	0,41678719	NA	NA	NA	NA	NA
Control 15.7	0,00022774	0,00592432	0,05250068	NA	NA	NA	NA
Control 16.7	1,34E-05	0,00064825	0,00917256	0,49204104	NA	NA	NA
Control 26.7	3,03E-05	0,00123467	0,01537696	0,61792016	0,85052843	NA	NA
Control 5.7	2,26E-07	2,73E-05	0,00072745	0,15082852	0,46435691	0,35484292	NA
Control 6.8	1,24E-07	1,33E-05	0,00034808	0,08424536	0,29186004	0,2154333	0,71497752
dIRNAi 5.7	6,69E-09	1,93E-06	8,89E-05	0,05532857	0,23735644	0,166517	0,66599688
dIRNAi, starved 16.7	5,48E-07	4,55E-05	0,00098429	0,15186034	0,44778437	0,34510446	0,95109649
High fat 6.8	0,03311721	0,20853307	0,64481079	0,14989888	0,03589038	0,05509534	0,00439967
starved 15.7	0,00256685	0,03777415	0,21265973	0,4605428	0,15176917	0,21387855	0,02691897
starved 16.7	0,00273313	0,04682304	0,27134124	0,30562457	0,07725264	0,11787579	0,0089038
sucrose 26.7	0,03465321	0,00354452	0,0001711	7,47E-09	1,24E-10	3,95E-10	2,80E-13
tIRNAi 5.7	2,54E-08	4,51E-06	0,00015819	0,06375825	0,24971197	0,1789064	0,6648216
tIRNAi, starved 16.7	6,99E-08	9,00E-06	0,00026126	0,07601933	0,27533535	0,2010457	0,6953487
virgins 15.7	0,25676875	0,7994508	0,56782545	0,01097325	0,0013031	0,00243059	6,02E-05

	Control 6.8	dIRNAi 5.7	dIRNAi, starved 16.7	High fat 6.8
dIRNAi 5.7	0,97730174	NA	NA	NA
dIRNAi, starved 16.7	0,77110197	0,72956687	NA	NA
High fat 6.8	0,00215253	0,0007879	0,00528492	NA
starved 15.7	0,01356636	0,00625982	0,02977539	0,45331363
starved 16.7	0,00424888	0,00134011	0,01091717	0,56697384
sucrose 26.7	2,81E-13	1,07E-15	2,02E-12	3,23E-05
tIRNAi 5.7	0,96175615	0,98183447	0,72498522	0,00118243
tIRNAi, starved 16.7	0,98470787	0,99359299	0,7529512	0,00172914
virgins 15.7	2,91E-05	4,52E-06	9,71E-05	0,30377742

	starved 15.7	starved 16.7	sucrose 26.7	tIRNAi 5.7	tIRNAi, starved 16.7
starved 16.7	0,81060361	NA	NA	NA	NA
sucrose 26.7	2,55E-07	1,11E-07	NA	NA	NA
tIRNAi 5.7	0,00846555	0,00217958	1,53E-14	NA	NA
tIRNAi, starved 16.7	0,01142155	0,00334039	9,91E-14	0,97707185	NA
virgins 15.7	0,06387422	0,08036617	0,00122131	1,02E-05	1,99E-05

List of academic teachers

My academic teachers during the study of dentistry in Marburg were:

Adamkiewicz, Althaus, Arweiler, Auschill, Brandt, Czubayko, Feuser, Frankenberger, Gente, Hoch, Huster, Höffken, Kinscherf, Kirschbaum, Korbmacher-Steiner, Lill, Lotzmann, Mandic, Milani, Mittag, Moll, Neff, Neumüller, Plant, Ramaswamy, Roggendorf, Schütz, Weber, Weihe, Westermann, Wilhelm, Winter, Ziebart

Acknowledgements

First of all, I would like to thank the powerhouse of a woman Dr. Katja Rust for making this possible, her great supervision, her open ear for every question and the constant support during all the stages of the project. I really appreciate all the time you took for me and the patience and trust you have shown me. Furthermore, I am very happy to have accepted a challenge with this project, because otherwise I would not have gotten to know the great team of the Bogdan Lab. Thank you to each and every one for the great time, patient help (whenever I drowned in panic fear of breaking the Sp8), assistance, advice, profound conversations, company on countless weekends in the lab, all the laughs and encouragement. Special thanks to my “Schubi” for your constant care and “Stimmung!!”. To all of my friends, thanks for your enthusiasm even when you have not heard the word "Drosophila" before. My dear Jannik and "Residenz", you guys have vividly seen me struggle with two jobs and a full time PhD program. Thank you for your cooking, encouragement, interest, care, hype and patience - you are simply amazing. Last but certainly not least, the greatest tribute goes to my family. To my sister and how I like to call him "Schwiegerbruder", thank you for your unconditional support and belief and the fact that you always had an open door (and fridge) for me. To my parents, who through their hard work and in spite of all the hardships we have faced made it possible for me to get the education I have enjoyed to this day. Your strength and endurance keeps inspiring me and I stand in awe of you and everyone who has lost lives and existence to struggle to rebuild it. I deeply respect and value the courage, risk and effort that comes with it. In this sense a very special thanks to my mother, your self-sacrifice is indescribable, I owe you everything and in a lifetime full of gifts I could not repay you. My love is hardly describable in words. Moj uspjeh će zauvijek biti vaš uspjeh.