Characterizing new players involved in iron homeostasis during *Drosophila* larval development: Shifting the classic paradigm of iron metabolism

by

Nhan Thanh Huynh

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences

University of Alberta

Abstract

Iron plays a critical role in many biological processes. The ability of iron to gain and lose electrons allows it to be involved in a wide variety of biochemical reactions. Organismal iron is commonly found in two types of protein cofactors: heme and iron-sulfur clusters (ISCs). Iron is critical for many cellular processes, including oxygen transport, energy production, steroid hormone synthesis. Iron's capacity for electron transfer is also a double-edged sword that can generate cell-damaging radicals. Iron-homeostasis has been linked to many diseases including hemochromatosis, the porphyrias, Friedreich ataxia, and sideroblastic anemia. Thus, iron levels have to be tightly regulated at the systemic as well as cellular levels. Many aspects of cellular iron biology remain unexplored, and many genes functions in iron metabolism still stay hidden.

This thesis focuses on increasing understanding the regulatory mechanism by which iron and heme metabolism is coupled with steroid hormone production. I first investigated the role of *AGBE* in iron regulation. Prothoracic gland (PG)-knock down of this gene results in a porphyria phenotype that can be rescued in an iron-supplemented medium. Further investigation allows us to establish the relationship between *AGBE*, *Cisd2*, and *IRP1A* in *Drosophila* iron metabolism. In *Drosophila*, Cisd2 is required to maintain the intact ISC in IRP1A, and AGBE will act as a bridge to strengthen the interaction between Cisd2 and oxidatively damage IRP1A for the repair process.

I also report the nuclear localization of IRP1A. This hitherto undocumented localization is tissues-specific and iron-sensitive. The unexpected presence in the nucleus suggests a function of IRP1A in gene regulation. Further work allowed me to propose a model where the nuclear IRP1A might participate in expression regulation of iron-related genes by regulating nuclear citrate level, a substrate of nuclear acetyl-CoA synthesis for histone acetylation. This finding has added an entirely unexpected aspect not explained before of holo-IRP1A in iron metabolism.

I also characterized functions for a gene called *ppk20* in iron or heme metabolism. This gene was identified from earlier genome-wide screening. *ppk20* is a member of *Drosophila* epithelial sodium channel (ENaC). PG-knock down of *ppk20* results in porphyria phenotype and trachea necrosis, both can result from abnormal iron homeostasis. These phenotypes can be rescued by hemin, injected ferritin, and human transferrin receptor. Thus, these data suggest a role for *ppk20* in iron metabolism.

I generated two CRISPR/Cas9 toolkits, which allow spatial and temporal gene manipulation. With these, one can generate somatic mutations, interfere with transcription or induce gene expression in the tissue of interest and at the desired time points. I also evaluated the efficiency and potential applications of another CRISPR system, Cas13, in *Drosophila*. Unlike Cas9, which is used to target DNA, Cas13 target RNA with high efficiency, and current preliminary data suggests its great potential in RNA targeting.

Preface

This thesis is an original work by Nhan Huynh.

The majority of results discussed in chapter 3 has been published:

Huynh N., Ou Q., Cox P., Lill R. and King-Jones K. Glycogen Branching Enzyme controls cellular iron homeostasis via Iron Regulatory Protein 1 and mitoNEET. Nature communications. (2019)10:5463. doi: 10.1038/s41467-019-13237-8.

In this manuscript, I performed and analyzed most of the experiments, generated the original versions and revised figures with Dr. King-Jones' suggestion. I also wrote portions and proof-read the manuscript. Dr. King-Jones performed RNA-sequencing analysis in chapter 3, while the analysis of mass-spectrometry in this chapter was done by Jack Moore at the Mass Spectrometry facility in Biochemistry Department. I am responsible for all other figures and tables.

Two aspects of chapter 6 have been published in two papers:

Huynh N, Zeng J, Liu W, King-Jones K. A *Drosophila* CRISPR/Cas9 toolkit for conditionally manipulating gene expression in the prothoracic gland as a test case for polytene tissues. G3. 2018;8:3593–3605. doi: 10.1534/g3.118.200539.

Huynh N, Wang S, King-Jones K. Spatial and temporal control of gene manipulation in *Drosophila* via drug-activated Cas9 nucleases. Insect Biochemistry and Molecular Biology. 2020;120:103336. doi: 10.1016/j.ibmb.2020.103336.

In these two manuscripts, all figures came from my work. I also wrote the manuscripts, generated, and revised the figures with Dr. King-Jones' suggestion.

A manuscript corresponding to chapter 7 has been submitted for publication. Like chapter 6, I carried out most of the work with assistance from two undergraduate students whose names are listed as co-authors. I wrote the manuscript, generated, and revised figures with Dr. King-Jones' suggestions.

This thesis is dedicated to the two people who most inspired me, despite never having the

privilege of seeing them in person.

Fanis Missirlis, the National Polytechnique Institute, Mexico

Feng Zhang, McGovern Institute, MIT, United States

Acknowledgements

I wish to acknowledge the support and great love of my family in Vietnam. I wish I could invite them to my thesis defense, but it might not be feasible with the current situation of Covid-19. I would like to thank my wife for her tireless support. I still remember our video calls when I did overnight experiments, wrote my manuscripts, proposals, or thesis. Thanks for listening to my everyday stories, complaints, and helping me with my stress. You are the only one who knows how often I failed an experiment before being able to make it work. Thanks for sharing your thoughts from a non-science perspective, which helped me a lot in my research. Marrying you is the best decision I have ever made. Thanks for being the right person at the right time.

Special thanks to the Pokemon Go community at the University of Alberta and team Instinct at Edmonton for making my Ph.D. life more colorful. I have had a chance to meet different people from different backgrounds, and many of them in turn helped me a lot with my projects. Ph.D. life is tough so having a chance to know you guys is a treasure for me. I still remember times we run as fast as we can between gyms during three-hour raids or tell others about the perfect IV Scyther near Tower of Flame. I miss you all.

Table of content

xi

xiii

Table of Figures

Table of tables

Abbreviations

20E	20-hydroxyecdysone
7dC	7-dehydrocholesterol
Aco1	human iron regulatory protein 1
ACO2	mammalian mitochondrial aconitase
Acon	*Drosophila* mitochondrial aconitase
ACL	ATP-citrate lyase
act	*actin* gene
ADAR2	Adenosine Deaminase Acting on RNA 2
AGBE	*1,4-Alpha-Glucan Branching Enzyme* gene
ALA	aminolevulinic acid
ALAD	ALA dehydratse
ALAS	ALA synthase
APBD	adult polyglucosan body disease
ASICs	acid-sensing ion channels
BPS	Bathophenanthroline Sulfate
BRGC	brain-ring gland complex
BzoCas13b	*Bergeyella zoohelcum* Cas13b
CA	corpus allatum
CAI	codon adaption index
Cas	CRISPR-associated proteins
CasRX	*Ruminococcus flavefaciens* XPD3002 Cas13d
CC	corpora cardiaca

cDNA	complementary DNA
CHX	cycloheximide
CIA	cytosolic Iron-sulphur protein Assembly
CLIP	cross-link immunoprecipitation
CO	carbon monoxide
coIP	coimmunoprecipitation
COMMD10	copper metabolism Murr1 domain containing 10
CP	ceruloplasmin
CPOX	coproporphyrinogen oxidase
CRAPOME	Contaminant Repository for Affinity Purification
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DBSs	double strand breaks
dCas9	nuclease-dead Cas9
dib	*disembodied* gene
DmTfR	*Drosophila* transferrin receptor
DR	direct repeat
DCYTB	duodenal cytochrome b
DMT1	divalent metal-ion transporter 1
DEER	double electron-electron resonance spectroscopy
dU6:3	*Drosophila* U6:3 promoter
eCFP	enhanced cyan fluorescent protein
ECF	extracellular fluid

EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
EMSA	electrophoretic mobility shift assay
ENaC	epithelial sodium channel
ENC	effective number of codons
EPR	electron paramagnetic resonance
ER	endoplasmic reticulum
FAC	ferric ammonium citrate
FB	fat body
FECH	ferrochelatase
FKBP12	12-kDa FK506 binding protein
FKBP	12-kDa FK506 binding protein
FLP	flippase
FpeCas13c	*Fusobacterium perfoetens Cas13c*
FPN1	ferroportin 1
FRB	FKBP-rapamycin binding domain of mTOR
FRT	flippase recombinase target
FtMt	mitochondrial ferritin
G1	one generation on BPS
G2	two generations on BPS
G3	three generations on BPS
GBE1	glycogen branching enzyme 1
gG-Cas9	general gateway Cas9

GlyS	glycogen synthase
gRNA	guide RNA
GS	geneswitch
GSD IV	glycogen-storage disease type IV
GSD	geneswitch-dependent disruption
HCP1	Heme carrier protein-1
HDV	Hepatitis delta virus ribozyme
HEPN	higher eukaryotic and prokaryotic nucleotide
HH	HH ribozyme
HHCS	hereditary hyperferritinemia-cataract syndrome
hIRP1	human iron regulatory protein 1
hIRP2	human iron regulatory protein 2
HO	heme oxygenase
HP	hephaestin
Hp	haptoglobin
hPR-LBD	human Progesterone-Receptor ligand-binding domain
Hx	hemopexin
HyNaC	Hydra Na$^+$ channels
ID	iron deficiency
IDT	Integrated DNA Technologies
IF	immunofluorescence
IP	immunoprecipitation
IR	RNAi, RNA inteference

IRE	iron responsive element
IRP	iron regulatory protein
ISC	iron sulfur cluster
ISCU	iron sulfur cluster assembly enzyme
KD	knock-down
KKJ	Kirst King-Jones
KO	knock-out
L1	first instar larvae
L2	second instar larvae
L3	third instar larvae
LshCas13a	*Leptotrichia shahii* Cas13a
LwaCas13a	*Leptotrichia wadei* Cas13a
MCO	multi-copper oxidase
MCP	MS2 coat protein
MFRN1	mitoferrin 1
MFRN2	mitoferrin 2
MRCKα	myotonic-dystrophy-related CDC42-binding kinase α
MNs	meganucleases
mRNA	messenger RNA
MS	mass spectrometry
mt:CoI	*mitochondrial cytochrome c oxidase subunit I*
mt:CoII	*mitochondrial cytochrome c oxidase subunit II*
mtDNA	mitochondrial DNA

mTOR	mammalian target of rapamycin
mvl	*malvolio* gene
Nedd4	neuronal precursor cell-expressed developmentally downregulated 4
NES	nuclear export signal
NIG	National Institute of Genetics (Japan)
NLS	nuclear localization signals
NRAMP1	natural resistance-associated macrophage protein 1
Nsf1	cysteine desulfurase
nt	nucleotide
NT	non-targeting
NUC	nuclease domains
nvd	*neverland* gene
PAM	protospacer adjacent motif
PBGD	porphobilinogen deaminase
PBGS	porphobilinogen synthase
PBS	phosphate buffered saline
PBST	PBS with 0.1% Triton
PBS3T	PBS with 0.3% Triton
PFS	protospacer flanking sequence
PG	prothoracic gland
PHD	HIF Prolyl Hydroxylases
phm	*phantom* gene
Pol II	polymerase II

ppk	pickpocket
PPNO	propylamine NONOate
PPOX	protoporphyrinogen oxidase
PspCas13b	*Prevotella buccae* Cas13b
RBC	red blood cell
RBCK1	RanBP-type and C3HC4-type zinc finger-containing protein 1
RDO	rapamycin-dependent gene overexpression
RDX	rapamycin-dependent gene disruption
REC	recognition lobe
rEMSA	RNA electrophoretic mobility shift assay
REPAIR	RNA editing for programmable A to I replacement
RESCUE	RNA editing for specific C to U exchange
RG	ring gland
RIP	RNA-immunoprecipitation
RNAi	RNA interference
rpk	*ripped pocket* gene
RT	room temperature
S2	Schneider's cell
SdhB	*Drosophila succinate dehydrogenase subunit B*
SG	salivary gland
SNP	single nucleotide polymorphism
spz5	*spatzle5* gene
SECIS	selenocysteine insertion sequence

SGK1	serum and glucocorticoid-regulated kinase 1
SHERLOCK	<u>s</u>pecific <u>h</u>igh-sensitivity <u>e</u>nzymatic <u>r</u>eporter un<u>lock</u>ing
SLC11A2	*divalent metal transporter 1* gene
SLC40A1	*ferreportin* gene
STEAP3	six-transmembrane epithelial antigen of prostate 3
TALENs	transcription activator-like effector nucleases
TCA	citric acid cycle
TIM2	T cell immunoglobulin and mucin domain protein 2
tim23	*translocase of the inner mitochondrial membrane 23*
TfR	transferrin receptor
Tsf	transferrin
TSS	transcription start site
TR	transcription regulation
UROS	uroporphyrinogen synthase
UROD	UROIII decarboxylase
UTR	untranslated region
VDRC	Vienna *Drosophila* Resource centre
WB	whole-body
YAco1	yeast aconitase 1
ZFNs	zinc finger nucleases
Zip99C	zinc/iron regulated transporter-related protein 99C

Chapter 1   Introduction

1.1   The importance of iron in living organisms

Iron plays a critical role in many biological processes, making life without it nearly impossible. This trace element, combined with other trace metals like chromium, copper or zinc, accounts for less than 0.006% of total body mass with approximately 2.3 to 3.0 grams in a human body [1]. Iron is used in two major types of proteins, iron-sulfur cluster (ISC)-bound and heme-bound proteins. It participates in many biological processes like oxygen transport, energy production, translation regulation, steroidogenesis and heme synthesis [2].

Despite its crucial roles, iron is a double-edged sword. At the cellular level, free iron in solution is quite toxic due to its capacity for electron transfer. Excess iron undergoes the Fenton reaction and generates cell-damaging radicals, including hydroxial radicals (HO·), which at high accumulation will attack cellular compartments and interfere with their physiological functions [3,4]. Iron misregulation results in severe consequences at the systemic level and was shown to be associated with diseases in human like hemochromatosis, anemia and Parkinson's, making iron homeostasis a vital process in living animals [1].

Iron deficiency (ID) is one of the world's most common nutritional deficiencies that affects more than 2 billion people [5]. ID refers to the reduction of iron stores that is considered as the most common cause of anemia. ID anemia is a severe condition in which low levels of iron are associated with anemia and the presence of microcytic hypochromic red cells [6]. ID also results in erythropoiesis deficiency, where there is insufficient mobilized iron for the production of new red blood cells (RBCs).

Having too much iron also results in severe consequences. Iron overload can be classified into two categories, primary and secondary. Most primary iron overload has a genetic basis. Hemochromatosis is the group of iron-overload syndromes and often linked to mutations in genes

that are involved in iron regulation [7]. In contrast, secondary iron overload usually comes from transfusions that are used to treat certain inherited disorders such as the iron-loading anemias [7]. Even though the syndrome is quite rare, it is relatively common in populations of Northern European origin where it affects one person in every 80 to 200 people [7]. Clinical consequences of iron overload include hepatic fibrosis and cirrhosis, increased risk of cellular carcinoma, cardiomyopathy, arthritis, and diabetes [7].

1.2 Overview of iron uptake and systemic iron regulation in mammals

Iron enters the human body from the diet either as heme or nonheme iron. Nonheme iron is abundant in foods of both animal and plant origins and the dominant form of iron in plants (Figure 1.1). It is found in a wide variety of forms including soluble iron, iron in low-molecular-weight complexes, storage iron in ferritin, and iron in the catalytic centers of a wide range of other proteins [8–10]. Much of this iron is not tightly sequestered, and its bioavailability can be affected by a range of dietary constituents and luminal factors. The low pH of the stomach and proximal small intestine helps to keep iron in a soluble form, making it available for absorption. Small organic acids, such as citric acid and ascorbic acid, also help to increase the absorption of iron [8–10].

In contrast, heme iron is tightly sequestered within a protoporphyrin ring and is not accessible to the factors that influence nonheme iron. As a consequence, heme iron tends to be absorbed more efficiently, and its absorption is less dependent on the composition of the diet. Most heme iron in the human diet is from myoglobin and hemoglobin and is animal-derived.

Dietary iron is absorbed by mature enterocytes of the mid-upper villus and mainly in the small intestine [11]. Nonheme iron transports across the brush-border membrane via DMT1. This transporter requires ferrous iron (Fe^{2+}) as a substrate. However, most dietary iron is in the ferric

(Fe^{3+}) form. Thus, iron needs to be reduced before it can be absorbed. In mammals, the duodenal cytochrome B (DCYTB) is a brush-border reductase involved in this conversion [12].

On the other hand, very little is known about the absorption of heme-bound iron. It is hypothesized that heme-bound iron binds to the enterocyte brush border intact and is likely endocytosed via the Heme Carrier Protein-1 (HCP1) [11]. Once located within the enterocyte, heme is degraded by the heme oxygenase to release ferrous iron for later iron-dependent processes.

Excess cellular iron is often toxic, so cells maintain a storage system called cytosolic ferritin. It is a ubiquitous protein complex with the ability to store up to 4500 atoms of iron [13]. Each ferritin complex consists of 24 subunits from the heavy chain (H) or the light chain (L). The H:L ratio in the complex can vary by cell types, tissues, and organisms. The H-ferritin utilizes ferroxidase activity that is responsible for the oxidation of ferrous to ferric iron. On the other hand, L-ferritin contains the acidic residues on the surface cavity that facilitate ferroxidase turnover and are crucial for the nucleation of ferric iron within the core of the protein complex [14] (Figures 1.1 and 1.2). In addition to the commonly known cytosolic ferritin, previous studies report the detection of ferritin in mitochondria, plant plastids, nucleus, extracellular space in serum, and the cerebrospinal fluid [15–21]. These different subcellular localizations suggest unique functions of ferritin.

1.3 Systemic iron transport and its delivery to tissues in mammals

Newly absorbed iron is bound to plasma transferrin (Tsf) for distribution to target tissues. Each transferrin protein can bind up to two ferric iron atoms, with approximately 30% of the iron-binding sites on the plasma transferrin pool are occupied under normal iron condition [22]. Thus, humans body maintains a considerable buffering capacity for urgent situations like the dynamic requirements of iron mobilization or excess non-transferrin-bound iron. Indeed, in iron-loading

4

diseases, transferrin often becomes saturated, thus limiting the ability to respond to overload non-transferrin-bound iron in serum (Figure 1.2).

Diferric transferrin delivers iron to cells by interaction with the transferrin receptor (TfR) on the plasma membrane [22]. The Tsf-TfR complex is internalized via clathrin-mediated endocytosis. Then, a combination of endosome acidification, conformational change and reduction of Tsf-bound ferric will release iron from the complex. The newly released iron moves into the cytoplasm across the endosomal membrane via DMT1 [23]. The later fate of iron depends on iron requirements of the cells. If iron demand is high, newly imported iron transport directly to sites of utilization. If iron is not immediately required, it will be stored in ferritin complexes until needed. In some conditions, cells also export iron through the Ferroportin 1 (FPN1) and ceruloplasmin (CP) in most cell types, or the iron oxidase hephaestin (HP) in enterocytes [24] (Figure 1.2).

Although diferric transferrin is an iron source used by many cell types or tissues, it is likely not the only source of cellular iron, especially when iron is required in high amounts. There is some evidence that ferritin can deliver its iron to target tissues. Ferritin secretion has been demonstrated in many cell types, including macrophages, hepatocytes, Kuffer cells of the liver [25–30]. The lack of glycosylation of extracellular ferritin suggests a non-classical secretion route. Recent discoveries have demonstrated the secretion of ferritin can occur through the lysosomal secretory pathway or multivesicular body-exosome pathway [26,31] (Figure 1.2). It will take time to understand mechanisms underlying the secretion and import of ferritin. Eventually, these findings will provide insight to iron transport across the body to target tissues. Unlike cytosolic ferritin with the storage capacity of 4500 iron atoms, serum ferritin can only contain roughly 700 iron atoms [26]. This capacity is still relatively higher than the iron amount can be stored in transferrin and can participate in iron delivery as a significant iron source.

1.4   Regulation of cellular iron homeostasis in mammals

Cellular iron homeostasis is tightly regulated to maximize the iron supply for iron-dependent processes while still restricting the bioavailable iron under toxic levels. At the center of cellular iron regulation is a protein called iron regulatory protein 1 (IRP1). This protein has two different conformations and plays two distinct functions depending on cellular iron levels. Under normal iron condition, IRP1 binds to an [4Fe-4S] iron-sulfur cluster (ISC) and acts as a cytosolic aconitase (holo-form) which interconverts between citrate and isocitrate. When iron levels are low, IRP1 releases its ISC and switches to another conformation (apo-form). Apo-IRP1 can interact with the iron-responsive element (IRE) found on the messenger RNAs (mRNAs) encoding various iron-related proteins (Figure 1.3) [32]. Depending on the location of IRE on mRNAs, the interaction will affect their translation, and as a consequence, will affect cellular bioavailable iron. The binding of apo-IRP1 to the IRE on 5' UTR will interfere with the recruitment of mRNA to the ribosome for translation, thus reduce protein production. On the other hand, when binding an IRE on a 3' UTR, apo-IRP1 will stabilize the transcript and as a consequence, increase protein levels. IREs have been found on at least nine transcripts involved in mammalian iron metabolism, including the ferritin heavy chain, TfR1, DMT1, FPN1 or transcripts involved in iron-related biological pathways like ALAS2 in heme biosynthesis, HIF2α in hypoxia [32]. Based on this classic paradigm, apo-IRP1 plays a direct function in iron homeostasis while holo-IRP1 seems to have no role in this regulation. Besides the activity of IRP1, cells may use other modes of regulation, most of which are less well studied. Biological pathways such as hypoxia, cytokines, and hormone regulation have been shown to regulate various iron-related genes [33–36].

1.5   Comparing iron regulation between mammals and insects

Most of what we know to this day regarding iron biology comes from studies in mammalian and yeast systems. Since iron regulation can deviate between species depending on their cellular and systemic organization, it is necessary to note the similarities and differences of iron metabolism across species. This section focuses on mechanisms underlying iron regulation in insects, the largest and most diverse group of organisms on earth. Most of what I will describe in this section is based on studies in *Drosophila melanogaster*. However, we will also mention other insect species along the way.

Similar to mammals, insects absorb iron mainly through the intestinal epithelial cells. In *Drosophila*, *malvolio* (*mvl*) is the ortholog of human DMT1 and also involved in the transportation of iron through the brush border [37]. Mutations in *mvl* result in taste perception defects and early lethality, which can be rescued by dietary manganese, iron, or by human DMT1 [38–40]. In theory, this iron intake is accompanied by a ferric reductase that converts dietary ferric (Fe^{3+}) to ferrous (Fe^{2+}) like Dcytb in mammals. However, despite there are two potential *Drosophila* homologs of Dcytb based on genomic studies, there is currently no functionally verified candidate in insects [41–43].

Once absorbed, iron can either be stored in ferritin or exported to target tissues, primarily via Tsf. In mammals, iron efflux across the enterocyte basolateral membrane via the transport protein complex composed of ferroportin (FPN1) and hephaestin (Hp), a multi-copper oxidase (MCO) [44]. There is no reported FPN1 orthologue in insects [45]. Interestingly, bioinformatic analyses across species only identified the presence of FPN1 in six organisms, including human, mouse, rat, zebrafish, worm, and rice plant [46,47]. There are four *MCO* (1-4) genes have been identified in *Drosophila*. Among the four *MCO* genes, MCO1 is an essential ferroxidase for iron

7

homeostasis, while MCO2 ferroxidase activity remains questionable [48,49]. Like MCO1, MCO3 oxidase activity has been confirmed, and its localization in the cellular secretory pathway suggests this enzyme plays a crucial role in iron and copper homeostasis (Table 1.1)[50,51]. On the other hand, little is known about MCO4. The *Drosophila MCO4* is the ortholog of a high-affinity iron importer in yeast (*Fet3*). Yeast Fet3 acts as a multicopper ferroxidase and works together with the iron permease FTR1[52]. There is no identified fly ortholog of FTR1 and the function of *Drosophila* MCO4 remains unclear.

Transferrin have been found in more than 34 invertebrate species, suggesting the system is highly conserved across species [53,54]. In *Drosophila*, there are three Tsf (1-3) homologs with Tsf1 being the most studied transferrin and the only putative hemolymph transferrin [55]. However, there is currently no identified TfR in this species, raising the question of whether there is an alternative mechanism for transferrin endocytosis in *Drosophila*. In addition to transferrin-based iron transport, recent studies in insects also suggest ferritin as an iron delivery source in *Drosophila* as well as *Aedes aegypti*, similar to with the secretory pathway in mammalian systems (Table 1.1) [49,56–58]. Cytoplasmic iron is transported into endoplasmic reticulum via zinc/iron regulated transporter-related protein 99C (Zip99C, aka dZip13) where it will be inserted to ferritin for either storage or secretion purposes.

The IRP proteins predominantly controls cellular iron homeostasis. In humans, there are two IRP proteins, IRP1 and IRP2. IRP1 is the phylogenetically older protein that can interconvert between the holo-form that binds ISC and acts as cytosolic aconitase, and the apo-form that binds IRE found on iron-related transcripts. IRP2, on the other hand, only stays as apo-form and is only activated under low cellular iron conditions. IRP has been found in many species, from mammals like human, mouse, rat, to invertebrates like *Drosophila melanogaster*, *Caenorhabditis elegans*.

Interestingly, IRP2 is only found in vertebrates, while the function of invertebrates IRP1 orthologs seems to shift toward cytosolic aconitase. Roundworm IRP1/Aco1 only exhibits aconitase activity while *Drosophila* harbours two IRP1 orthologs, namely IRP1A and IRP1B [59–62]. IRP1B only acts as a cytosolic aconitase, while IRP1A appears to be more similar to mammalian IRP1 with the ability to interconvert between the aconitase holo-form and RNA-binding apo-form. IRP1 is also found in plant, however, current evidence suggest that the protein cannot bind RNA and regulates iron homeostasis [63,64]. Lastly, no IRP was detected in yeast.

Overall, most of the key players in iron metabolism are conserved between mammals and insects. However, the lack of certain orthologs raises the question as to whether there is an alternative mechanism for iron regulation in insects. Unlike mammals, where most of the iron is used for hemoglobin synthesis, oxygen delivery in insect does not rely heavily on hemoglobin and thus, insect can act as a model to study iron homeostasis in other tissues (Table 1.1).

1.6 Using the *Drosophila* prothoracic gland to study cellular iron homeostasis

Current models for studying cellular iron homeostasis are limited in the sense that they have either static, or at best, linearly increasing iron requirements (cell cultures and developing erythrocytes, respectively) [2,65]. We introduce the *Drosophila* prothoracic gland (PG) as a model to study highly dynamic iron requirements. Together with nearby glands, including the corpora allata (CA) and corpora cardiaca (CC), the PG is part of a bigger endocrine tissue, called the ring gland (RG) (Figure 1.4A). What makes the PG unique is it accumulates a significant amount of iron. The accumulation of iron in the PG comes from its function in producing ecdysone, the principal steroid hormone in insects. In both vertebrates and insects, the synthesis of steroid hormones is mostly dependent on enzymes that require heme and Fe-S clusters [66,67]. Ecdysone is produced via a series of seven enzymatic steps with *neverland* (*nvd*) harbors an ISC while the

9

remaining, with the exception of *shroud*, are cytochrome P450s that require heme as a cofactor (Table 1.2) [67]. During *Drosophila* larval development, ecdysone is produced and released from the PG as pulses that trigger all the major and minor developmental transitions (Figure 1.4B). These pulses require relatively high amounts of enzymes involved in ecdysone synthesis. Based on transcriptome analysis, these genes are highly expressed prior to ecdysone pulses [67]. As a result, PG cells also require a high amount of iron and heme. However, having too much mobilized iron or free heme can be toxic. Therefore, I hypothesize the bioavailable iron and heme metabolism in the PG are coordinately controlled with ecdysone pulses (Figure 1.4B). This feature makes PG an ideal model to study tissue with highly dynamic iron requirements.

1.7 Cellular heme biosynthesis and regulation

Heme is a ubiquitous molecular complex of iron and the tetrapyrrole protoporphyrin IX. When bound to hemoproteins, heme plays an essential role in reversible oxygen binding, electron transport of the respiratory chain, detoxification, and steroid hormone production [68,69]. The biosynthesis pathway of heme is highly conserved from humans to *Drosophila*. Starting with glycine and succinyl-coA, heme is produced via a series of eight enzymatic steps with certain steps occuring in the cytoplasm, while the first and the last three steps happen in the mitochondria (Figure 1.5A). The first step, when glycine and succinyl-CoA are converted to aminolevulinic acid (ALA), is considered the rate-limiting step and carried out by the enzyme ALAS. In mammals, there are two ALAS enzymes responsible for this step, namely ALAS1 and ALAS2. ALAS2 is expressed in a high amount relative to ALAS1 and responsible for heme production in erythrocytes. Its transcript also carries an IRE and can be regulated in an iron-dependent manner. On the other hand, ALAS1 governs all other heme production [70–72]. Once ALA is produced, it is exported from the mitochondria to reach the second step enzyme, porphobilinogen synthase

(PBGS), also known as ALA dehydratase (ALAD). This enzyme combines two molecules of ALA to form one molecule of the monopyrrole PBG [73]. Then, the enzyme PBG deaminase (PBGD) catalyzes the formation of a linear tetrapyrrole hydroxymethylbilane from four PBG molecules [74]. Starting from the fourth to the seventh step, heme precursor is capable of giving red autofluorescence under UV exposure. This feature comes from its structure, which now carries a protoporphyrin ring upon circularization activity of uroporphyrinogen synthase (UROS) [74]. The porphyrin ring can be spontaneously oxidized, thus making the molecule sensitive to UV light (Figure 1.5B). It is important to note that heme intermediates are generally short-lived so red autofluorescence is not noticeable in normal conditions. The fifth step, also the last step in the cytoplasm, is carried out by UROIII decarboxylase (UROD) and results in the generation of coproporphyrinogen III [74]. This heme precursor is then transported into the mitochondrial intermembrane space where the coproporphyrinogen oxidase (CPOX) catalyzes oxidative decarboxylation of the A and B ring propionates to yield the vinyl groups of protoporphyrinogen IX [75,76]. At the very last step, protoporphyrinogen IX is oxidized by the protoporphyrinogen oxidase (PPOX) to produce protoporphyrin IX before the insertion of ferrous iron via activity of the enzyme ferrochelatase (FeCH) [74]. The heme molecule is no longer capable of giving red autofluorescence and ready for export to the cytoplasm, where it is incorporated as a cofactor for hemoproteins (Table 1.3).

As a cofactor of hemoproteins, heme is an essential component involved in multiple biological processes. However, like iron, heme can be harmful in the pathogenesis of various diseases. Excess free-heme can promote oxidative stress and lipid peroxidation, thus leading to membrane injury and cell apoptosis [77]. Besides, heme is strongly pro-inflammatory since it

induces the recruitment of leukocytes, platelets, and RBCs to the vascular endothelium, oxidizes low-density lipoproteins, and consumes nitric oxides, thus impairing vascular function [78–81].

Disruption of heme biosynthesis results in severe consequences due to the accumulation of heme precursors. Impairment of the process has been linked to a group of human diseases called porphyria. This is a rare genetic disorder that affects one person in every 5,000 people even though the frequency varies depending on which enzyme is affected. Disruption in later steps of heme synthesis can be recognized by the accumulation of red autofluorescent heme precursors to an amount that can be visualized under UV. The congenital porphyria, or Gunther's syndrome, is the most severe form, where patients suffer from skin lesions. Heme precursors from busting cells can accumulate in other tissues, such as teeth [82–84]. Heme precursors accumulation is neurotoxic, because hemoproteins become impaired in these patients.

Mammals have various systems to monitor and regulate heme toxicity tightly. Free hemoglobin and heme in the extracellular space are controlled by haptoglobin (Hp) and Hemopexin (Hx), respectively. Once bound to Hp, hemoglobin is transported to the macrophages of the reticulo-endothelial system and further bound by the scavenger receptor CD163. Excess hemoglobin is quickly oxidized to methemoglobin and releases free heme [85]. Free heme is then bound by Hx at the molar ratio 1:1 and internalized by receptor-mediated endocytosis. This interaction limits the amount of free heme and contributes to the recycling of iron [86]. Interestingly, there are currently no identified *Drosophila* orthologs of Hp and Hx. Once residing in the cells, heme is either used by hemoprotein or degraded by the activity of heme oxygenase (HO). During heme catalyzation, the molecule is broken down into the antioxidant biliverdin, the vasodilator carbon monoxide (CO) and ferrous iron (Fe^{2+}). Biliverdin is further reduced to bilirubin by the enzyme biliverdin reductase.

1.8   Using CRISPR to study iron and heme homeostasis in *Drosophila*

Throughout many years of iron biology, scientists have exploited a combination of genetic and protein studies to answer critical questions. The development of technologies has enabled a more thorough understanding of iron roles in biological systems. The adaptation of chromatography and spectrometry has facilitated the detection of cellular iron [87], while protein crystallography allows us to visualize the metal inside protein [88]. In addition, protein engineering enables the purification of target proteins and allows us to investigate the role of iron atoms. On the other hand, advances in genomic studies have revealed new candidates with a potentially important role in iron homeostasis. One now has more tools to generate desired mutations or interfering with transcription of target genes and used them to study iron biology. The rapid increase of our knowledge, in turn also facilitates the development of these technologies.

The recent discovery of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and the generation of guide RNA- (gRNA-) dependent Cas9 endonucleases has been quickly adapted by *Drosophila* researchers [89–91] and we now possess a universal and powerful toolkit that can be used for both loss- and gain-of-function studies by using distinct versions of Cas9 [92,93]. As such, CRISPR-based techniques are ideal to replace, validate, and complement traditional approaches relying on conditionally expressing RNAi or cDNAs. Recent advances in CRISPR-based approaches include codon-optimizations of Cas9, utilizing Cas9 variants as a RNA-guided transcription factors that block or increase target gene transcription, and large-scale transgenic *Drosophila* gRNA collections launched at Harvard Medical School (https://fgr.hms.harvard.edu/vivo-crispr-0), the German Cancer Research Center in Heidelberg (https://www.crisprflydesign.org/library/) and the National Institute of Genetics in Mishima, Japan (https://shigen.nig.ac.jp/fly/nigfly/).

The ability to manipulate target genes at tissue of interest and at the desired time point will greatly facilitate the study of tissue-specific iron metabolism. One can generate somatic CRISPR or create a desired mutation in a reasonable time course. To identify genes with undiscovered roles in heme or iron homeostasis, the King-Jones lab performed a genome-wide RNAi screening in collaboration with two other labs, namely Michael O'Connor's lab at the University of Minnesota, US and Kim Furbo Rewitz lab at the University of Copenhagen, Denmark. In this approach, RNAi against each gene was crossed with *phm22*, a prothoracic gland specific Gal4, and observed for any developmental defects. Then the secondary screening was conducted based on the first screening, where genes with developmental defects were screened for abnormal RG morphology. This secondary screening yielded a list of 34 hits for further investigation. The overall list seemed to be quite diverse with respect to cellular functions, including transcription factors, enzymes, signaling factors, and receptors. Further evidence was required to make sure the observed phenotype can be recapitulated. As a result, I developed new genetics tools like CRISPR can be exploited, either at the DNA or RNA level, to validate this phenotype.

1.9   Outline of the thesis

Overall, this thesis focuses on understanding the regulatory mechanism by which iron and heme metabolism is coupled with steroid hormone synthesis regulation. My work consists of four major parts. Initially, I started my project by investigating the role of *AGBE* in iron regulation. PG-knock down of this gene resulted in a porphyria-like phenotype that I can rescued in iron-supplemented medium. Further investigation allowed me to establish the relationship between *AGBE*, *Cisd2*, and *IRP1A* in *Drosophila* iron metabolism. PG-knock down of either gene in iron depletion conditions also results in porphyria-like phenotypes. The biological functions of IRP1A have been mentioned earlier, while Cisd2 is the homolog of mammalian mitoNEET, a protein

responsible for repairing the oxidatively damaged ISC in IRP1. On the other hand, AGBE is the ortholog of the human branching enzyme GBE1, which prior to my study, was only linked to glycogen synthesis with no apparent role in iron homeostasis. From my project, I could show the physical and genetic interactions between AGBE, Cisd2, and IRP1A. In *Drosophila*, Cisd2 is also required to maintain the healthy ISC in IRP1A, and AGBE serves as a bridge to strengthen the interaction between Cisd2 and oxidatively damage IRP1A for the repair process.

I also report that IRP1A can localize to cell nuclei. This was observed in multiple tissues, including the PG, fat body (FB) while no nuclear signal was detected in the brain or salivary gland (SG). The unexpected result suggests an undocumented nuclear function of IRP1A. I later validated the interaction between IRP1A with histones and evaluated its nuclear aconitase activity. These data, together with transcriptome analysis, support a hypothesis that nuclear IRP1A might participate in regulation of iron-related genes by regulating citrate levels, a substrate of nuclear acetyl-CoA synthesis for histone acetylation. This finding has added an entirely unexpected aspect not explained before of holo-IRP1A in iron metabolism, since vertebrate holo-IRP was only believed to function as a dispensable aconitase.

The third part of my project focused on characterizing the undiscovered function *ppk20* in iron or heme metabolism. *ppk20* is a member of *Drosophila* epithelial sodium channel (ENaC). Like *AGBE*, it was identified from earlier genome-wide screening. PG-knock down of *ppk20* results in the porphyria-like. I also found that *ppk20*-impaired animals can be partially rescued by dietary hemin as well as by injected ferritin. On the other hand, *ppk20*-impaired animals also exhibit the trachea necrosis, raising the possibility that the animals suffered from low oxygen levels. Since hypoxia has been linked to iron regulation in mammalian systems, it suggests *ppk20* may play a role in this link.

Lastly, during my project, I developed a strong interest in adopting CRISPR tools for *Drosophila* research. Not only because CRISPR has greatly facilitated my research in iron metabolism but also because of the remarkable potential of this system in molecular research. I have generated two CRISPR/Cas9 toolkits, which allow spatial and temporal gene manipulation. With these toolkits, one can generate somatic mutation, interfere with transcription or induce gene expression in the tissue of interest and at the desired time points. Along with these toolkits, I also evaluated the efficiency and potential applications of another CRISPR system, Cas13, in *Drosophila*. Unlike Cas9 targets DNA, Cas13 targets RNA with high efficiency, and current preliminary data suggested its great potential to be a replacement of RNAi in RNA targeting.
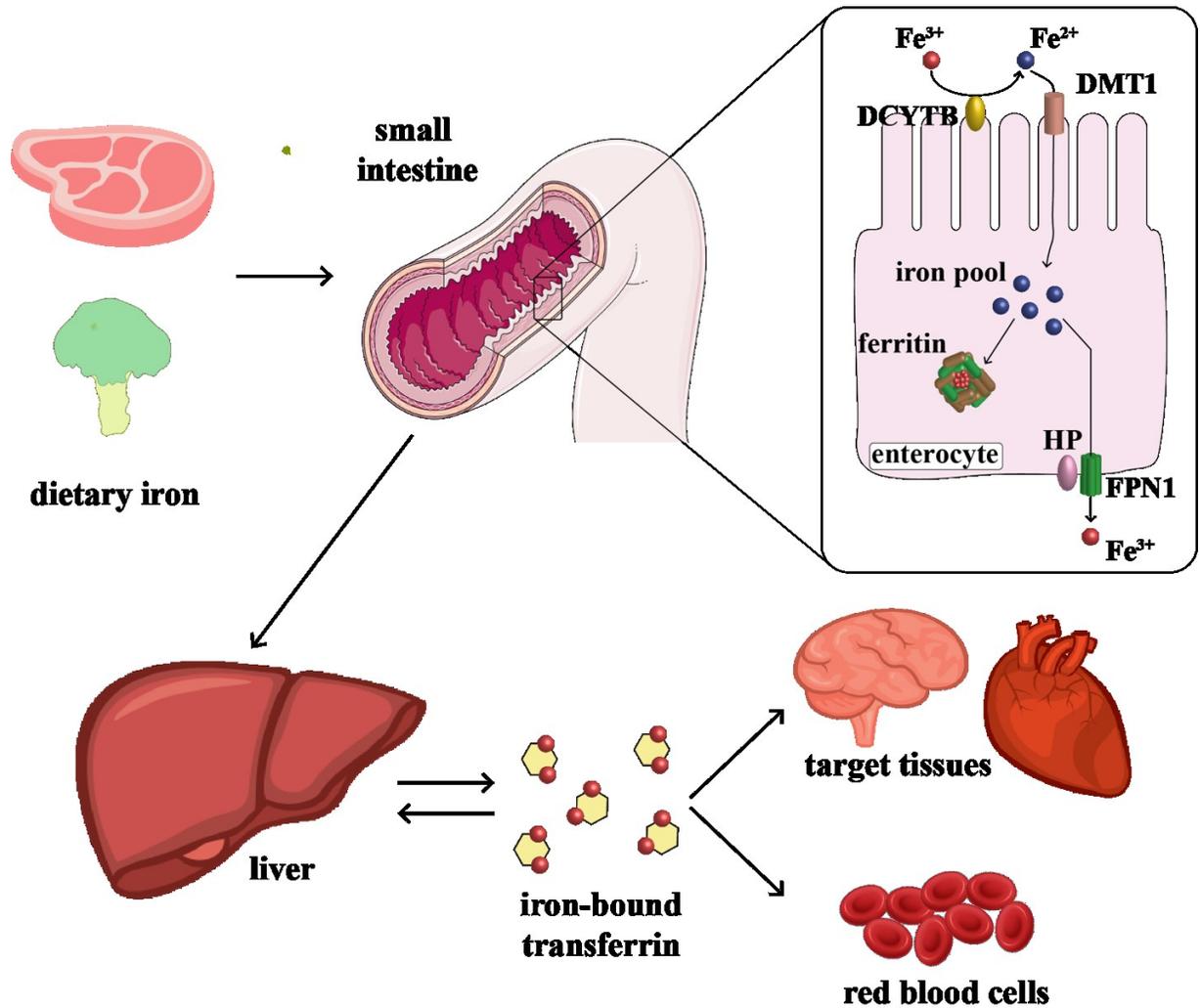
Figure 1.1 Overview of dietary iron uptake in human.

Iron enters the body from the diet either as heme or nonheme iron. Dietary iron is absorbed by the mature enterocytes of the mid-upper villus and mainly in the small intestine. Nonheme iron transports across the brush-border membrane via divalent metal transporter 1 (DMT1) with assistance from duodenal cytochrome b (DCYTB) to convert ferric iron (Fe^{3+}) to ferrous iron (Fe^{2+}). On the other hand, very little is known about the absorption of heme-bound iron. It is hypothesized that heme-bound iron binds to the enterocyte brush border intact and is likely

endocytosed via the Heme Carrier Protein-1 (HCP1). Absorbed iron is exported from enterocytes via ferreportin 1 (FPN1) and hephaestin (HP), and eventually resides in the liver for long term storage in ferritin complex or distributed to target tissues or red blood cells via the transferrin-bound form.
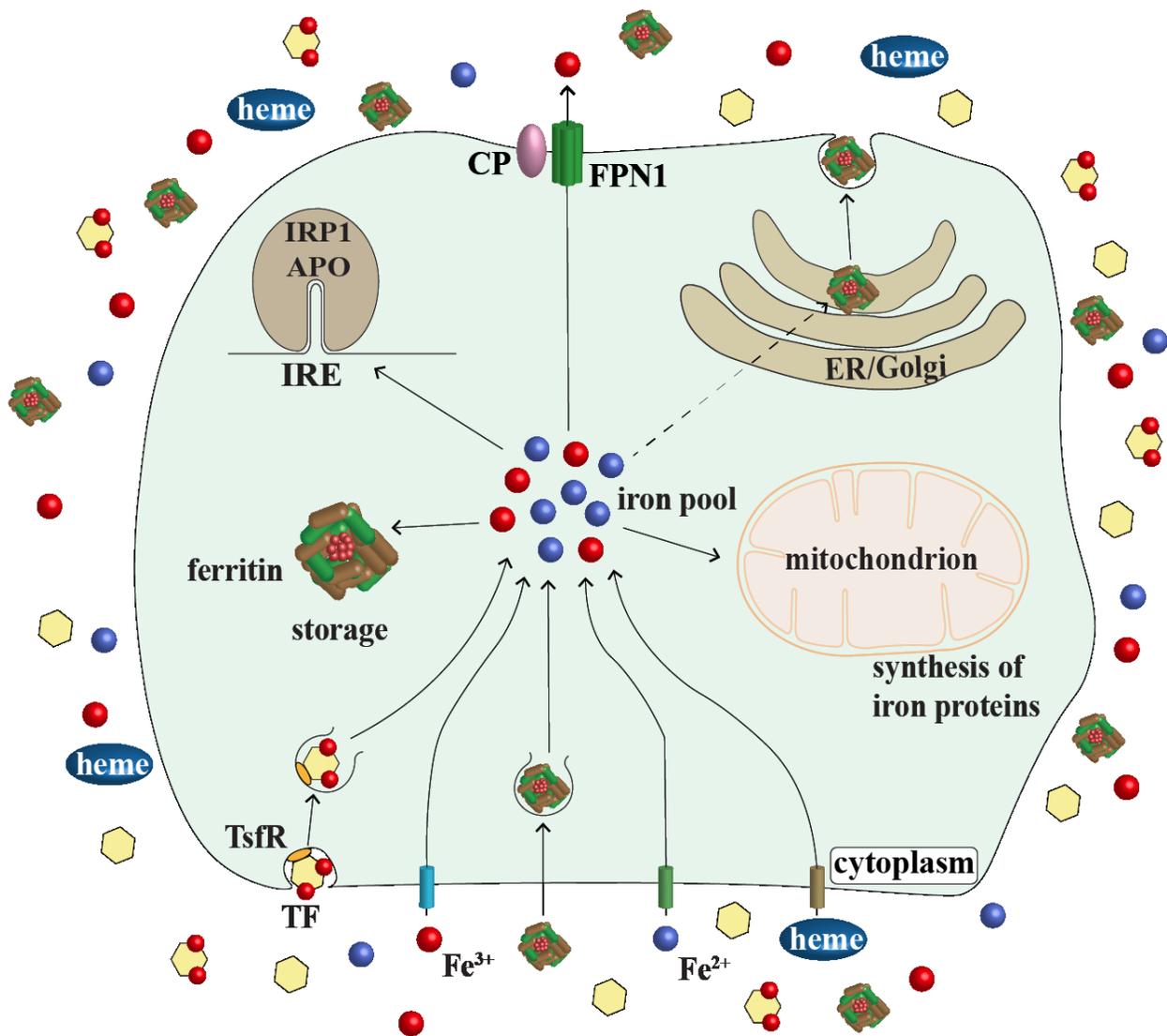
Figure 1.2 Overview of cellular iron homeostasis in mammalian cells

Iron is delivered to target tissues in a variety of forms. The most commonly used pathway requires the participation of iron-bound transferrin (TF) from serum or circulation pathways. Diferric TF binds to transferrin receptor (TfR) on the plasma membrane before being endocytosed. Later processing will release iron from TF. On the other hand, non-transferrin-bound iron can be imported via specific ion channels. Recent studies suggest that ferritin and heme can be secreted and acts as another iron source in cells with the urgent iron requirement. These mechanisms contribute to the cellular iron pool, which will be used for the synthesis of iron proteins. Excess

iron is stored in ferritin for exported out of the cells through ferreportin 1 (FPN1) and ceruloplasmin (CP). These processes are tightly monitored and regulated by the iron regulatory protein 1 (IRP1), mostly via binding with the iron-responsive element (IRE) found of iron-related transcripts and regulating their translation.
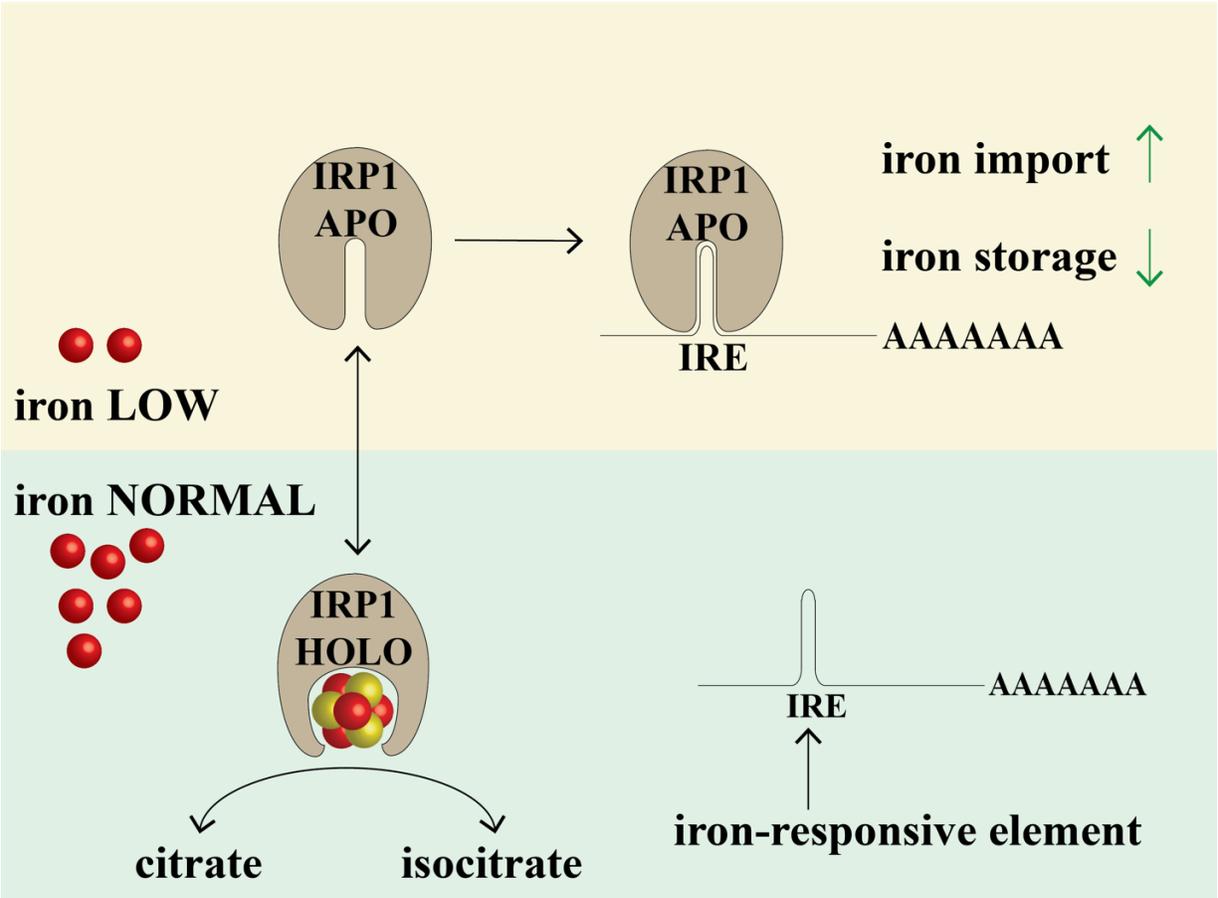
Figure 1.3 Classic paradigm of Iron Regulatory Protein 1 (IRP1) in cellular iron homeostasis.

IRP1 has two conformations that depend on cellular iron levels. Under iron-replete conditions, IRP1 binds to iron-sulfur cluster (ISC) and forms holo-IRP1, which acts as cytosolic aconitase and interconverts between citrate and isocitrate. Under low iron condition, IRP1 loses its ISC and switches to a different conformation, known as apo-IRP1. This conformation can bind an iron-responsive element (IRE) found in some mRNAs encoded for proteins involved in iron metabolism. The interaction will affect transcripts' stability as well as their translation.
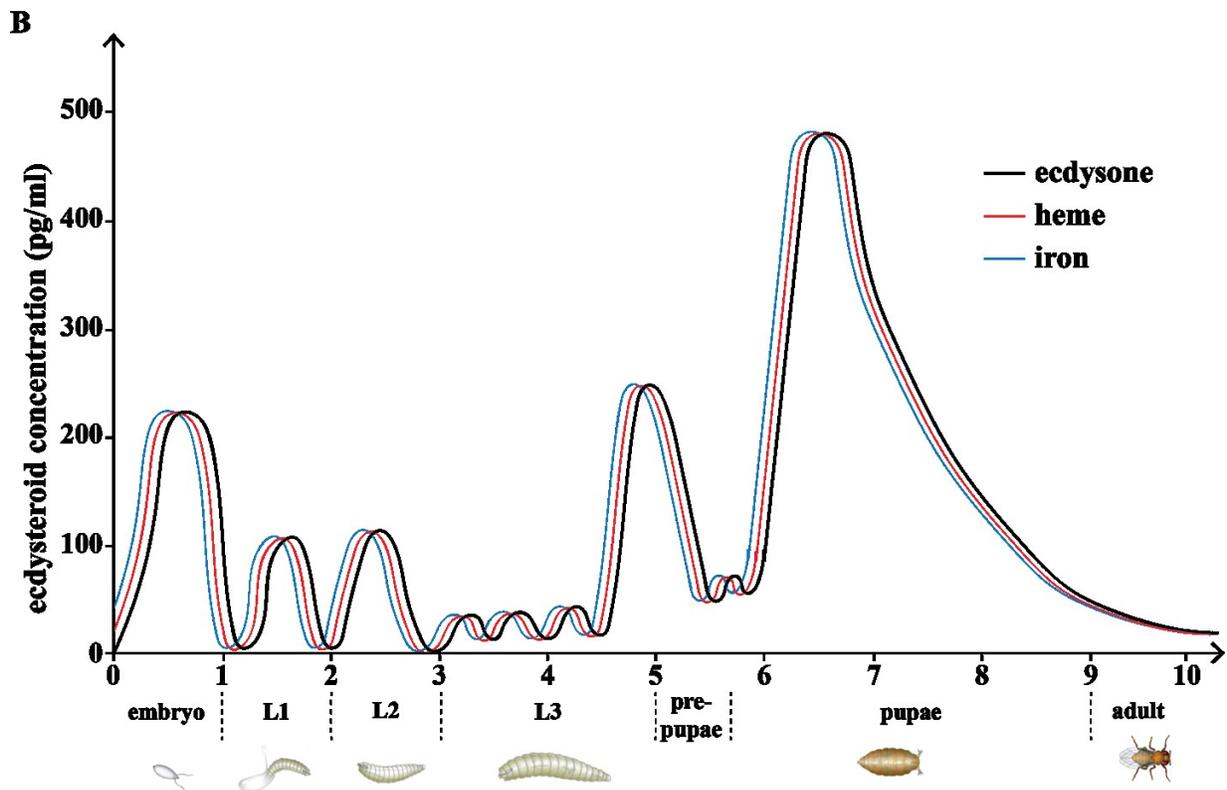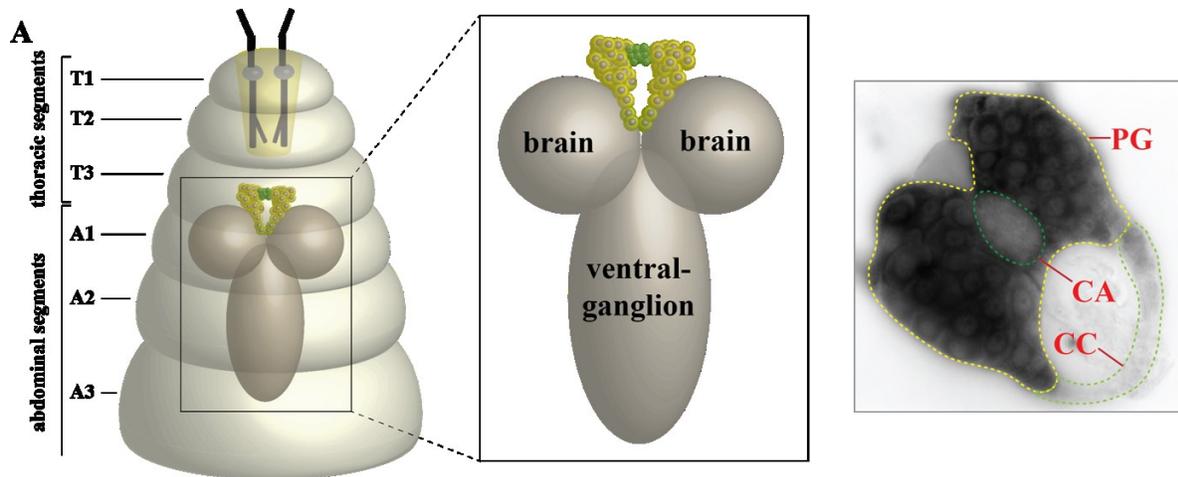
Figure 1.4 The larval prothoracic gland is an ideal model to study dynamic iron mobilization.

A. The prothoracic gland (PG) is part of a bigger gland, called ring gland (RG). PG accumulate a high amount of iron which can be visualized by staining for ferric iron, which represents stored or newly imported iron. **B**. During larval development, PG is responsible for producing ecdysone, the principle hormone in insects. This hormone is produced and released as pulses (black) that

trigger animals' developmental transitions, including embryogenesis, larval development, and metamorphosis. These pulses will require a high expression of enzymes involved in making ecdysone and as cofactors of those enzymes, iron, and heme are also in high demand. Since excess free iron and heme are toxic, mobilized iron (blue) and heme synthesis (red) are hypothesized to be coordinately controlled with ecdysone pulses. The ecdysone curve (black) is based on previous studies while mobilized iron (blue) and heme synthesis (red) lines are speculation. CA: corpora allata, CC: corpora cardiaca, L1: first instar larval stage, L2: second instar larval stage, L3: third instar larval stage.
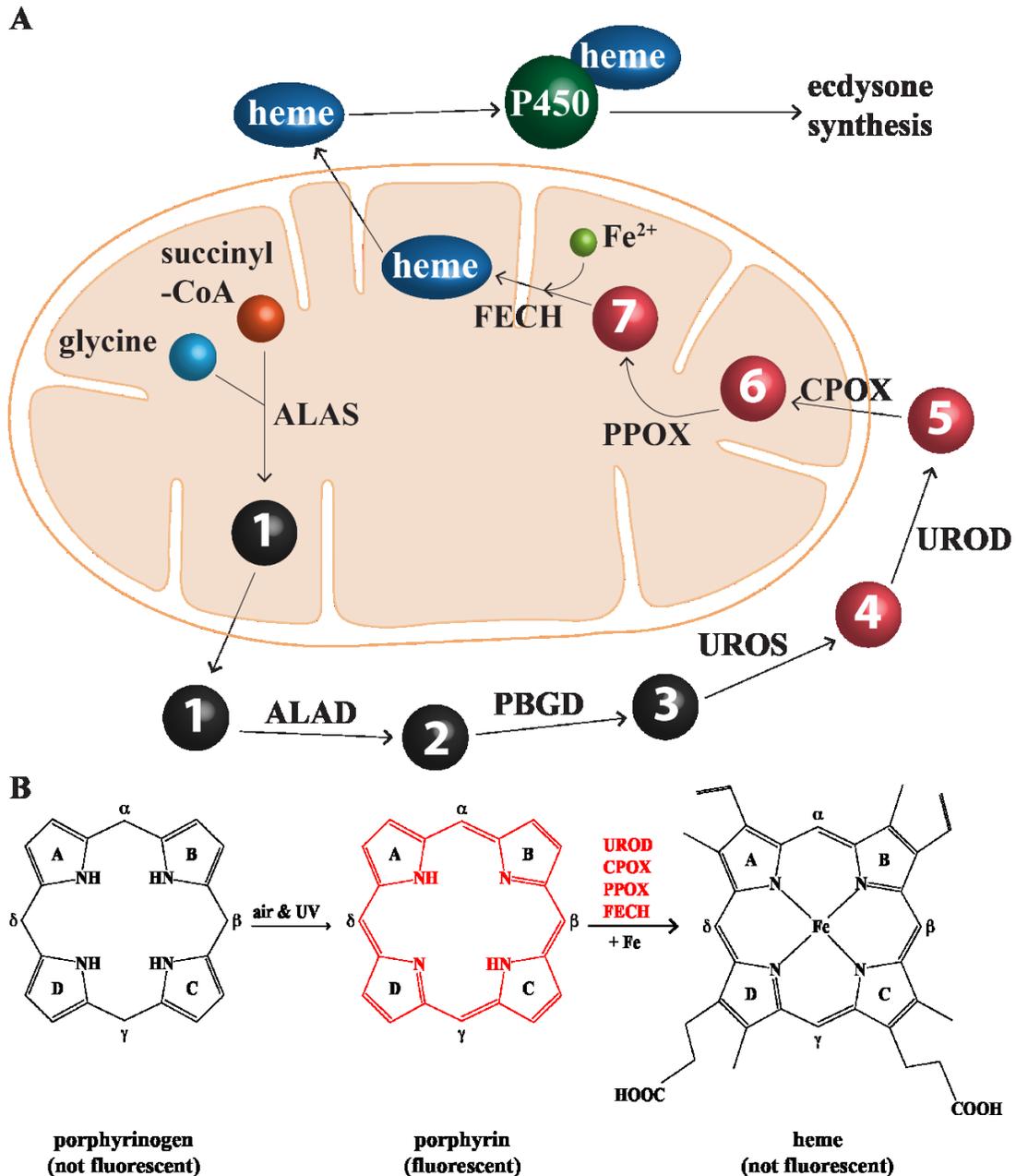
Figure 1.5 Overview of heme biosynthesis in mammals.

A. Heme is produced from glycine and succinyl-coA via a series of eight enzymatic steps. Steps 2-4 happen in the cytoplasm while step 1 and steps 5-8 are carried out in the mitochondria. Starting from step 4, heme precursors now carry the porphyrin structures that can be oxidized to give red autofluorescence (red circles). At the final step, iron is incorporated into protoporphyrin IX to

make heme, which now longer gives red autofluorescence. Heme synthesis is highly conserved from humans to fly. **B.** The oxidation of porphyrinogen under air and UV effects produces the red autofluorescence porphyrin ring.

1.11 Tables

Table 1.1 Comparison of proteins involved in iron metabolism between mammals and insects

protein	function in iron metabolism	mammal gene	*Drosophila* gene
divalent metal-ion tranporter 1 (DMT1)	transmembrane ferrous iron transporter	*SLC11A2*[a]	*malvolio (mvl)*[O]
natural resistance-associated macrophage protein 1 (NRAMP1)	transmembrane ferrous iron transporter	*SLC11A1*[a]	*malvolio (mvl)*[H]
duodenal cytochrome b (DCYTB)	membrane ferric reductase of enterocytes	*CYBRD1*[a]	*CG1275*[O]
six-transmembrane epithelial antigen of prostate 3 (STEAP3)	membrane ferric reductase of erythrocytes	*STEAP3*[a]	none
ferroportin 1 (FPN1)	transmembrane ferrous iron transporter (exporter)	*SLC40A1*[a]	none
ceruloplasmin (CP)	systemic multicopper ferroxidase	*CP*[a]	*MCO1*[O]
hephaestin (HP)	intestinal and central nervous system multicopper ferroxidase	*HEPH*[a]	*MCO1*[H] *MCO3*[H]
hepcidin	systemic regulator of iron metabolism	*HAMP*[a]	none
heme carrier protein 1 (HCP1)	putative membrane heme transporter	*SLC46A1*[a]	*CG30345*[H]
heme oxygenase 1 (HO1)	degradation of heme and release iron	*HMOX1*[a]	*heme oxygenase (HO)*[O]
H-ferritin	iron storage and delivery	*FTH1*[a]	*Fer1HCH*[O]
L-ferritin	iron storage	*FTL*[a]	*Fer2LCH*[O]
mitochondrial ferritin (FtMt)	mitochondrial iron storage and antioxidant	*FTMT*[a]	*Fer3HCH*[O]

protein	function in iron metabolism	mammal gene	*Drosophila* gene
T cell immunoglobulin and mucin domain protein 2 (TIM2)	H-ferritin receptor in mouse and rat	*TIM2*[b]	none
transferrin (TF)	plasma iron transport protein	*TF*	*Transferrin 1 (Tsf1)*[H] *Transferrin 2 (Tsf2)*[H] *Transferrin 3 (Tsf3)*[H]
transferrin receptor 1 (TfR1)	cellular uptake of transferrin bound iron, also binds H-ferritin	*TFRC*[a]	none
iron regulatory protein 1	aconitase, regulator of cellular iron homeostasis	*ACO1*[a]	*IRP1A*[O] *IRP1B*[O]
iron regulatory protein 2	regulator of cellular iron homeostasis	*IREB2*[a]	*IRP1A*[H]
mitoferrin 1 (MFRN1)	mitochondrial iron importer in erythropoietic tissues	*SLC25A37*[a]	*mitoferrin (mfrn)*[O]
mitoferrin 2 (MFRN2)	ubiquitous mitochondrial iron importer	*SCL25A28*[a]	*mitoferrin (mfrn)*[H]
HIF1α	hypoxia signaling pathway	*HIF1A*[a]	*similar (sima)*[H]
HIF1β	hypoxia signaling pathway	*ARNT*[a]	*tango (tgo)*[H]

[a] genes found in human

[b] gene found in mouse but not in human

[O] ortholog

[H] homolog

Table 1.2 Ecdysone synthesis pathway in *Drosophila melanogaster* larvae

gene	enzyme characteristics	product
neverland (*nvd*)[a]	oxidoreductase, harbors an iron-sulfur cluster, locates in endoplasmic reticulum (ER)	7-dehydrocholesterol
shroud (*sro*)	17β-estradiol 17-dehydrogenase, belongs to the ecdysteroidogenesis black box	unknown
cyp6t3 (*cyp6t3*)[b]	cytochrome P450 with oxidoreductase activity, belongs to the black box	unknown
spookier (*spok*)[b]	cytochrome P450 with oxidoreductase activity, belongs to the black box	5β-ketodiol
phantom (*phm*)[b]	cytochrome P450, ecdysteroid 25-hydroxylase, locates in ER	3β, 5β-ketodiol
disembodied (*dib*)[b]	mitochondrial cytochrome P450, ecdysteroid 22-hydroxylase	2-deoxyecdysone
shadow (*sad*)[b]	mitochondrial cytochrome P450, ecdysteroid 2-hydroxylase	ecdysone (E)
shade (*shd*)[b]	ecdysone 20-monooxygenase, cytochrome P450	20OH-ecdysone (20E)

[a] Enzyme harbours an iron-sulfur cluster (ISC)

[b] Enzyme harbours heme as cofactor

Table 1.3 *Drosophila* orthologs of the human heme biosynthesis enzymes

vertebrate enzyme	human genes	fly genes[a]	product
ALA synthase	*ALAS1*, *ALAS2*	*Alas*	1. aminolevulinic acid (ALA)
ALA dehydratase	*ALAD*	*Pbgs*	2. porphobilinogen (PBG)
PBG deaminase	*HMBS*	*l(3)02640*	3. hydroxymethylbilane
UROIII synthase	*UROS*	*Uros1*, *Uros2*	4. uroporphyrinogen III (UROIII)
UROIII decarboxylase	*UROD*	*urod* (*updo*)	5. coproporphyrinogen III
coproporphyrinogen III oxidase	*CPOX*	*Coprox*	6. protoporphyrinogen IX
protoporphyrinogen IX oxidase	*PPOX*	*Ppox*	7. protoporphyrin IX (PPIX)
ferrochelatase	*FECH*	*FeCH*	8. heme

[a] "fly" represents *Drosophila melanogaster*.

Chapter 2   Materials and methods[1]

---

[1] This chapter describes the general approach for each experiment used in my study. However, there are some deviations in each project for which I will briefly mention whenever needed.

2.1 *Drosophila* husbandry and survival study

List of fly strains that were used will be described in each corresponding chapter. For fly lines being generated during my program, see Appendix A.8. Overall, stocks were maintained on a standard cornmeal diet unless otherwise specified. Backup stocks were kept as two copies in vials at 18^0C while active stocks are maintained in bottles at room temperature or 25^0C.

Regular fly food refers to "NutriFly"-based media, which follows the standard recipe from the Bloomington *Drosophila* Stock Center (https://bdsc.indiana.edu/information/recipes/bloomfood.html). In brief, for every 100 mL of food, 18.0 grams of NutriFly powder was boiled for 3-5 minutes in autoclaved miliQ water for a total volume of 100 mL. When the mixture was cool down to approximately 40-50^{0}C, compounds can be added at the desired concentration (Table 2.1), followed by the addition of 450 μl of propionic acid. Food was mixed well and distributed to petri discs, vials or bottles depending on the experiment.

Active flies were reared at 25^0C and 60-70% humidity. Prior to any fly-based experiments, stocks were reared on NutriFly media for at least two generations. The whole experiment was conducted in a walk-in incubation chamber. For egg collections, flies were allowed to lay eggs for 3x one hour in order to reduce egg retention and minimize the presence of old embryos. For each vial, 50 embryos were then collected in 1-hour intervals. Embryos were counted and transferred to petri dics containing appropriate media. Larval survival was scored for every stage. At least three independent crosses (= three biological replicates), were carried out per experimental condition. For experiments where animals were embryonic lethal, the fresh embryos (0.5-1 hours old) were dechonrionated and immersed in 1x PBS containing desired supplement for 5 minutes (Table 2.1).

2.2 Generation of transgenic lines and CRISPR/Cas9 fly lines

For a list of main vectors being used during my program, see Appendix A.9. Vector backbones were amplified via PCR to generate two fragments per vector and fused to cDNA fragments via the Gibson reaction. For gRNA or crRNA constructs, vector backbone were predigested with BbsI restriction enzyme while fragment containing target site was generated via PCR. Mutations, if any, were generated via Q5 mutagenesis PCR (NEB #M0491S) following the standard protocol. Fused fragments were cloned into DH5α *E. coli* competent cells, and validated by Sanger sequencing.

For CRISPR/Cas9, we identified optimal target gRNA sites by relying on comparable results from two independent programs, "CRISPR Optimal Target Finder" (University of Wisconsin; http://tools.flycrispr.molbio.wisc.edu/targetFinder/index.php) and Harvard's "Find CRISPR" sgRNA design tool (http://www.flyrnai.org/crispr/index.html) [89]. Target sites were confirmed by sequencing corresponding loci in the *vas.Cas9* line (Bloomington #51323) that we used for embryo injections. CRISPR lines were generated via CRISPR/Cas9 homology-directed repair to replace endogenous alleles. Plasmids carrying gRNA target sites were cloned into pCFD3 (Addgene #49410) for *AGBE^{FCF}*, *AGBE^{FCM}*, *IRP1A^{3F}* and *IRP1B^{3F}* constructs, or pCFD5 [91,94] (Addgene #73914) for the remaining CRISPR constructs (Appendix A.8). All donor template fragments were amplified from genomic DNA via PCR and cloned into the pDsRed-attP vector (Addgene #51019) [91].

2.3 Generating the general gateway Cas9 (gG-Cas9) collection

The gG-Cas9 collection is based on the pBPGUw plasmid, which we modified to produce different Cas9 versions. This vector contains a Gateway Cassette, a synthetic core promoter and a Gal4-coding sequence [95,96]. The pBPGUw backbone was amplified to remove the Gal4

sequence and combined with the different Cas9 versions amplified from corresponding Addgene or DGRC plasmids mentioned above using Gibson reaction (Appendix A.9) [97]. Constructs were then transformed into competent DH5α cells and validated by Sanger sequencing.

2.4   Generating the prothoracic gland-specific Cas9 collection (PG-Cas9)

To generate different PG-Cas9 constructs, we used PhiC31 vectors from the above-described gG-Cas9 collection. Vector backbones were amplified via PCR and fused with a 1.45kb fragment containing the *spok* regulatory region amplified from pCRII-TOPO Spok plasmid (a kind gift from Michael O'Connor) or *Act5C* (*act*) promoter amplified from Ac5-Stable2-neo (Addgene 32426) via the Gibson reaction. For the corresponding UAS-Cas9 variants, the Cas9 fragments were amplified via PCR and inserted into pBID-UASc plasmid (Addgene 35200) via Gibson reaction (Appendix A.8) [98]. Constructs were then transformed into competent DH5α cells and validated by Sanger sequencing.

2.5   gRNA selection and cloning

Target gene sequences were obtained from FlyBase and analyzed for optimal target gRNA sites by selecting sequences that showed consensus between two programs, namely "CRISPR optimal target finder" (http://tools.flycrispr.molbio.wisc.edu/targetFinder/) and "Harvard CRISPR gRNA design tool" (http://www.flyrnai.org/crispr/). Optimal target sites were then confirmed by sequencing the loci from genomic DNA extracted from corresponding fly lines we used for plasmid injection. The pCFD5 and PG-gRNA plasmids were pre-digested with BbsI (NEB R3539S) and fused to appropriate gRNA-containing PCR fragments via the Gibson reaction, followed by Sanger sequencing. For PG2-gRNA, which only has a single BbsI cutting site, we

used PG1 as a template to extend the scaffold for multiple gRNA sequences in the PG2 vector. For more information regarding the whole procedure, see Appendix A.2.

2.6   Generation of *Drosophila*-optimized Cas13 orthologs

To generate the fruit fly codon-optimized Cas13, the original Cas13 nucleotide sequences were evaluated using two independent web tools: i) ATGme (https://atgme.org) and ii) OPTIMIZER (http://genomes.urv.es/OPTIMIZER) [99,100] with the customized codon usage frequency specific for *Drosophila* [101–103]. The two indices, namely Codon Adaption Index (CAI) and Effective Number of Codons (ENC), are used to measure the optimized sequences. CAI has a value ranges from 0 to 1 and is used to evaluate the similarity between codon usage of a gene and codon usage of the reference group [104], thus the higher CAI value is, the higher theoretical gene expression [105,106]. On the other hand, ENC is a measure of codon usage bias with the values between 20 and 61. Since the expression of a gene is usually dependent on the availability of the tRNA species, one would expect that gene with higher expression will use smaller subset of codons that are recognized by the most abundant tRNAs, result in lower ENC values [107]. Taking these two factors into consideration, we picked the top 10 orthologs per Cas13 subtypes for further investigation. I reasoned that it was not ideal to just looking at the ortholog with top scoring but rather to compare the activity of those that also have the almost top scores. Generation of selected orthologs were done via a combination of mutagenesis on original Cas13 sequences as well as fusion of gBlocks gene fragments from Integrated DNA Technologies (IDT).

2.7   Design and generation of target crRNA

The very first Cas13 proteins being characterized in bacterial requires a sequence constraint, namely protospacer flanking sequence (PFS) to ensure target cleavage efficiency, this includes *Leptotrichia shahii* Cas13a (LshCas13a), *Bergeyella zoohelcum* Cas13b (BzoCas13b)

and *Prevotella buccae* Cas13b (PspCas13b) [108,109]. However, further investigation of PspCas13b in mammalian cells and plant as well as other Cas13 orthologs being investigated showed a high target RNA degradation efficiency with no PFS required [110–113]. This gives some flexibility on target site selection. However, it is necessary to take the secondary structure of target transcripts into consideration since there have been evident that this factor is negatively correlated with knockdown efficiency [108,110]. Even though the original studies did not investigate this association in detail, I consider it is safe to pre-evaluate secondary structure of transcripts being targeted in our study before making crRNA construct. For this, I use two independent online tools, namely RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) and RNAstructure (https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html) [108,110,114–117]. In addition, we also use the siRNA design tool RNAxs (http://rna.tbi.univie.ac.at/cgi-bin/RNAxs/RNAxs.cgi) to find the regions of transcripts with good accessibility to narrow down the target region space for designing gRNAs [118]. For the case of Cas13a orthologs, we compared the target sequences with the online CRISPR-RT tool (http://bioinfolab.miamioh.edu/CRISPR-RT/interface/C2c2.php) [119]. The crRNA cassette was amplified to contain target site and cloned into pre-digested expression backbone via Gibson assembly. All gRNA used in this study were driven by *Drosophila* U6:3 promoter (dU6:3). For more information regarding crRNA cloning, see Appendix A.3.

2.8   Embryo injection

PhiC31 constructs were injected at 500-600 ng/μl concentrations, while CRISPR plasmids were used at a concentration of 100-150 ng/μl for double gRNA plasmid and 500-600 ng/μl for the donor template. Injections were performed either at the University of Alberta (Canada), Da Lat University (Vietnam), or via GenetiVision (U.S.) using standard procedures [120]. 300-500

embryos were injected per construct. Surviving adults were backcrossed to w^{1118} (for classic CRISPR, Cas9, Cas13 and cDNA transgenes) or $y^2cho^2v^1$ (for gRNA, crRNA transgenes) and used to generate independent lines.

2.9   Larvae injection

Horse spleen (F4503) and human liver ferritin (F6754) were purchased from Sigmad Aldrich. Ferritin was injected at the final concentration of 4.219 pM into late L1 larvae following previously established protocols [121] with some modifications. In brief, the experiment was done in a 18^0C walk-in incubator unit at Biological Sciences building, University of Alberta. Animals were raised on Nutri-Fly food at 25^0C with 60-70% humidity and late L2 larvae were washed in 1xPBS for 3x 3 minutes in a mesh basket (Diamed GEN46-101) before being dried on a Kimwipe. Larvae were quickly aligned on the glass slide with double side tape. Slides were allowed to dry in the desiccation chamber left at 0^0C for 3-5 minutes depending on humidity and animals survival at the time of experiment. This process will also anaesthetize larvae for later injection. Aligned larvae were then covered with a thin layer of 200S halocarbon oil (Cedarlane 25073-50) to maintain proper humidity. Ferritin solution was injected into the dorsal side of the second abdominal segment of each larva (Figure 1.4A) and the whole injection was done for no more than 3 minutes per slide. A thin layer of broken up Nutri-Fly food were then added close to the anterior of the injected larvae. Injected animals were allowed to recover on slides for 2 hours at 25^0C and carefully transferred to Nutri-Fly media for later developmental studies.

2.10   Immunostaining

Brain-Ring Gland Complexes (BRGC), fat body (FB), and salivary gland (SG) were isolated from 40-42 hour 3rd instar larvae (L3) and transferred to 1x PBS. Samples were fixed in 1x PBS 4% formaldehyde (ThermoFisher #28906) for 20 min at room temperature (RT) followed

by washing in 1x PBS 0.3% Triton (Sigma #T9284) (PBS3T) for 3x 10 min. Samples were blocked at RT for 1 hour in blocking solution (1x PBS3T 5% normal goat serum (Abcam #ab138478)) and incubated in primary antibody dilution buffer (antibody diluted in 1x PBS3T and 1% BSA) overnight at 4°C with gentle shaking. Samples were then washed in 1x PBS3T for three times with 10 min each, incubated in secondary antibody dilution buffer for 1 hour at RT, washed in 1x PBS3T and 1:50,000 DAPI (Cell Signaling #4083) for three times. Samples were mounted in Vectashield mounting medium (#VECTH1000). Pictures were taken on Nikon Eclipse 80i Confocal C2+ microscope/camera. For list of antibodies and concentrations being used, see Table 2.2.

2.11 Ferric iron staining

This protocol was modified from Perl's staining for iron with to reduce background noise, a common issue with iron-staining techniques. 42-hour L3 larvae were washed in 1x PBS for three times and dissected for BRGC. Samples were fixed with 1x PBS/4% formaldehyde for 20 min at RT. BRGC were washed 1x 10 min, 1x 20 min and 1x 30 min in 1x PBS/0.3% Triton. Samples were incubated at RT for 1 hour in fresh staining solution (2% $K_4Fe(CN)_6$ + 2% HCl) and briefly washed in 1x PBS/0.3% Triton for 5x 2 min. Samples were then incubated in 0.01 NaN_3 / 0.3% H_2O_2 for 30 min at RT and washed 3x 10 min in 0.1M Phosphate buffer pH 7.0 (57.75 mM Na_2HPO_4 and 42.25 mM NaH_2PO_4). Samples were then incubated for 10 min with fresh intensification buffer (0.1M phosphate buffer pH 7.0 containing 0.00125% DAB and 0.0025% $CoCl_2$) to reduce background staining, followed by 3x10 min wash steps in 0.1M phosphate buffer pH 7.0. Images were taken using epifluorescence camera (Nikon Digital Sight DS-U3).

2.12 RNA-sequencing and analysis

Animals were reared on standard NutriFly media (Diamed). For a single biological replicate, 50 ring glands were manually dissected in 1x PBS, transferred to Trizol (ThermoFisher #15596026), and flash-frozen in liquid nitrogen for long-term storage. RNA was extracted with the RNAeasy kit (Qiagen #74106) coupled to an on-column DNA digestion step using RNA-free DNAse (Qiagen #79254). Extracted RNA was examined on a Bioanalyser using Agilent RNA 600 nano kit (#5067-1511) to confirm RNA integrity. 100 ng total RNA from each sample was used for generating strand-specific RNA-Seq libraries based on the Ovation Drosophila RNA-Seq System 1-16 (Nugen #0350-32). cDNA quality was analysed on a Bioanalyser using the high sensitivity DNA analysis kit (Agilent #5067-4626). 100 ng cDNA in 25 μl nuclease-free water with 3 replicates per condition was used for RNA-Seq analysis (Genome Quebec Innovation Center at McGill University).

Sequencing data was analysed using Arraystar 4.0 (DNAstar), MS Access and DAVID GO Tools[122]. All RNA-Seq data has been deposited with GEO (entry # GSE130103) (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130103).

2.13 *Ex vivo* culturing of ring glands

In the first approach, BRGC were isolated from w^{1118} L3 larvae just after the L2/L3 moult, transferred to culture medium (Schneider insect medium with 10% heat-inactivated FBS, 1% streptomycin-penicillin, 10 μg/ml insulin and 2 μg/ml ecdysone), and incubated at 25°C. These conditions efficiently mimicked *in vivo* conditions and allowed physiological functions to be studied for up to 48 hours[123].

To reduce available iron (Chapter 3), BPS was added to the culture medium at a final concentration of 100 nM[124]. For evaluation o Cas9 variants (Chapter 6), BRGCs were

transfected with plasmids carrying corresponding transgenes for 24 hours, following manufacturer's protocol (Invitrogen). After 24 hours, ring glands (50 per replicate) were transferred to Trizol for later qPCR analysis.

2.14 Quantitative real-time PCR (qPCR)

Extracted RNA (Qiagen RNeasy extraction kit) was reverse transcribed via ABI High capacity cDNA synthesis kit (ThermoFisher #4368814). Synthesized cDNA was used for qPCR (QuantStudio 6 Flex) using KAPA SYBR Fast qPCR master mix #Sigma KK4601). For each condition, three biological samples were each tested in triplicate. Samples were normalized to *rp49* based on the $\Delta\Delta$CT method, with the exception of RNA-immunoprecipitation, where we normalized results to immunoprecipitated IRP1A protein levels.

2.15 Constructs for co-immunoprecipitation (S2 cells)

Fragments carrying *Drosophila* AGBE and human GBE1 cDNAs were cloned into pAMW while *Drosophila* IRP1A and human IRP1 cDNAs were cloned into pAFW. The *Drosophila* Cisd2 cDNA was cloned into pAHW, and eGFP was cloned into pAFW as well as pAMW. This approach allowed for the generation of in-frame tagged cDNAs. We used an approach that allows for the co-expression of two cDNAs that are separated by a viral-derived 2A-like peptide, which is then cleaved post-translationally to yield equal amounts of both proteins. Appropriate pairwise combinations of cDNAs encoding wild type or modified versions of 6x Myc-tagged AGBE cDNA, 6x Myc-tagged human GBE1, 3x FLAG-tagged IRP1A, 3x HA-tagged Cisd2 and 3x FLAG - tagged eGFP (the latter served as a control) were cloned into pAc5-STABLE2-Neo (Addgene #32426). For the triple transfection of IRP1A, AGBE and Cisd2, 3x FLAG -tagged IRP1A was cloned together with 6x Myc-tagged AGBE as well as 3x HA-tagged Cisd2 into pAc5-STABLE2-Neo, separated by viral-derived 2A-like peptides.

2.16   Transfection, co-immunoprecipitation, and western blotting

Cells were grown in Schneider Insect medium with 10% heat-inactivated FBS, 1% Streptomycin-Penicillin following standard procedures and transfected by the Calcium Phosphate-based method (Invitrogen). Transfected cells were lysed, and Myc-tagged bait proteins were immunoprecipitated using Myc-trap agarose beads (Chromotek Myc-Trap®-A). FLAG-tagged bait proteins were immunoprecipitated using M2 FLAG agarose beads (Sigma-Aldrich #A2220) following instructions of the manufacturer. Pulled-down samples were analysed via western blotting. For list of antibodies and concentrations used, see Table 2.2. Blots were scanned for image acquisition with a ChemiDoc imaging system (Bio-Rad).

2.17   Cell immunostaining

On the first day of the transfection experiment, coverslips were pre-cleaned in 70% ethanol and placed into transfection plate (Sigma CLS3516). Cells were then seeded and transfected following the standard procedure (Invitrogen). This allows the adherence of cells onto the coverslips for later immunostaining. Later procedure was done as described in cell culture transfection section. 7 days after transfection, coverslips were transferred to a clean transfection plate for immunostaining while cells in the supernatant were collected for cell lysis and protein extract.

Samples were fixed in 1x PBS 4% formaldehyde (ThermoFisher #28906) for 15 min at room temperature (RT) with gentle shaking followed by washing in 1x PBS 0.3% Triton (Sigma #T9284) (PBS3T) for 3x 10 min. Samples were blocked at RT for 30 minutes in blocking solution (1x PBS3T 5% normal goat serum (Abcam ab138478)) and incubated in primary antibody dilution buffer (antibody diluted in 1x PBS3T and 1% BSA) for 1 hour at RT. Samples were then washed in 1x PBS3T for three times with 10 min each, incubated in secondary antibody dilution buffer for

1 hour at RT, and then washed in 1x PBS3T with 1:50,000 DAPI (Cell Signaling #4083) for 3x 10 min. Samples were mounted in Vectashield mounting medium (#VECTH1000). Pictures were taken on Nikon Eclipse 80i Confocal C2+ microscope/camera. For list of antibodies and concentrations used, see Table 2.2.

2.18   Mass spectrometry of whole larvae

Our whole-body mass spectrometry (MS) approach was adapted from an *in vivo* cross-linking procedure developed for Drosophila embryos [125,126]. We collected 150-200 L3 larvae (40-42-hour after the moult), and washed them in 1x PBS 3x 5 min. Animals were then incubated in 1x PBS with 0.1% Triton (PBS1T) 2x 5 min before fixing in fresh fixative solution (1x PBS1T with 0.2% Formaldehyde) for 10 min. The fixing solution was removed and replaced by fresh quenching solution (0.25 M glycine in 1x PBS1T). Animals were washed in 1x PBS1T three times before being flash-frozen in liquid nitrogen for long term storage at -80°C. Larvae were homogenized in 1 ml of 1x lysis buffer (25 mM Na-HEPES pH 7.5, 75 mM NaCl, 0.5 mM EDTA, 10% glycerol, 0.1% Triton X-100, proteinase inhibitor cocktail (Sigma #11873580001)) using a Dounce homogenizer. Lysates were centrifuged at 16,000 g for 30 min at 4°C. Protein concentrations of supernatants were determined with the Qubit^{TM} Protein assay (Invitrogen #Q33212) and served to equalize protein amounts for subsequent co-immunoprecipitation assays. The supernatants were then transferred to spin columns (Chromotek sct-50) and incubated with 40 μl of anti-FLAG M2 affinity gel (Sigma #A2220) on a rotating shaker for two hours at 4°C. Columns were centrifuged and treated with wash buffer 1 (25 mM Na-HEPES pH 7.5, 75 mM NaCl, 0.5 mM EDTA, 10% Glycerol, 0.1% Triton X-100) and wash buffer 2 (25 mM Na-HEPES pH 7.5, 75 mM NaCl, 0.5 mM EDTA, 10% Glycerol) for three times each. At the last step, 40 μl of loading buffer (0.125 M Tris-HCl pH 6.8, 5% SDS, 0.004% Bromophenol blue, 20% glycerol,

1.43 M β-mercaptoethanol) was added and tubes were incubated at 95°C for 5 min before collecting samples. Samples were then loaded on a 12.5% SDS-gel, stained with Coomassie Blue and submitted for MALDI-TOF MS analysis (carried out by the Alberta Proteomics and MS Facility, University of Alberta).

In brief, we performed overnight in-gel trypsin digestion following standard procedures. Gel bands were excised and destained twice in 100 mM ammonium bicarbonate (Sigma #09830-500g)/acetonitrile (Sigma #271004) at a ratio of 50:50 (v/v). Samples were then reduced using 10 mM β-mercaptoethanol (Sigma #M6250) in 100 mM bicarbonate, followed by alkylation in 55 mM iodoacetamide (Sigma #I11490) in 100 mM bicarbonate. After dehydration, a trypsin solution (Promega #V5111) was added to cover the gel pieces at a final concentration of 6 ng/μl, and digested overnight (~16 hours) at RT. Tryptic peptides were first extracted from the gel using 97% water/2% acetonitrile/1% formic acid followed by a second extraction using 50% of the first buffer and 50% acetonitrile.

Fractions containing tryptic peptides were resolved and ionized by using nanoflow HPLC (Easy-nLC II, Thermo Scientific) coupled to an LTQ XL-Orbitrap hybrid mass spectrometer (Thermo Scientific). Nanoflow chromatography and electrospray ionization were carried out with a PicoFrit-fused silica capillary column (ProteoPepII, C18) with a 100 μm inner diameter (300Å, 5 μm, New Objective). Peptides were loaded onto the column at a flow rate of 3000 nl/min and resolved at 500 nl/min using a 60 min linear gradient from 0 to 45% v/v aqueous acetonitrile in 0.2% v/v formic acid. The mass spectrometer was operated in data-dependent acquisition mode, recording high-accuracy and high-resolution survey Orbitrap spectra using external mass calibration, with a resolution of 30,000 and m/z range of 400-2000. The fourteen most intense multiply charged ions were sequentially fragmented by using collision-induced dissociation, and

the spectra of their fragments were recorded in the linear ion trap. After two fragmentations, all precursors selected for dissociation were dynamically excluded for 60 seconds.

Data was processed using Proteome Discoverer 1.4 (Thermo Scientific). Search parameters included a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.8 Da. Peptides were searched with carbamidomethyl cysteine as a static modification and oxidized methionine and deamidated glutamine and asparagine as dynamic modifications.

2.19 Mass spectrometry of the prothoracic gland (PG)

For PG-specific MS we separated hand-dissected BRGC into individual cells. We used larvae that expressed Venus-tagged CD8 (*UAS-CD8.Venus*, Bloomington stock #65609) in a PG-specific manner. CD8 localises to the cell membrane, allowing purification of PG cells from unlabelled cells [127]. Approximately 15,000 BRGC were collected in *ex vivo* medium (Schneider insect medium with 10% heat-inactivated FBS, 1% streptomycin-penicillin, 10 µg/ml insulin and 2 µg/ml ecdysone) containing proteinase inhibitor cocktail (Sigma #11873580001). Dissection times were limited to one hour to minimize physiological changes. Samples were incubated in 1x PBS1T for 2x 5 min before being fixed in fresh fixing solution (1x PBS1T containing 0.2% formaldehyde) for 10 min. Fixing solution was removed and replaced by fresh quenching solution (0.25 M glycine in 1x PBS1T). Samples were washed three times in 1x PBS1T, followed by immersion in 1x PBS1T / 25% glycerol and flash-frozen in liquid nitrogen for long-term storage at -80°C. For knock-in derived proteins we collected 1.0 ml containing the equivalent of ~15,000 BRGCs isolated from 40-42 hr old L3, while roughly half the amount was used for flies with transgenically produced protein. Samples were removed from -80°C and thawed gradually for 15 min at -20°C followed by 15 min at 4°C until completely thawed. Tissue samples were then incubated in cell dissociation buffer (CMF buffer with 1mg/ml collagenase, 1 mg/ml papain) for

30 min at 30°C. The digestion was terminated by adding 4x volumes of CMF to the dissociation reaction. Samples were left at RT for 5 min before being centrifuged at 1,000 x g for 1 min. Cells were 3x washed in PBS1T and incubated with IgG beads that had been cross-linked with mouse CD8 antibody (#ab82005) for 30 min, followed by three brief washes in PBS1T, and an elution step (0.1 M citrate pH 2.3) to release PG cells from beads. All subsequent steps for protein extraction and immunoprecipitation were as described for whole-body MS. All MS proteomics data have been deposited to the ProteomeXchange Consortium [127] via the PRIDE [128] partner repository with identifier #PXD013499.

2.20 Quantitative RNA-immunoprecipitation (RIP)

Our *in vivo* RIP approach was adapted from different cell culture protocols[129–131]. As controls, we used w^{1118}, which is the parental strain for our transgenic and mutant lines and thus harbours no tagged genes. To immunoprecipitate IRP1A, we used 3x FLAG-tagged CRISPR/Cas9-generated knock-in alleles, namely $IRP1A^{3F}$, $IRP1A^{C450S.3F}$, and $IRP1A^{3R3Q.3F}$, representing tagged wild type, constitutively RNA-binding and non-RNA-binding forms of IRP1A, respectively. We collected 200 L3 larvae (staged at 40 hours after the L2/L3 moult) per sample. Larvae were washed for 3x 5 min in PBS, flash-frozen in liquid nitrogen and stored at -80°C. Larvae were homogenized in 1 ml lysis buffer (150 mM KCl, 25 mM Tris-HCl pH 7.4, 5 mM EDTA, 0.5% v/v Nonidet P-40, 1x proteinase inhibitor cocktail, 100 U/ml RNAse inhibitor (NEB #M0314S) using a Dounce homogenizer. Lysates were centrifuged at 12,000 g for 30 min at 4°C. Supernatants were transferred and filtered through a 0.45 µm syringe filter (Sigma #CLS431225-50EA). Flow-through samples were incubated with 300 µl equilibrated anti-FLAG M2 affinity gel solution on a rotating platform for 4 hours at 4°C followed by centrifugation at 12,000 g for 30 min at 4°C. The supernatant was removed, and the affinity gel washed in 10x

volume of lysis buffer for 2x 5 min. 5% of the final volume was saved for Western Blotting to determine IRP1A levels to adjust sample input for RIP. Western Blots were scanned with the ChemiDoc imaging system (Bio-Rad) and bands were quantified using ImageJ following standard procedures. The remaining 95% was used for Trizol-based RNA extraction followed by qPCR for *SdhB*, which harbours a validated IRE [89,98,128].

2.21 Measuring IRP1A and IRP1B aconitase activity (S2 cells)

S2 cells were grown in Schneider Insect medium with 10% heat-inactivated FBS/1% Streptomycin-Penicillin and transfected by the Calcium Phosphate-based method (Invitrogen). Transfected cells were lysed and IRP1 protein levels were evaluated as follows: From each sample, 50% of the lysate was used to immunoprecipitate IRP1A or IRP1B, and proteins were separated via SDS-PAGE. This was followed by Coommassie Blue staining of the gel to evaluate IRP1A and IRP1B protein levels, and cell lysate amounts used for aconitase assays were normalized accordingly. Aconitase activity was determined by measuring the rate of NADPH production via absorbance at 340 nm every 5 min (Aconitase-340TM kit, OxisResearch 21041, DU-730 UV/Vis Spectrophotometer). The absorbance rate was normalized relative to the rate of untransfected S2 cells, which served as a negative control for background aconitase activity.

2.22 Measuring IRP1A and IRP1B aconitase activity *in vivo*

We measured aconitase activity from both transgenically produced IRP1 as well as from knock-in alleles. For the former, we collected 200-250 L1 larvae that ubiquitously expressed transgenic IRP1A or IRP1B alleles (tub-Gal4>UAS-cDNA) and washed them 3x 5 min in 1x PBS. To measure the IRP1 aconitase activity in an AGBE mutant background, we generated lines carrying either transgenic tub>IRP1A or tub>IRP1B together with transgenic UAS-FLP-cDNA and the AGBEFCF knock-in allele to remove AGBE ubiquitously. For corresponding controls, we

used the same combination, except that we replaced UAS-FLP-cDNA with a UAS>eGFP-cDNA transgene. To evaluate the aconitase activity produced by knocked-in IRP1A alleles, we collected 200 L3 larvae (staged at 42 hours after the L2/L3 moult) carrying different IRP1A alleles in an IRP1B null mutant background to eliminate IRP1B aconitase activity. Larvae were homogenized in 1 ml of 1x Lysis buffer (25 mM Na-HEPES pH 7.5, 75 mM NaCl, 0.5 mM EDTA, 10% glycerol, 0.1% Triton X-100, proteinase inhibitor cocktail). To reduce the contribution of mitochondrial aconitase (Acon) we removed the mitochondrial fraction via ultracentrifugation at 20,000 x g and 4°C. We normalized samples based on immunoprecipitated tagged protein levels (i.e., IRP1A variants), which we evaluated via western blotting, as described in the quantitative RIP section above. Aconitase activity was determined as described for S2 cells. For the knocked-in CRISPR alleles of IRP1A, we used IRP1B null mutants as controls, which harbour a wild type copy of IRP1A.

2.23 Surveyor nuclease assay for genome modification evaluation

S2 cells were transfected as previously described (section 2.14). Cells were growth at 27-28°C for 48 hours post-transfection and washed in 1x cold PBS. Cells were collected by centrifugation at 1,000 x g for 10 min at 4°C and all wash buffer was carefully removed. Cell pellets were filled with 20 μL DNA extraction buffer (10mM Tris-HCl pH 8.2, 25mM NaCl, 1 mM EDTA pH 8.0, 0.2% v/v Triton X-100, 1x proteinase K(AM2546)) and vortexed vigorously for 3x 30 seconds. Cells lysates were incubated at 37°C for 30 min before heat-inactivation with proteinase K at 95°C for 5 min. Cell lysis solution was centrifuged at 12,000 x g at 4°C and the supernatant was transferred to a new collection tube. 1 μL of supernatant was used for PCR amplification of the genomic region spanning the target sites. PCR products were purified by HighPrep[TM] PCR reagent from MagBio (AC-60005) following manufacturer's protocol.

Test sample DNA was then combined with wild-type reference DNA at a concentration of 45-50 ng/μL with total amount of 500ng. Combined samples were run on thermocycler with a ramping temperature from 95°C to 85°C at the rate of 2°C per second, then from 85°C to 25°C at a rate of 0.3°C per second. Post-ramping samples were then treated with surveyor nuclease (Integrated DNA technologies 706025) following the manufacturer's protocol and analyzed using electrophoresis agarose gel. Band intensity was measured via ImageJ and the mutation efficiency was analyzed using the following formula [132,133]:

$$\text{indel}(\%) = 100 \text{ X} \left(1 - \sqrt{\frac{b + c}{a + b + c}} \right)$$

(a is undigested band intensity, b and c are intensities of cleavage products)

2.24 Tissue-specific DNA extraction

To analyze mutagenesis efficiency in the prothoracic gland (PG) conditional CRISPR, 15-20 ring glands were hand-dissected in collection buffer (10 mM Tris-HCl pH 8.2, 25 mM NaCl, 1 mM EDTA, 0.2% v/v Triton X-100 and 200 μg/mL proteinase K (AM2546)) and incubated for 40 min at 37°C before heat-inactivating proteinase K at 95°C for 5 min. The target region was amplified from the extracted genomic PG DNA via PCR and cloned into the pUC19 vector (NewEngland Biolabs N3041S), which was pre-digested with EcoRI and XbaI. Products were transformed into DH5α competent cells and colonies were randomly selected for Sanger sequencing.

2.25 Electron paramagnetic resonance (EPR)

To analyze the peak of intact ISC in IRP1A and human IRP1 (hIRP1), plasmids expressing *IRP1A^{WT}* or *hIRP1* cDNA were transfected into S2 cell culture following standard

protocol. 72 hours after transfection, superoxide (H_2O_2) was added at two different final concentrations of 0.6 mM or 6.0 mM and incubated for 6 hours. The proteins were then purified following the above-described procedure. The whole purification step was conducted in a glove box to avoid any oxidatively damage to ISC. For endogenous IRP1A EPR, 200 40-42 hours post L2/L3 molting larvae were collected as described above. In all cases, samples were packed into quartz EPR tubes (Sigma Z566535) by centrifugation at 2,000 x g for 10 minutes at 0^0C. Excess suspension buffer was removed, and the samples were frozen by slowly immersion in liquid nitrogen. EPR spectra was recorded using a double electron-electron resonance (DEER) spectroscopy at a temperature 11-14K in liquid Helium. Condition of the spectroscopy was at follows: microwave power of 1 milliwatt, microwave frequency of 9.235 GHz, modulation amplitude = 5 millitesla, modulation frequency = 100 kHz, time constant = 0.064s, scanning time = 2 minutes. There were 5 replicates per condition tested and data was analyzed using Xenon software (obtained from the Bruker Corporation, Massachusetts, United States).

2.26 Tables

Table 2.1 Concentration of supplements being used

supplement	experiment	solvent	stock concentration	final concentration
cholesterol	sterol rescue	ethanol	25 mg/ml	25 µg/ml[a,b,c]
7-dehydrocholesterol	sterol rescue	ethanol	125 mg/ml	125 µg/ml[a,b,c]
ecdysone	sterol rescue	ethanol		
20-hydroxyecdysone	sterol rescue	ethanol	250 mg/ml[a,b] 330 mg/ml[d,e]	250 µg/ml[a,b,c] 330 µg/ml[c,d,e] 110 µg/ml[d,e,f]
ferric ammonium citrate	iron manipulation	nuclease-free water	1M	1 mM[a,b,c]
bathophenanthroline sulfate	iron manipulation	nuclease-free water	100 mM	100 µM[a,b,c]
hemin	iron manipulation	NaOH	38mM	1mM[a,b,c]
zinc protoporphyrin	iron manipulation	DMSO	50mM	1mM[a,c]
equine (horse) ferritin	iron manipulation	NaCl	42.19 nM	42.19 pM[b,c] 14.06 pM[b,g]
human ferritin	iron manipulation	NaCl	42.19 nM	42.19 pM[b,c] 14.06 pM[b,g]
$ZnCl_2$	metal rescue	nuclease-free water	1M	1mM[a,c]
$CuSO_4$	metal rescue	nuclease-free water	150mM	150µM[a,c]
$MnCl_2$	metal rescue	nuclease-free water	1M	1mM[a,c]
glucose	glucose rescue	nuclease-free water	1M	1mM[a,c]

[a] concentration used in chapter 3

[b] concentration used in chapter 5

[c] dietary concentration

[d] concentration used in chapter 6

[e] concentration used in chapter 7

[f] embryo immersion concentration

[g] larval injected concentration

Table 2.2 Antibodies used during my study

antibody	source	product ID	experiment	dilution ratio
monoclonal mouse anti-FLAG	Cell Signaling	8146S	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal rabbit anti-FLAG	Cell Signaling	14793S	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal mouse anti-myc	Cell Signaling	2276S	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal rabbit anti-myc	Cell Signaling	2278S	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal mouse anti-HA	Abcam	18181	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal rabbit anti-HA	Cell Signaling	3724S	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal mouse anti-GFP	Invitrogen	MA5-15256	endogenous^{IF}	1:400
			transgene^{IF}	1:1000
			endogenous^{WB}	1:1000
			transgene^{WB}	1:2500
monoclonal rabbit anti-GFP	Invitrogen	G10362	endogenous^{IF}	1:200
			transgene^{IF}	1:500

			endogenous[WB]	1:200
			transgene[WB]	1:1000
monoclonal mouse anti-Cas9	Abcam	191468	transgene[IF]	1:1000
mitotracker green	Cell Signaling	9074S	mitochondrial label[IF]	400 nM
DAPI	Cell Signaling	4083	nuclear label[IF]	1:50000
monoclonal mouse anti-MTCO1	Abcam	14705	endogenous[IF]	1:2000
			endogenous[WB]	1:500
monoclonal rabbit anti-MTCO2	Abcam	79793	endogenous[IF]	1:2000
			endogenous[WB]	1:500
monoclonal mouse anti-DsRed	Santa Cruz	sc-390909	transgene[WB]	1:1000
monoclonal mouse anti-β-tubulin	Sigma Aldrich	05-661	endogenous[WB]	1:10,000
rabbit polyclonal anti-H2A	Abcam	1764	endogenous[WB]	1:1000
			endogenous[IP]	1:50
monoclonal mouse anti-Su(var)3-10	Spierer P	not applicable	endogenous[WB]	1:500
monoclonal mouse anti-Su(var)205	DSHB	C1A9	endogenous[WB]	1:100
goat anti-mouse IgG H&L Alexa Fluor 488	Abcam	150113	IF	1:2000
goat anti-mouse IgG H&L Alexa Fluor 555	Abcam	150114	IF	1:2000
goat anti-mouse IgG H&L Alexa Fluor 647	Abcam	150079	IF	1:2000
goat anti-rabbit IgG H&L Alexa Fluor 488	Abcam	1500777	IF	1:2000
goat anti-rabbit IgG H&L Alexa Fluor 555	Abcam	150078	IF	1:2000
goat anti-mouse IgG H&L HRP	Abcam	97023	WB	1:20000
goat anti-rabbit IgG H&L HRP	Abcam	97051	WB	1:20000

[IF] Immunofluorescence / Immunostain

[WB] Western Blot

[IP] Immunoprecipitation

Chapter 3   Glycogen branching enzyme controls cellular iron homeostasis

via Iron Regulatory Protein 1 and mitoNEET[2]

3.1 Introduction

3.1.1 Iron is an essential element for ecdysone biosynthesis in *Drosophila*

Iron is an essential trace element for nearly all organisms and cells, because iron co-factors, most commonly in the form of heme and iron-sulphur (Fe-S) clusters [2,134], are required for a wide range of protein functions. While essential, free iron is also highly reactive and cytotoxic, and thus its acquisition requires tight regulation by cells [2]. IRP1 is a bifunctional protein because it can reversibly bind to Fe-S clusters [135]. Under iron-replete conditions, the protein forms holo-IRP1 and acts as a cytosolic aconitase, an enzyme that interconverts citrate and isocitrate. When cellular iron levels drop, IRP1 loses its Fe-S cluster and assumes a different conformation [136], apo-IRP1, which then binds to and regulates specific target mRNAs containing iron-responsive elements (IREs). This action can either block translation or enhance transcript stability, depending on the location of the IRE, with a net outcome that promotes increased cellular iron availability and trafficking [137].

3.1.2 Prothoracic gland is a model to study cellular iron homeostasis

Current models for studying cellular iron homeostasis are limited in the sense that they have either static, or at best, linearly increasing iron requirements (cell cultures and developing erythrocytes, respectively) [2,65]. We introduce here the *Drosophila* prothoracic gland (PG) as a model to study highly dynamic iron requirements. The PG is an endocrine tissue mainly devoted to the production of steroid hormones in developing insects. In both vertebrates and insects, the synthesis of steroid hormones is largely dependent on enzymes that require heme and Fe-S clusters [66,67]. In *Drosophila*, the PG is the principal steroid-producing gland, and part of a larger endocrine tissue called the ring gland. During larval development, the PG produces pulses of the steroid hormone ecdysone, which acts as a recurring systemic signal that controls gene expression

in target tissues to coordinate hatching, molting and metamorphosis (Figure 1.4B). The first step of ecdysone biosynthesis is carried out by Nvd, which harbours an Fe-S cluster [138], while all but one of the following reactions are carried out by cytochrome P450 proteins, which require heme as an obligate co-factor (Table 1.2). The last larval stage of *Drosophila* development is accompanied by exceedingly high expression of ecdysone-producing enzyme ("Halloween") genes (Figure 1.4B) [67], indicating that the PG requires substantial amounts of iron, which can be visualized by staining for ferric iron (Figure 1.4A). PG cells have fluctuating iron and heme demands because they must match the rise and fall of Halloween enzyme levels with appropriate production rates of iron co-factors. Thus, the PG represents a powerful model to study mechanisms by which cells control dynamic changes in cellular iron and heme requirements.

3.1.3 The link between glycogen synthesis and iron homeostasis

In *Drosophila*, the *1,4-Alpha-Glucan Branching Enzyme* (*AGBE*) gene encodes an ortholog of human glycogen branching enzyme 1 (GBE1) with approximately 61% sequence identity. There is only a copy of the gene in each species and both proteins share a very similar domain architecture (Figure 3.1A). The GBE1 polypeptides carry three main domains involved in glycogen synthesis: the immunoglobulin-like, the glycosyl hydroxylase catalytic, and the alpha-amylase domains [139]. Human GBE1 acts as a hydrolase that often works together with the activity of glycogen synthase to ensure proper branching during glycogen synthesis. Indeed, this branching process not only increases the solubility of glycogen but also allows the easy break down of the molecule whenever needed (Figure 3.1B).

Abnormal GBE1, be it through gene mutations or enzyme misexpression, has been linked to two glycogen-related diseases, namely glycogen-storage disease type IV (GSD IV), also known as Andersen disease, and the adult polyglucosan body disease (APBD). Both of them are rare

disease that affect one person per every 80,000 people worldwide for GSD IV or one in every 800,000 people for APBD [140,141]. These two diseases both result from the abnormal activity of GBE1 that interfere with glycogen synthase function and thus, change the structure of glycogen to either starch in APBD with seldom branches or amylose in GSD IV [140,141]. Patients with GSD IV often suffer from a group of phenotypes, including progressive liver cirrhosis, hypotonia and cardiomyopathy. Most of the GSD IV patients die during early childhood while the APBD patients only exhibit phenotypes at adult stage [142–144].

Interestingly, literature and databases have hinted at some potential links between glycogen synthesis and iron homeostasis. The regulatory sequence of GBE1 carries a recognition sites for the hypoxia response regulator, HIF1α, to bind and regulate its transcription [145–147]. In addition, GBE1 was reported to be upregulated in hypoxic cell culture [148,149]. It is interesting that hypoxia mimics anemia since both result in low hemoglobin levels. In vertebrates, hypoxia and iron metabolism are tightly linked with findings that hypoxia results in higher iron absorption in cell cultures [150–153]. Perhaps most intriguingly, Nrf2, a transcription factor controlling mitochondrial biogenesis and important iron metabolism genes, was shown to bind directly to the GBE1 promoter [154,155], raising the idea that this glycogen enzyme is coordinately controlled with other key iron genes.

We show here that the *Drosophila* Glycogen Branching Enzyme (AGBE) is a novel regulator of iron homeostasis. AGBE interacts physically with the holo-form of Iron Regulatory Protein 1A (IRP1A) and Cisd2, an ortholog of vertebrate mitoNEET. This synergistic interaction ensures that holo-IRP1A remains functional. Further, we show that holo-IRP1A has a surprising new role in the nucleus, where it transcriptionally downregulates genes acting in steroid hormone biosynthesis as well as iron and heme metabolism.

3.2 Modified materials and methods

3.2.1 *Drosophila* stocks and husbandry

We obtained the following stocks from the Bloomington Drosophila Stock Center: w^{1118} (#3605), *UAS-AGBE^{IR2}* (#42753), *Cisd2^{G6528}* (#30170), *tubulin-Gal4/TM3, Sb^1 Ser^1* (#5138), *UAS-eGFP* (#5431), *UAS-FLP* (#4539), *UAS-CD8.Venus* (#65609), *vas.Cas9* (#51323).

Stocks *UAS-AGBE^{IR1}* (#108087), *UAS-IRP1A-RNAi* (#105583), *UAS-IRP1B-RNAi* (#110637), *UAS-Cisd2-RNAi* (#104501) were obtained from the Vienna Drosophila Resource Center.

We used CRISPR/CAS9 to generate the following knock-in and knock-out alleles (Figure 3.2, Appendix A.8): *AGBE^{FCF}*, *AGBE^{FCM}*, *IRP1A^{3F}*, *IRP1A$^{C450S.3F}$*, *IRP1A$^{3R3Q.3F}$*, *IRP1B^{3F}*, *IRP1A^{FCF}/TM3 Ser.GFP*, *IRP1A^{KO}/TM3 Ser.GFP*, *IRP1B$^{KO/KO}$*. We also generated transgenic lines based on the PhiC31 system: *UAS-3xFlag-IRP1A^{WT}*, *UAS-IRP1A^{WT}*, *UAS-3xFlag-IRP1A^{C450S}*, *UAS-IRP1A^{C450S}*, *UAS-3xFlag-IRP1A^{3R3Q}*, *UAS-3xFlag-IRP1B^{WT}*, *UAS-IRP1B^{WT}*, *UAS-3xFlag-IRP1B^{C447S}*, *UAS-IRP1B^{C447S}*, *UAS-3xFlag-IRP1B^{3R3Q}*, *UAS-Yeast Aco1^{WT}*, *UAS-Yeast Aco1$^{\Delta Sp}$*, *UAS-3xFlag-hIRP1^{WT}*, *UAS-3xFlag-hIRP2^{WT}*, *IRP1A^{gRNA}* (Table 3.1). *y^{1}w*P(nos-PhiC31\int.NLS)X; P(carryP)attP40*(II) and *y^{1}w*P(nos-PhiC31/int.NLS)X; P(carryP)attP2* (III) were gifts from BestGene Inc.

phm22-Gal4 was a kind gift from Michael O'Connor's lab. Stocks were maintained on a standard cornmeal diet unless otherwise specified.

3.2.2 Other experiments

Other experiments, including survival studies, generation of CRISPR/Cas9 strains, embryo injection, mass spectrometry of whole body larvae or ring gland, immunostaining, ferric iron staining, RNA-sequencing and analysis, aconitase assay, RNA immunoprecipitation, cell culture

transfection, immunoprecipitation, real-time PCR, and western blot were done as described in chapter 2.

3.3   Results and Discussions

3.3.1   Loss of Glycogen Branching Enzyme causes porphyria-like Phenotypes

The heme biosynthetic pathway is highly conserved in metazoans and fungi and comprises eight enzymatic steps that convert glycine and Succinyl-CoA to mature heme (Figure 1.5A) [156]. Exposure to air and UV light isomerizes porphyrinogen rings, first produced in step 4, into autofluorescing porphyrins, but incorporation of iron into protoporphyrin IX results in non-fluorescing heme (Figure 1.5B) [156,157]. We noticed the presence of red autofluorescence in the PG when we exposed larvae from four RNAi lines to UV light (targeting *Updo*, *Ppox*, *spz5* and *AGBE*), all of which had been identified in two unrelated PG-specific RNA interference (RNAi) screens[66,67]. Also, the ring glands were enlarged compared to time-matched controls and had a red-brown appearance under brightfield light (Figure 3.3). A fifth RNAi line, *Nos^{IR-X}*, targeting the *Nitric Oxide Synthase* (*Nos*) gene, had been reported to produce large red-brown PGs [158], and when re-examined by us, also showed red autofluorescence (Figure 3.4A). Consistent with their role in heme biosynthesis, depleting Updo (= vertebrate UROD, Table 1.3) and Ppox caused protoporphyrin accumulation in the PG. This is equivalent to what occurs in patients afflicted with porphyria, a group of rare diseases impairing heme biosynthesis[159]. PG-specific *Alas*-RNAi, on the other hand, disrupts heme synthesis prior to porphyrinogen ring formation, and therefore lacked the autofluorescence, but causes enlarged ring glands (Figure 3.4A). We then sought to validate the three remaining lines, *Nos^{IR-X}*, *AGBE^{IR1}* and *spz5^{IR}*, since their relationship to heme biosynthesis was intriguing. We were unsuccessful in finding independent evidence for the *Nos^{IR-X}* and *spz5^{IR}* lines, suggesting that the phenotypes were caused by off-target effects. However, a

second, non-overlapping RNAi line targeting *AGBE*, *AGBE^{IR2}*, caused similar phenotypes (Figure 3.4B). *AGBE* encodes a glycogen branching enzyme, which is an essential enzyme that acts in glycogen biosynthesis [160]. There are, however, no reports that link glycogen branching enzymes to iron or heme homeostasis. Therefore, we further validated these results by using CRISPR/CAS9 to replace the endogenous *AGBE* gene with a genomic copy that was flanked by Flippase (FLP) Recombinase Target (FRT) sites, and where the last exon extended into a FLAG- or Myc-tag (*AGBE^{FCF}* and *AGBE^{FCM}*, Figure 3.2). Excision of the conditional *AGBE^{FCF}* allele via PG-specific expression of FLP confirmed the *AGBE*-RNAi phenotypes, as we observed strong autofluorescence with overall higher penetrance than the two RNAi lines, since no adults eclosed compared to 4.6% in homozygous *PG>AGBE^{IR1}* animals (Figures 3.4 B,C).

We reasoned that a lack of cellular or mitochondrial iron could disrupt heme production and may explain the porphyria phenotype in *AGBE*-loss-of-function lines. Remarkably, upon rearing *PG>AGBE^{IR1}* and *PG>AGBE^{FCF}* larvae on an iron-supplemented diet, the autofluorescence was absent (Figure 3.4B), and we observed that ~40-50% of the larvae now developed into phenotypically normal adults (Figure 3.4C). In agreement with this, adding the iron chelator Bathophenanthroline Sulfate (BPS) slightly decreased survival rates, while adding both iron and BPS to the diet reversed the rescue seen by iron alone, confirming that BPS is an effective tool to reduce available iron. *AGBE*-RNAi larvae were uniquely rescued by dietary iron, since neither *Alas*-, *Updo*-, *Ppox*-, *Nos*- or *spz5*-RNAi lines benefited from iron supplementation. *AGBE* expression was moderately upregulated under iron-chelating conditions, consistent with the idea that the gene partakes in cellular iron homeostasis (Figure 3.4D). Next, we tested whether disrupting four other glycogen biosynthesis genes via PG-specific RNAi would phenocopy AGBE-depletion (Figure 3.5). This neither caused autofluorescence nor significant lethality.

58

However, ubiquitous expression of RNAi targeting these glycogen biosynthesis genes caused widespread larval lethality, confirming that all RNAi transgenes in these lines were expressed, which suggested that the disruption of glycogen biosynthesis per se in the PG did not cause any iron- or heme-related phenotypes, but was a unique feature of AGBE.

3.3.2   Glycogen Branching Enzyme physically interacts with Iron Regulatory Protein 1

AGBE is the single orthologue of vertebrate GBE1 (Glycogen Branching Enzyme 1), and the two proteins are 61% identical. A search of protein-protein interaction databases [161–163] revealed that human GBE1 physically interacts with IRP1[2]. Vertebrates have two *IRP* genes, IRP1 and IRP2, but only IRP1 can switch between the aconitase and the RNA-binding form, while IRP2 lacks the Fe-S cluster and is constitutively RNA-binding[135]. Based on sequence comparison, *Drosophila* does not have the *IRP2* gene, but harbours two *IRP1* genes (IRP1A and IRP1B) (Figure 3.6). Only IRP1A has been shown to switch from holo- to the IRE-binding apoform, while IRP1B is believed to act only as an aconitase, as it failed to bind canonical IREs[61]. The reported interaction between GBE1 and IRP1 raised the possibility that *Drosophila* AGBE and human GBE1 function in the regulation of iron homeostasis by modulating IRP1 activity. Using a cell culture approach, we established that the interaction also occurred in *Drosophila*, namely between AGBE and IRP1A (Figure 3.7A), and we recapitulated the interaction between human IRP1 and GBE1 in the same system (Figure 3.7B). AGBE interacted robustly with wild type IRP1A, as well as with IRP1A^{3R3Q}, which carries three point mutations (R549Q, R554Q, R793Q, see Figure 3.2) predicted to disrupt RNA-binding [164]. Strikingly, a single point mutation that prevents Fe-S cluster binding to IRP1A (predicted to generate constitutively RNA-binding IRP1A^{C450S}) abolished the interaction with AGBE (Figure 3.7A), suggesting that holo-IRP1A is the *in vivo* binding partner of AGBE. This was paradoxical, as

AGBE mutations caused iron-deficiency phenotypes, but holo-IRP1A has no known roles in iron homeostasis as it is believed to only act as an aconitase. To examine whether the IRP1A^{3R3Q} and IRP1A^{C450S} forms acted as predicted, we tested their *in vivo* ability to bind an IRE-containing mRNA (*SdhB*) and whether either of them had aconitase activity (Figures 3.7 C,D). As expected, the IRP1A^{3R3Q} form displayed strongly reduced mRNA-binding compared to both the wild type and the IRP1A^{C450S} forms (Figure 3.7C). Likewise, both knocked-in and transgenic alleles of *IRP1A^{C450S}* resulted in significantly reduced aconitase activity compared to wild type IRP1A and the IRP1A^{3R3Q} form (Figures 3.7D and 3.8).

To further substantiate the interaction between AGBE and holo-IRP1A, we carried out a series of MALDI-TOF-based MS experiments (see later section). We also sought to validate this interaction by genetic means. For the latter, we tested whether animals with PG-specific loss of *AGBE* function could be rescued by expressing transgenic wild type *IRP1A* or *IRP1A^{C450S}* (Table 3.1). Remarkably, wild type *IRP1A* rescued both the larval lethality (Figures 3.7E, F) as well as the porphyria-like phenotype of animals that lacked functional AGBE (Figure 3.7G), while *IRP1A^{C450S}* was completely ineffective (Figures 3.7E, G). It was possible that the *IRP1A^{C450S}* allele was not functional, despite differing only in a single point mutation from *IRP1A*. However, when we expressed *IRP1A^{C450S}* in other genetic backgrounds, we observed dramatic rescue of *PG>NOS^{IR-X}* RNAi animals with respect to both the lethality (Figure 3.7F) and protoporphyrin accumulation. This demonstrated that *IRP1A^{C450S}* was fully active but not sufficient to compensate for the iron deficiency in AGBE-depleted animals, suggesting that holo-IRP1A has functions beyond the aconitase that are important for iron homeostasis. To test whether the aconitase function of holo-IRP1A had unexpected essential functions, we attempted rescuing *AGBE^{FCM}* mutants with the aconitase-only form of IRP1A (IRP1A^{3R3Q}), as well as a cytosolic and

mitochondrial version of yeast aconitase (*YAco1^WT* and *YAco1^{ΔSp}*, respectively), neither of which can switch to the RNA-binding form [164,165]. None of these approaches rescued the loss-of-AGBE-function phenotypes (Figure 3.7G), indicating that both holo- and apo-IRP1A were required for survival. Lastly, we crossed human *IRP1* (*hIRP1*) and *IRP2* (*hIRP2*) into the *AGBE^{IR1}*-RNAi strain. Consistent with the above findings, only hIRP1 (equivalent to IRP1A, Figure 3.6) could fully rescue *AGBE^{IR1}* larvae, while constitutively RNA-binding hIRP2, was much less effective (Figure 3.7G), albeit more efficient than *IRP1A^{C450S}*, suggesting partial rescue by hIRP2.

Since no null mutations were available for *IRP1A* or *IRP1B*, we needed to establish that a) these genes had indeed roles in *Drosophila* iron regulation and if so, b) whether *IRP1A* and *IRP1B* had distinct roles in controlling cellular iron levels, c) whether these genes were required in the PG, and d) whether this would phenocopy AGBE-depletion in the PG. We first disrupted both *IRP1A* and *IRP1B* in the *PG* via RNAi (*PG>IRP1A^{IR}* and *PG>IRP1B^{IR}*). On regular fly food, neither RNAi line resulted in obvious phenotypes. However, when flies were reared on iron-depleted fly food for three generations, *PG>IRP1A^{IR}* animals displayed significant larval lethality, with a concomitant appearance of red autofluorescence in the larval PG (Figure 3.9A). Control and *PG>IRP1B^{IR}* populations did not exhibit lethality until the 5th generation and larvae never showed any autofluorescence (Figure 3.10). To confirm these data, we used two approaches. First, we generated a Flag-tagged and FRT-flanked knock-in allele of endogenous *IRP1A* (*IRP1A^{FCF}*), allowing us to excise the gene via PG-specific expression of *FLP* (Figure 3.2). This approach resulted in red-fluorescing PGs in *PG>FLP;IRP1A^{FCF}* larvae that were switched from iron-replete to BPS-containing media (Figure 3.9B). However, homozygous *IRP1A^{FCF}* flies were not viable on regular fly media, indicating that the inserted *FRT* sites had generated a loss-of-*IRP1A*-function allele. Therefore, we employed a second CRISPR strategy, where we crossed flies that specifically

expressed *CAS9* in the PG [199] to flies expressing two gRNAs that targeted *IRP1A* (Table 3.1). The resulting F1 progeny also displayed PG-specific autofluorescence and 100% lethality on regular fly media (Figure 3.9B). To complement these PG-specific lesions with classic mutant analysis, we examined an existing mutant *IRP1A* line (Bloomington # 30181). However, we did not consider this allele further as when we analyzed its phenotype and function, it turned out to be a weak hypomorph with still detected normal IRP1A activity. We therefore generated deletion mutants for *IRP1A* and *IRP1B* using CRISPR/CAS9, designated here as *IRP1A^{KO}* and *IRP1B^{KO}*. On regular fly food, *IRP1A^{KO}* mutants died as first (L1) and second instar larvae (L2), but were able to develop into phenotypically normal adults when reared on an iron-supplemented diet (Figure 3.9C), indicating that IRP1A was essential for responding to iron-poor conditions. In contrast, *IRP1B* mutants revealed no obvious phenotypes under any of the tested conditions. Taken together, these results showed that IRP1A is the principal regulator of cellular iron homeostasis in *Drosophila* and that IRP1A depletion phenocopied the iron-dependent porphyria seen in *AGBE* mutants.

3.3.3   Subcellular localization of apo- and holo-IRP1

We then addressed whether holo-IRP1A had novel roles in the regulation of cellular iron homeostasis, since only the holoform interacted with AGBE, and was required to rescue *AGBE* mutants. When we carried out immunolocalisation of *PG>3xFLAG-IRP1A^{WT}* and *PG>3xFLAG-IRP1B^{WT}* transgenic lines (Table 3.1), we found that both IRP1A and IRP1B were enriched in PG nuclei (Figure 3.9D). In stark contrast, expressing the single-point mutation variants *IRP1A^{C450S}* or *IRP1B^{C447S}* (which abolishes Fe-S-binding in IRP1B, Table 3.1) resulted in predominantly cytoplasmic accumulation of either fusion protein (Figure 3.9D). Similarly, we found that human IRP1 localises to PG nuclei as well, while human IRP2 failed to do so (Figure 3.9E). This

behaviour is consistent with the absence of an Fe-S cluster in hIRP2, indicating that nuclear accumulation strongly favours holo-IRP proteins.

Given the interaction between AGBE and holo-IRP1A, we wondered whether entry of IRP1A into nuclei was dependent on AGBE. For this, we crossed the FLAG-tagged *IRP1A^{3F}* and *IRP1B^{3F}* knock-in alleles (Figure 3.2) into *AGBE* mutants. This approach revealed that IRP1A, but not IRP1B, was dependent on *AGBE* for nuclear translocation (Figure 3.9F), suggesting that IRP1A requires AGBE for maintaining Fe-S clusters, which in turn are needed for nuclear entry. Finally, we determined the subcellular localisation of IRP1A^{3F} and IRP1B^{3F} in other tissues. Both IRP1A and IRP1B showed strong cytoplasmic and negligible nuclear presence in the larval salivary gland, while the adjacent fat body displayed predominantly nuclear IRP1A and IRP1B (Figure 3.9G). This strongly suggests that nuclear translocation of IRP1 proteins is highly tissue-specific, and not solely a function of systemic iron load, and hence may reflect tissue-specific iron requirements. Given that human IRP1 also localises to *Drosophila* nuclei, this raises the question of whether vertebrate IRP1 may also enter nuclei in specific tissues or during specific developmental/physiological conditions.

We next sought to identify proteins that would physically interact with IRP1A and IRP1B in order to shed light on the interaction with AGBE and the presence of both IRP1s in nuclei. For this, we immunoprecipitated endogenous or transgenic versions of FLAG-tagged AGBE, IRP1A and IRP1B and subjected ring gland and whole-body samples to MS, for a total of 17 conditions (Table 3.2). As controls, we used a total of five wild type samples (which lack Flag-tagged proteins), processed them in parallel to the experimental samples, and removed all proteins found in the controls from the experimental data sets [166]. Briefly, the interactome for IRP1A indicated extensive interactions with ribosomal proteins and eukaryotic initiation factors, consistent with

previous findings [167,168] and IRP's role in regulating translation. We also identified four histone proteins for IRP1A (H4, H2A, H2B, and H2Av) and two for IRP1B (H4 and H2A), consistent with the presence of both IRP1s in nuclei (Figure 3.11A) [166]. Importantly, IP-MS endogenously tagged AGBE[FCF] specifically pulled down a total of 22 proteins from whole-body samples (Figure 3.11B), which included IRP1A, IRP1B and Cisd2, an Fe-S protein. Vertebrate Cisd1, Cisd2, and Cisd3 comprise a small CDGSH iron sulfur cluster domain (Cisd) family referred to as the NEET proteins [169], which harbour an unusual 2Fe-2S cluster that enables these proteins to transfer their cluster to other proteins [170–172]. *Drosophila* encodes only two NEET proteins, Cisd2 and CG3420 (aka Cisd3), where Cisd2 lies evolutionary between human mitoNEET (encoded by Cisd1) and Naf-1 (encoded by Cisd2) [173], and as such, fly Cisd2 may be functionally related to both.

We also used endogenously tagged IRP1A (IRP1A[3F]) as bait, which co-immunoprecipitated 166 proteins that included AGBE, Cisd2, and IRP1B, as well as two ferritins (iron storage proteins), Fer1HCH and Fer2LCH (Figure 3.11A, B) [166]. Both AGBE and IRP1A interacted with another glycogen enzyme, glycogen synthase (GlyS), further corroborating that cellular iron homeostasis and glycogen metabolism are physically linked. IRP1B pulled down AGBE and the histones H2A and H4, but not Cisd2. For IRP1A, all above interactions, with the exception of GlyS, were validated by PG-specific MS (Figure 3.11A) [166]. Finally, we further validated these MS data with samples from four fly strains that expressed one of the following transgenes: apo-IRP1A (*IRP1A^C450S*); apo-IRP1B (*IRP1B^C447S*); non-RNA-binding IRP1A (*IRP1A^3R3Q*) and IRP1B (*IRP1B^3R3Q*) (Tables 3.1, 3.2). This approach confirmed the results seen with the knock-in alleles, and, importantly, showed that IRP1A[C450S] failed to interact with AGBE, while Cisd2 interacted with both IRP1A variants, but none of the IRP1B proteins (Figure 3.12).

64

Since iron-depletion triggers the switch from holo- to apo-IRP1A, one would predict that this results in cytoplasmic accumulation of IRP1A and should therefore alter the interactome of this protein. When we reared flies for three generations on BPS-containing media, we noticed that it takes two generations to purge IRP1A and IRP1B from PG nuclei (Figure 3.13). In the fat body, however, it takes only one generation for IRP1A to become entirely cytoplasmic, while IRP1B is still nuclear after three generations of iron-depletion (Figure 3.14), suggesting that IRPs have tissue-specific behaviours. To test whether iron-depletion affected protein-protein interactions of IRP1A, we reared endogenously tagged IRP1A^{3F} flies for two generations on BPS-supplemented food, and conducted MS from whole-body larval samples. This strategy reduced the number of co-immunoprecipitated proteins from 166 (no BPS) to 117 (in G1 = one generation BPS) and 30 (in G2 = two generations of BPS) [166]. Consistent with the cytoplasmic localisation of IRP1A on BPS media, the interaction with histone proteins was reduced in G1 and undetectable in G2, similar to co-immunoprecipitation results with the cytoplasmic IRP1A^{C450S} protein (Figure 3.11C). Further, binding to AGBE was lost in G1 and G2, consistent with our finding that AGBE only interacts with holo-IRP1A. The interaction between IRP1A and IRP1B was lost in G2, while binding to Fer2LCH was detectable in all conditions. Importantly, the top-scoring protein in G2 was Cisd2/mitoNEET (Figure 3.11D) [166], suggesting that this interaction is particularly important when iron levels are low.

3.3.4 *MitoNEET* mutants phenocopy *IRP1A* and *AGBE* mutants

MitoNEET is a homodimeric Fe-S protein that resides in the outer mitochondrial membrane, with the Fe-S cluster facing the cytosol [169]. While the exact range of functions for mitoNEET remains unclear, the protein has been shown to act in the repair of oxidatively damaged vertebrate IRP1 Fe-S clusters [169,174]. Since the MS data suggested that Cisd2/mitoNEET

represents an important link between AGBE and IRP1A, we sought validation by molecular and genetic means. First, we validated the physical interaction via co-immunoprecipitation assays in *Drosophila* S2 cells. This showed that both AGBE and IRP1A physically interact with Cisd2 (Figure 3.11E). However, this interaction was much enhanced when AGBE and IRP1A were co-transfected together with Cisd2 (Figures 3.11E, F), suggesting synergistic interactions between the three proteins.

Next, we examined Cisd2 function by genetic means. When we depleted Cisd2 via RNAi in the PG, or examined a *Cisd2* mutant, flies survived on regular food, and displayed no protoporphyrin accumulation (Figures 3.11G, H). On BPS media, however, most (*Cisd2* mutant) or all (*Cisd2*-RNAi) animals arrested development during the third instar and displayed red autofluorescence in the PG (Figures 3.11G, H). We then tested whether Cisd2 and IRP1A interacted genetically, and therefore analysed RNAi lines targeting *Drosophila Cisd2*, *IRP1A* and *IRP1B* alone and in combination. None of the individual PG>RNAi larvae displayed any overt phenotypes when reared on regular fly media. However, when we combined *IRP1A*- with *Cisd2*-RNAi, we observed strong synthetic lethality, where none of the larvae reached adulthood, and importantly, all larvae displayed protoporphyrin accumulation in the PG (Figure 3.11I), indicating that both proteins participate in the same process. In contrast, the combination of *IRP1B*- with *Cisd2*-RNAi was as ineffective as the individual lines alone. We concluded that the functional importance of the IRP1-mitoNEET interaction is conserved between vertebrates and *Drosophila*, and that this process is essential, at least in *Drosophila.* Finally, we tested whether IRP1A and IRP1B could localise to nuclei in a *Cisd2*-mutant background. To test this, we crossed flies that harboured FLAG-tagged *IRP1A^{3F}* or *IRP1B^{3F}* knock-in alleles into a *Cisd2*-mutant background and reared them on fly food in the presence or absence of BPS. On regular fly food, both proteins

were nuclear, while exposure to BPS shifted their subcellular distribution to the cytoplasm of the PG (Figure 3.11J). Control flies reared for one generation on BPS-containing food still show predominantly nuclear IRP1A and IRP1B in the PG (Figure 3.13). Taken together, these data indicate that AGBE, Cisd2/mitoNEET and IRP1A act together to ensure that holo-IRP1A remains functional and can enter the nucleus.

3.3.5 AGBE and Cisd2 mutations affect the stability of the IRP1A iron-sulfur cluster

We showed AGBE and Cisd2 are required for IRP1A nuclear localization and the predominant nuclear form of this protein is the holo-form (Figures 3.9D, F and 3.11H). We also showed that AGBE, Cisd2 and IRP1A physically interact (Figures 3.7A, B and 3.11C). Since Cisd2 is the homolog of mammalian mitoNEET, a protein acts in the repair of oxidatively damaged ISC in IRP1, we hypothesized that in *Drosophila*, Cisd2 might also act to repair the damaged ISC in IRP1A and the presence of AGBE stabilizes the interaction between Cisd2 and the damaged IRP1A for the repair process. If the hypothesis is correct, one can expect the accumulation of damaged IRP1A when either AGBE or Cisd2 function is impaired. This logic is reasonable since the nuclear localization of holo-IRP1A is dependent on AGBE and Cisd2. To this end, we tested the status of ISC of IRP1A in *AGBE* and *Cisd2* mutant animals using electron paramagnetic resonance (EPR) (Figure 3.15A). In mammals, the ISC in IRP1 contains four ferrous ions (Fe^{2+}) and four sulfur ions (S^{2-}). When iron level is low, IRP1 initiates its conformational switch by first releasing the outermost ferrous ion, results in the $[3Fe-4S]^{2-}$ cluster (IRP1$^{[3Fe-4S]}$) [175]. Our understanding of the subsequent cluster disassembly steps remain unclear. Recent studies have suggested the phosphorylation of serine residue 138 (S138) and oxidation of [4Fe-4S] ISC are both involved in this process [176,177]. IRP1$^{[3Fe-4S]}$ can be considered as an intermediate form between holo- and apo-IRP1. Since IRP1$^{[3Fe-4S]}$ is generally short-lived, it is not easily captured by

crystallization. Instead, by using EPR, we can rely on the presence of unpaired electrons in IRP1 in the magnetic field and detect this transition conformation. Previous studies have shown that the intact [4Fe-4S] in human IRP1 purified from yeast results in a single EPR peak at g factor of 2.015 while IRP1$^{[3Fe-4S]}$ exhibits two peaks at 2.033 and 2.015. Since other researchers in the field have successfully used EPR, I wondered if I can use this approach to detect the accumulation of damaged IRP1A.

Even though IRP1A is the ortholog of human IRP1 (hIRP1), their sequences are not 100% identical. Thus, their conformations might be slightly different and the ISC in this protein maybe different from its vertebrate counterpart and may not show EPR peaks at the similar g factors as hIRP1. We first used the *Drosophilia* S2 cells to determine the peaks of healthy and damaged IRP1A. In this approach, we tested the purified IRP1A^{WT} treated with different amounts of H_2O_2 (0, 0.6 and 6.0 mM H_2O_2). We also tested the hIRP1 under similar conditions as a reference. In our hands, without H_2O_2 treatment, the ISC cluster of purified hIRP1 has an EPR peak at a g value of 2.014, and two EPR peaks at g values of 2.014 and 2.032 when treated with 0.6 mM H_2O_2. These values are very close to the reported g values of earlier studies and show a similar trend that the g factor of [4Fe-4S] cluster has a smaller value than the [3Fe-4S] cluster [176,177]. Interestingly, hIRP1 treated with 6.0 mM H_2O_2 shows no peaks at all (Figure 3.15B). This suggests the complete removal of ISC, and the protein is now predominantly in the apo-form.

Similar to hIRP1, [4Fe-4S] IRP1A also exhibits a distinct peak from [3Fe-4S] IRP1A. Intact ISC in this protein generates an EPR peak at a g value of 2.019 while proteins exposed to a low dose of H_2O_2 is a mixture of intact and damage IRP1A and thus carry two peaks at 2.019 and 2.038 (Figure 3.15C). We then tested the status of ISC in IRP1A in the *AGBE*- and *Cisd2*- impaired animals. In this approach, we used a whole-body lysate from mutant animals and detected these

distinct peaks of IRP1A at the desired g values. Since IRP1A and IRP1B are two *Drosophila* orthologs of hIRP1, we reason that the ISC of IRP1B may exhibit similar peaks as IRP1A and thus, interfere with our results. We combined the desired mutants with *IRP1B^{KO}* allele that we generated. For *Cisd2* mutant, since the animals only show porphyria-like phenotype on iron-depleted medium, we also raised and collected animals in the iron-rich medium as the control. In both approaches, the loss of either *AGBE* or *Cisd2* shows the accumulation of damaged IRP1A via two EPR peaks (Figure 3.15D). We also used EPR and tested IRP1A's ISC in *Nos*-RNAi. Earlier work by Pendleton Cox, a previous graduate student in King-Jones lab, showed that the RNAi could be rescued by IRP1A^{C450S} overexpression and suggested a link between this RNAi and IRP1A. However, we did not notice any significant accumulation of damaged IRP1A in these animals. These data together suggest that *AGBE* and *Cisd2* are both required to protect the integrity of IRP1A's ISC, which further supports our hypothesis (Figure 3.16).

3.3.6 Nuclear Iron Regulatory Protein 1

An intriguing possibility is that holo-IRP1 has additional roles in the nucleus that contribute to tissue-specific cellular iron homeostasis. This is supported by the MS data, which indicates distinct but overlapping binding behaviours by IRP1A and IRP1B to histone proteins. To examine this further, we carried out genome-wide transcript profiling of hand-dissected ring glands (which contain the PG) that expressed one of six FLAG-tagged transgenes in a PG-specific manner: *IRP1A^{WT}* and *IRP1B^{WT}*, which are both wild type; *IRP1A^{C450S}* and *IRP1B^{C447S}*, both of which can only assume the apo-form and are predominantly cytoplasmic; as well as *IRP1A^{3R3Q}* and *IRP1B^{3R3Q}*, both of which are presumed to be non-RNA-binding and can enter nuclei (Table 3.1). The design of this approach was based on the idea that the transcriptional changes elicited by IRP1A^{3R3Q} and IRP1B^{3R3Q} should largely result from their nuclear function, as they are predicted

to have lost RNA-binding capability. When we examined the 234 most significantly downregulated genes by IRP1A^{3R3Q} in comparing with controls, we noticed strong enrichment of genes involved in iron-dependent processes, most notably steroid hormone biosynthesis (Tables 3.3 and 3.4). The results for IRP1B^{3R3Q} were very similar (Pearson correlation 0.896, P<0.001), and will not be discussed separately here. Specifically, six of the seven known Halloween enzymes were found among the 60 most strongly downregulated genes. Furthermore, other genes involved in ecdysone biosynthesis, such as transcription factors, sterol transporters, heme biosynthesis and iron-sulphur cluster assembly proteins were also significantly enriched in this set (Table 3.4). Remarkably, the fold changes for these genes were highly consistent with the predicted functions of these IRP1A variants. In particular, wild type IRP1A displayed the same trend as IRP1A^{3R3Q}, but fold changes were less severe. This is consistent with the idea that wild type IRP1A is still capable of binding to mRNAs, effectively reducing nuclear IRP1A levels, resulting in similar, but reduced responses. IRP1A^{C450S} is mostly cytoplasmic, but interacted weakly with histones (Figure 3.11), suggesting some nuclear presence. However, most gene expression changes were not significant, suggesting that IRP1A^{C450S} had little influence on altering the expression profiles of this gene set. In conclusion, the use of different IRP1A variants allowed us to distinguish the different subcellular roles of IRP1A, and we could show that IRP1A^{3R3Q}, but not IRP1A^{C450S}, dramatically and significantly altered the expression of genes involved in iron-dependent processes.

3.4   Discussion

In this chapter, we demonstrated that the *Drosophila* glycogen branching enzyme, AGBE, has hitherto undiscovered and essential roles in the regulation of cellular iron homeostasis. We expect that AGBE's role in iron is not limited to the PG, since genome-wide expression profiling

indicates that AGBE is widely expressed [178]. While AGBE has not been directly linked to iron homeostasis, a possible indirect link exists because mutations in *RBCK1* (RanBP-type and C3HC4-type zinc finger-containing protein 1), a gene that encodes an E3 ubiquitin ligase, cause Polyglucosan Body Myopathy, a recently described glycogen storage disorder [179]. Intriguingly, RBCK1 was shown to control cellular iron homeostasis by degrading the oxidized form of IRP2 [180], raising the idea that glycogen and iron processes are linked on multiple levels.

The finding that AGBE regulates cellular iron homeostasis led to another surprising discovery, namely that IRP1, in a tissue-specific manner, enters nuclei in its holoform to transcriptionally downregulate iron-intensive processes. Further, both AGBE and IRP1A interact with Cisd2, a close homolog of vertebrate mitoNEET, which is known to repair oxidatively damaged IRP1. We conclude that the glycogen metabolism enzyme AGBE has a "moonlighting" function in aiding Cisd2 in this repair process, and that loss of either Cisd2- or AGBE-function results in the accumulation of damaged IRP1A, which interferes with nuclear entry (Figures 3.9F, 3.11H and 3.16) and IRP1A aconitase activity (Figure 3.8). This is consistent with our finding that only holo-IRP1A can translocate to nuclei, since both BPS-treatment and a mutation in a critical cysteine required for Fe-S binding (IRP1A^{C450S} and IRP1B^{C447S}) impairs nuclear access (Figure 3.9D). Thus, cells possess two mechanisms by which functional IRP1A can be generated. One is by "*de novo*" insertion via the Cytosolic Iron-sulphur protein Assembly (CIA), a highly conserved machinery that assembles and inserts [4Fe-4S] clusters into client proteins [181]. Once inserted, cells require a second "maintenance" process via the mitoNEET/AGBE proteins to repair oxidatively damaged clusters to functional units (Figure 3.16). This elegantly explains as to why *AGBE*-loss-of-function animals can only be rescued by expressing a wild type *IRP1A* transgene, but not by the constitutively RNA-binding form (IRP1A^{C450S}): Sustained transgenic expression of

wild type *IRP1A* allows cells to produce sufficient functional IRP1A before oxidative damage occurs, simply because the CIA machinery is able to maintain critical levels of holo-IRP1A, despite the absence of a functioning AGBE/mitoNEET repair machinery (Figure 3.16). In contrast, IRP1A^{C450S} fails to rescue, since it cannot assume the holoform that is required to enter nuclei.

What could be the biological context that requires IRP1A and IRP1B entering the nucleus? In the PG, iron demands are not only exceedingly high, but they also must change dramatically as the need for Halloween enzyme production changes during development (Figures 1.4B and 3.16). It is therefore plausible that once production of ecdysone has peaked, PG cells need to downregulate all processes that are tied to the synthesis of steroids. Since all but one of the ecdysone-producing Halloween enzymes require iron co-factors, it is necessary to downregulate iron-cofactor production in concert with the proteins that require them. We hypothesize that peak levels of bioavailable iron correlate with maximal nuclear activity of holo-IRP1, resulting in a downregulation of iron-dependent processes, in particular steroid hormone biosynthesis. As such, holo-IRP1 appears to have a novel role in iron regulation: Apo-IRP1, as a cytoplasmic mRNA-binding protein, responds to a drop in cellular iron and facilitates an increase of bioavailable iron, yet nuclear holo-IRP1 transcriptionally downregulates iron- and heme-dependent processes once peak iron demand is over (Figure 3.16).

How does IRP1 cause the coordinated transcriptional downregulation of iron-dependent processes? An attractive model is that IRP1 proteins interact with a subset of modified histone tails, rather than binding to their histone partners in a non-discriminate fashion. We see two possible scenarios from here. First, histone-bound IRP1 could directly recruit repressive chromatin factors such as histone deacetylases or chromatin remodellers, and simply act as a co-factor that serves as a readout for cellular iron concentrations. The second and perhaps more intriguing

possibility is that holo-IRP1 controls nuclear citrate levels via its aconitase function to indirectly regulate histone acetylation. Nuclear citrate is converted to acetyl-CoA and oxaloacetate by nuclear ATP-citrate lyase (ACL), a metabolic enzyme with critical roles in histone acetylation [182]. Acetyl-CoA is the principal substrate for histone acetylation, and is considered a highly regulated nuclear metabolite that controls histone acetylation status [183,184]. Histone-bound IRP1A and IRP1B could then act by converting citrate into isocitrate and deplete Acetyl-CoA levels, thus negatively impacting gene expression by promoting histone de-acetylation.

Mutations in human GBE1 cause Andersen disease, also known as Glycogen Storage Disease Type IV (GSD IV) [185], but the gene has not been linked to iron homeostasis yet. There are strong indications, however, that GBE1 has hitherto undocumented roles in vertebrate iron metabolism as well. Besides the earlier mentioned interaction with IRP1, GBE1 was identified by whole-exome sequencing as a novel mitochondrial disorder locus [186], consistent with a study that found abnormal mitochondria in GSD IV patients [187]. Furthermore, GBE1 is transcriptionally upregulated in response to hypoxia and one of the most strongly induced genes upon nickel exposure [145,149,188]. Nickel exposure elicits hypoxic responses, and at least in vertebrates, hypoxia and iron metabolism are tightly linked[43]. Perhaps most intriguingly, Nrf2, a transcription factor controlling mitochondrial biogenesis and important iron metabolism genes, was shown to bind directly to the GBE1 promoter [154,155], raising the idea that this glycogen enzyme is coordinately controlled with other key iron genes. Taken together, our findings strongly suggest that the disease etiology of GSD IV needs to be re-assessed from the perspective that GBE1 has a key role in cellular iron homeostasis, and that there must be a re-evaluation of current therapeutic strategies in the future.

The loss-of-function phenotypes for Cisd2/mitoNEET, AGBE and IRP1A are very similar, since they display protoporphyrin accumulation that disappears under iron-replete conditions. This is consistent with the idea that depleting Cisd2/mitoNEET or AGBE equates the loss of IRP1A function, as both appear to act in concert to replace damaged Fe-S clusters in IRP1A. We have shown that IRP1A is an essential protein required for responding to low dietary iron levels, but *IRP1A* null mutants survive on an iron-rich diet. Therefore, our data strongly suggests that Cisd2/AGBE are gatekeepers that ensure proper functioning of IRP1A, a function that becomes non-essential in iron-replete conditions. Vertebrates encode three mitoNEET-like proteins, Cisd1-3[169]. *Drosophila* lacks a direct Cisd1 orthologue, but harbours copies of *Cisd2* and *CG3420* (*Cisd3*). Of the two, Cisd2 is more similar to Naf-1 and mitoNEET [189]. We were unable to identify any defects when disrupting *CG3420* function via RNAi, suggesting that Cisd2 is the functional equivalent of mitoNEET in *Drosophila*. The fact that a) fly Cisd2 interacts physically and genetically with IRP1A, and b) that mutations in either gene resulted in comparable phenotypes strongly supports the notion that IRP1A function depends on Cisd2, consistent with the finding in vertebrates that mitoNEET is involved in repairing oxidatively damaged Fe-S clusters. Similar to *Drosophila* IRP1A, null mutations of mouse *IRP1* or *IRP2* are non-lethal under normal conditions, however, the double knockout is embryonic lethal [190–194]. *IRP1* null mutants exhibit increased blood hemoglobin levels (polycythemia) [190,192,194] and one lab reported [192] that these mice developed also pulmonary hypertension that was exacerbated by exposure to a low iron diet, causing premature death.

The existing parallels between vertebrate IRP1 and *Drosophila* IRP1A raise the interesting question as to whether vertebrate IRP1 has a nuclear role as well, and whether it is conceivable that such a function has been hitherto overlooked. Consistent with this idea, a search of a human

protein-protein interaction database [161,162] found that IRP1 interacts with Histone 2Ab [195] (out of 19 reported proteins in total). In addition, we found that the presence of IRP1A in nuclei varies with tissue and nutritional conditions, raising the possibility that nuclear translocation occurs only under certain circumstances. This may be controlled by physiological parameters, depending on whether a tissue has high or normal iron requirements, and may be temporally regulated during development, as is the case for the PG. Further, we showed that the vertebrate IRP proteins use the same principles as their *Drosophila* counterparts for nuclear entry, since only human IRP1 has the ability to translocate to *Drosophila* nuclei, while IRP2, which lacks an Fe-S cluster, does not. It should also be noted that our findings were aided by the fact that the tissues we investigated are polytene, and consequently harbour, compared to most human cells, very large nuclei that allow easy visualization of nuclear proteins. Finally, we searched the literature for studies that had examined the subcellular localisation of IRP1 in more detail. To the best of our knowledge, the existing data relies solely on cell culture experiments with SW1088 and HepG2 cells, which reported IRP1 to mainly reside in the cytosol, but also found IRP1 associated with the endoplasmic reticulum and the Golgi apparatus [152,196]. While the effects of hypoxic and iron-deprived conditions on IRP1 localisation were tested, iron-rich conditions were not. Taken together, we believe that IRP1, at least in certain circumstances, behaves like its *Drosophila* counterpart, and enters nuclei where it is physiologically relevant. Future studies will have to revisit this issue in vertebrates.

3.5   Figures
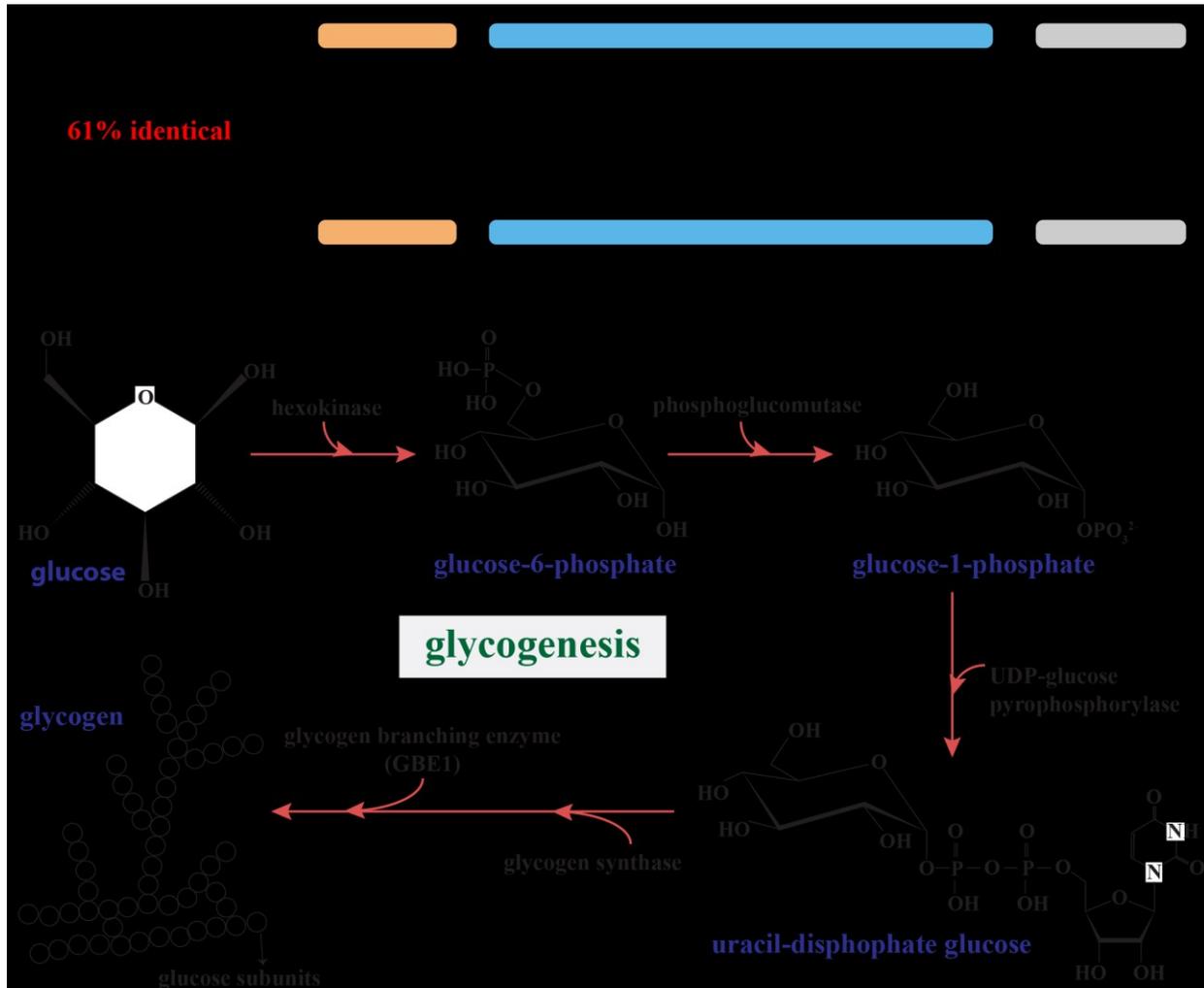
Figure 3.1 *Drosophila* AGBE is an ortholog of human GBE1.

A. *Drosophila* 1,4-alpha-glucan branching enzyme (AGBE) is 61% identical to human glycogen branching enzyme (GBE1). There is only a copy of the gene in each species with very similar domain architecture. **B.** The function of human GBE1 in glycogenesis. During glycogen synthesis, GBE1 catalyzes the final step to add side branches to the newly synthesized glycogen molecule. Adding side branches not only increases solubility of the molecule but also allows it to be easily broken down whenever needed.

C450S

R793Q   R554QR549Q

C450S

3R3Q

77

Figure 3.2 CRISPR/Cas9-mediated knock-in lines.

Three genes were targeted, *AGBE*, *IRP1A*, and *IRP1B*, each of which encodes a single mRNA isoform, respectively. Allele naming: WT: wild type; FCF: FRT/CRISPR/3xFlag; FCM: FRT/CRISPR/3xMyc. KO: knockout. 3F: 3xFlag. Note: Homozygous $IRP1A^{FCF}$ flies die on normal fly medium, but are viable when the diet is supplemented with iron, indicating that the FRT site insertions disrupt IRP1A function. Excision of *IRP1A* via prothoracic gland-specific expression of Flippase (FLP) caused protoporphyrin accumulation when second instar larvae were transferred from iron- supplemented to iron-depleted food.
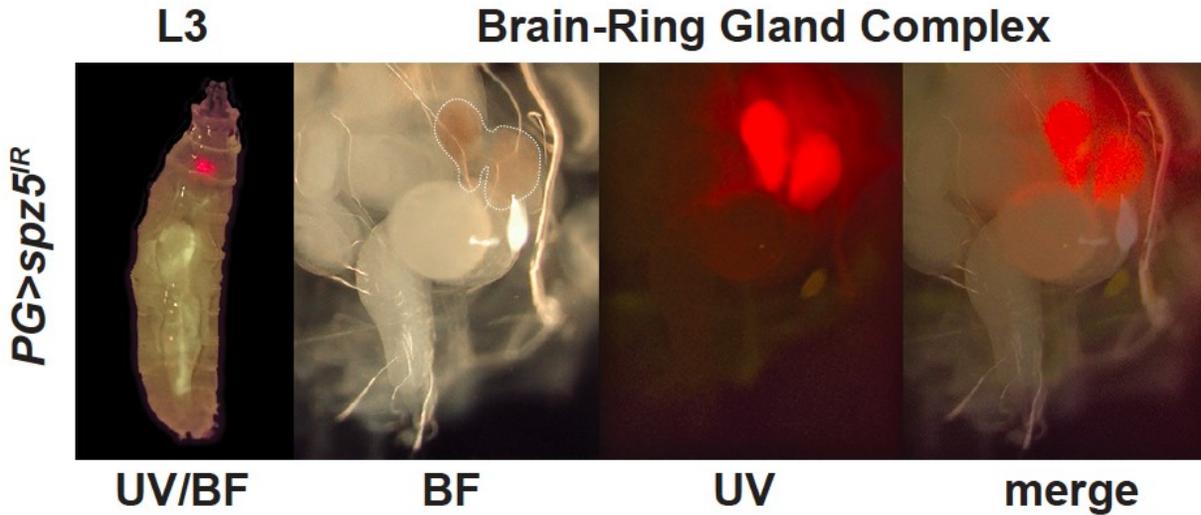
Figure 3.3 Porphyria-like phenotype in the *Drosophila* prothoracic gland.

Red autofluorescence is visible in whole larvae when exposed to ultraviolet (UV) light (left). Dissection shows an enlarged ring gland of red-brownish colour (dotted line) in brightfield (BF) light. The three glands that comprise the ring gland are not discernible in this image, but RNAi expression is limited to the prothoracic gland (PG). "*PG>*" refers to the "*phm22-Gal4*" driver, which mediates PG-specific expression. Pictures show results for VDRC line #102389 targeting the *spatzle5* gene (*spz5*^{IR}). UV exposure shows autofluorescence caused by accumulated protoporphyrins. Image credit: Ou Qiuxiang.
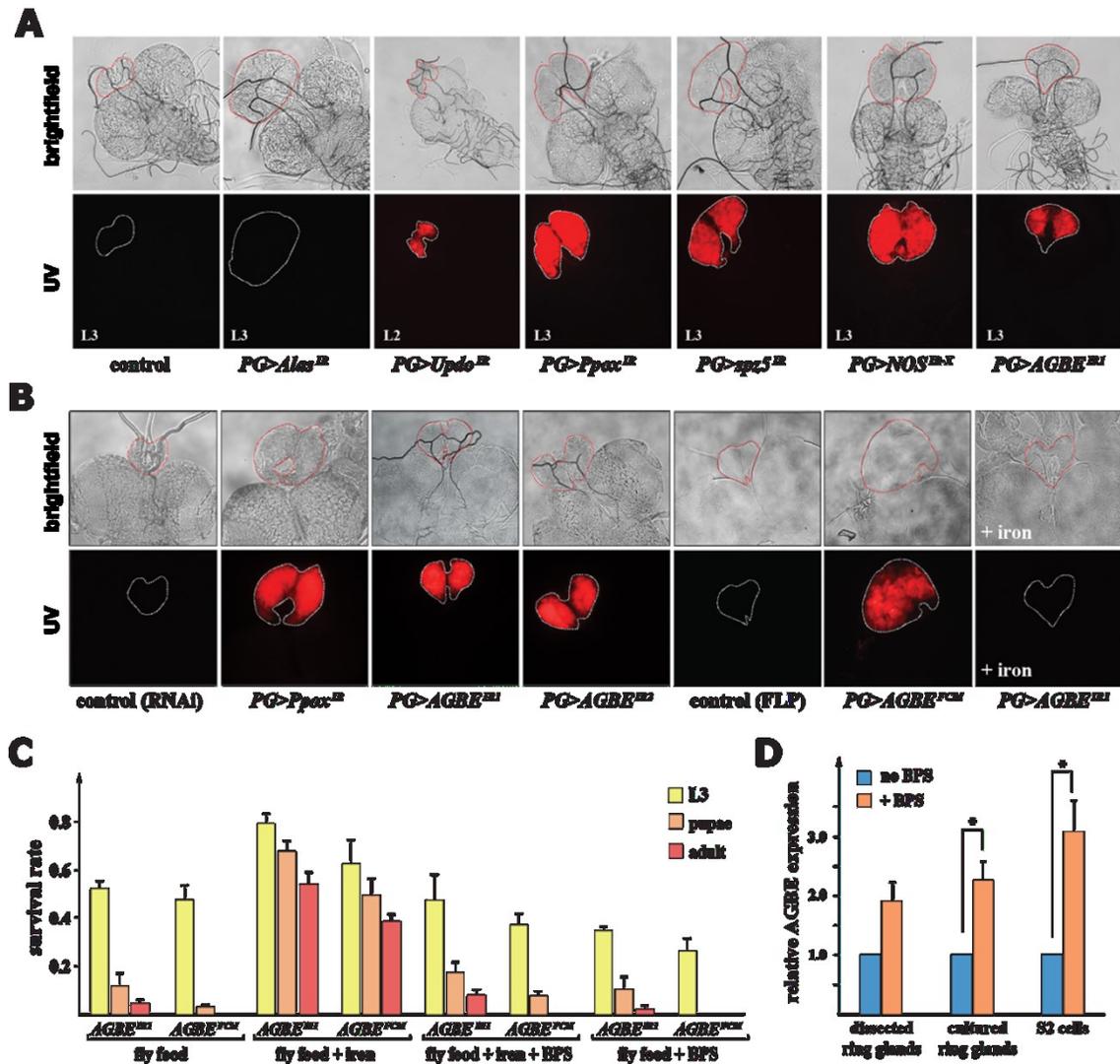
Figure 3.4 Disruption of heme biosynthesis in the *Drosophila* prothoracic gland (PG).

A. UV exposure of dissected ring glands from RNAi lines (designated as gene[IR]) from second (L2) or third (L3) instar stages. *Alas*, *Updo*, and *Ppox* encode heme-synthesizing enzymes. spz5: spaetzle5, Nos: nitric oxide synthase, AGBE: 1,4-Alpha-Glucan Branching Enzyme. Scale bar = 250 μm. **B.** UV exposure of dissected ring glands isolated at 40 h after the L2/L3 moult (~8 h prior to pupariation in controls). RNAi lines *AGBE*[IR1] and *AGBE*[IR2] target distinct regions of the AGBE mRNA. *AGBE*[FCM] is a conditional CRISPR-knock-in allele that can be excised in a tissue-specific

manner via the expression of Flippase (FLP) recombinase. + iron: larvae were reared on a diet containing ferric ammonium citrate (FAC) as an iron supplement. Scale bar = 250 μm. **C.** Survival of $AGBE^{IR1}$ and $AGBE^{FCM}$ larvae fly food supplemented with iron (FAC) or an iron chelator, bathophenanthroline sulfate (BPS). Error bars represent standard deviation. Three biological replicates, with each sample containing 50 individuals. **D.** Relative AGBE mRNA expression levels. Dissected ring glands: isolated from L3 reared on media ± BPS. Cultured ring glands: isolated from L3 reared on normal media, but then transferred to buffer containing ± BPS. S2 cells: Schneider 2 cells grown on medium ± BPS. mRNA levels were analysed via quantitative real-time PCR. Asterisk indicates a P-value < 0.05 based on the student's t test. Error bars represent 95% confidence intervals. Each of the three biological replicates was tested three times.

pathway & metabolites	vertebrate enzyme	fly enzyme	PG phenotype	WB phenotype
glucose				
⋯ hexokinase	hexokinase	Hex A (CG3001)	10-20% pupae lethal	embryonic lethal
glucose-6-phosphate				
⋯ phospho glucomutase	phospho glucomutase	PGM-1 (CG5165)	normal	embryonic lethal
glucose-1-phosphate				
⋯ UDP-glucose pyrophosphorylase	UDP-glucose pyrophosphorylase	UGP (CG4347)	normal	L2 lethal
uracil-disphophate glucose				
⋯ glycogen synthase	glycogen synthase	GlyS (CG6904)	normal	L2 & L3 lethal
⋯ glycogen branching enzyme (GBE1)	glycogen branching enzyme (GBE1)	AGBE (CG33138)	L3 arrested porphyria PG	embryonic & L1 lethal
glycogen				

Figure 3.5 Phenotypes associated with prothoracic gland -specific and ubiquitous depletion of enzymes acting in the glycogen biosynthetic pathway of *Drosophila*.

Names for vertebrates and *Drosophila* enzyme orthologues are shown in blue. PG = prothoracic gland-specific expression (*phm22-Gal4 x UAS*-RNAi). WB = whole body expression (*tubulin-Gal4 x UAS*-RNAi). Red indicates the appearance of red autofluorescence in the *Drosophila* prothoracic gland.
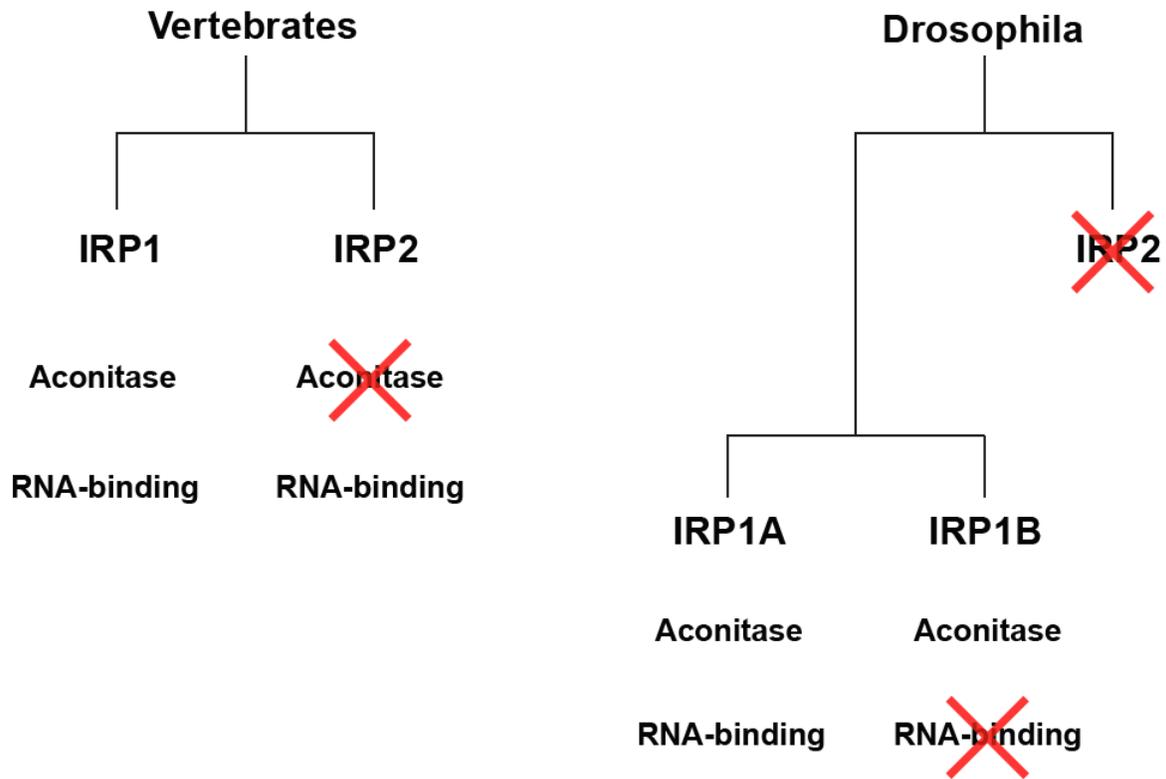
Figure 3.6 Comparison of genes encoding iron-regulatory proteins (IRPs) between vertebrates and *Drosophila melanogaster*.

Flies lack IRP2 but have two IRP1 genes. Only IRP1A was shown to bind to canonical iron responsive elements (IREs) in target mRNAs.

A

AGBE ↔ IRP1A   AGBE ↔ IRP1A^C450S   AGBE ↔ IRP1A^3R3Q

eGFP^M ↔ IRP1A   eGFP^M ↔ IRP1A^C450S   eGFP^M ↔ IRP1A^3R3Q

B

hGBE1 ↔ hIRP1

eGFP^M ↔ hIRP1

C

relative *SdhB* mRNA binding

1.0   93.5   393.0   2.5

w^1118, IRP1A^3F, IRP1A^C450S-3F, IRP1A^3R3Q-3F

D

relative aconitase activity

IRP1A^3F, IRP1A^C450S-3F, IRP1A^3R3Q-3F

IRP1B^-/-

E

survival rate

L3   Pupae   Adult

AGBE^FCM   AGBE^FCM IRP1A   AGBE^FCM IRP1A^C450S

F

control   IRP1A   IRP1A^C450S   AGBE^FCM   AGBE^FCM IRP1A   AGBE^FCM IRP1A^C450S   Nos^IR-X   Nos^IR-X IRP1A   Nos^IR-X IRP1A^C450S

PG-specific expression of FLP (*AGBE^FCM*), cDNA (*IRP1A* and *IRP1A^C450S*), dsRNA (*Nos^IR-X*)

G

Brightfield

UV

AGBE^IR1   AGBE^IR1 IRP1A   AGBE^IR1 IRP1A^C450S   AGBE^IR1 IRP1A^3R3Q   AGBE^IR1 YAco1^WT   AGBE^IR1 YAco1^ΔSp   AGBE^IR1 hIRP1   AGBE^IR1 hIRP2

84

Figure 3.7 AGBE is in a protein complex with IRP1A

A. Co-transfection of S2 cells with plasmids encoding Flag-tagged IRP1A variants (IRP1A*) and Myc-tagged AGBE followed by immunoprecipitation via anti-Myc antibodies and western blotting. Names shown in red indicate the protein used as bait. IRP1A: wild type IRP1A, IRP1A^{C450S}: constitutively RNA-binding IRP1A, IRP1A^{3R3Q}: non-RNA-binding form of IRP1A (Table 3.1). Myc-tagged enhanced GFP (eGFP^M) served as a negative control. Input lane represents 10% of the sample. Presence of co-immunoprecipitated proteins were tested with anti-Flag antibodies. **B.** Like A, but co-transfection of S2 cells with plasmids encoding Flag-tagged human IRP1 (aka Aco1) and Myc-tagged human GBE1, as well as eGFP^M as a negative control. **C.** Quantitative RNA-immunoprecipitation (RIP). Samples from larvae carrying Flag-tagged knock-in alleles of *IRP1A* (*IRP1A^{3F}*, *IRP1A$^{C450S.3F}$*, and *IRP1A$^{3R3Q.3F}$*) (Figure 3.2) were normalized via Western blotting to visualize Flag-tagged proteins followed by ImageJ quantification. Western blot of adjusted samples shown below graph. Untagged IRP1A (control line *w^{1118}*) served as a negative control and calibrator (normalized expression = 1). *SdhB* mRNA harbours a validated IRE[197,198]. Co-immunoprecipitated *SdhB* mRNA was quantified via qPCR. Error bars represent 95% confidence intervals. *** = p<0.001, ** = p<0.01, * p<0.05. **D.** Aconitase activity. Same *IRP1A* alleles and normalization procedure as described in C, except that *IRP1A^{3F}* served as the control (normalized to 1). All alleles were crossed into an *IRP1B$^{-/-}$* mutant background to eliminate the aconitase activity of IRP1B. Further, we removed mitochondria via ultracentrifugation to reduce the contribution of mitochondrial aconitase. Error bars represent standard deviation. ** = p<0.01, * p<0.05. **E.** Survival rates of *PG>FLP; AGBE^{FCM}* animals (Figure 3.2), which causes Flippase-mediated excision of the *AGBE* transcription unit specifically in the prothoracic gland (PG). Tested in either the presence or absence of the *IRP1A* and

IRP1A^C450S transgenes that are also expressed in a PG-specific manner. Error bars represent standard deviation. **F.** Larval and adult phenotypes of *PG>FLP*; *AGBE^FCM* and *PG>Nos^IR-X* animals expressing *IRP1A^C450S* or wild type *IRP1A* transgenes. Arrows point to red-stained PG. **G.** Ring glands dissected from *PG>AGBE^IR1* larvae in the presence or absence of the following transgenic cDNAs: *IRP1A* (wild type *IRP1A*); *IRP1A^C450S* (constitutively RNA-binding); *IRP1A^3R3Q* (non-RNA-binding); *YAco1^WT*: wild type yeast aconitase (mitochondrial); *YAco1^ΔSp* (cytoplasmic); *hIRP1 & hIRP2*: human IRP1 & IRP2. All transgenes are expressed in a PG-specific manner via the Gal4-UAS system.
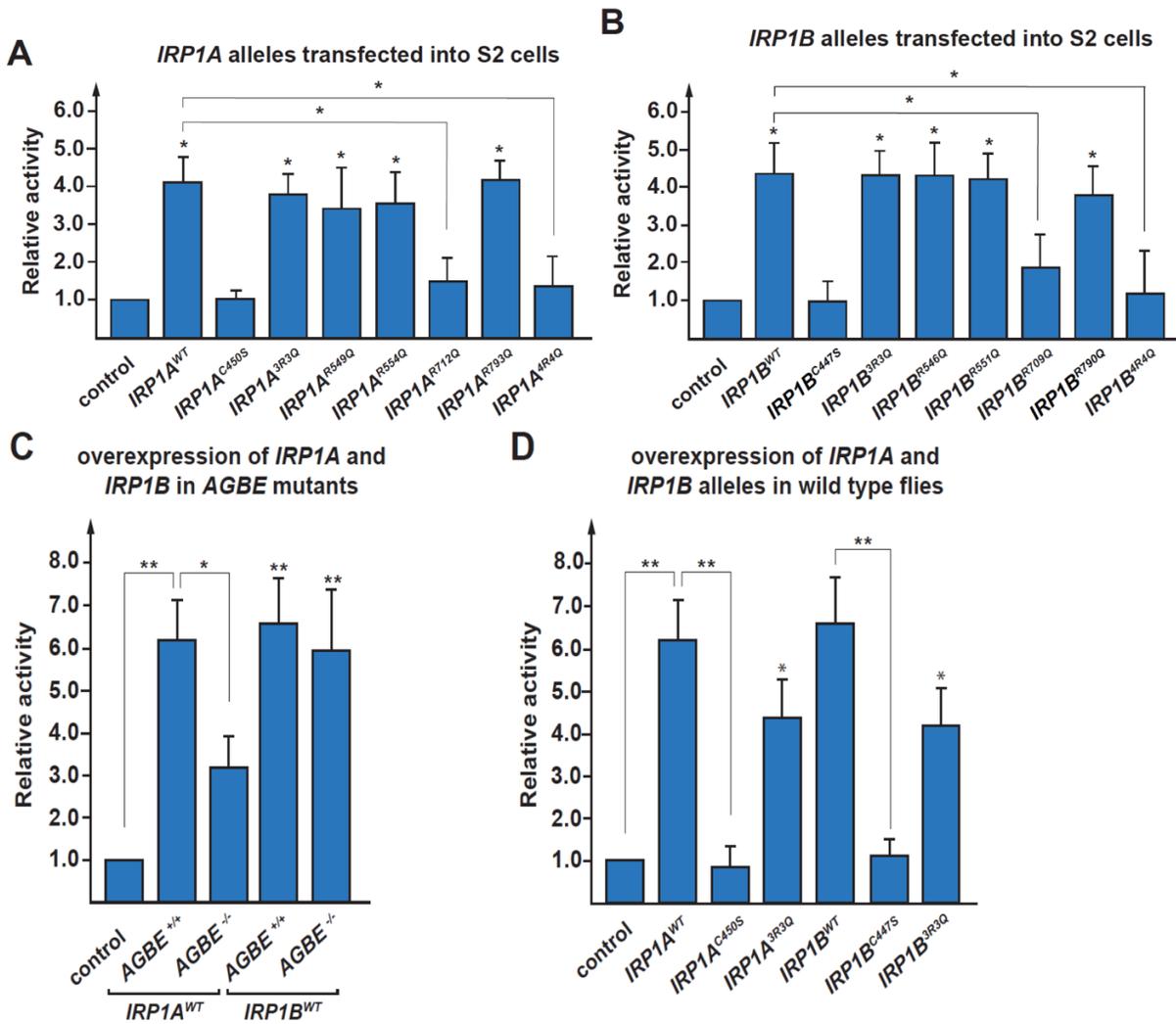
Figure 3.8 Aconitase activity of IRP1 variants.

A-B. Relative aconitase activity of IRP1A (A) and IRP1B (B) variants in S2 cell culture. Results were normalized to untransfected S2 cells. Transfection efficiency was evaluated by SDS-PAGE to normalize protein levels for aconitase assay. IRP1A^{3R3Q}: non-RNA binding (Figure 3.7C), has amino acid substitutions (R→Q) in positions 549, 554 and 793. IRP1A^{4R4Q}: predicted to be non-RNA binding, has amino acid substitutions (R→Q) in positions 549, 554, 712 and 793. IRP1A^{C450S}: constitutively RNA-binding, has amino acid substitution (C→S) in position 450. The IRP1B^{3R3Q}, IRP1B^{4R4Q} and IRP1B^{C447S} variants have corresponding substitutions in IRP1B. C.

Relative aconitase activity of whole body extracts from first instar larvae (L1) that ubiquitously expressed *IRP1A* or *IRP1B* in either wild type or *AGBE^{FCF}* (Figure 3.2) mutant backgrounds. Ubiquitous excision of *AGBE* results in L1 lethality. Data were normalized to control animals that lack *IRP1* transgenes. D. Relative aconitase activity of whole-body extracts from L1 that ubiquitously expressed one of the following transgenes in a wild type background: $IRP1A^{WT}$, $IRP1A^{C450S}$, $IRP1A^{3R3Q}$, $IRP1B^{WT}$, $IRP1B^{C447S}$ and $IRP1B^{3R3Q}$. For allele properties, see A-B. A-D. Asterisks indicate a *P*-value <0.05 (*) or <0.01 (**) relative to the control, or relative to the indicated reference sample.
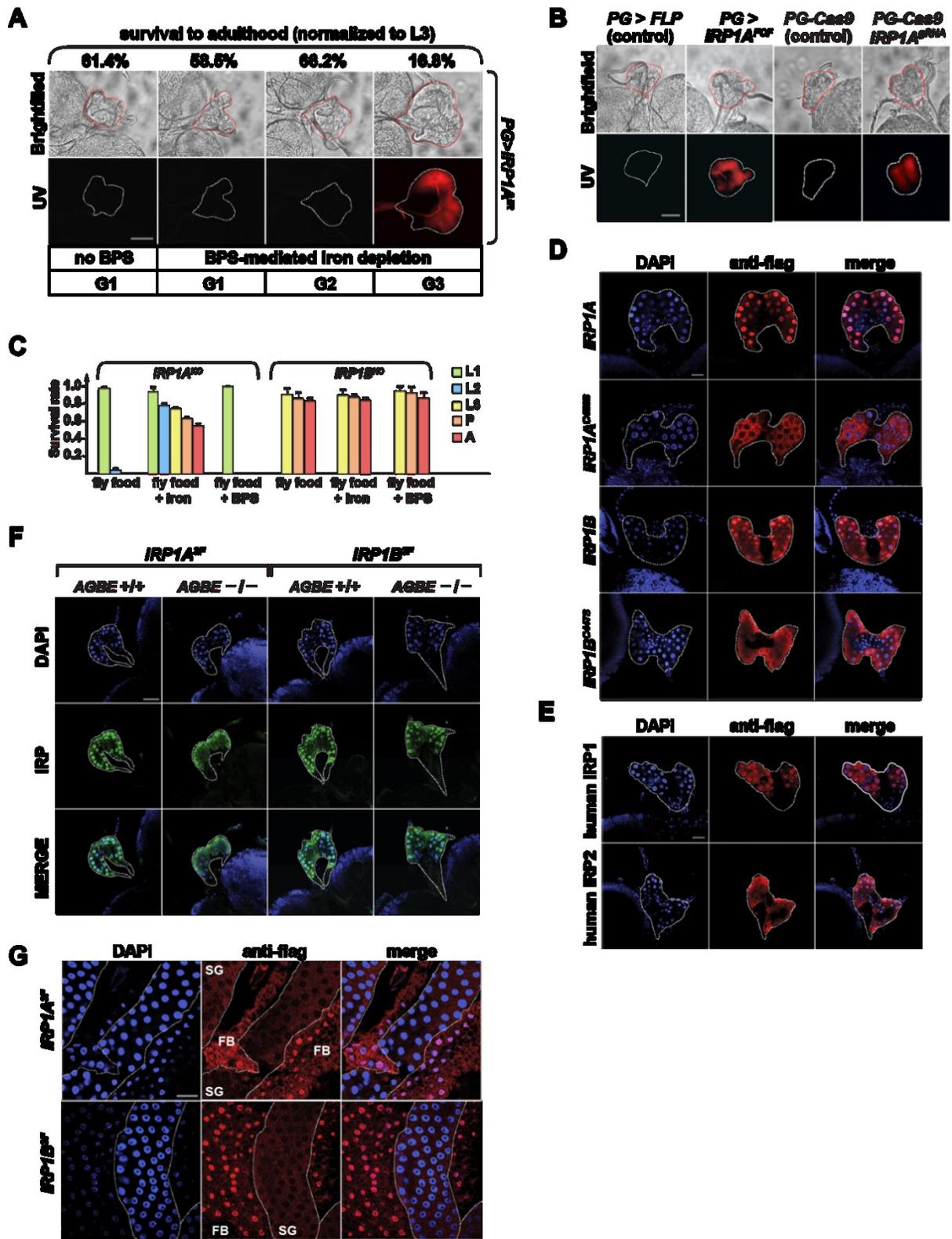
Figure 3.9 IRP1 localizes to nuclei.

A. Protoporphyrin accumulation/autofluorescence in prothoracic glands (PG) from *IRP1A*-RNAi (*IRP1A^{IR}*) flies reared on iron-depleted (= BPS) media for three generations. Adult survival relative to last larval stage (surviving third instar larvae = L3 = 100%). B. Protoporphyrin accumulation/autofluorescence in PGs from *IRP1A^{FCF}* animals (tissue-specific excision of *IRP1A*, Figure 3.2) reared on iron-rich medium until L2, after which larvae were switched to BPS-supplemented food. C. Survival of *IRP1A* and *IRP1B* null mutants (KO = knockout, Figure 3.2). D. Subcellular localization of PG-specific, Flag-tagged IRP1A and IRP1B (*PG>IRP1A / PG>IRP1B* transgenic lines, Table 3.1). DAPI was used to stain DNA/nuclei. E. Subcellular localization of Flag-tagged transgenic human IRP1 and IRP2 (*PG>hIRP1* and *PG>hIRP2*, Table 3.1) expressed specifically in the PG. F. Subcellular localization of Flag-tagged proteins encoded by *IRP1A^{3F}* and *IRP1B^{3F}* knock-in alleles (Figure 3.2) in control or *AGBE* mutant backgrounds (*AGBE$^{+/+}$* = *PG>FLP*. *AGBE$^{-/-}$* = *PG>FLP*; *AGBE^{FCM}*, Figure 3.2). G. Subcellular localization of Flag-tagged proteins encoded by *IRP1A^{3F}* and *IRP1B^{3F}* knock-in alleles (Figure 3.2) in the fat body (FB) and salivary gland (SG).
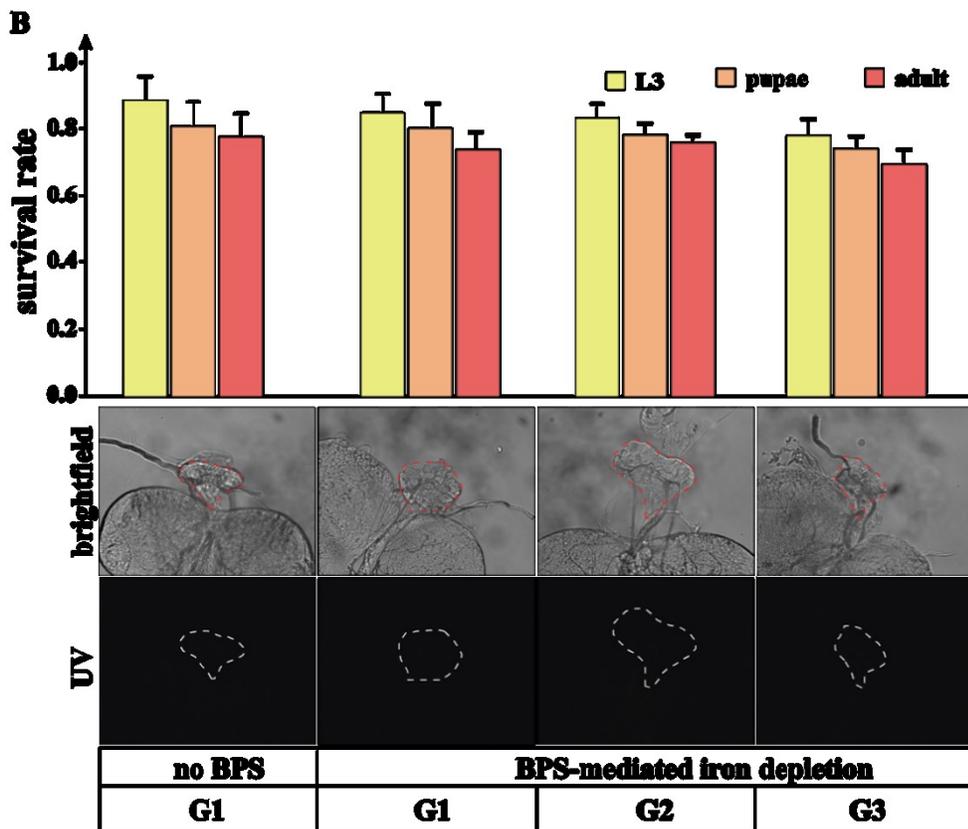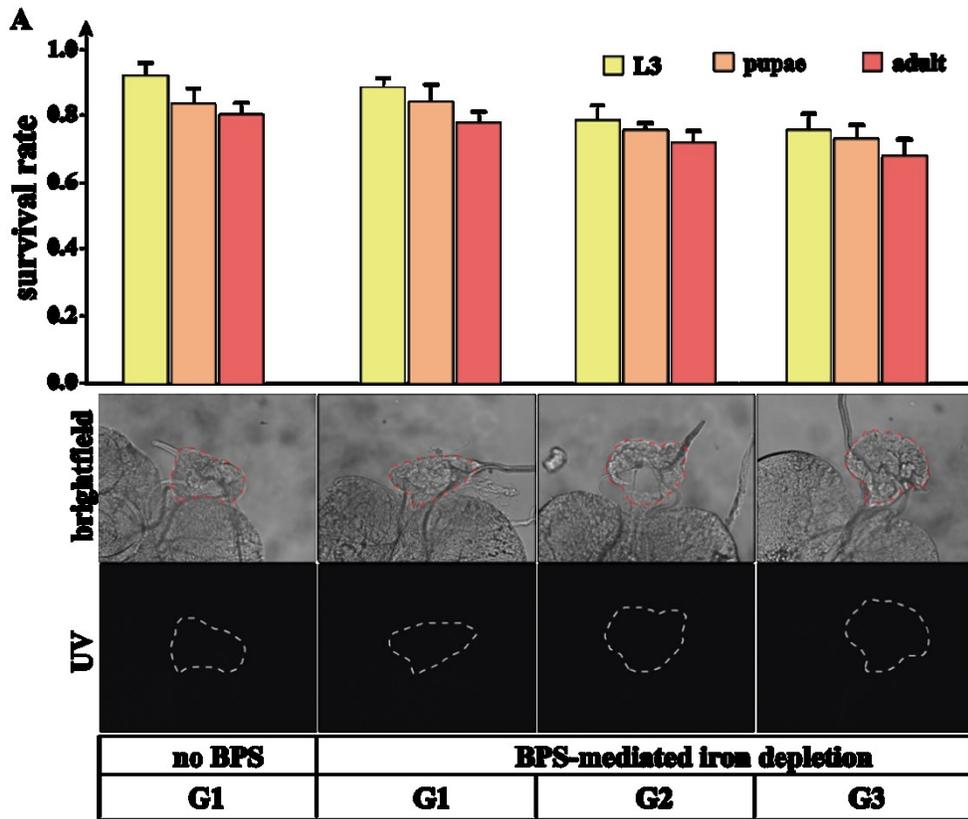
Figure 3.10 Ring gland phenotypes in control and *IRP1B^{IR}* lines are normal when raised for three generations on iron-depleted media.

Control (*w^{1118}*) (A) and PG-specific *IRP1B*-RNAi (*PG>IRP1B^{IR}*) animals (B) were kept continuously on Bathophenanthroline Sulfate (BPS)-containing food to deplete cellular iron stores over three consecutive generations. Third instar (yellow), pupal (orange) and adult (red) survival was scored for each generation. Ring glands were dissected during the L3 stage and examined under brightfield and UV light.
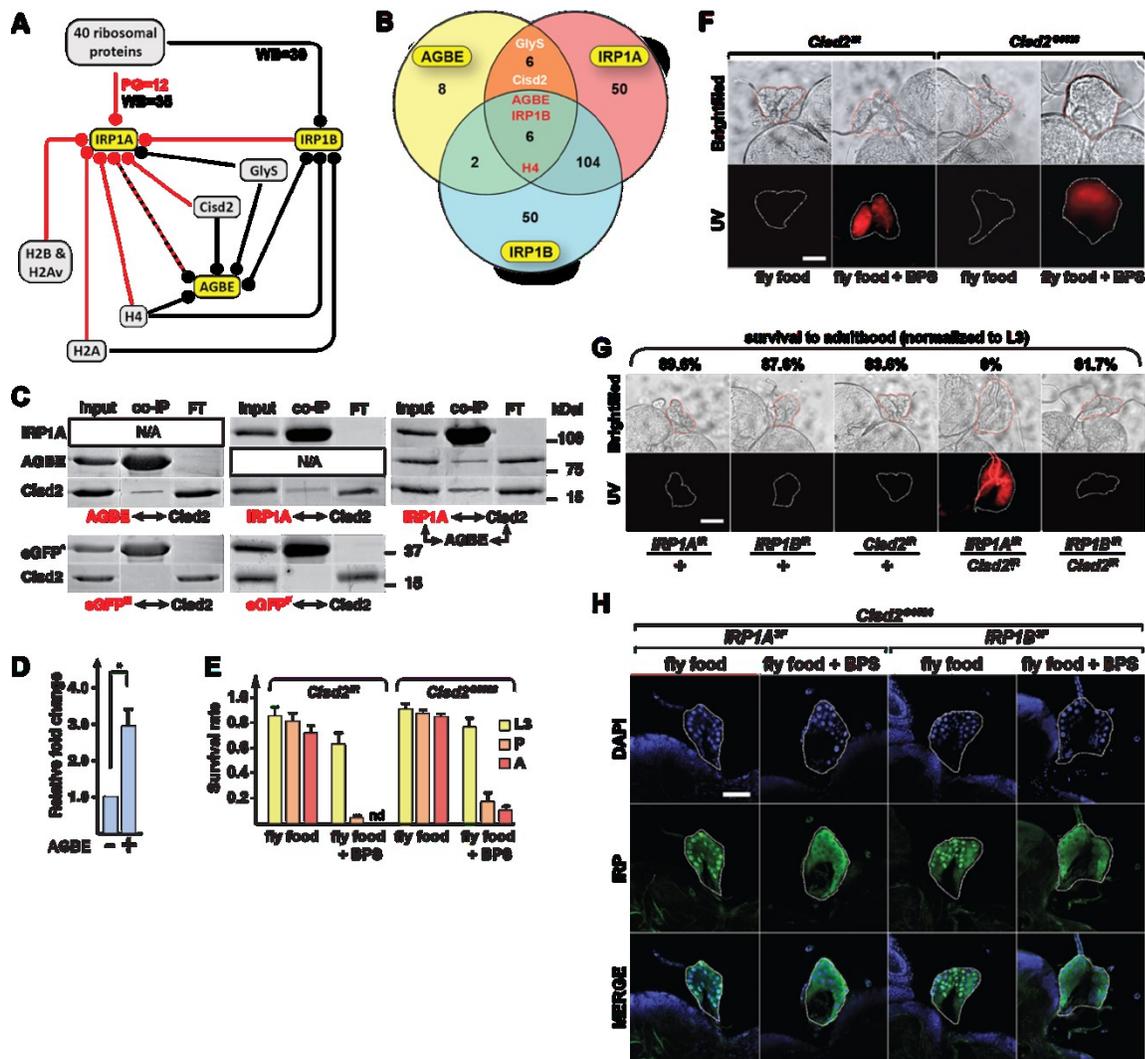
Figure 3.11 Cisd2 interacts with IRP1A and AGBE.

A. Protein-protein interaction map. Lines carrying knock-in alleles encoding Flag-tagged AGBE, IRP1A and IRP1B (yellow boxes, Figure 3.2) were used to produce bait (circle) for immunoprecipitation followed by mass spectrometry (MS) to identify physically bound proteins to the bait. Whole-body (WB, black) and prothoracic gland samples (PG) were used. Red: detected in both PG and WB samples. Dashed line: Only WB samples were tested for AGBE. H2Av, H2A, H2B and H4 are histone proteins. GlyS = Glycogen Synthase. **B.** Venn diagram depicting overlaps of immunoprecipitated proteins from endogenously tagged proteins (WB samples). H4 & GlyS

93

see A. **C.** Co-transfection of Schneider 2 cells with plasmids encoding Myc-tagged AGBE, Flag-tagged IRP1A and HA-tagged Cisd2, followed by immunoprecipitation via anti-Myc or anti-Flag antibodies and Western blotting. Names shown in red indicate the protein used as bait. Myc-tagged and Flag-tagged enhanced GFP (eGFP^M and eGFP^F, respectively) served as negative controls. Presence of co-immunoprecipitated proteins were tested with anti-HA antibodies and anti-Myc antibodies. **D.** Quantification of immunoprecipitated Cisd2 in the triple co-transfection experiment shown above in C. Graph shows relative fold change of co-immunoprecipitated Cisd2 with Flag-IRP1A as bait in the presence or absence of AGBE. Data was normalized to the amount of Cisd2 protein in the absence of co-transfected AGBE. The asterisk indicates a *P*-value < 0.05 according to the Student's *t*-test. Error bars represent standard deviation. **E.** Survival rates of *Cisd2^{IR}*-RNAi animals and *Cisd2^{G6528}* mutants on fly food ± BPS. nd = not detected. Error bars represent standard deviation. **F.** Autofluorescence/protoporphyrin accumulation in prothoracic glands (PG) of *PG>Cisd2^{IR}* and *Cisd2^{G6528}* larvae reared on fly food ± BPS. **G.** Genetic interaction between *Cisd2* and *IRP1A* on regular (= iron-replete) fly food based on autofluorescing PGs and survival of the corresponding RNAi lines. All lines express RNAi via a PG-specific Gal4 driver (*phm22-Gal4* = *PG>*). **H.** Subcellular localization of Flag-tagged IRP1A and IRP1B proteins expressed from knock-in alleles (Table 3.1) in *Cisd2^{G6528}* mutants reared on fly food ± BPS. For control larvae, see Figure 3.10.
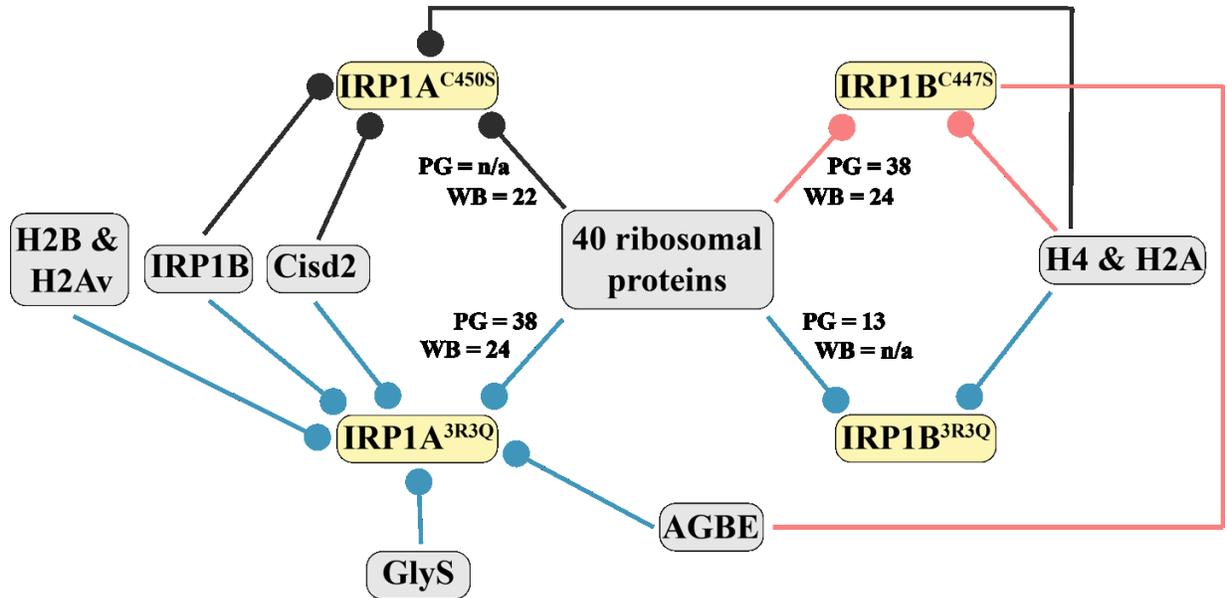
Figure 3.12 Interaction map for transgenic *IRP1A* and *IRP1B* lines

C450S: abolishes Fe-S cluster incorporation into IRP1A. 3R3Q: Replacement of three arginines implicated in RNA-binding (Table 3.1). Red: Interaction detected in prothoracic gland (PG) and whole body (WB) samples. Black: Detected in WB samples only. Blue: Tested in PG samples only. Yellow boxes represent transgenically expressed 3xFlag-tagged protein that was immunoprecipitated with anti-Flag antibodies. Circles next to yellow boxes indicate whether the interaction was reciprocal (circles on both ends) or not (circle on one end). In total, 40 ribosomal proteins were common to all tested PG and WB sets [166], numbers indicate how many were detected in either PG or WB samples. H4, H2A, H2Av and H2B are histone proteins. Cisd2 = mitoNEET, GlyS = Glycogen Synthase, AGBE = 1,4-Alpha-Glucan Branching Enzyme.

	no BPS	BPS-mediated iron depletion		
	G1	G1	G2	G3

96
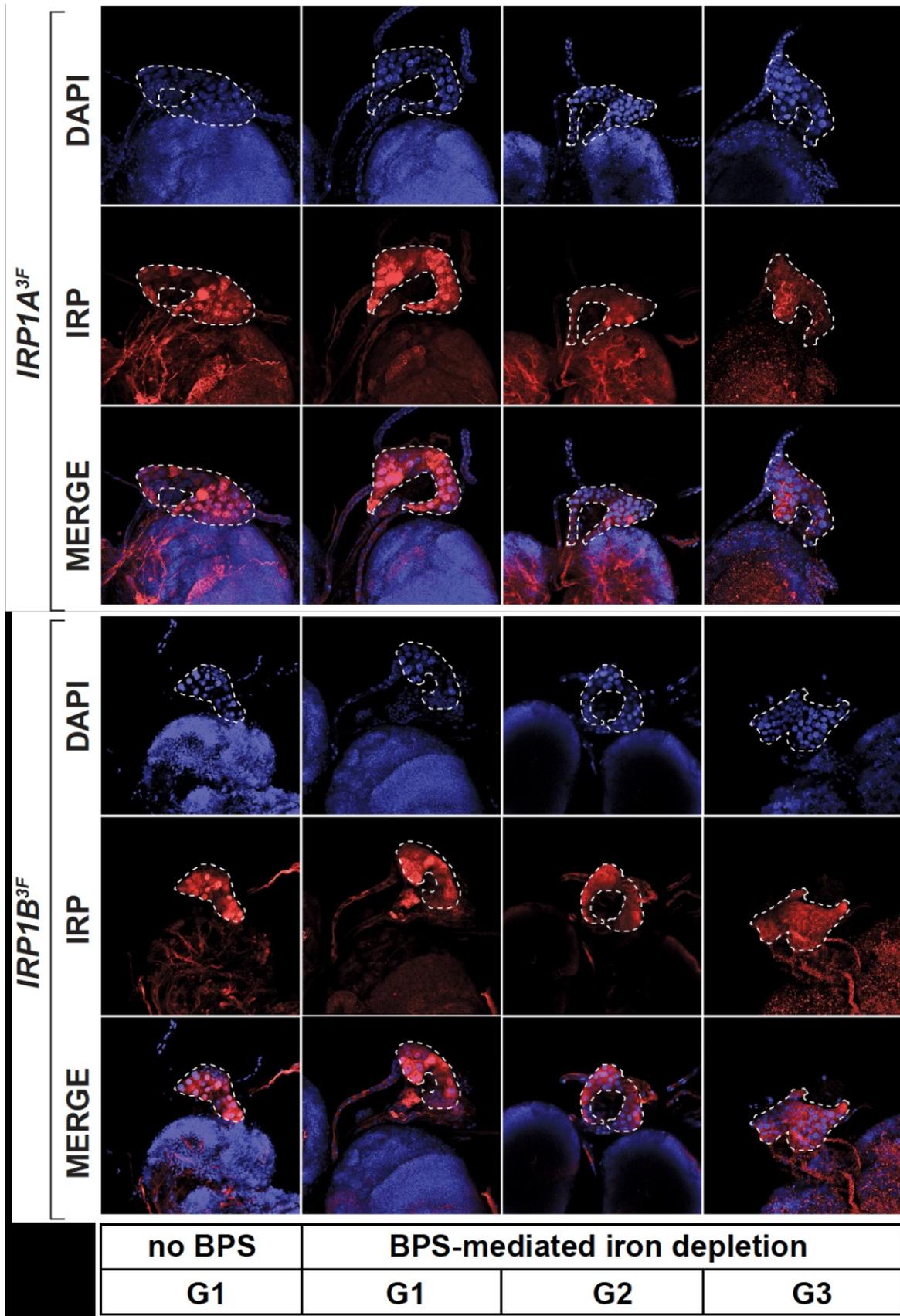
Figure 3.13 Nuclear localization of IRP1A and IRP1B in the prothoracic gland when reared under sustained iron-depleted conditions.

Flies were reared on Bathophenanthroline Sulfate (BPS)-supplemented media for three generations. Each generation (G1-G3), ring glands were stained for the subcellular localization of Flag-tagged proteins produced from knock-in alleles of *IRP1A* and *IRP1B* (*IRP1A^{3F}* and *IRP1B^{3F}*, Figure 3.2). DAPI was used to stain DNA/nuclei.

	no BPS	BPS-mediated iron depletion		
	G1	G1	G2	G3

98
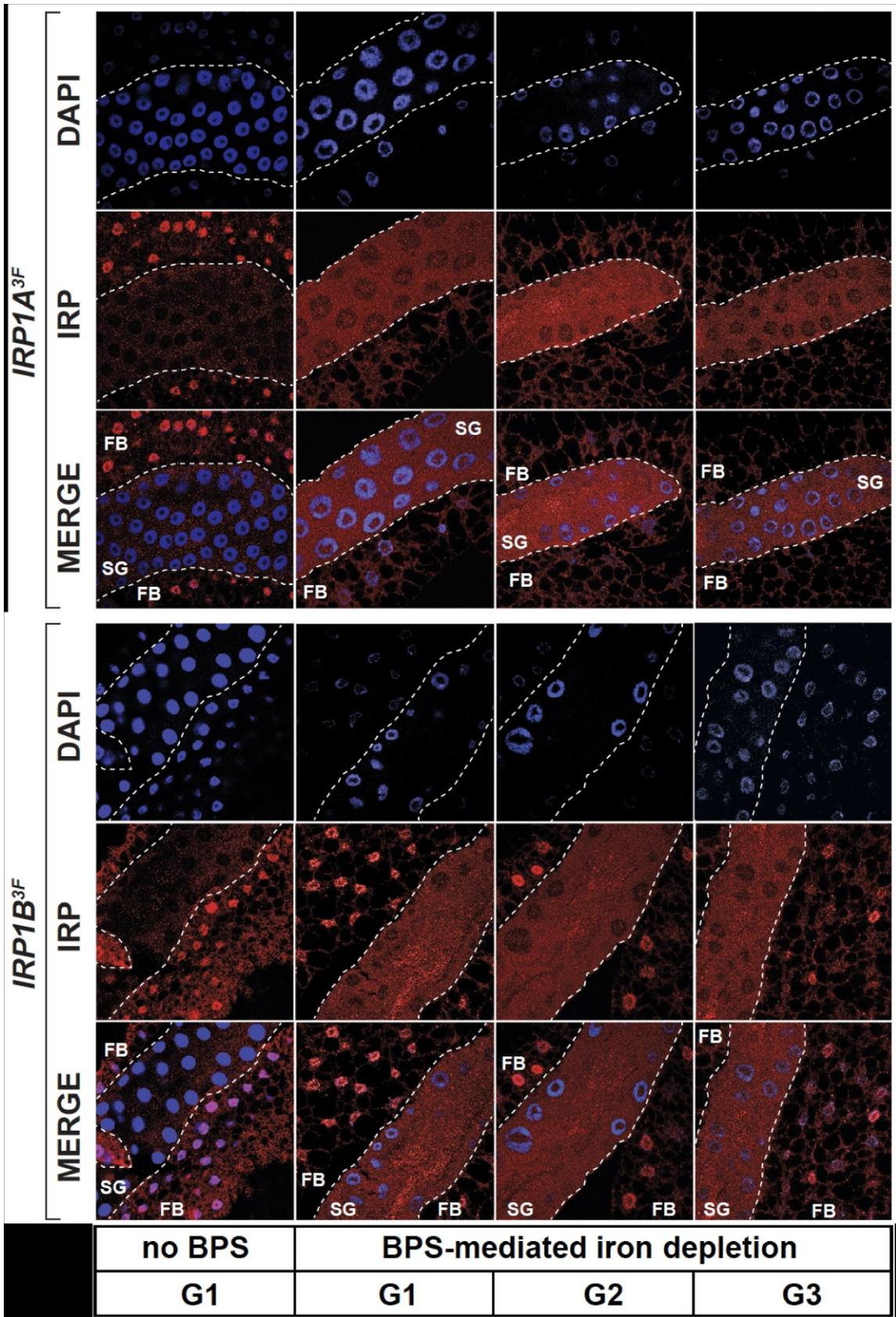
Figure 3.14 Nuclear localization of IRP1A and IRP1B in salivary gland and fat body tissue when reared under sustained iron-depleted conditions.

Flies were reared on Bathophenanthroline Sulfate (BPS)-supplemented media for three generations. Each generation (G1-G3), ring glands were stained for the subcellular localization of Flag-tagged proteins produced from knock-in alleles of *IRP1A* and *IRP1B* (*IRP1A^{3F}* and *IRP1B^{3F}*, Figure 3.2). DAPI was used to stain DNA/nuclei. SG: salivary gland. FB: fat body.
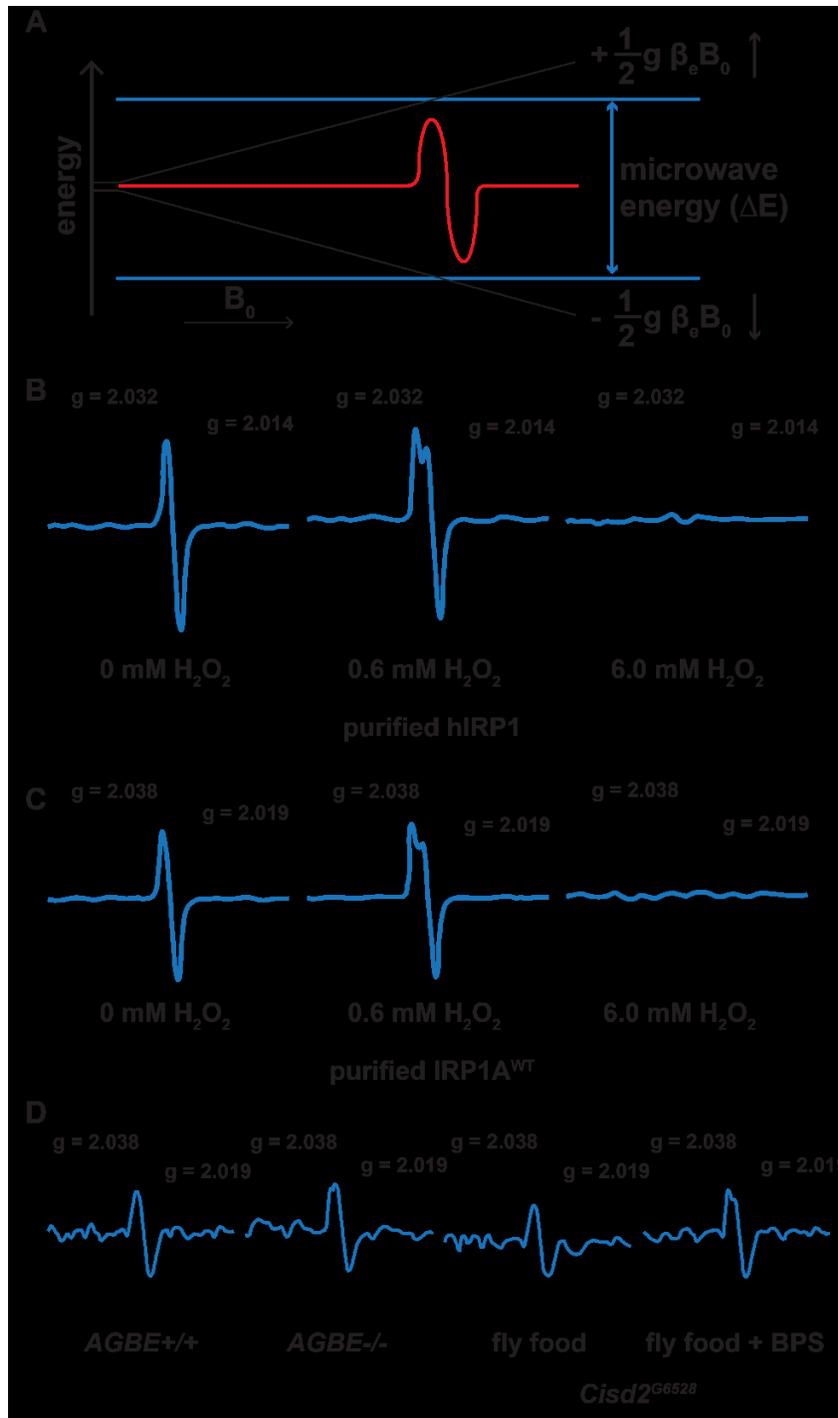
Figure 3.15 IRP1A requires AGBE and Cisd2 to maintain intact iron-sulfur cluster (ISC)

A. Principle of electron paramagnetic resonance (EPR). EPR is a magnetic resonance technique used to measure the transition of unpaired electrons. By exposing the samples in a fixed frequency

of microwave irradiation and looking for changes in the energy level, we can evaluate the status of target molecule. **B.** EPR of human IRP1 (hIRP1) purified from *Drosophila* S2 cells treated with hydrogen peroxide. **C.** EPR of *Drosophila* IRP1A purified from S2 cells treated with hydrogen peroxide. **D.** EPR of IRP1A from whole body lysate of *AGBE*$^{-/-}$ mutant and *Cisd2*^{G6528} mutant. All samples also carry and *IRP1B*^{KO} mutant background.
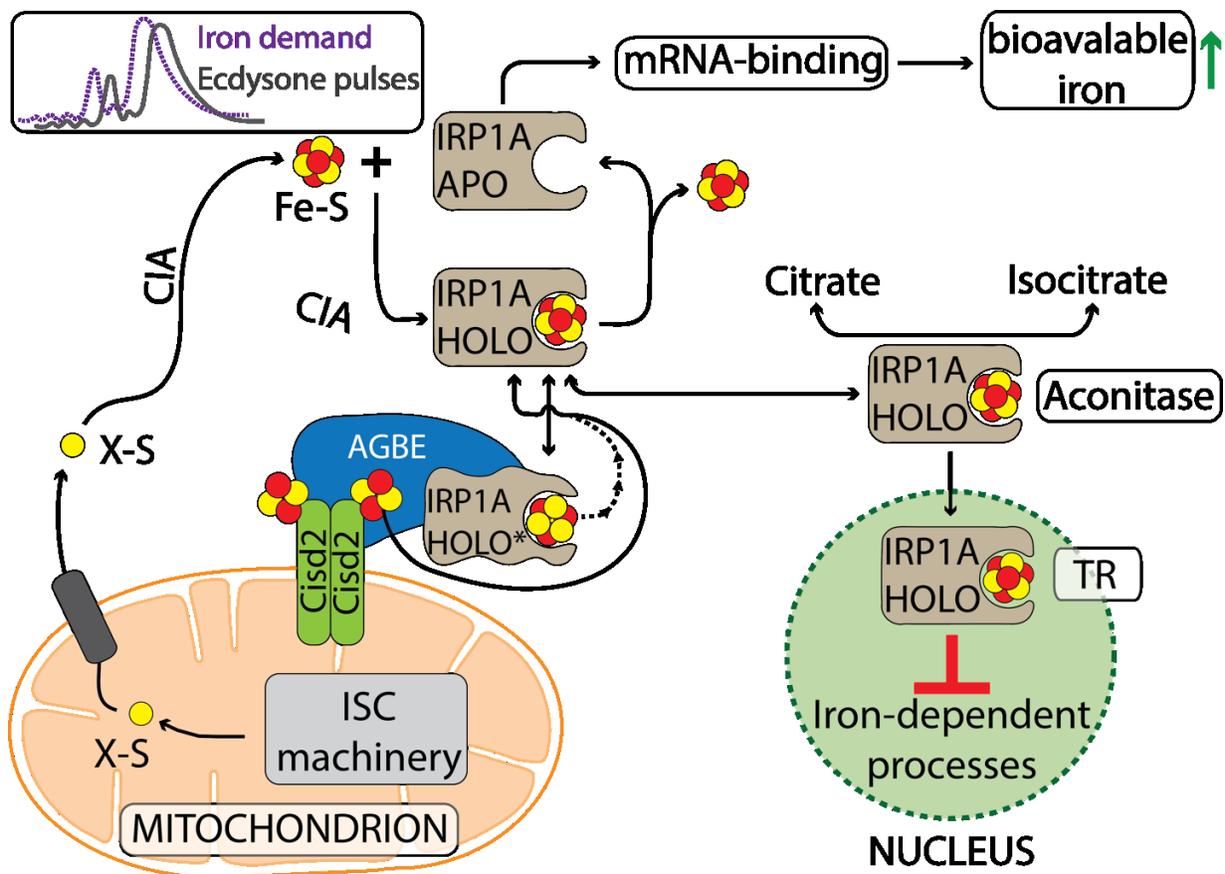
Figure 3.16 Model for AGBE-mitoNEET/Cisd2-IRP1A function in tissues with dynamic iron requirements.

In *Drosophila*, iron demand peaks prior to maximal ecdysone production to equip ecdysone-synthesizing enzymes with iron cofactors. Newly synthesized IRP1A receives Fe-S clusters (red circles) from the Cytosolic Iron-sulfur cluster Assembly (CIA) machinery, which produces [4Fe-4S] sulphur clusters from an unidentified mitochondrial precursor molecule, X-S (produced in mitochondria by ISC = Iron-Sulphur Cluster Assembly Machinery). Oxidatively damaged IRP1A (IRP1A-HOLO*) requires the mitoNeet/Cisd2 and AGBE proteins to replace impaired clusters with functional units. Holo-IRP1A is both needed for the aconitase function as well as nuclear entry. Loss of AGBE or mitoNEET/Cisd2 function results in a depletion of holo-IRP1A, and the concomitant loss of nuclear IRP1A, explaining why *AGBE* mutants cannot be rescued with the

IRP1A^{C450S} form, which only assumes the apo-form since it cannot incorporate an Fe-S cluster. The model suggests that once iron demand has peaked and nuclear holo-IRP1A levels become maximal, the protein acts to throttle expression of genes acting in steroid, heme and iron metabolism in anticipation of falling iron demands. As such, IRP1A has two functions: 1. As known from mammalian cells it acts in response to low cellular iron levels as an RNA-binding protein that promotes increased iron availability, and 2. as a new function, it responds to peak iron levels as a nuclear protein to promote downregulation of processes depending on iron and heme (TR = Transcriptional Regulation).

3.6 Tables

Table 3.1 Transgenic constructs and properties

transgene	description
UAS-3xFlag-IRP1A^{WT}	express wildtype 3xFlag tagged *IRP1A* cDNA under *Gal4/UAS* system
UAS-IRP1A^{WT}	express wildtype untagged *IRP1A* cDNA under *Gal4/UAS* system
UAS-3xFlag-IRP1A^{C450S}	expresses mutant 3xFlag-tagged *IRP1A^{C450S}* cDNA under *Gal4/UAS* control, single point mutation abolishes Fe-S-binding and forces protein to assume apo-form [136,164]. Predicted to abolish aconitase function and render protein constitutively RNA-binding.
UAS-IRP1A^{C450S}	expresses mutant untagged *IRP1A^{C450S}* cDNA under *Gal4/UAS* control, single point mutation abolishes Fe-S-binding and forces protein to assume apo-form [136,164]. Predicted to abolish aconitase function and render protein constitutively RNA-binding.
UAS-3xFlag-IRP1A^{3R3Q}	expresses mutant 3xFlag-tagged *IRP1A^{3R3Q}* cDNA under *Gal4/UAS* control, converting three arginine into three glutamine residues (R549Q, R554Q and R793Q). Predicted to result in non-RNA-binding holo-IRP1A [136,164].
dU6-3-IRP1A^{gRNA}	ubiquitously expresses two *IRP1A* gRNAs for somatic disruption when crossed to CAS9-expressing stock (PG-specific CAS9, causes red autofluorescence) [199].
UAS-3xFlag-IRP1B^{WT}	expresses wild type 3xFlag-tagged *IRP1B* cDNA under *Gal4/UAS* control.
UAS-IRP1B^{WT}	expresses wild type untagged *IRP1B* cDNA under *Gal4/UAS* control.
UAS-3xFlag-IRP1B^{C447S}	expresses mutant 3xFlag-tagged *IRP1B^{C447S}* cDNA under *Gal4/UAS* control, mutation at the same cysteine residue as in *IRP1A^{C450S}*.
UAS-IRP1B^{C447S}	expresses mutant untagged *IRP1B^{C447S}* cDNA under *Gal4/UAS* control, single point mutation affecting equivalent cysteine residue as in *IRP1A^{C450S}*. Predicted to abolish aconitase function.

UAS-3xFlag-IRP1B^{3R3Q}	expresses mutant 3xFlag-tagged *IRP1B^{3R3Q}* cDNA under *Gal4/UAS* control, point mutations equivalent to those in *IRP1A^{3R3Q}* . Predicted to interfere with RNA-binding, however, no RNA-binding has been reported for IRP1B.
UAS-Yeast Aco1^{WT}	expresses wild type untagged yeast *Aco1* cDNA under *Gal4/UAS* control, predicted to cause cytoplasmic and mitochondrial localization of the enzyme [165].
UAS-Yeast Aco1^{7Sp}	expresses mutant untagged yeast *Aco1* cDNA under *Gal4/UAS* control. The mutation removes the mitochondrial targeting sequence [165].
UAS-3xFlag-hIRP1^{WT}	expresses 3xFlag-tagged human *IRP1* cDNA under *Gal4/UAS* control.
UAS-3xFlag-hIRP2^{WT}	expresses 3xFlag-tagged human *IRP2* cDNA under *Gal4/UAS* control.

Table 3.2 Samples tested in mass spectrometry

genotype	allele type	prothoracic gland (regular fly food)	whole body (regular fly food)	whole body (BPS – G1)	whole body (BPS – G2)
w^{1118}	wild type	2 samples	1 sample	1 sample	1 sample
$IRP1A^{3F}$	knock-in	1 sample	2 samples	1 sample	1 sample
$IRP1A^{C450S}$	transgene	not tested	2 samples	not tested	not tested
$IRP1A^{3R3Q}$	transgene	1 sample	not tested	not tested	not tested
$IRP1B^{3F}$	knock-in	1 sample	1 sample	not tested	not tested
$IRP1B^{C447S}$	transgene	1 sample	1 sample	not tested	not tested
$IRP1B^{3R3Q}$	transgene	1 sample	1 sample	not tested	not tested
$AGBE^{FCF}$	knock-in	not tested	2 samples	not tested	not tested

Table 3.3 Transcriptional responses of expressing IRP1A alleles in the prothoracic gland

rank (n/234)	symbol	description	FC 3R3Q:C	P	FC IRP1A:C	P	FC C450S:C	P
6	*dib*	ecdysone biosynthesis/heme binding/P450	-13.7	4.0E-03	-2.4	>0.05	-1.7	>0.05
7	*phm*	ecdysone biosynthesis/heme binding/P450	-11.7	1.7E-02	-2.2	>0.05	1.1	>0.05
12	*sad*	ecdysone biosynthesis/heme binding/P450	-9.7	3.2E-02	-1.7	1.7E-02	1.2	>0.05
13	*Start1*	sterol transport	-9.6	4.7E-03	-3.1	3.3 E-02	-2.3	>0.05
15	*CG7322*	short-chain dehydrogenase	-9.0	1.4E-02	-1.8	>0.05	-1.1	>0.05
23	*Cyp6g2*	heme binding/P450	-7.2	1.7E-02	-1.1	>0.05	-1.2	>0.05
26	*spidey*	short-chain dehydrogenase	-7.0	1.8E-02	-1.6	2.9 E-02	-1.6	1.6E-02
34	*nvd*	ecdysone biosynthesis/iron sulfur cluster	-6.1	1.7E-02	-2.1	8.9 E-02	-1.6	>0.05
35	*sro*	ecdysone biosynthesis/short-chain dehydrogenase	-6.0	2.1E-03	-4.2	1.3 E-02	-1.2	>0.05
41	*GstE14*	ecdysone biosynthesis/glutathione S transferase	-5.6	2.8E-02	-2.6	>0.05	-1.4	>0.05
44	*ND-15*	NADH:ubiquinone oxidoreductase, iron-sulfur subunit 5	-5.2	2.8E-02	-3.3	>0.05	-2.7	8.9E-03
47	*ouib*	ecdysone biosynthesis/zinc finger	-4.9	2.4E-02	-2.1	>0.05	-1.8	>0.05
51	*spok*	ecdysone biosynthesis/heme binding/P450	-4.8	3.6E-03	-2.4	2.9 E-03	-1.2	>0.05
52	*scu*	short-chain dehydrogenase/reductase	-4.7	1.1E-02	-1.3	>0.05	1.2	>0.05
55	*Cyt-b5*	cytochrome b5-like heme/steroid binding domain	-4.5	4.4E-02	-1.8	>0.05	-1.4	1.0E-02
56	*CG17928*	cytochrome b5-like heme/steroid binding domain	-4.5	1.4E-02	-1.8	>0.05	-2.3	3.00E-03
64	*Tig*	heme oxygenase-like	-4.2	6.0E-03	-1.7	>0.05	-1.2	>0.05

74	*ND-19*	NADH:ubiquinone oxidoreductase	-3.9	4.0E-02	-1.5	>0.05	-1.2	>0.05
98	*Fdx2*	ferredoxin/iron-sulfur cluster assembly	-3.4	1.3E-03	-1.6	>0.05	1.3	>0.05
107	*Npc2a*	sterol transport	-3.3	8.4E-03	-1.9	>0.05	-1.6	>0.05
116	*Pbgs*	heme biosynthesis	-3.2	6.4E-03	-1.5	>0.05	-1.2	>0.05
125	*Npc1a*	sterol transport	-3.1	3.2E-04	-2.3	2.4E-02	-1.2	>0.05
146	*Drat*	response to hypoxia	-2.9	1.6E-02	-1.4	>0.05	-1.6	1.1E-02
170	*CG31548*	short-chain dehydrogenase/reductase	-2.7	3.1E-02	-1.8	>0.05	-1.4	>0.05
173	*CG32857*	Nfu homolog/iron-sulfur cluster assembly	-2.7	3.7E-02	-1.2	>0.05	-1.2	>0.05
184	*ance*	ecdysone biosynthesis/zinc finger	-2.7	3.7E-02	-2.0	>0.05	-1.2	>0.05
202	*Alas*	heme biosynthesis	-2.6	1.5E-02	-1.0	>0.05	12.9	2.6E-04
212	*CG12056*	cytochrome b5-like heme/steroid binding domain	-2.6	1.7E-02	-1.5	>0.05	1.1	>0.05
218	*CG2254*	short-chain dehydrogenase/reductase	-2.6	6.7E-03	1.2	5.1E-05	-1.1	>0.05
221	*Vhl*	response to hypoxia	-2.6	2.0E-02	-1.6	>0.05	1.1	>0.05

FC = fold change

108

Table 3.4 Term enrichment analysis via DAVID tools

category	term	P	E
KEGG	insect hormone biosynthesis	9.9E-07	17.8
keywords	oxidoreductase	2.9E-06	2.8
biological Process	ecdysone biosynthetic process	5.1 E-05	22.5
interPro	NAD(P)-binding domain	2.5 E-04	3.9
cellular component	mitochondrion	4.0E-03	2.1
interPro	short-chain dehydrogenase/reductase, conserved site	9.0E-03	9.1
keywords	iron	4.0E-02	2.5
interPro	cytochrome b5-like heme/steroid binding domain	4.0E-02	9.3
molecular function	iron-sulfur cluster binding	4.0E-02	8.9
molecular function	heme binding	4.6E-02	2.7
keywords	metalloprotease	6.2E-02	4.4
biological process	positive regulation of ecdysteroid biosynthetic process	9.8E-02	19.5

Chapter 4   Investigating the functions of known IRP1A conformations

4.1   Introduction

4.1.1   Aconitase function of holo-IRP1

Mammalian IRP1 and IRP2 are two homologs of mitochondrial aconitase (mAco) [200,201]. Similar to mAco, IRP1 assembles an aconitase-type [4Fe-4S] cluster in its active site. *de novo* IRP1 receives the ISC via a complex pathway that involves several cofactors, such as cysteine desulfurase (Nfs1/ISCS), frataxin, iron-sulfur cluter assembly enzyme (ISCU), glutaredoxin 5 (*GLRX5*) and others in the cytosolic iron-sulfur cluster assembly (CIA) machinery [202] (Figure 3.16). Assembly of the [4Fe-4S] cluster alters the conformation of IRP1 and precludes IRE-binding under iron-replete condition [88,136]. At the same time, holo-IRP1 acquires enzymatic activity as a cytosolic aconitase, comparable to its mitochondrial counterpart. In this process, the enzyme catalyzes the reversible stereo-specific isomerization of citrate to isocitrate via cis-aconitate as an intermediate in a non-redox-active manner (Figure 4.1) [203]. Citrate is a crucial intermediate in several major pathways of energy and intermediary metabolism. In the mitochondria, citrate is an intermediate in the citric acid cycle, which converts acetyl-CoA to two molecules of CO_2 with concomitant generation of NADH and $FADH_2$. Reoxidation of NADH and $FADH_2$ via the electron transport chain yields ATP. Citric acid cycle (TCA) flux is in part controlled by the activity of isocitrate dehydrogenase through allosteric inhibition by ATP and product inhibition by NADH [204]. Thus, when the need for ATP synthesis is low, citrate accumulates and can be transported across the inner mitochondrial membrane via the tricarboxylate carrier [205]. In the cytosol, citrate is the substrate for ATP-citrate lyase (ACL), which generates acetyl-CoA, the building block for cholesterol and fatty acid biosynthesis in the liver and adipose tissues. Cytosolic citrate is metabolized via holo-IRP1 and cytosolic $NADP^+$-dependent isocitrate dehydrogenase to generate α-ketoglutarate (Figure 4.1). This process reduces

NADP$^+$ to NADPH, which is an essential cofactor for many enzymatic reactions involved in glutathione metabolism and lipid and cholesterol biosynthesis [206,207]. Similarly, citrate is also used in the nucleus as a substrate for acetyl-CoA during histone acetylation (Figure 4.1).

Besides being an intermediate, citrate also has regulatory roles in glycolysis, fatty acid synthesis, and oxidation. Citrate is a negative regulator of the glycolytic enzyme phosphofructokinase [208,209]. Citrate can chelate divalent cations such as Fe^{2+}, Ca^{2+}, and Zn^{2+}, and various studies have indicated that citrate has complex functions in the homeostasis of these divalent metal ions. Serum citrate is thought to be one of the carriers of nontransferrin-bound iron [210] that can contribute to hepatic iron loading in hemochromatosis [211]. Cytoplasmic citrate might be needed for transport of iron into mitochondria [212], while *in vitro* experiments and yeast genetic studies have shown that citrate–iron complexes [213,214] can promote autooxidation of ferrous iron and may contribute to iron-dependent toxicity [215]. Being a dehydratase, holo-IRP1 might be involved in regulating citrate levels in an iron-dependent manner. On the other hand, under an iron-depleted condition, holo-IRP1 releases its iron-sulfur cluster via a non-well-characterized mechanism and switch to apo-IRP1 with the capability to interact with IRE found in some iron-related transcripts (Figure 4.1).

4.1.2 Overview of Iron Responsive Element (IRE) and its interaction with IRPs

Local RNA structures in the untranslated regions (UTRs) of mRNA play essential roles in translation regulation. For example, the selenocysteine insertion sequence (SECIS) found in both bacteria and eukaryotes is responsible for incorporation of selenocysteine in proteins [216,217]. Similarly, the riboswitch found in bacteria is a regulatory segment of mRNA molecules that binds a small molecule, resulting in a change in the production of the proteins encoded by the mRNA [218,219]. Another RNA structural element is the IRE in the UTR of the mRNAs encoding

proteins involved in iron metabolism like ferritin and TfR1 [220–222]. Proteins with iron-containing cofactors participate in DNA synthesis, oxygen transport, electron transfer, nitrogen fixation, photosynthesis, scavenging of reactive oxygen species, and the metabolism of organic toxins.

The posttranscriptional IRE/IRP control system for iron and oxygen metabolism has been observed in vertebrates, invertebrates, other eukaryotes, and prokaryotes. In plants, iron metabolism is regulated at the transcriptional level only and neither IRE-containing RNA nor RNA-binding of IRPs has been detected [223–226]. The IREs constitute binding sites of two cytoplasmic iron regulatory proteins, IRP1 and IRP2. Depending on the IRE location on the mRNA, IRE/IRP interaction can regulate gene expression in two completely different ways [227,228] (Figure 4.2). One mechanism is by controlling the abundance of mRNAs with rapid-turnover elements through IRE-dependent intervention in degradation; the other mechanism is by regulating rates of translation initiation of a discrete group of transcripts with IRE structures. In iron-deficient cells, the interaction between IRP and the IRE motif in the 5' UTR can interrupt the interaction between the target mRNA and ribosome for the initiation of translation [227]. Meanwhile, in iron-replete cells, iron can bind with IRPs to induce conformational change, which promotes the dissociation of IRPs from target mRNA, results in the facilitation of translation [227,229]. Some transcripts carry one or more IRE motifs in their 3' UTR. These transcripts are susceptible to endonuclease attack and degradation, which limits their translation ability. Therefore, the interaction between IRPs and 3' UTR IRE can stabilize these transcripts and induce translation of target mRNAs [227] (Figure 4.2).

The best-characterized IRE structures include human ferritin H- and L- chains as well as TfR1 mRNAs. Later computational and experimental approaches provide more insights into the

IRE-containing mRNA family. IRE structures have been detected in several mRNAs encoding proteins related to iron utilization, including heme biosynthesis enzyme *ALAS2*, mitochondrial aconitase *ACO2*, *Drosophila* succinate dehydrogenase (*Sdh-B*), divalent metal transporter 1 (*SLC11A2*) and ferroportin (*SLC40A1*) [200,230–235]. All these mRNAs have one IRE in their 5′ UTR and are therefore regulated at the translational level with the exception from one splicing form of DMT1, which contains an IRE-like structure in its 3' UTR and is upregulated by iron deficiency [236]. Interestingly, different findings in the last decades suggest that the influence of IRPs extends over several regulatory pathways not directly related to iron homeostasis. In particular, different bioinformatic as well as biochemical approaches led to the identification of novel IRE-containing genes (Figure 4.3). Myotonic-dystrophy-related CDC42-binding kinase α (MRCKα) is a kinase that acts downstream of small GTPases known to be involved in cytoskeletal regulation and has an IRE in its 3' UTR. This IRE may mediate a similar response to iron as TfR1, although with lower intensity [237].

4.1.3   IRE structure

IREs are highly conserved hairpin structures of 25-30 nucleotides [228]. Although all IRE sequences bind IRPs and have high nucleotide conservation for each mRNA (> 95%), differences among the IREs in a single organism range from 36% to 85% [238]. Secondary and tertiary structures of IRE RNAs have been determined by mutagenesis, nuclease and chemical probing *in vitro*, *in vivo*, and by solution NMR [239–244]. The conserved structural features of all members of the IRE family are a double-stranded RNA helix of 9 – 10 base pairs with a terminal hexaloop (N_{14}-N_{19}), and an unpaired C_8 residue in the helix creating a five-base-pair upper stem and a lower stem of variable length (Figure 4.3). The apical loops of many reported IREs follow a canonical sequence of CAGUGN with N_{19} can be A, C, U or G [245]. In the terminal loop, N_{14} and N_{18} are

base-paired and create a pseudotriloop of N_{15}-N_{17} [239–244] (Figure 4.3 and Table 4.1). However, despite their high conservation, derivations have been reported for each residue. Besides a cysteine, N_{14} can also be G or U. Thus, N_{18} can be derived and carries either G, C, or A accordingly to form a temporal hydrogen bond with N_{14}. Meanwhile, N_{15}-N_{17} residues also have derivations (Table 4.1).

Sequence and base-pairing around the unpaired C_8 residue varies among different IREs, separating IRE mRNAs into two groups, one with an isolated, unpaired C, and the other with an internal loop constructed from the unpaired C_8, an unpaired base at position 6, 5' UTR to the unpaired C, and an additional paired base between them (Figure 4.3). Helix distortion around the conserved C_8 residue plays an important role in selective repressor binding, especially for IRP2 [236,242,246]. These features allow proper interaction between IRPs and IREs, and ensure their functions in cellular iron metabolism. Any disruption in IRP/IRE interaction often causes severe consequences and has been linked to human disorders (Table 4.2).

4.1.4 Current status of IRE-containing transcripts in *Drosophila*

In *Drosophila*, there are two characterized IRE-containing mRNAs, including *Fer1HCH RA* transcript, a splice form that encodes a ferritin heavy chain isoform, and *SdhB* mRNA, which encodes the subunit B of succinate dehydrogenase [43,197,198,247]. Unlike mammals, *Drosophila* and other insect species have a distinct respiratory system which allows them to deliver a sufficient amount of oxygen to target tissues independent of circulating hemoglobin, they still require a high amount of iron for steroid hormone development, detoxification and cuticle development [43,166]. Furthermore, IRP1A, the *Drosophila* ortholog of human IRP1, still plays a vital role in iron metabolism in this species [166,248] with the capability to switch between two

conformations, including the constitutively RNA-binding apo-form. The low number of known IRE-transcripts in *Drosophila* is insufficient to explain the systemic and cellular iron homeostasis in this species. A computational study in Tsetse fly *Glossina morsitans* has reported about 150 transcripts with candidate IREs, suggesting *Drosophila melanogaster* might have uncharacterized IRE-containing transcripts [249]. Since the search of canonical IRE structures only detects *Fer1HCH RA* and *SdhB* transcripts (Figure 4.3A and Table 4.1), any unidentified IRE-containing transcript will likely represent noncanonical structures.

4.2 Results and Discussions

4.2.1 *Drosophila* requires both IRP1A conformations for normal development

The dual conformations of IRP1A provide flexibility over citrate regulation as well as cellular iron homeostasis. In *Drosophila,* holo-IRP1A is not the only dehydratase that interconverts between citrate and isocitrate. Two other hydratases, namely mitochondrial aconitase (Acon) and IRP1B, can also carry out this reaction. Especially IRP1B, similar to IRP1A, is also an ortholog of human IRP1 but it does not have RNA-binding activity [166,248]. On the other hand, apo-IRP1A is the only RNA-binding IRP in *Drosophila*. These features suggest that among any mutation affects IRP1A conformations, mutations abolish the apo-form should, in theory, show more severe phenotypes while the holo-form mutations should not be as severed due to the redundancy of hydratase activities from Acon and IRP1B.

Using classic CRISPR, I generated two main IRP1A mutations that result in only one of the conformations, namely the holo-form IRP1A^{3R3Q} and apo-form IRP1A^{C450S} (Figure 3.2, Appendix A.8). I tested the survival of these mutants as a function of dietary iron concentration as well as IRP1B function. I have shown previously that wild-type animals' development is affected when being raised on iron-depleted medium for five generations (Figure 3.10A) while *IRP1B^{KO}*

116

animals can survive without issue in the absence of functional IRP1B (Figure 3.9C). In our hands, both $IRP1A^{3R3Q}$ and $IRP1A^{C450S}$ mutant exhibit different levels of developmental effect. For example, approximately 50% of animals homozygous for $IRP1A^{C450S}$ allele (lacks ISC) could not develop to adulthood. The survival rate is even worse when animals were raised in iron-depleted medium with less than 10% surviving to adult (Figure 4.4). The developmental defect seems to be worse in the presence of $IRP1B^{KO}$ allele, animals homozygous for $IRP1A^{C450S}$ and also carries a copy of $IRP1B^{KO}$ allele shows a more severe defect in both regular as well as iron-depleted media. Interestingly, these animals have better survival rates in iron-rich medium (Figure 4.4). A similar trend can be observed in animals homozygous for both $IRP1A^{C450S}$ as well as $IRP1B^{KO}$ (Figure 4.4). On the other hand, animals homozygous for $IRP1A^{3R3Q}$ (lacks apo-form) shows more dramatic developmental defects with only 6% adult on regular medium and no adult in an iron depleted medium. Similar to $IRP1A^{C450S}$, $IRP1A^{3R3Q}$ mutant exhibit a better development on iron-rich medium (Figure 4.4).

Interestingly, animals heteroallelic for $IRP1A^{C450S}$ and $IRP1A^{3R3Q}$ show higher survival with 50-60% of larvae forming adults, independent of dietary iron levels or functional IRP1B (Figure 4.4). These data demonstrate that animals require both IRP1A conformations for normal growth and missing either form result in iron-sensitive development. The survivability of each mutation is only partially dependent on functional IRP1B, while the presence of both mutations could ensure proper development, indicating the redundancy of IRP1B in *Drosophila* cellular iron metabolism.

4.2.2   Holo-IRP1A interacts with histone H2Av and regulates nuclear citrate levels

The nuclear localization of holo-IRP1A is intriguing since it provides an insight into the nuclear functions. Since holo-IRP1A acts as an aconitate hydratase, we hypothesize that the protein

still has aconitase activity and helps regulate nuclear citrate level. Under the catalytic activity of ACL, nuclear citrate is converted to acetyl-CoA, a substrate for histone acetylation. The consequence of this process is the alteration of gene expression. We have detected IRP1A interaction with histones H2A and H2Av in IRP1A mass spectrometry samples (Figure 3.12). Interestingly, these histones were not detected in MS samples of IRP1A^{WT} animals raised on low iron medium for three generations [166], thus we can rule out the possibility of accidental detection of non-specific histones (CRAPome) [250,251]. Based on the above data, we hypothesize that holo-IRP1A localizes to the nuclei and regulate local citrate levels, and thus affects local gene expression. This hypothesis was supported by performing PG-specific RNA-sequencing to evaluate the transcription profile of *IRP1A^{3R3Q}* overexpression in comparing with *IRP1A^{WT}* and *IRP1A^{C450S}* overexpression. Many genes down-regulated in *IRP1A^{3R3Q}* overexpression were involved in iron-dependent processes like ecdysone synthesis, heme production, or ISC assembly (Table 4.3).

To further investigate the role of nuclear IRP1A, I first validated the interaction between IRP1A with histone H2Av. For this approach, I used whole-body larvae and performed coimmunoprecipitation (coIP). The interaction between IRP1A with histones, if exists, can be explained by four scenarios: (i) IRP1A interacts physically with histones, (ii) IRP1A interacts physically with histones as well as DNA, (iii) IRP1A interacts with DNA but not with histones, the detection of both protein in this situation is indirect and happens via DNA-coIP, and (iv) IRP1A interacts with a DNA/histone-binding protein. To validate and characterize the interaction, I modified the coIP procedure by incubating lysate samples with micrococcal nuclease. This enzyme is an endonuclease that preferentially digests single-stranded as well as double-stranded DNA [252]. By incubating samples with micrococcal nuclease to remove any intact DNA, we expect to

rule out the indirect interaction in scenario (iii). As a positive control, I coIP H2Av with Su(var)205, a nuclear protein with reported interaction with H2Av [253]. As a negative control, I used Su(var)3-7, another member of Su(var) family. However, unlike Su(var)205, no interaction between Su(var)3-7 and H2Av was reported to the best of my knowledge. Since Su(var)3-7 is involved in heterochromatin organization, it can interact with DNA and as such, be accidentally pulled down. In our hand, coIP H2Av could pull down both Su(var)205 and Su(var)3-7. However, in samples with micrococcal nuclease pre-incubation before coIP, we could only detect Su(var)205 while no Su(var)3-7 was observed (Figure 4.5A). These data indicate the approach is sufficient to validate the interaction between H2Av and candidate proteins. Interestingly, coIP H2Av could successfully pull down IRP1A in both samples, even in samples incubated with micrococcal nuclease (Figure 4.5A). These results suggest IRP1A and H2Av stays in a same complex.

I then used the same approach to validate the interaction between histone H2A and IRP1A. For this experiment, I used H2Av as positive control based on the previously reported interaction with H2A [254] and still used Su(var)3-7 as a negative control. In my hands, IRP1A can interact physically with histone H2A, similar to the interaction with H2Av (Figure 4.5B). Since H2Av has been shown previously to have a physical interaction with IRP1A, I wondered if IRP1A and H2A interact directly or via H2Av. There are other proteins reported to interact with H2A [254–257], however, none of them can bind both histone H2A as well as DNA. I tested whether H2Av is required for the interaction between H2A and IRP1A by interrupting H2Av and testing if any interaction is detected. Using CRISPR Cas13/crRNA to target H2Av mRNA (chapter 7), I was able to significantly reduce the expression of H2Av. This reduction also interfered with the interaction between H2A and IRP1A since I could not detect any band corresponding to IRP1A in this experiment (Figure 4.5C). On the other hand, I have not tested if H2A is required for the

interaction between H2Av and IRP1A. A similar approach can be used to answer this question in the future. Overall, these data suggest that IRP1A, histones H2A, and H2Av interact.

In addition to histone interaction, I also tested whether the nuclear IRP1A is involved in regulating the nuclear citrate level. To answer this question, I aim to look for any difference in nuclear citrate when there is an increase in nuclear IRP1A amount by comparing the nuclear aconitase assay between wild-type samples with ubiquitously overexpressed IRP1A. In my hands, nuclear aconitase activity in overexpressed $IRP1A^{WT}$ samples is about 2.66 times higher than wild-type samples. In contrast, samples with overexpressed $IRP1A^{C450S.NLS}$, an apo-IRP1A variant with nuclear localization sequence (NLS), did not show any significant difference from wild-type samples (Figure 4.5D).

Furthermore, when evaluating the nuclear citrate levels in wildtype samples during L2 and L3 development, I noticed a negative correlation between amount of nuclear IRP1A and nuclear citrate levels. Comparing to nuclear lysate at12hr L2 larvae, I noticed that at 4-8 hr, 20hr, 28hr and 44-48hr after L2/L3 molting, along with the accumulation of nuclear IRP1A, the nuclear citrate level is significantly reduced. Interestingly, these time points are also corresponding to either minor or major ecdysone peaks when cells already have enough enzymes responsible for iron-dependent pathways like heme synthesis or ISC assembly. The low nuclear citrate level will presumably reduce histone acetylation and thus suppress genes expression (Figure 4.5E, Tables 3.3 and 3.4). These data support the hypothesis that nuclear holo-IRP1A is involved in regulating nuclear citrate levels, thus, controlling histone acetylation.

4.2.3 Apo-IRP1A overexpression reveals highly stabilized transcripts

Drosophila was one of the first insects shown to have IRP/IRE binding activity with the detection of IRE in *SdhB* mRNA [197,198]. However, since the discovery of an IRE in *SdhB*, only

one more IRE-containing transcript was confirmed in this species, namely *Fer1HCH-RA* [258–260]. Unlike mammals, which have two IRP proteins capable of binding IREs, among two IRPs identified in *Drosophila*, only IRP1A can retain binding activity to IRE structures [166,248]. Recent studies have shown that IRPs have other, unknown mRNA targets. In a transcriptome-wide screening to identify IRP/IREs candidates in mouse, Sanchez et al. have reported 25 novel mRNAs with IRP binding activity [261]. This study showed that noncanonical IREs exist and require further characterization, and also raise the necessity of establishing a reliable screening procedure to validate potential candidates of IRPs. Previously, Pendleton Cox, a former graduate student in King-Jones lab, has attempted to identify additional IRE-containing transcripts via SIRES, a bioinformatic program that detects IRE-like motifs based on present and previous studies [240,241,262]. In this work, Cox et al. only focused on the 20 genes related to iron homeostasis and did not confirm the existence of IRE structures by any other means. As an effort to continue his work, I carried out two independent approaches: (i) via RNA-sequencing to evaluate transcription profile of $IRP1A^{C450S}$ overexpressed animals, and (ii) via RNA-immunoprecipitation, followed by sequencing (RIP-seq). In this section, I will only focus on the first approach, while the second approach is currently developing and will only be briefly mentioned.

We performed RNA-sequencing experiments using PG-specific overexpression of IRP1A variants and filtered any transcripts uniquely upregulated in $IRP1A^{C450S}$ but not significantly misregulated in $IRP1A^{WT}$ or $IRP1A^{3R3Q}$ overexpressed animals. The upregulation of these transcripts can result from either transcription misregulation due to having too much apo-IRP1A or transcript stabilization as a result of IRP/IRE interaction. This approach resulted in the 23 transcripts for further investigation. Among these transcripts, interestingly, two encode enzymes involved in heme biosynthesis, namely ALAS (rank 15/23) and PBGD (rank 23/23). We also

detected *Mco4* (rank 2/23), a transcript encoding the multi-copper oxidase MCO4, a member of the MCO family with a proposed role in iron trafficking across the cell membrane (Table 4.3). Using SIRES search against these transcripts failed to detect the canonical IRE structure in 19/23 hits except for PBGD (23/23), NaPi-T (14/23), Mur18B (7/23), and Mal-A8 (3/23) (Figure 4.6 and Table 4.4). Among the four transcripts with detected canonical IREs, the latter three all carry potential IRE-motifs in the CDS regions, while PBGD carries IRE in 5' UTR, which contradicts what we know about the IRE location in transcripts stabilized by the /IRE interaction. This issue suggests these might not be the real IREs, or there are some unknown derivations in the IRP/IRE mechanism.

Despite the above issues, the 23 transcripts stabilization can be explained by two possible scenarios: (i) apo-IRP1A directly interact with a noncanonical IRE-motif and stabilize these transcripts, (ii) some unknown reasons that cause the upregulation of these transcripts. To test these scenarios, I manipulated IRP1A activity in the PG using *ex vivo* culture and monitored target gene expression. By adding propylamine propylamine NONOate (PPNO), a NO donor, I can destabilize the ISC in holo-IRP1A, resulting in the formation of the apo-form and induce mRNA-binding [263]. In another sample, PPNO was added to the culture together with cycloheximide (CHX), a translation inhibitor [264]. While IRP1A was mostly switched to mRNA-binding, the translation machinery was also blocked by CHX, thus preventing transcriptional up regulation via *de novo* translated transcription factors. 24 hours later, I used 50 RG per replicate for qPCR analysis. If the target genes are enriched with PPNO alone but not under PPNO + CHX, their upregulation in $PG{>}IRP1A^{C450S}$ depends on unknown reasons. If the expression is up in both cases, the effect should be due to direct binding of apo-IRP1A. I also included samples exposed to CHX only or cell medium only as controls. Among 23 candidates, I found that 10 transcripts were

upregulated independently of the presence or absence of CHX while the other 13 candidates were only induced in the absence of CHX (Figure 4.7). On the other hand, the expression of candidate genes in samples exposed to CHX only shows little or no, but not significantly different from control samples without exposure to either CHX or PPNO. The only exception was *Fst*, which showed upregulation even in the presence of CHX alone. *Fst* was also upregulated in the presence of PPNO only or a combination of PPNO and CHX. However, when comparing *Fst* expression level in PPNO with samples in CHX only or PPNO+CHX, Fst is still significantly upregulated (Figure 4.7).

In an independent approach, I performed RNA-immunoprecipitation (RIP), followed by qPCR to detect transcripts potentially interact with apo-IRP1A. Using *SdhB* as a positive control, I compared the expression levels of 23 candidates in RNA extracted from IRP1A^{C450S}, IRP1A^{WT}, and IRP1A^{3R3Q} animals. Since IRP1A^{3R3Q} is predicted to abolish RNA-binding activity, I used this samples as a negative control for any random coIP of mRNA. In my hands, none of the 23 candidates were detected in IRP1A^{3R3Q} samples, indicating this approach is sufficient to test the binding of apo-IRP1A with RNA. Interestingly, only eight candidates were detected in IRP1A^{C450S} RIP-qPCR (Figure 4.8). Remarkably, all of these eight candidates were found upregulated in the NO-induced qPCR samples above, and importantly, even in samples incubated with both PPNO and CHX (Figures 4.7 and 4.8). *Mur18B* and *Mal*-A8 are the two transcripts stabilized in NO-induced apo-IRP1A switch but were not pulled down in IRP1A RIP-qPCR. This difference can be explained by the weak and sensitive interaction between IRP1A and these transcripts, making them not feasible for coIP. These data suggest there are most likely transcripts with unidentified IRE motif in *Drosophila*. Taken together, based on RNA-seq, RIP-qPCR, and *ex vivo* induction of apo-IRP1A, I came up with a list of at least eight candidates with potential IRE-motifs. On the other

hand, the drawback of this approach is it only allows me to identify transcripts with IRE-motifs locate on the 3' UTR while IRE on 5' UTR will remain missing because when looking at RNA-seq data, I could only filter transcripts potentially stabilized by IRP/IRE interaction while transcripts with IRE motifs on the 5' UTR are only regulated at the translation levels, thus could not be captured with this approach. A more reliable method is needed to investigate the IRP/IRE interaction in *Drosophila*.

4.3   Conclusion and future directions

4.3.1   The nuclear function of IRP1A

Using coIP, I have validated the interaction between IRP1A with histones H2Av and H2A. I also showed that increasing nuclear holo-IRP1A affects the nuclear citrate level. These data, together with the earlier RNA-seq result, support a model in which nuclear holo-IRP1A acts to regulate nuclear citrate levels and thus controls histone acetylation. The ultimate consequence of this activity is the transcriptional regulation of genes involved in iron-dependent processes. The classic paradigm of cellular iron homeostasis will be revised by adding another aspect of holo-IRP1 in regulating iron-dependent gene expression. On the other hand, apo-IRP1 still plays an essential role in cellular iron metabolism. My preliminary data suggest apo-IRP1 can localize into the nuclei of anterior larval midgut cells. However, nuclear apo-IRP1 is still not detected at a significant amount in other tissues like fat body, salivary gland or prothoracic gland. On the other hand, histone H4 was detected in IRP1A^{C450S} mass spectrometry. It is possible that the nuclear localization of apo-IRP1A is tissue-specific, and it would be interesting to know why apo-IRP1A needs to localize into nuclei. Since transcription occurs in the nucleus, one hypothesis is the nuclear localization of apo-IRP1A is to ensure the quick binding of the protein to target transcripts, especially in the gut where iron influx is dramatically altered dependent on nutrient conditions and

thus, requires rapid and precise regulation on iron-related protein synthesis. In addition, the anterior midgut might not be the only tissue with detectable nuclear apo-IRP1A. In all cases, *Drosophila* requires both IRP conformations for its normal development suggesting missing either conformation results in different degrees of developmental defects.

4.3.2   The nuclear translocation of IRP1A

It remains unclear under which mechanism IRP1A can shuttle between cytoplasm and nucleus. Based on mass spectrometry data, I detected Ran and the nuclear transport factor-2 (Ntf2) in IRP1A pull-down assay. Ntf2 is a cytosolic protein responsible for nuclear import of Ran, a small Ras-like GTPase involved several critical cellular processes, including cell cycle regulation, chromatin organization, or controlling the directionality of nucleocytoplasmic transport [265,266]. One hypothesis is IRP1A serves as cargo during Ran-dependent nuclear transportation. On the other hand, we also detected chickadee (chic), the *Drosophila* ortholog of human profilin that is responsible for the nuclear export process. Interestingly, human profilin transcript also carries an IRE motif. Furthermore, among 34 hits identified from earlier genome-wide screening, there is a gene encodes for a nuclear export factor, namely RanBP3 [267,268]. I hypothesize Ran and Ntf2 work together to regulate the nuclear translocation of IRP1A while Ntf2, RanBP3 and chic are involved in nuclear export of the protein. The nuclear translocation of IRP1A is being studied in detail by another student in KKJ lab.

4.3.3   Identification of non-canonical IRE in *Drosophila*

The lack of a reliable screening system has limited our understanding on IRE-containing transcripts in *Drosophila*. To continue previous work by Pendleton Cox, we have generated a list of potential IRE-motif candidates using RNA-seq and RIP-qPCR. However, the candidates require further investigation since little is known about the non-canonical IRE-motif on their mRNAs. A

future study using a combination of RNA electrophoretic mobility shift assay (rEMSA) and mutagenesis is needed. We have generated two CRISPR/Cas9 toolkits for temporal and spatial manipulation of target genes (chapter 6) as well as CRISPR/Cas13 for RNA editing (chapter 13). These tools will be a suitable approach to validate the IRE on target transcripts.

RNA-seq is not the best approach to identify IRE-containing candidates since I can only identify transcripts stabilized by IRP/IRE interactions while cannot detect transcripts with IRE at 5' UTR. A RNA-immunoprecipitation sequencing (RIP-seq) or cross-linking immunoprecipitation sequencing (CLIP-seq) approach will provide better insight. My first effort using RIP-seq and CLIP-seq resulted in a big pool of RNA being pulled down together with IRP1A^{C450S}. Many of them are also detected in IRP1A^{3R3Q} samples, suggesting that the current method requires further optimization. Even though my procedure was based on RIP-qPCR experiment, there are some modifications, including longer incubation, the addition of RNAse and cross-linking step like UV-crosslinking (in CLIP-seq). These modifications, especially the incubation time, may affect the RNA pool being pull-down. Future optimization is needed to ensure the reliability of the outcome.

4.4   Figures
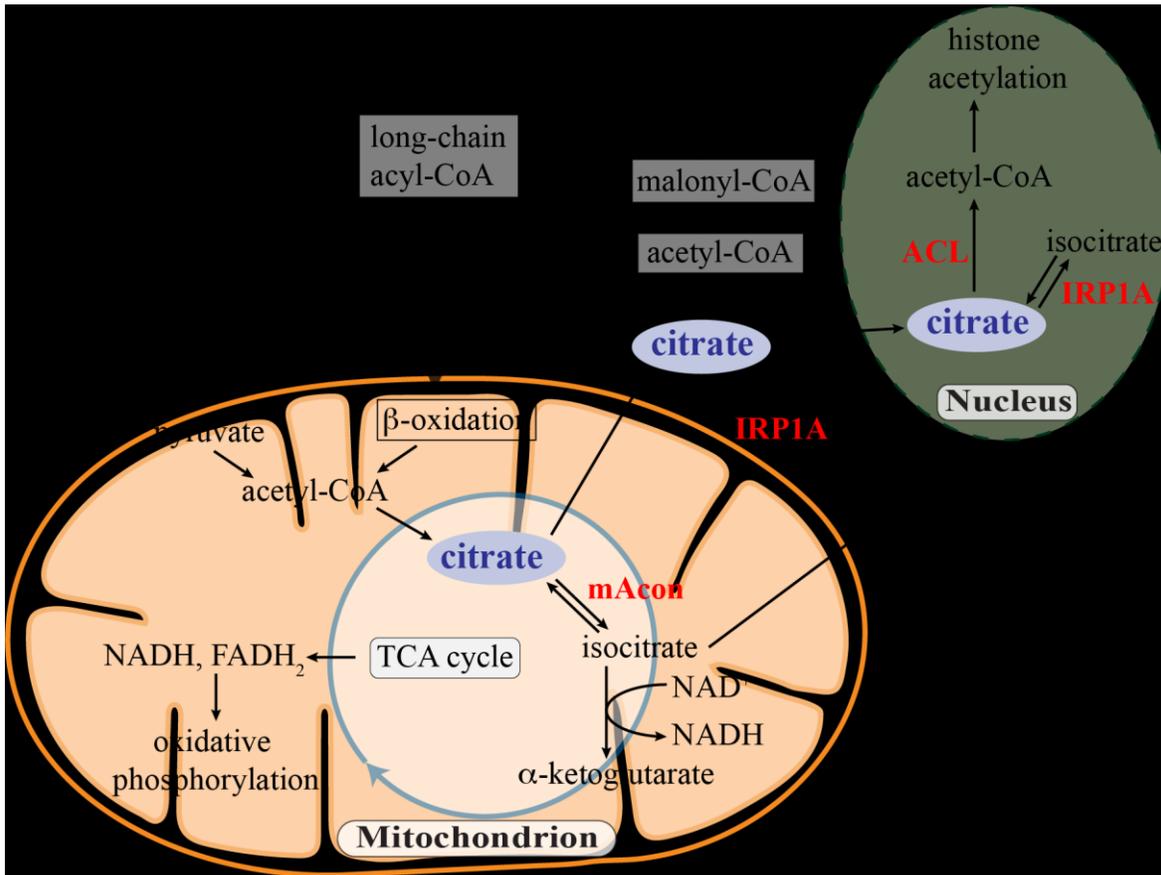
Figure 4.1 Overview of citrate and aconitase roles in the major energy metabolism pathways

Citrate is a key intermediate that interconnects the metabolic pathways of the citric acid cycle (TCA) and oxidative phosphorylation in the mitochondria to glycolysis and fatty acid synthesis in the cytosol. When ATP synthesis is in high demand, citrate is metabolized through the TCA cycle to generate NADH and $FADH_2$, the reducing equivalents needed for ATP production. When the need for ATP synthesis is low, citrate is exported into the cytosol and channeled into fatty acid biosynthesis for energy storage. Inhibition of holo-IRP1A due to iron deficiency or oxidative damage to the iron-sulfur cluster (ICS) could decrease ATP production, promote fat accumulation, decrease glycolysis, and decrease fatty acid oxidation. In addition, decreased citrate flux through

holo-IRP1A may decrease production of cytosolic NADPH, a major source of reducing equivalents for fatty acid synthesis and an important defense against cytosolic oxidative stress. Nuclear citrate is the substrate of acetyl-CoA, which will be used for histone acetylation and regulation of gene expression, holo-IRP1A may also act as aconitase to regulate nuclear citrate level.
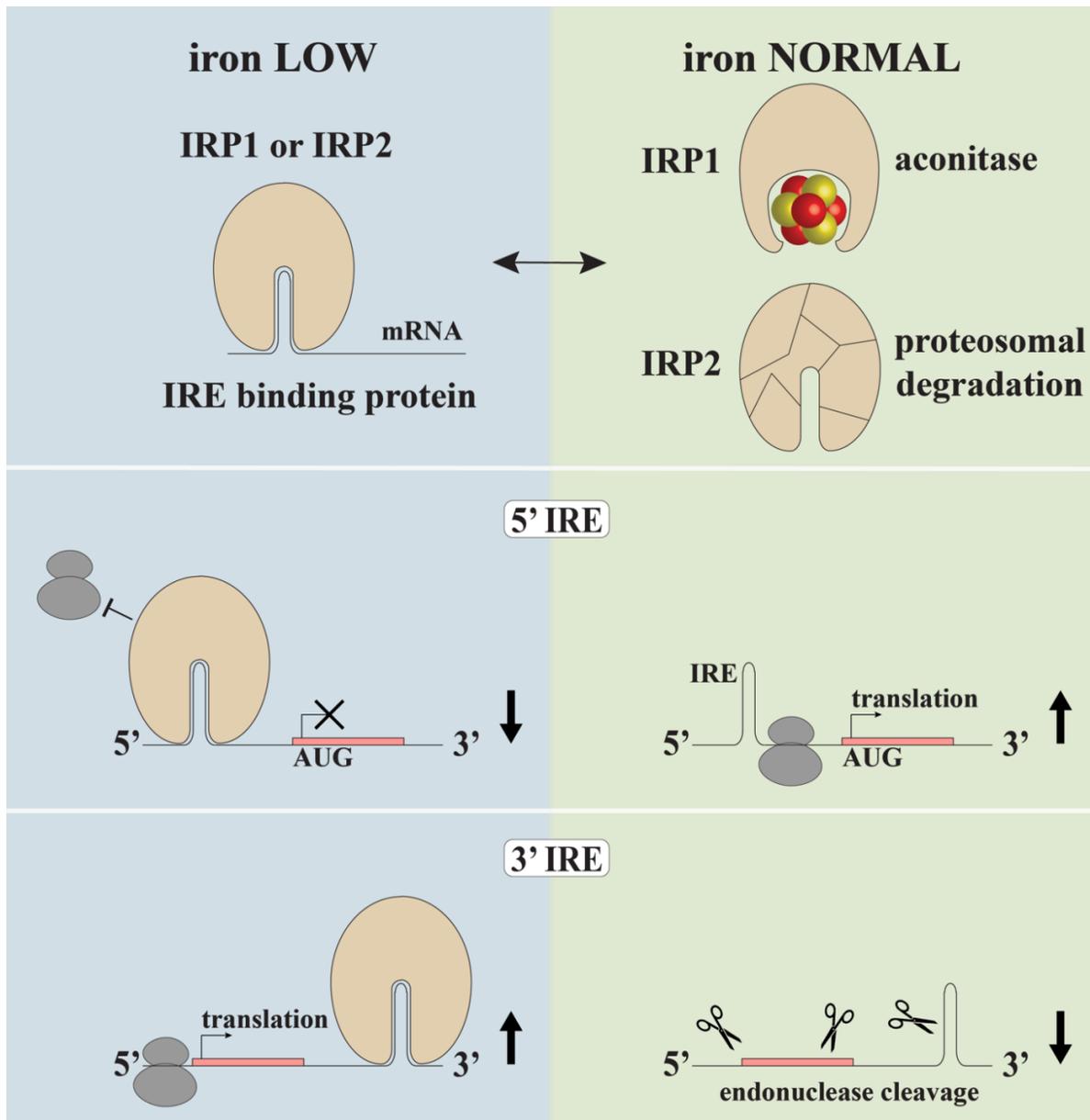
Figure 4.2 The iron regulatory protein (IRP)/ iron responsive element (IRE) mechanism of translation modulation

IRP1 and IRP2 bind to IRE in iron deficient conditions and mediate the translation repression in those mRNAs with an IRE at the 5' UTR, decreasing their protein levels. If IRE locates in the 3' UTR, binding with IRP will enhance target mRNA stabilization by preventing an endonuclease cleavage. In iron replete cells, iron can bind IRPs and abolish the IRP/IRE interaction. In this

condition, IRP1 binds to the iron-sulfur cluster and acts as cytosolic aconitase while IRP2 is degraded. Transcripts with 5' IRE can proceed to translation pathway while transcripts with 3' IRE are cleaved by endonucleases.
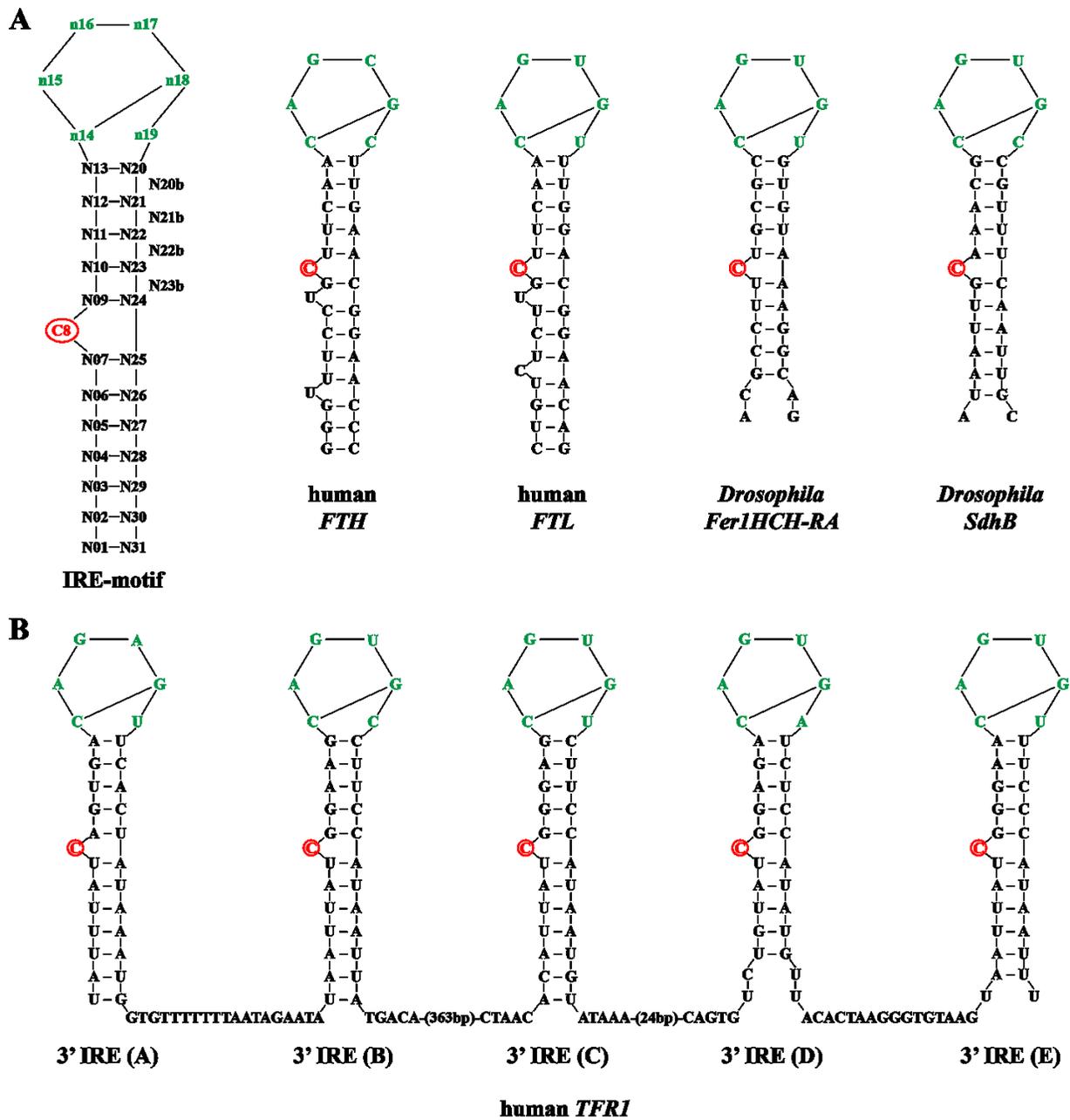
Figure 4.3 Schematic of iron responsive element (IRE) structures

A. General structure of known IRE-motif. Most motifs contain a characteristic C-bulge (C8, red) in the stem motif and a 6-nucleotide (nt)-CAGAGU-apical loop (green). Shown here are two human ferritin transcripts and two *Drosophila* transcripts with IRE structure, including *Fer1HCH RA*, which encodes an isoform of ferritin heavy chain, and *SdhB*, which encodes the isoform B of

succinyl dehydrogenase. All of these transcripts carry the IRE-motif at the 5' untranslated region (UTR). **B.** The structure of penta-IRE-motif in human transferrin receptor (TFR1). These IREs all locate at the 3' UTR of the transcript.
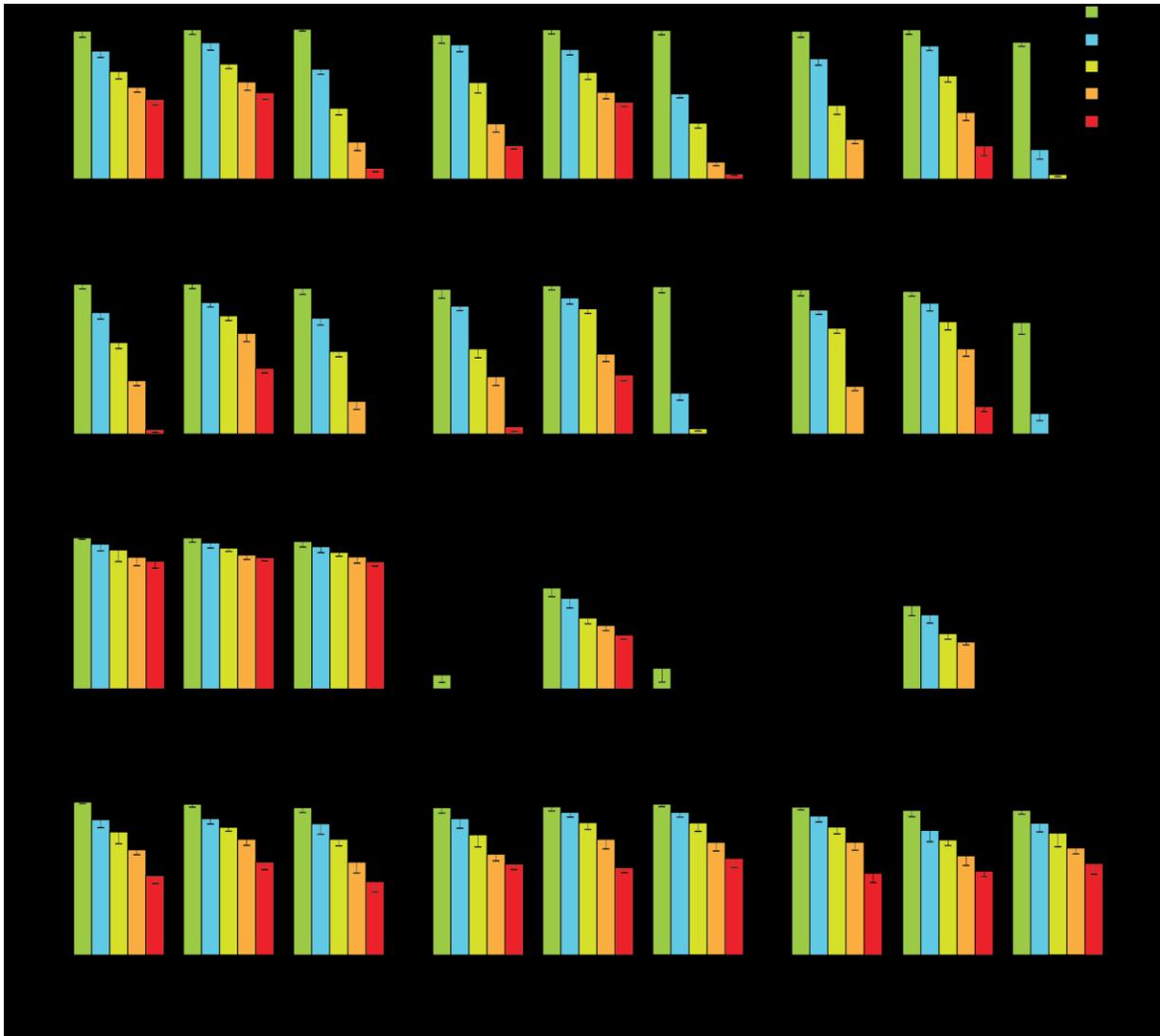
Figure 4.4 Survival rates of animals with different *IRP1A* and *IRP1B* allele combinations

The $IRP1A^{C450S}$ and $IRP1A^{3R3Q}$ alleles are CRISPR-mediated replacements of the endogenous loci, while "+" indicates a wildtype and a "-" is a CRISPR-mediated removal of the entire coding region. Survival rates relative to starting population of embryos, L1 = first instar larvae, L2 = second instar larvae, L3 = third instar larvae. Error bars = standard deviation. The $IRP1A^{C450S}$ allele carries a single point mutation that abolishes iron-sulfur cluster (ISC) binding of *IRP1A*, resulting in the constitutively RNA-binding apo-IRP1A. The $IRP1A^{3R3Q}$ allele harbors three-point mutation that covert three arginines (R) critical for mRNA-binding to glutamines (Q). The resulting IRP1A^{3R3Q}

133

protein is predicted to have normal holo-form function (aconitase and nuclear entry), but impaired or no RNA-binding capability. Two findings are noteworthy here: (i) double nulls of *IRP1A* and *IRP1B* are not viable, while all other combinations can be kept alive on iron-rich media, (ii) we have shown the IRP1 switch mechanism can be bypassed in a heteroallelic combination (*IRP1A$^{C450S/3R3Q}$*), where one allele provides the RNA-binding aspect of IRP1A, while the other provides the aconitase/nuclear holo-form. This provides genetic evidence that the holo-form is essential. Note that neither homozygous combination of these alleles form adults under iron-depleted conditions (=> interallelic complementation of *IRP1A^{3R3Q}* and *IRP1A^{C450S}*).
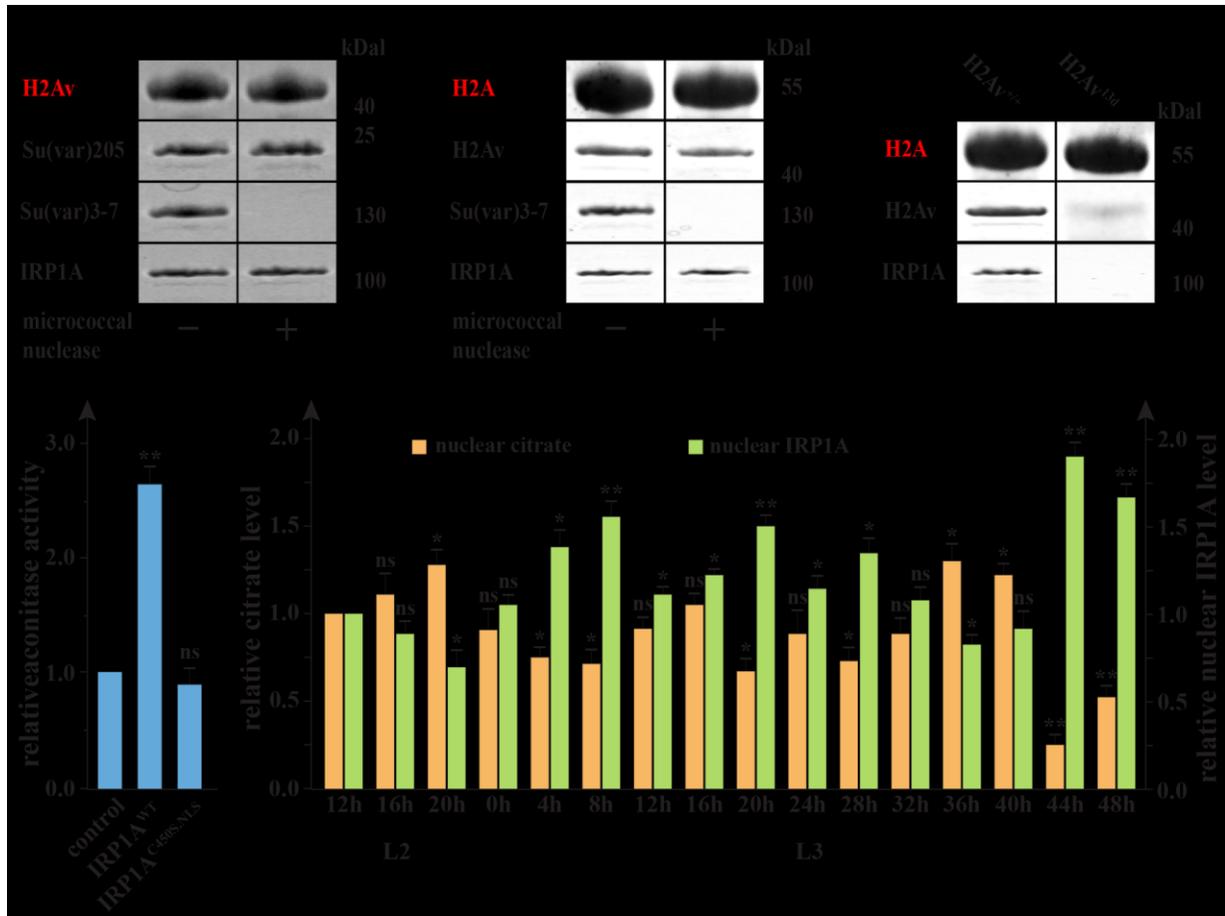
Figure 4.5 Nuclear function of holo-IRP1A

A. Validation of histone H2Av and IRP1A physical interaction. Histone H2Av was used as bait for coimmunoprecipitation (coIP) from whole larval body lysates. To distinguish protein interaction from DNA binding possibility, one lysate was pre-incubated with micrococcal nuclease for 1 hour prior to coIP. Su(var)205 was tested as a positive control while Su(var)3-7 was used as a negative control. **B.** Validation of histone H2A and IRP1A physical interaction. Histone H2A was used as bait for coIP with H2Av served as positive control and Su(var)3-7 was negative control. **C.** The relationship between H2Av- and H2A-IRP1A physical interaction. By using H2A as bait for coIP and manipulating H2Av expression level, we can compare the pull-down level of IRP1A. **D.** Relative nuclear citrate level of overexpressed $IRP1A^{WT}$ and $IRP1A^{C450S.NLS}$ comparing

with control sample. Error bars represent standard deviation from three biological replicates, ** $p<0.01$. **E.** Relative nuclear citrate level (orange) in relationship with endogenous nuclear IRP1A level (green). Samples were tested at different time points during second instar (L2) and third instar (L3) larval development. Samples were normalized to nuclear citrate and IRP1A level at 12h L2 stage which was determine based on relative ratio of nuclear IRP1A over histone H3, a common nuclear loading control. Error bars represent standard deviation from three biological replicates, ** $p<0.01$, * $p<0.05$.
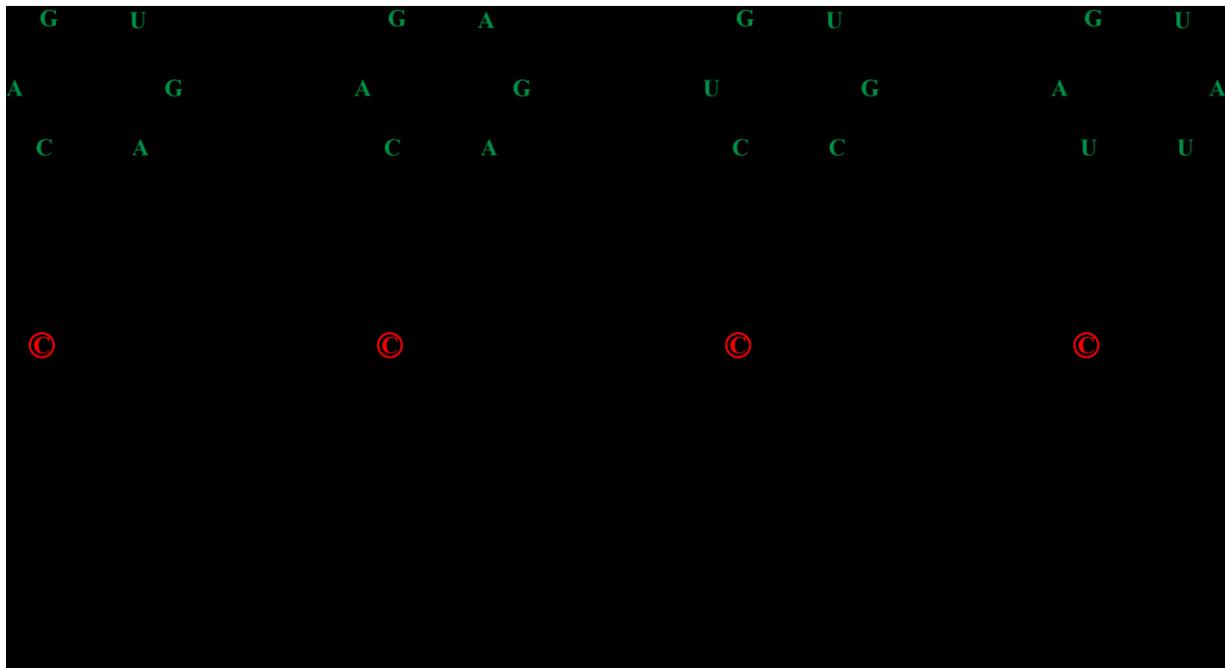
Figure 4.6 SIRES analysis of 23 IRE-containing candidates from IRP1A^{C450S} RNA-sequencing

Sequences of 23 transcripts were analyzed via SIRES platform. Canonical IRE-like structures were identified in four transcripts. All motifs carry a characteristic C-bulge (C8, red) in the stem motif and a 6-nucleotide (nt)-apical loop (green).
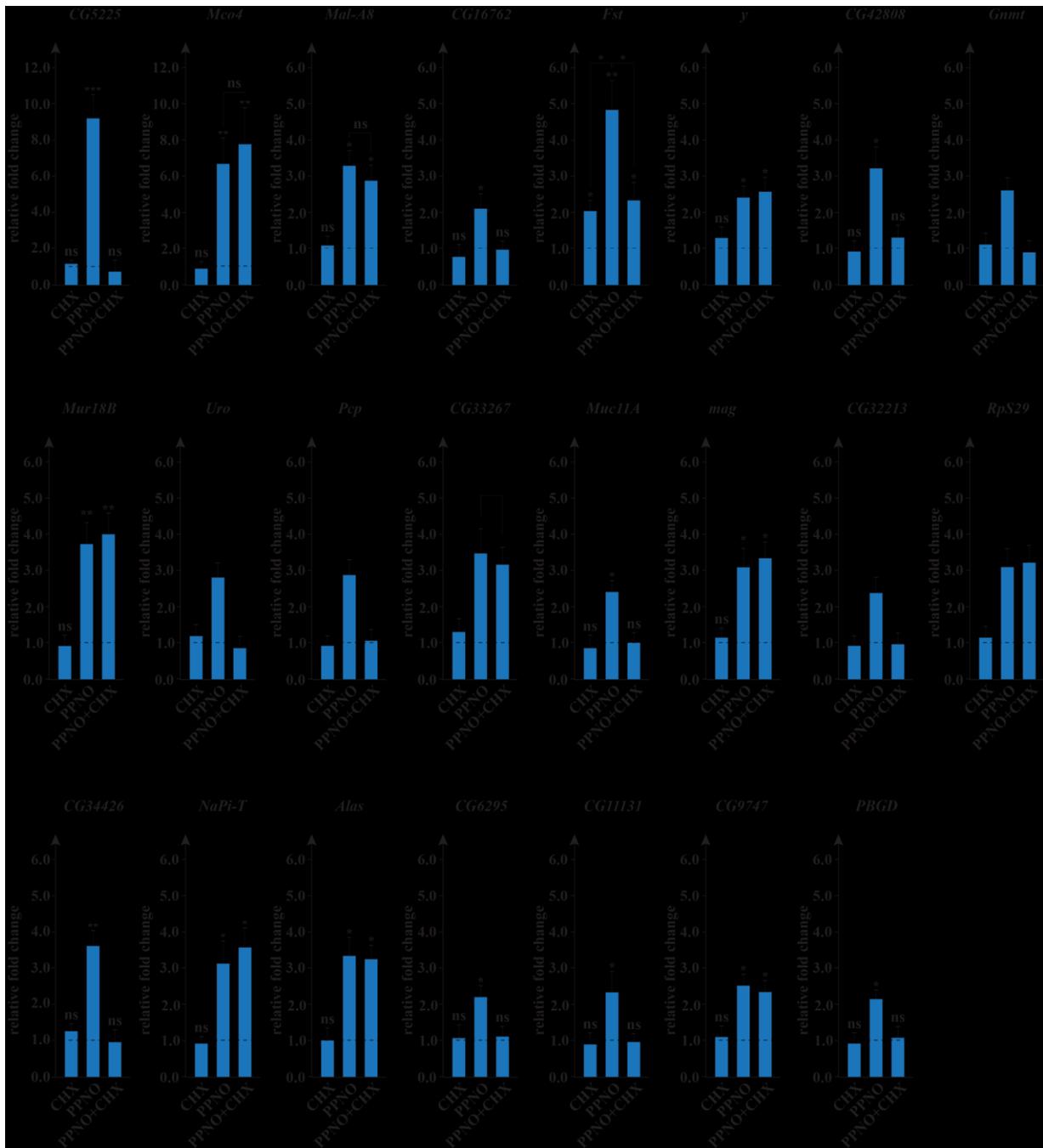
Figure 4.7 qPCR evaluation of transcripts upregulated in IRP1A^{C450S} RNA-sequencing

Ex vivo brain-ring gland complexes (BRGCs) are incubated with propylamine propylamine NONOate (PPNO), a nitric oxide (NO) donor. NO destabilizes the iron-sulfur cluster (ISC) in holo-IRP1A, resulting in the formation of the apo-form and induce mRNA-binding. In another

sample, PPNO is added to culture together with cycloheximide (CHX), a translation inhibitor. While IRP1A is mostly switched to mRNA-binding, translation is also blocked by CHX. 24 hours later, 50 ring glands (RGs) per replicate were used for qPCR analysis. If the target genes are upregulated with PPNO alone but not under PPNO + CHX, their upregulation in *PG>IRP1A^{C450S}* depends on a transcription factor. On the other hand, if the expression is up in both cases, the effect should be due to direct binding of apo-IRP1A. Samples exposed to CHX only or cell medium (dotted line) were included as controls. Error bars represent 95% confidence intervals from three biological replicates, ***$p<0.001$, ** $p<0.01$, *$p<0.05$.
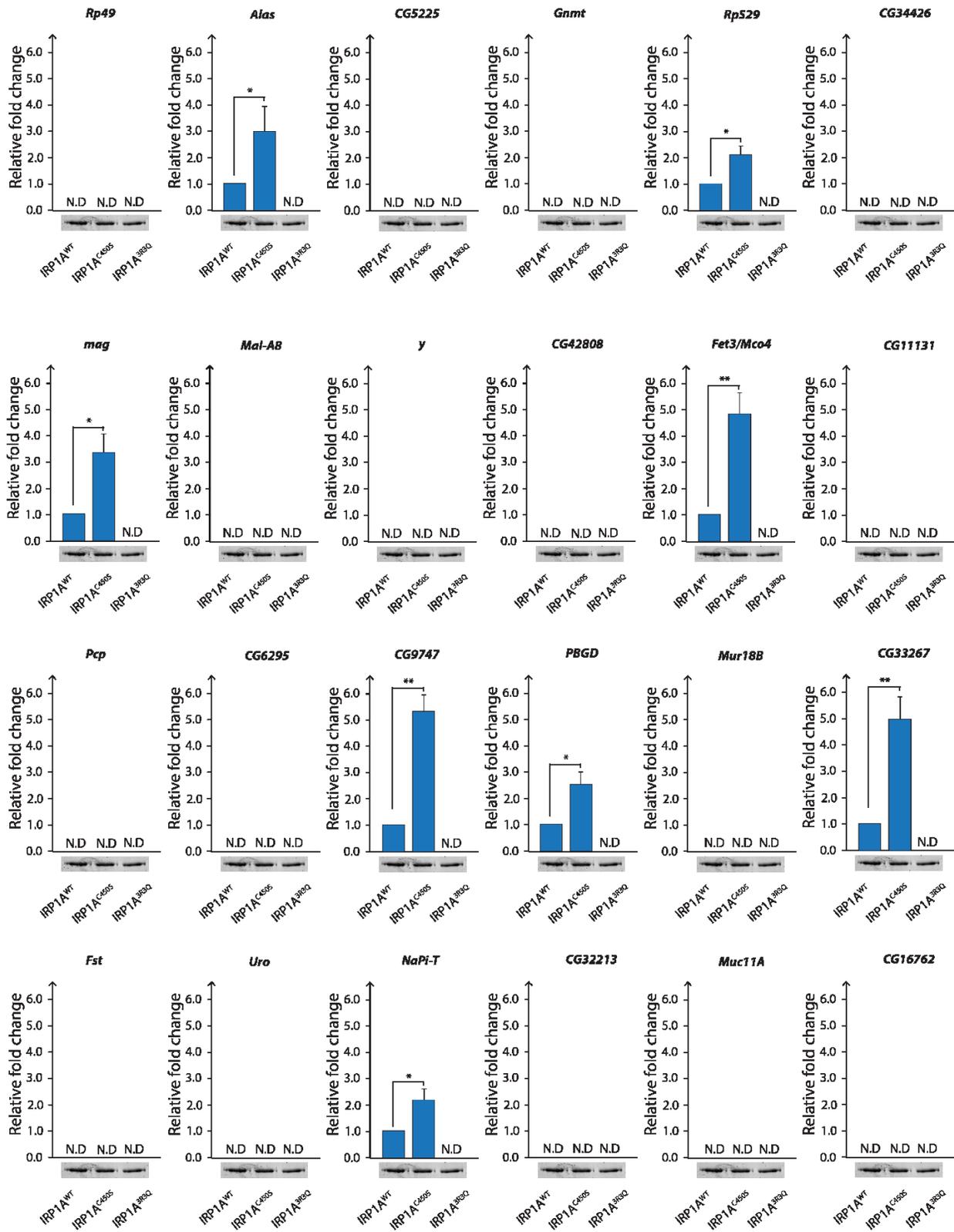
Figure 4.8 RIP-qPCR results for 23 candidate transcripts

140

We used CRISPR/Cas9-generated Flag-tagged knock-in alleles of IRP1A, namely IRP1A^{3F} ("wild type"), IRP1A^{C450S} (constitutively RNA-binding) and IRP1A^{3R3Q} (non-RNA-binding). *Rp49*, a control commonly used for qPCR in Drosophila, served here as a negative control for an abundant transcript (it encodes a ribosomal protein). The 23 candidate transcripts were obtained from RNA-Seq results overexpressing IRP1A^{C450S} and IRP1A^{3R3Q}.15 transcripts were not immunoprecipitated by IRP1A, and results are therefore pooled into one panel. Western blots below panel (all identical) show amount of IRP1A variant in sample. Samples were normalized so that all samples contained the same amount of IRP1A protein. All primers were validated prior to the experiment. Error bars are 95% confidence intervals. ND = not detected. * => $p<0.05$; ** => $p<0.01$.

Table 4.1 Summary of IRE derivations

classification	motif	sequence		position	specificity
		N_8	N_{14}-N_{19}	N_{14}-N_{18}	
canonical	1	C-----CAGUG(C/U/A/G)		C-G	
	2	C-----CAGAG(C/A/U)		C-G	
noncanonical	3	C-----CUGUG(U/C)		C-G	
	4	C-----CCGUG(A/U/C)		C-G	IRP2
	5	C-----CCGAGA		C-G	IRP2
	6	C-----CUUAGC		C-G	
	7	C-----CAAUGC		C-G	
	8	C-----CAGGG(A/C/U/G)		C-G	
	9	C-----UAGUA(A/C/U/G)		U-A	IRP1
	10	C-----UAGGAU		U-A	IRP1
	11	C-----UAGAA(U/C)		U-A	IRP1
	12	C-----UAGCAG		U-A	IRP1
	13	C-----GAGUC(G/A)		G-C	
	14	C-----GAGCC(G/A)		G-C	
	15	C-----GAGAG(U/G)		G-G	
	16	C-----GGGAG(A/C/G/U)		G-G	IRP2
	17	C-----GAGUG(U/A)		G-G	IRP1
	18	G-----CAGUGA		C-G	

Table 4.2 Humans disorders that are causatively linked to defects in the IRE/IRP system

human disease	mutation	phenotype
hereditary hyperferritinemia-cataract syndrome (HHCS)	mutations in the IRE of L-chain ferritin mRNA that impairs IRP binding	overexpression of serum ferritin in the absence of systemic iron overload or inflammation. Tendency for the development of bilateral cataract.
iron overload disorder with autosomal dominant transmission	point mutation in the IRE loop of H-chain ferritin mRNA that increase IRP binding	suppression of ferritin leading to an iron overload disorder phenotypically related to hemochromatosis.
sideroblastic-like anemia with iron overload	*GLRX5* deficiency leading to increased IRE-binding activity of IRP1 and suppression of *ALAS2* mRNA translation	development of a sideroblastic-like anemia with microcytosis and systemic iron overload.

Table 4.3 Enriched transcripts in IRP1A^{C450S} overexpression RNA-sequencing

gene name	FC IRP1A^{C450S} vs control	P-value	FC IRP1A^{WT} vs control	P-value	FC IRP1A^{3R3Q} vs control	P-value
CG5225	**375.4**	4.4E-04	2.66	0.04	2.3	0.14
Mco4	**164.5**	0.05	2.24	0.62	6.1	0.26
Mal-A8	**68.5**	0.03	11.96	0.08	7.0	0.07
CG16762	**66.9**	0.26	8.18	0.01	5.0	0.07
Fst	**64.0**	0.09	3.98	0.24	2.0	0.41
y	**39.6**	0.03	25.86	0.03	4.2	3.6E-03
Mur18B	**39.1**	0.09	8.61	0.08	1.8	0.51
Uro	**37.2**	0.10	2.73	0.08	3.6	0.11
Pcp	**20.6**	0.06	2.01	0.14	1.6	0.47
CG33267	**19.2**	0.45	-2.72	0.13	-3.0	0.16
Muc11A	**18.7**	0.21	3.19	0.26	2.0	0.46
mag	**17.8**	0.03	7.04	0.11	-2.9	0.08
CG34426	**13.7**	0.22	2.43	0.55	1.7	0.72
NaPi-T	**13.7**	0.13	2.50	0.64	1.6	0.66
Alas	**12.9**	2.6E-04	-1.01	0.92	-2.6	0.02
CG6295	**12.6**	0.07	2.72	0.50	1.1	0.86
CG11131	**10.7**	0.24	1.27	0.58	1.2	0.84
CG9747	**8.8**	0.07	2.79	0.11	1.1	0.86
CG42808	**8.7**	0.04	1.82	0.02	-4.5	3.5E-03
Gnmt	**6.6**	9.5E-03	1.00	1.00	-2.3	0.59
CG32213	**4.9**	0.19	1.54	0.85	-2.3	0.22
RpS29	**4.7**	0.03	3.51	0.23	-3.4	0.23
PBGD	**4.3**	0.08	1.52	0.22	-2.1	0.01

Table 4.4 SIRES analysis of top 23 candidates from *IRP1A^C450S* RNA-sequencing

gene name	potential IRE sequence from SIRES analysis (5' – 3')	IRE position	IRE score
CG5225	not detectable		
Mco4	not detectable		
Mal-A8	TGCAGAG**C**TCGTC**CAGTGA**AACGACCAAGGA	CDS (94-125)	high
CG16762	not detectable		
Fst	not detectable		
y	not detectable		
Mur18B	GCTCCTG**C**AGAGC**CAGAGA**GTTCTAGCACTT	CDS (781-812)	high
Uro	not detectable		
Pcp	not detectable		
CG33267	not detectable		
Muc11A	not detectable		
mag	not detectable		
CG34426	not detectable		
NaPi-T	GTCGAGG**C**GCGCA**CTGTGC**TGTGGTACATGA	CDS (113-144)	low
Alas	not detectable		
CG6295	not detectable		
CG11131	not detectable		
CG9747	not detectable		
CG42808	not detectable		
Gnmt	not detectable		
CG32213	not detectable		
RpS29	not detectable		
PBGD	CTTTGCT**C**GGTCG**TAGTAT**CGGCCAGCAGCT	5' UTR (20-51)	high

Chapter 5   **Examining the role of *ppk20* in iron and heme homeostasis during**

Drosophila larval development

5.1   Introduction

5.1.1   Epithelial sodium channel physiological roles

Water is one of the major components of living organisms. In humans, it accounts for approximately 68% of total body mass, with about 2/3 of water resides within the cell while the remaining 1/3 can be found in the extracellular space [269]. In many species, extracellular and intracellular fluids exhibit the osmolarity, which is determined mainly by the concentration of major ions like Na^+, K^+, Ca^{+2}, Mg^{+2}, Cl^-, HCO_3^-, PO_4^{3-}, and SO_4^{2-}. Osmolarity-dependent volume changes may affect cell morphology and stability [270]. The processes of absorption, secretion and excretion of water and solutes take place in epithelial cell layers that cover the body's internal and external surfaces.

In mammals, the epithelial sodium channel (ENaC) acts as a major participant in sodium and water homeostasis [271–273]. As a constitutively active channel, ENaC allows the flow of sodium ions from the lumen into the epithelial cell, across the apical cell membrane [272,274,275]. It is predominantly selective for sodium ions and exhibits a low single-channel conductance [276]. The absorbed Na^+ ions are then pumped out of the cell into the interstitial fluid by the action of Na^+/K^+ ATPase located on the basolateral membrane. This process allows ENaC to modulate the amount of Na^+ in the extracellular fluid (ECF), and thus, has a central role in the regulation of ECF volume and blood pressure [277,278]. The sodium-potassium ATPase therefore drives the overall process. This basic scheme accounts for sodium absorption in many epithelia, especially the distal nephron, the colon, and the lungs [275], where the channel participates in fundamental physiological functions. In the lungs, sodium absorption contributes to the maintenance of the normal composition of the surface liquid in the airways [279,280]. It is interesting that ENaC seems to be the only channel mediating apical sodium transport in some tissues, like colon [281],

cortical collecting duct of the kidney [278], and airway epithelia [282]. Several signaling pathways participate in the regulation of ENaC activity, including the renin-angiotensin-aldosterone system [277,278,283,284] and a complex variety of extracellular factors including Na^+, Cl^-, protons, shear stress and proteases [274,275,284]. Several inherited diseases involving alterations in sodium homeostasis, such as Liddle's syndrome and pseudo-hypoaldosteronism type I, are a consequence of ENaC mutations [278,285]. With its physiological importance, this characteristic may explain the very complex regulatory mechanisms controlling the channel activity.

5.1.2 ENaC structure and interaction

The subunits that form ENaC constitute a family within the ENaC/Degenerin superfamily. In addition to ENaC, this superfamily includes acid-sensing ion channels (ASICs)

[275,286–289], the pickpocket (ppk) genes in insects like fruit flies or mosquitoes [290], degenerin subunits involved in sensory transduction in nematodes such as worms [291], and peptide-gated Hydra Na^+ channels (HyNaC) in hydrozoans [292]. Overal, ENaC is composed of three to four homologous subunits: α, β, δ, and γ [278,285,293]. Each ENaC subunit possesses a large extracellular domain, which is the site of many regulatory effects. The intracellular domains correspond to the carboxyl and amino terminals of each subunit. Channel subunits undergo post-translational modifications by terminal processing of N-glycans and by proteolytic cleavage at defined sites within their extracellular domains [294–296] (Figure 5.1). The resolved crystal structure of ASIC1 has provided essential insights into the structural organization of ENaCs[297]. ASIC1 is a homotrimer, suggesting that ENaCs have an α1:β1:γ1 subunit stoichiometry. The extracellular region of ASIC1 is a highly ordered structure that resembles an outstretched hand containing a ball and has clearly defined domains termed wrist, finger, thumb, palm, knuckle, and β-ball. The extracellular regions within ENaC subunits likely have a similar domain organization.

ENaC interacts with a multiprotein regulatory complex that includes, among others, the ubiquitin protein ligase Nedd4 (neuronal precursor cell-expressed developmentally downregulated 4) and the serum and glucocorticoid-regulated kinase 1 (SGK1) [298] (Figure 5.2). The ENaC chains appear to assemble in the endoplasmic reticulum and then transport from the endoplasmic reticulum (ER) to Golgi, where they are processed by proteases [299]. Previous studies suggest the activation and processing of ENaC require steroid hormones like aldosterone. This hormone induces the expression of SGK1, which phosphorylates the ubiquitin ligase Nedd4-2. This phosphorylation will abolish the interaction between Nedd4-2 and ENaC and result in longer lifetimes of the ENaC subunits at the surface [300–302]. The whole process of ENaC trafficking requires the involvement of Rab proteins. [303–306] (Figure 5.2).

5.1.3 The *Drosophila* pickpocket genes are members of epithelial sodium channel (ENaC) family

In native tissues, ENaC is a rare protein with low expression levels [307,308]. Scientists could only carry out electrophysiological and limited immunofluoresence analyses but still struggle with biochemical studies on such low abundance proteins. Several research groups have investigated aspects of ENaC trafficking and protein stability in cell lines and heterologous expression systems, primarily in *Xenopus* oocytes, *Xenopus* A6 cells, and mammalian fibroblasts transiently expressing the ENaC subunits [309–311].

The mammalian genomes encode only eight to nine independent DEG/ENaC subunits, while the worms *Caenorhabditis elegans* and different *Drosophila* species harbor a significantly larger number of DEG/ENaC-like genes with 31 in *D. melanogaster* and 30 in *C. elegans* [312–316]. In *Drosophila*, the ENaC-like genes are referred to as the pickpocket genes and represent one of the largest ion channel families in the *Drosophila* genome. The pickpocket protein

sequences in *Drosophila* share approximately 30-40% similarity with their homologs in humans and only 25-33% between pickpocket polypeptides. The remarkable diversity of *ppk* genes in *Drosophila* suggests two potential hypotheses: i) DEG/ENaC ion channels serve a wider range of physiological functions in *Drosophila* than their mammalian counterparts, ii) it is possible that each pickpocket channel in *Drosophila* evolved to serve highly specialized functions and might be responsible for more narrow physiological functions than the mammalian ENaC. However, identifying physiological and functional homology between family members across distant species is often impossible due to the poor overall protein sequence conservation of the extracellular loop domains. Thus, protein alignment analyses alone are typically not sufficient to draw physiological homology conclusions.

5.1.4 Steroid-regulated ENaC expression

Most of what we know to this day about the ENaC family comes from studies in mammals. Overall, ENaC is regulated by a variety of extrinsic and intrinsic factors through changes in plasma membrane abundance of inherent activity [284,296]. Its plasma membrane abundance is controlled primarily through changes in intracellular trafficking, and ENaC activity is controlled mostly by changes in open probability, which is profoundly influenced by proteolytic processing and by interactions with cytoplasmic domains with specific membrane acidic phospholipids. Regulatory effects on open probability and membrane abundance are not mutually exclusive and raise the exciting possibility that both could be operative and linked through trafficking-dependent events [317]. The expression and activation of the ENaC family is regulated by steroid hormones like aldosterone, glucocorticoid or progesterone [283,298,318–321]. Transcriptional regulation of ENaC subunit expression constitutes another crucial regulatory mechanism, particularly for the consolidation of aldosterone-stimulated Na^+ transport [322].

In mouse, renal Na^+ reabsorption is highly controlled by the mineralocorticoid aldosterone, which acts primarily through the mineralocorticoid receptor to alter the transcription of a set of target genes [323]. As with other steroid-regulated processes [324], two major classes of target genes have been identified: early and late. Early response genes appear to be required for initiation of the response, whereas late response genes participate in consolidation [325]. The latter include components of the ion transport machinery itself, including ENaC and Na^+,K^+-ATPase subunits. The late response genes also encode regulatory proteins that likely act to limit the extent of the aldosterone response, such as activators of the MAPK cascade, including the epidermal growth factor receptor [326]. Early response genes include primarily signaling molecules implicated in pathways that control ENaC activity and trafficking [323].

5.1.5 Relationship between ENaC, hypoxia and iron/heme homeostasis

Certain tissues in animal's bodies experience a reduction in oxygen level in different conditions, including high altitude or certain disease states, such as myocardial infarction and stroke in human patients. To survive, cells, tissues, and organisms have developed various strategies to adapt to such changing oxygen conditions. There are indeed major differences between different organisms and cells in their ability to survive reduced environmental O_2. For example, turtle neurons are very tolerant to low oxygen and can survive without oxygen for hours and days [327,328]. On the other hand, mammalian neurons are sensitive to reduced oxygen and cannot survive for even minutes under similar conditions [148]. However, the mechanisms underlying survival in such extreme hypoxic conditions are not clear at present, although there have been several interesting observations in this regard in the past few decades. For instance, some hypoxia-tolerant animals like the red-eared slider *Pseudemys scripta* and *Crucian Carp* fish reduce their oxygen consumption during hypoxia as a way to minimize the mismatch between

151

oxygen supply and demand [329,330]. Similar phenomena were observed in *Drosophila melanogaster* [331–333] and newborn mammals [334]. Many questions, however, remain unsolved. For instance, we do not have an adequate understanding of the mechanisms responsible for reducing metabolic rate during low O_2 conditions. Similarly, the mechanisms responsible for coordinating the suppression of these metabolic processes remain mostly unknown.

Hypoxia is involved in a wide variety of biological processes, including heme and iron homeostasis. In a recent study, Schwartz *et al*. elucidated more precisely the mechanisms by which HIF-2 mediates regulation of iron absorption using the hypoxia inducible mouse models. They investigated the importance of the well-described hepcidin/FPN axis to stabilize HIF-2 in the intestine. In a context of a HIF-2 knockout mouse model for hepcidin in the liver, duodenal FPN was stabilized, and serum iron increased (leading to iron overload in case of sustained hepcidin deficiency mimicking the hereditary hemochromatosis) [335]. FPN stabilization leads to a decrease in intracellular iron in the duodenal enterocytes. Under low cellular iron concentrations, the HIF Prolyl Hydroxylases (PHDs) no longer hydroxylate HIF-2 and thus destabilizes the protein, results in the consecutive transcription of the iron absorption genes (DMT1, DcytB, FPN). Both iron overload and anemia models support the crucial role of the hepcidin/FPN axis in intestinal HIF-2 regulation [335]. Using inducible intestine-specific FPN and DMT1 knockout model, Schwartz *et al*. confirmed that enterocyte iron flux was the major mechanism by which the hepcidin/ FPN axis regulated HIF-2α and demonstrated *in vitro* that the FPN-mediated efflux of iron triggers the stabilization of HIF-2α in a cell-autonomous manner. Interestingly, the use of a recently developed HIF-2 antagonist decreases systemic iron accumulation in hepcidin-deficient mice, confirming previous studies using mice lacking HIF-2 in the intestinal epithelium [336,337]. Noteworthy, besides PHD-mediated posttranslational regulation, HIF2-α is subjected to IRP-

mediated translational regulation due to the presence of an IRE in its 5′UTR [153]. Altogether, these studies confirm that HIF-2α is a potential pharmacological target downstream of the hepcidin/FPN axis in patients with iron overload.

As a "feeding machine", the *Drosophila* larvae spend most of their time inside the food. Thus, they experience a low oxygen environment filled with fly food and water throughout most of their larval stages. Therefore, hypoxia also has a big impact on *Drosophila* development. Similar to other insects, *Drosophila* larvae establish a trachea system that carries air-filled tubes that allow direct delivery of oxygen and removal of carbon dioxide to and from target tissues. This system is similar to the trachea branching system in the vertebrate's lungs. In mammals, the sodium transport in bronchial epithelial cells is involved in the accumulation of iron in the tissue during the liquid clearance process [338]. Similar to their vertebrate counterparts, the *Drosophila* pickpocket genes also regulate tracheal fluid clearance [315], suggesting they also work under the same mechanism regarding regulating iron trafficking in air-filled systems.

5.2 Modified materials and Methods

5.2.1 *Drosophila* stocks and husbandry

I obtained the following stocks from the Bloomington *Drosophila* Stock Center: *w^1118* (#3605), *UAS-ppk*-RNAi (#29571), *UAS-rpk*-RNAi (#39053), *UAS-ppk3*-RNAi (#61995), *UAS-Nach*-RNAi (#62894), *UAS-Nach*-RNAi (#27262), *UAS-ppk5*-RNAi (#25816), *UAS-ppk6*-RNAi (#53010), *UAS-ppk6*-RNAi (#25880), *UAS-ppk7*-RNAi (#25292), *UAS-ppk8*-RNAi (#25814), *UAS-ppk9*-RNAi (#25892), *UAS-ppk10*-RNAi (#27256), *UAS-ppk11*-RNAi (#26253), *UAS-ppk12*-RNAi (#27092), *UAS-ppk13*-RNAi (#25817), *UAS-ppk14*-RNAi (#27091), *UAS-ppk15*-RNAi (#28012), *UAS-ppk16*-RNAi (#25890), *UAS-ppk17*-RNAi (#58557), *UAS-ppk18*-RNAi

(#61949), *UAS-ppk18*-RNAi (#27240), *UAS-ppk19*-RNAi (#58203), *UAS-ppk19*-RNAi (#25887), *UAS-ppk19*-RNAi (#58203), *UAS-ppk19*-RNAi (#25887), *UAS-ppk20*-RNAi (#25897), *UAS-ppk21*-RNAi (#25849), *UAS-ppk21*-RNAi (#62487), *UAS-ppk22*-RNAi (#61821), *UAS-ppk22*-RNAi (#28706), *UAS-ppk23*-RNAi (#28350), *UAS-ppk24*-RNAi (#26006), *UAS-ppk25*-RNAi (#28088), *UAS-ppk26*-RNAi (#25825), *UAS-ppk27*-RNAi (#27239), *UAS-ppk28*-RNAi (#31878), *UAS-ppk29*-RNAi (#27241), *UAS-ppk30*-RNAi (#25810), *UAS-ppk31*-RNAi (#44013), *UAS-ppk31*-RNAi (#27087), *Tubulin-Gal4/TM3, Sb^1 Ser^1* (#5138), *UAS-hTfR^{WT}.GFP* (#36858), *UAS-hTfR^{7C}.GFP* (#36858), *UAS-FLP* (#4539), *Vas.Cas9* (#51323). Please note that for *Tubulin-Gal4* line, this is not the line being used by other lab members due to lack of origin information. Instead, a new tubulin Gal4 line was obtained from Bloomington stock center and the *TM3, Sb^1 Ser^1* balancer was switched to *TM3, Sb^{1}Ser^{1}GFP* which allows selection at the larval stage.

I obtained the following stocks from the Vienna *Drosophila* Resource Center: *UAS-ppk*-RNAi (#108683), *UAS-rpk*-RNAi (#105463), *UAS-rpk*-RNAi (#8549), *UAS-ppk3*-RNAi (#104593), *UAS-Nach*-RNAi (#106647), *UAS-Nach*-RNAi (#45920), *UAS-Nach*-RNAi (#45921), *UAS-ppk5*-RNAi (#101664), *UAS-ppk5*-RNAi (#48289), *UAS-ppk5*-RNAi (#48290), *UAS-ppk6*-RNAi (#101091), *UAS-ppk7*-RNAi (#100643), *UAS-ppk7*-RNAi (#7900), *UAS-ppk8*-RNAi (#47047), *UAS-ppk8*-RNAi (#47048), *UAS-ppk9*-RNAi (#104952), *UAS-ppk9*-RNAi (#23391), *UAS-ppk9*-RNAi (#17213), *UAS-ppk11*-RNAi (#107741), *UAS-ppk11*-RNAi (#330319), *UAS-ppk12*-RNAi (105131), *UAS-ppk13*-RNAi (#110084), *UAS-ppk13*-RNAi (#9494), *UAS-ppk14*-RNAi (#110258), *UAS-ppk14*-RNAi (#7903), *UAS-ppk14*-RNAi (#7904), *UAS-ppk15*-RNAi (#109855), *UAS-ppk15*-RNAi (#42523), *UAS-ppk16*-RNAi (#22989), *UAS-ppk16*-RNAi (#22990), *UAS-ppk17*-RNAi (#109927), *UAS-ppk17*-RNAi (#8596), *UAS-ppk18*-RNAi (#13209), *UAS-ppk19*-RNAi (#107638), *UAS-ppk19*-RNAi (#33660), *UAS-ppk20*-RNAi (#33659), *UAS-*

154

ppk21-RNAi (#107892), *UAS-ppk21*-RNAi (#1345), *UAS-ppk22*-RNAi (#106384), *UAS-ppk23*-RNAi (#106873), *UAS-ppk23*-RNAi (#39580), *UAS-ppk23*-RNAi (#39581), *UAS-ppk24*-RNAi (#102923), *UAS-ppk24*-RNAi (#30196), *UAS-ppk25*-RNAi (#101808), *UAS-ppk25*-RNAi (#7343), *UAS-ppk26*-RNAi (#100834), *UAS-ppk26*-RNAi (#5509), *UAS-ppk27*-RNAi (#7470), *UAS-ppk28*-RNAi (#100946), *UAS-ppk28*-RNAi (#44412), *UAS-ppk29*-RNAi (#106888), *UAS-ppk29*-RNAi (#330294), *UAS-ppk30*-RNAi (#105896), *UAS-ppk30*-RNAi (#1351), *UAS-ppk31*-RNAi (#106385), *UAS-ppk31*-RNAi (#1269).

I used CRISPR/Cas9 to generate the following knock-in and knock-out alleles (Figure 5.3): $ppk20^{FCH}$, $ppk20^{KO}$/*TM6B, Hu Tb*. I also generated transgenic lines based on the PhiC31 system: *UAS-3xHA-ppk20^{PA}*, *UAS-ppk20^{PA}-mVenus, UAS-ppk20^{PC}-3xHA*, *UAS-mVenus-ppk20^{PC}*, $ppk20^{gRNA}$ (Table 5.1, Table 5.2). $y^1w^*P($nos-PhiC31\int.NLS$)X;$ $P($carryP$)attP40(II)$ and $y^1w^*P($nos-PhiC31/int.NLS$)X; P(carryP)attP2(III)$ were gifts from BestGene Inc.

Phm22-Gal4 was a kind gift from Michael O'Connor's lab. Stocks were maintained on a standard cornmeal diet unless otherwise specified.

5.2.2 Generation of *ppk20* cDNAs

Unlike many other genes for which cDNAs are available at *Drosophila* Genomics Resources Center (DGRC), no cDNA clone was available for *ppk20* by the time this project was done. Even though according to FlyBase, there are three isoforms of *ppk20* with provided sequences, to the best of our knowledge, the lack of established cDNAs makes me suspect if the database in FlyBase is still reliable. To solve this problem, I generated the *ppk20* cDNAs. Total RNA was extracted from five third instar larvae at the age of 40-42 hours after molting, followed by cDNA synthesis reaction using AMV Reverse Transcriptase (NEB M0277). The cDNA product was then amplified using standard PCR procedure and cloned into pENTR/D backbone via Gibson

155

reaction [97]. Gibson reaction products were transformed into DH5α *E.coli* competent cells. 20 colonies were picked for further analysis by Sanger sequencing. Any colony with different sequence was kept separately as different potential isoforms.

5.2.3 Generation of transgenic constructs

For transgene properties, see Table 5.1. cDNAs generated in section 4.2.4 gave me three isoforms for *ppk20* with similar sequences as to those reported in Flybase. As a result, I will refer to each isoform at the same label as Flybase. Since three *ppk20* isoforms were identified, it is possible that each isoform either plays the same roles or different roles depending on tissues and developmental time points. As a result, it worth to generate tools allow us to distinguish the function for each individual isoform. To generate equivalent expression of transgenic constructs, I used PhiC31 vectors pUAST.attB (DGRC #1419) to generate *ppk20* RA cDNA with 3xHA N-terminal tag and pBID-UASC-GV (Addgene #35204) to generate *ppk20* RA and RC cDNAs with mVenus C-terminal tag. I also used pUASg-HA.attB to generate *ppk20* RC cDNA with 3xHA C-terminal tag. pUAST.attB and pUASg-HA.attB vectors were a kind gift from Johannes Bischof, Basler lab, University of Zurich, Germany [339], pBID-UASC-GV was a gift from Brian McCabe lab at EPFL Brain Mind Institute, Switzerland [98]. Vector backbones were amplified via PCR to generate two fragments per vector and fused to cDNA fragments via the Gibson reaction. Mutagenesis were done to generate 3xHA tag via Q5 based mutagenesis PCR (NEB M0491S) following the standard procedure. Fused fragments were cloned into DH5α *E.coli* competent cells, and validated by Sanger sequencing.

5.2.4   Other experiments

Other experiments, including qPCR, RNA-sequencing, embryo injection, larval injection and western blot were done as described in chapter 2 with no major modifications.

5.3   Results

5.3.1   Pickpocket 20 is required for *Drosophila* developmental transition

ppk20 was among the hits from the described above genome-wide screening which aimed to identify genes with undiscovered function in heme or iron homeostasis. Similar to what happened when knocking down *PPOX*, a gene involved in heme biosynthesis, PG-specific knockdown (KD) of *ppk20* also resulted in developmental defects with the majority animals arrested at either L2 or L3 larval stage with very few pupae and around 1% can get to adult. The arrested animals also display a porphyria-like phenotype with red autofluorescence ring gland (RG) (Figure 5.4A). Since the first screening was done using PG-driven of a single RNAi, it is important to validate the observed phenotype. I tried to validate this phenotype by looking for another independent RNAi (IR) targeting *ppk20* coding sequence. By the time I conducted this study, there was only another RNAi transgene available from the Bloomington *Drosophila* stock center (# 25897). This second RNAi (hereafter refers as RNAi 2 or IR^2), however, has a partial overlapped region with the first RNAi. Even though PG>IR^2 still resulted in red RG, it is not convincing enough to conclude if the porphyria phenotype was real. Besides, the developmental defects of the animals appeared to be milder than the first RNAi being used, with the majority arrested at L3 larval stage or pupariation delay (Figure 5.4C). As a result, I decided to use a different strategy and used CRISPR/Cas9 to generate 1) a classic deletion of the whole *ppk20* gene (namely *ppk20^{KO}*) and 2) a conditional mutant where I replaced the endogenous *ppk20* with a

ppk20 allele flanked by FRT sites on both ends of the gene (namely *ppk20^{FCH}*). For the second line, by expressing FLP specifically in the PG, I can remove the whole gene in this tissue only. Both CRISPR approaches, intriguingly, resulted in the red RG (Figures 5.3 and 5.4B). These data together confirmed that the porphyria-like phenotype observed in *ppk20* impaired animals is real.

In addition to the porphyria phenotype in RG, PG-knockdown of *ppk20* also resulted in the trachea necrosis (Figure 5.5A). The Gal4 transgene I used for this study was *phm22-Gal4*, which is supposed to be explicitly expressed in the PG. Through personal communication with Michael O'Connor's lab, where this line was generated, I know *phm* is also expressed at a low level in the trachea. Thus might explain why I observed phenotype outside of the tissue. This trachea necrosis is unique to *ppk20* since it is the only gene among our 34 candidates which RNAi resulted in trachea necrosis. The phenotype can be linked to the known function of its family in liquid clearance [315]. Like mammalian airways, the *Drosophila* tracheal system is also a branching network of tubular epithelia with an important role in delivering oxygen to the organism. Especially during *Drosophila* larval development, these animals are feeding constantly and often immerse themselves deeply into the food. The trachea system in this organism has to be properly regulated to ensure enough oxygen is delivered to target tissues throughout the body. Previous studies have shown that in mammals, the ENaCs are involved in liquid clearance from the airways from the time of birth and remove liquid from the airspaces in adults [338]. Earlier efforts by researchers confirmed that the *Drosophila ppk* family also plays an vital role in the organism tracheal liquid clearance [315]. Whether this trachea necrosis is linked to the porphyria phenotype of *ppk20* loss-of-function remains unknown. However, other studies showed the activation mechanism for ENaC family, which requires oxygen and heme oxygenase, thus there is a potential role of ENaC, in this case is *ppk20*, which links hypoxia and heme production in *Drosophila*.

5.3.2 pickpocket 20 is the only validated *Drosophila* ENaC member with a porphyria phenotype

Since *ppk20* has never been studied in detail before, its functions are unknown. We can only predict its functions based on studies of other ppk genes. PG-specific KD of *ppk20* resulted in a porphyria phenotype, which can be linked to heme biosynthesis. Three possible scenarios can explain this phenotype: i) related to the known functions of ENaC family in sodium absorption and tracheal liquid clearance, ii) a moonlighting function of ppk20 in heme/iron biosynthesis or iii) the combination of i) and ii). I first hypothesized that if the porphyria-like phenotype in *ppk20*-impaired animals is related to the known functions of ENaC family, knockdown other ENaC family members will result in porphyria phenotype. However, the first screening done in collaboration with Michael O'Connor's and Kim Rewitz's labs did not suggest any other pickpocket candidates [66,67]. Since additional genetics tools were generated since that RNAi screen, I later repeated this experiment with the addition of other RNAi from Vienna *Drosophila* Resource centre (VDRC), Bloomington *Drosophila* stock center and Japan National Institute of Genetics (NIG). In my hands and in agreement with the previous screening, PG-specific KD of other ppk genes did not result in any significant developmental as well as tissue morphological defects (Table 5.3), with the exception of one RNAi targeting *ripped pocket* (*rpk*). *PG>rpk*-RNAi resulted in L3 arrested with a mild porphyria phenotype. However, this phenotype was only captured in one *rpk*-RNAi while none of the other RNAi exhibited any phenotype in RG. Meanwhile, two RNAi against *ppk20* result in porphyria phenotype as described above (Figure 5.4A).

Considering that the RNAi approach might not be a reliable approach to study genes functions, I further tested a different approach using somatic CRISPR. In this line of experiment,

I generated a PG-specific Cas9 and combined it with a transgenic line that carries ubiquitous gRNA target individual pickpocket genes [199] (chapter 6). I first examined if the gRNAs can efficiently recruit Cas9 and generate double-strand breaks (DSBs) by crossing with an *act-Cas9* line in combination *lig4* mutant (*lig4^{169}*) (Bloomington #58492). *lig4* is a gene located on the X chromosome and encodes a lig4 enzyme responsible for repairing the DSBs generated by Cas9 nuclease activity [340]. The homozygous *lig4^{169}* females are viable under normal conditions. However, combining these females with a male carries gRNA will result in the male progeny hemizygous for *lig4* mutant allele. The DSBs generated by Cas9 cannot be repaired in males while the female progeny can do this. Therefore, if the sgRNA works efficiently to generate DSBs, one would expect no male progeny can survive to adulthood (Table 5.4). Similar to the RNAi approach, only *PG-Cas9>ppk20*-gRNA showed the porphyria phenotype while no other genes in ppk family exhibit any phenotype (Table 5.4).

These data together suggest the porphyria phenotype is unique to *ppk20* loss-of-function animals among the *Drosophila* ENaC family. However, this does not rule out the possibility that the porphyria phenotype was generated partially due to the known function of ENaC family. Especially members of the ENaC family are reported to form a trimer or tetramer subunit to function, interruption of only one gene at a time might not be sufficient to fully understand each gene's importance in this process. A future investigation testing multiple pickpocket genes is necessary and can be done by combining different RNAi or gRNA transgenes into the same animals.

5.3.3 Characterization of *ppk20* and *ppk* family in *Drosophila* larval development as well as porphyria animals

The expression pattern of *ppk* genes family were mined from FlyAtlas, FlyBase, FlyMine. Ou Qiuxiang, a previous Ph.D. student as well as previous PostDoc in King-Jones lab, has conducted transcriptome analysis of *Drosophila* genes in the whole body (WB) as well as ring gland during third instar larval stage. In all approaches, I noticed a significantly low expression level of pickpocket genes family (Table 5.5). In particular, the *ppk20* gene is among the genes with the lowest expression level in this family, only equal 1/10 to 1/30 of the pickpocket gene with the highest expression level (*ppk13*) (Table 5.5). In agreement with this low expression profile, I also observe a low expression level of *ppk20* polypeptide. Using roughly 750 L3 *ppk20^{FCH}* larvae, I could slightly detect a band corresponding to ppk20 polypeptide size (Figure 5.5B). I know this is the right band because the protein extract from the same number of *ppk20^{KO}* larvae failed to detect this band (Figure 5.5B). Interestingly, when running these protein lysates on the native gel (non-denaturing), result in the protein band of approximately 200kDa, three times the expected polypeptide weight. This result suggests that ppk20^{FCH} more likely to form a complex of three ENaC monomers in the cell membrane (Figure 5.5C). It remains unclear whether they are identical monomers (all ppk20^{FCH}) or require polypeptides of other pickpocket proteins.

On the other hand, the expression of most ppk genes are often mis-regulated in porphyria animals (Table 5.6). This table is a summary of RNA-seq done by me as well as other lab members during their programs, including *PG>AGBE*-RNAi, *PG>AGBE^{FCF}*, *PPOX* mutant, *PG>FeCH*-RNAi, *PG>spz5*-RNAi, *PG>NOS*-RNAi, *PG>DHR51*-RNAi, *PG>ZFRP8*-RNAi, *PG>CG8145*-RNAi, all were reported with the porphyria phenotype.

5.3.4 Dietary rescue of *PG>ppk20*-RNAi animals

PG>ppk20-RNAi animals were mainly arrested at either L2 or L3 larvae with only 3.5% of animals can become pupae and 1% of the original embryos can develop to adult (Figure 5.4C). Since PG is the tissue responsible for ecdysone production, I wondered if knocking down *ppk20* in this tissue affects ecdysone synthesis pathway and thus, affects the animal development. I fed the *ppk20*-impaired animals with different supplements from ecdysone synthesis pathway, including (i) cholesterol as the starting sterol in the diet, (ii) 7dC, an intermediate sterol and (iii) 20E, the active form of ecdysone. Animals supplemented with cholesterol showed a similar developmental trend as in the control medium, with the majority of L2 and L3 larval arrested, with 1.5% pupae and 0% adults. 7dC and 20E supplementation, on the other hand, can partially rescue the animals to later developmental stages, with 12% pupae and 7% adult in 7dC or 16% pupae and 10% adult in 20E (Figure 5.4D). However, these data can only provide evidence for insufficient ecdysone production in *ppk20*-impaired animals. Since adding ecdysone intermediates, especially 20E could not successfully rescue the developmental defect, it is more likely that ecdysone production is not the main issue in these animals.

On the other hand, the porphyria phenotype in ppk20 loss-of-function can result from two possible scenarios: (i) failure to synthesize the protoporphyrin IX, an essential component of heme or (ii) incapability to incorporate iron into porphyrin ring. To test which scenario fits into *ppk20*-impaired animals, I tried to supplement animals with different components that are produced as intermediates in the heme biosynthesis pathway. Starting with the same number of embryos, I raised animals on different supplements and quantified survival throughout development. First, I supplemented *ppk20*-impaired animals with hemin, a compound similar to heme, which contains a porphyrin ring as well as iron. However, heme carries ferrous ion while hemin carries ferric ion.

Suprisingly, this approach can improve animals survival from 3% pupation in control media to 42% pupae in hemin supplemented (Figure 5.4D), and from 1% adult in control media to 24% adult in hemin supplemented media (Figure 5.4D). This result further confirms the developmental defect in *ppk20*-impaired animals comes from the failure to complete heme biosynthesis. The incomplete rescue rates can be explained because I only tried at a concentration of 1mM hemin. Even though this concentration was sufficient to rescue *PPOX* loss-of-function animals, this dose might not be optimal for *ppk20*-impaired animals. Different hemin concentrations might result in varying levels of rescue.

I also raised animals on iron-rich media. This supplement, however, could not improve the survival rate of animals since most of them were still arrested at L2/L3 larval stage. Interestingly, when I raised these animals on media supplemented with iron chelator, the survival rate did not improve but the trachea necrosis rate increased. All of these data indicated that dietary iron could not rescue *ppk20*-impaired animals. However, these results do not indicate that the porphyria-like phenotype in these animals is unrelated to iron homeostasis. Studies in ENaC family indicates members of ENaC membrane proteins. If ppk20 is a key component in iron trafficking, loss of ppk20 may interfere with the ability of the cells to intake enough iron for their needs. Thus, feeding iron will not be sufficient since there is no way for the cell to import this metal inside. Further investigation is needed to figure out the right function of *ppk20*.

5.3.5 Injected ferritin can partially rescue the porphyria phenotype in *ppk20* loss-of-function animals

Data from supplemental rescue reject the scenario where porphyria phenotype in *ppk20*-impaired animals comes from a failure to synthesize porphyrin ring during heme synthesis as hemin but not zinc porphyrin could rescue the animals (Figure 5.4D). On the other hand, failure to

163

rescue *ppk20*-impaired animals with dietary iron could not reject the potential function of this gene in iron regulation since if ppk20 was important for iron trafficking, excess iron in the diet would not be delivered properly to target tissues and cannot help relieve the phenotype, especially when ppk20 is a member of ENaC family, which is known to localize to the cell membrane. One hypothesis is that disruption of ppk20 function also interferes with the cell's ability to import iron from the diet. On the other hand, the ability to rescue mutant animals with hemin can be explained that the heme pathway is still functional in these animals. This pathway allows a partial increase of iron for any iron-dependent processes, including heme biosynthesis. I decided to investigate the potential role of this gene in iron/heme regulation by other approaches. One of them was to feed the animals with ferritin in the diet. In my hands, adding ferritin extracted from horse or human spleen (namely ferritin^{ho} and ferritin^{hu}, respectively) could not rescue *ppk20^{KO}* animals (Figure 5.5E). However, one issue with this approach is these ferritin are proteins so they might be degraded when preparing media, or ferritin may not be stable in the diet for prolonged times. Furthermore, in this approach, ferritin will first be taken through the digestive system, where the protein is exposed to different digestive enzymes. Indeed, ferritin supplements failed to rescue *Fer1HCH* mutant animals (Figure 5.5D), suggesting this is not the proper way to test the effect of dietary ferritin on animals' survival.

In an attempt to deliver ferritin into the body cavity of larvae, I injected ferritin directly to the larvae (see methods section). As a positive control, I injected ferritin^{ho} and ferritin^{hu} into *Fer1HCH* mutants. Without injected ferritin, 60-70% of *Fer1HCH* mutants were arrested at L1 with no animals can survive to the next developmental stage. Interestingly, injected ferritin^{ho} and ferritin^{hu} can partially rescue *Fer1HCH* mutant with 27% and 43% animals can reach adulthood, respectively (Figure 5.5D). These data suggest that injection of ferritin is a promising approach. I

then tried the same approach with *ppk20^{KO}* animals. Surprisingly, both ferritin sources can partially

rescue *ppk20^{KO}* animals to adulthood but with different efficiencies of 21% for ferritin^{ho} and 33%

for ferritin^{hu} (Figure 5.5E). This result is intriguing because it suggests another direct link between

ppk20 and iron/heme homeostasis. However, the rescue efficiency is still moderate and indicates

that the rescue is not complete.

This low rescue efficiency from injected ferritin can be explained by two scenarios: i)

ferritin is not the only factor that affects the animal's survival and ii) there are some differences in

functions of mammalian ferritin and insect ferritin. The first scenario can be tested by generating

a ferritin^{ho} and ferritin^{hu} transgenes and drives their expression in *ppk20*-impaired animals to see if

they can recapitulate the earlier reported rescue. Each ferritin transgenic line carries both heavy

chain and light chain of ferritin cDNA from the same origin (horse or human) so that both

polypeptides will be produced in sufficient amounts and can form a functional complex.

Interestingly, ubiquitous overexpression of either ferritin^{ho} or ferritin^{hu} failed to rescue the

developmental defect in *ppk20^{KO}* (Figure 5.5E). I know the ferritin transgenes are functional since

they can rescue the *Fer1HCH* mutant (Figure 5.5D), indicating that mammalian ferritin can form

a functional complex in *Drosophila* cells. Failure to rescue *ppk20^{KO}* by ferritin transgenes can be

explained by two scenarios: (i) the amount of iron being stored in the ferritin complexes remains

low and not sufficient for the rescue, and (ii) the ability to utilized the stored iron in these ferritin

complexes for any iron-dependent process. This explanation can be linked to the second scenario

for the partial rescue effect from injected ferritin. Mammalian ferritin acts mostly as a cellular iron

storage while there are some evidence support the function of insect ferritin in iron secretion and

trafficking [31,341]. In *Drosophila*, Fer1HCH consists of 9 different isoforms with different

characteristics regarding the presence of IRE in transcript sequence or the secretion signaling

peptide [124,342]. These isoforms may play different roles during iron homeostasis. Thus, the difference in function makes mammalian ferritin not the ideal ferritin to rescue *ppk20*-impaired animals. Indeed, even though they can significantly rescue *Fer1HCH* mutant, the effect is not complete either and can come from the missing function in these complex, including iron trafficking (Figure 5.5D). A future study in which injection or overexpression of each *Drosophila* ferritin isoform in *ppk20*-impaired animals might give us a better understanding for this rescue effect. Furthermore, one can also test other insect ferritins since they are predicted to work under a similar mechanism for iron storage and trafficking.

5.3.6 *ppk20*-impaired animals can be rescued by overexpression of human transferrin receptor but not by IRP1A overexpression

I now investigate *ppk20* function in (a) cellular iron homeostasis, iron trafficking, either (b) free iron or (c) via iron-bound proteins. I first wondered if the overexpression of IRP1A, an essential protein for this process, can rescue the phenotype. Earlier, I have shown that IRP1A plays an important role in *Drosophila* cellular iron homeostasis [166]. I also showed that IRP1A overexpression could rescue *AGBE* loss-of-function animals and study the function of this glycogen branching enzyme in cellular iron homeostasis (Figure 3.7) [166]. Interestingly, PG-overexpression of IRP1A, either wild-type (IRP1A[WT]) or apo-form (IRP1A[C450S]), or IRP1B, failed to rescue *PG>ppk20*-RNAi animals (Figure 5.6A-C). This result rules out the possibility that ppk20 has an essential role in cellular iron homeostasis. I have also shown that injected ferritin[ho] and ferritin[hu] can partially rescue the animals, suggesting a potential link between ppk20 and iron trafficking. Once injected, ferritin could circulate throughout the animal's serum. Later on, target cells can partially intake these ferritins via an unknown mechanism and use their stored iron for dependent processes. Most iron trafficking in mammals is taken care by the transferrin (Tf) and

166

transferrin receptor (TfR). Upon binding with iron, Tf is delivered to target cells and bound to TfR on the cell membrane before being imported. In *Drosophila*, there are three Tf (1-3) while no TfR was ever identified. Since ppk20 is a membrane-bound protein, I wondered if it may work together or in parallel with the unidentified *Drosophila* TfR. This rationale is based on the similar regulatory mechanisms between mammalian ENaC and TfR where both proteins trafficking are regulated by Rab proteins including Rab5, Rab7, Rab9 and Rab11. In addition, the mammalian copper metabolism Murr1 domain containing 10 (COMMD10) initiates the first step during cytosol localization of both TfR and ENaC [343,344]. I wondered if expressing human TfR (hTfR) can rescue *ppk20*-impaired animals (Figure 5.6 D,E). Intriguingly, PG>hTfR could significantly rescue *PG>ppk20*-RNAi animals. On the other hand, I also overexpressed a mutant form of hTfR which lacks the cytoplasmic signaling sequence, resulting in the protein only localizes in the cell membrane without trafficking to the cytosol. This mutant transgene could only partially rescue the animals (Figure 5.6 D,E). This data suggests a potential link between the uncharacterized *Drosophila* TfR and ppk20, where they might share a common trafficking pathway between cytoplasm and cell membrane.

5.3.7   Transcriptome analysis of *ppk20^{KO}* animals

ppk20 loss-of-function animals exhibit developmental defect with major larval arrest. Supplemental rescue of either steroid hormone intermediates, heme derivatives, or iron-related supplement could only partially rescue the animals. I suspect multiple biological processes in the animals were affected. Based on the above data, ecdysone synthesis and heme/iron homeostasis were disrupted while based on literature and some preliminary data, hypoxia-dependent biological processes might also be interfered. To have a general picture of how *ppk20* dysfunction can affect cellular pathway, I conducted a RNA-sequencing experiments where 75 ring gland of *ppk20^{KO}*

animals at 18-hour L2 stage were manually dissected per sample with three biological replicates. Extracted RNA from these samples were used to prepare libraries for sequencing, and transcription profile were compared with control samples. I filtered genes that are either significantly upregulated (>10 fold up) or downregulated (>10 fold down) and used David to analyze the functions of affected transcripts. There are 164 genes with high expression levels and 85 with low expression levels in $ppk20^{KO}$ animals relative to controls. Among them, 50% of those 164 upregulated genes and 85% of those 85 downregulated ones belong to different cellular signaling pathways including G-protein coupled photoreceptor, cell membrane receptor, cholesterol homeostasis or storage signaling. Among these genes, I can detect genes involved in heme, iron or ecdysone biosynthesis, for example *transferrin 1* (*Tsf1*), *spookier* (*spok*) or *shroud* (sro) with 12.83-, 7.23- and 14.93-fold enrichment, respectively. These data indicate that many cellular pathways are interfered in *ppk20* lost-of-function animals. Thus, looking at transcription profile can only give us a partial picture of cells with impaired ppk20. Future study comparing expression at protein level is needed to give us a better picture of ppk20 function.

5.4   Discussion

5.4.1   Porphyria phenotype in *ppk20* loss-of-function

The detection of porphyria phenotype in *ppk20*-impaired animals is intriguing since this can suggest a direct link between sodium absorption and iron homeostasis. Although a potential link between ENaC family and heme synthesis has been suggested in other studies, it is not seriously addressed anywhere to best of my knowledge. As a result, this study will add another light to this link. Failure to rescue *ppk20* mutants with iron but partial rescue with injected ferritin or hemin indicates that ppk20 may have a role in iron regulation rather than synthesis of the

porphyrin ring. Disruption in iron homeostasis will affect heme synthesis, thus interfere with other iron/heme-dependent processes, including steroid hormone synthesis.

5.4.2   A model of *ppk20* function in iron regulation

As a membrane protein, *ppk20* might work directly or indirectly with other proteins involved in iron trafficking. Other mammalian studies have shown that upon activation by steroid hormone, ENaC protein translocate to ER and Golgi before being delivered to cell membrane. This process requires Rab5 and Rab11 to move between cytoplasm and cell membrane. Interestingly, these two Rab proteins are also involved in the trafficking of transferrin/transferrin receptor complex during iron import in the vertebrate system. One hypothesis is this process is conserved in *Drosophila*, and ppk20 can translocate together with the *Drosophila* transferrin receptor. Via personal communication with another lab member, Sattar Soltani, who is actively studying this pathway, I know that PG-knock down of either Rab5, Rab11 or transferrin all resulted in porphyria phenotype. This further strengthens the hypothesis that ppk20 acts as a sensor to regulate the trafficking of the vesicles. Thus, without this polypeptide, cells mis-regulate the trafficking of these vesicles and affect other downstream pathways. Another hypothesis is ppk20 is the hitherto unidentified *Drosophila* TfR (DmTfR). However, with the low expression level, ppk20 might not likely exhibit this function. It might be one of many DmTfR and only functions in specific tissues or at specific time points during development. Future studies will need to address this by testing the interaction, if any, between ppk20 and the *Drosophila* transferrin. It is also desirable to test the interaction between ppk20 and *Drosophila* ferritin since current evidence suggests that this iron-storage complex can participate in iron trafficking across tissue as a way to deliver iron into target tissues.

5.4.3 Conclusion and significance

The potential function of ppk20 in iron/heme homeostasis will provide a further insight into the ENaC family physiology. As one of the most diverse channel family s species, any issue affects the normal function and physiological activity of ENaC can result in servere consequences. In fact, mutations in ENaC have been linked to different diseases. My study suggests a novel function of the family, which might be conserved in other species, and any future clinical approach might need to take function into consideration.

5.5   Figures
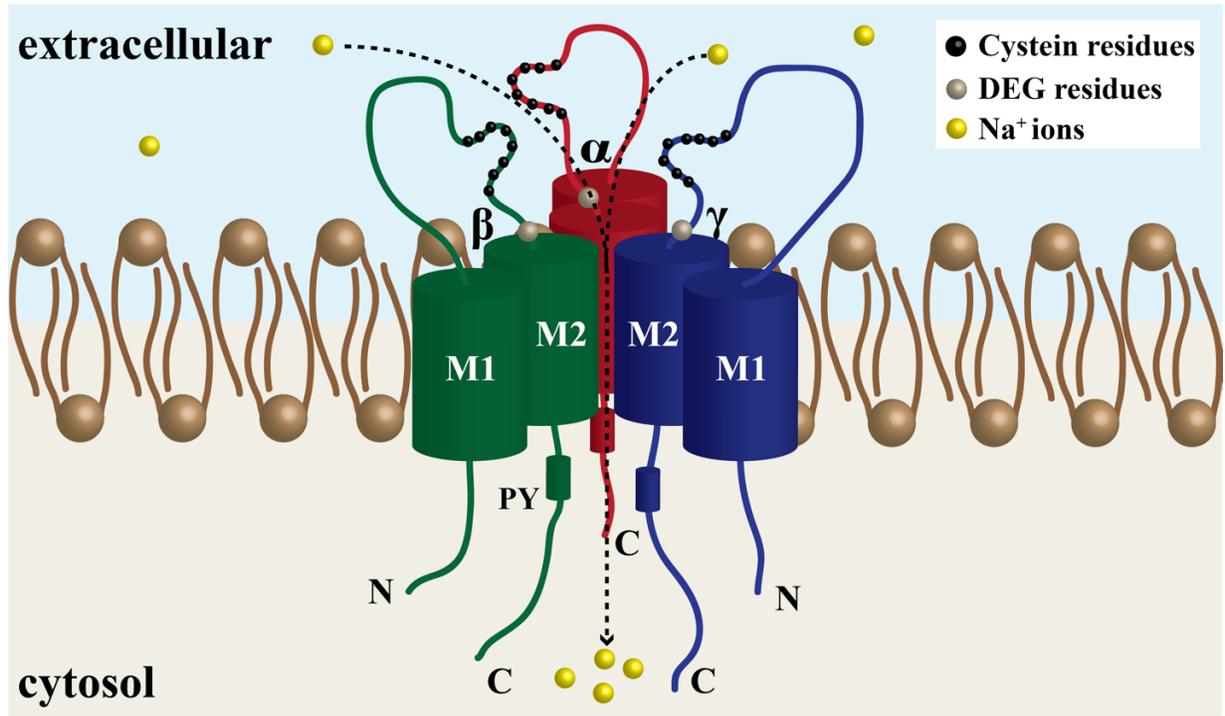
Figure 5.1 Typical DEG/ENaC channel structure.

Each channel comprises on average three subunits which can be either homomeric or heteromeric protein complexes. Each subunit comprises two transmembrane domains (M1 and M2), two short intracellular domains where the N terminus is normally longer than the C terminus domain, and an unusually large and highly structured extracellular domain. The C terminus domain in mammals carries a PY motif to facilitate ENaC trafficking between cytoplasm and cell membrane. The extracellular domain carries a cysteine-rich region that facilitates the sodium transport across the channel while the DEG residue acts as the key to ensure the channel is properly opened or closed whenever needed.
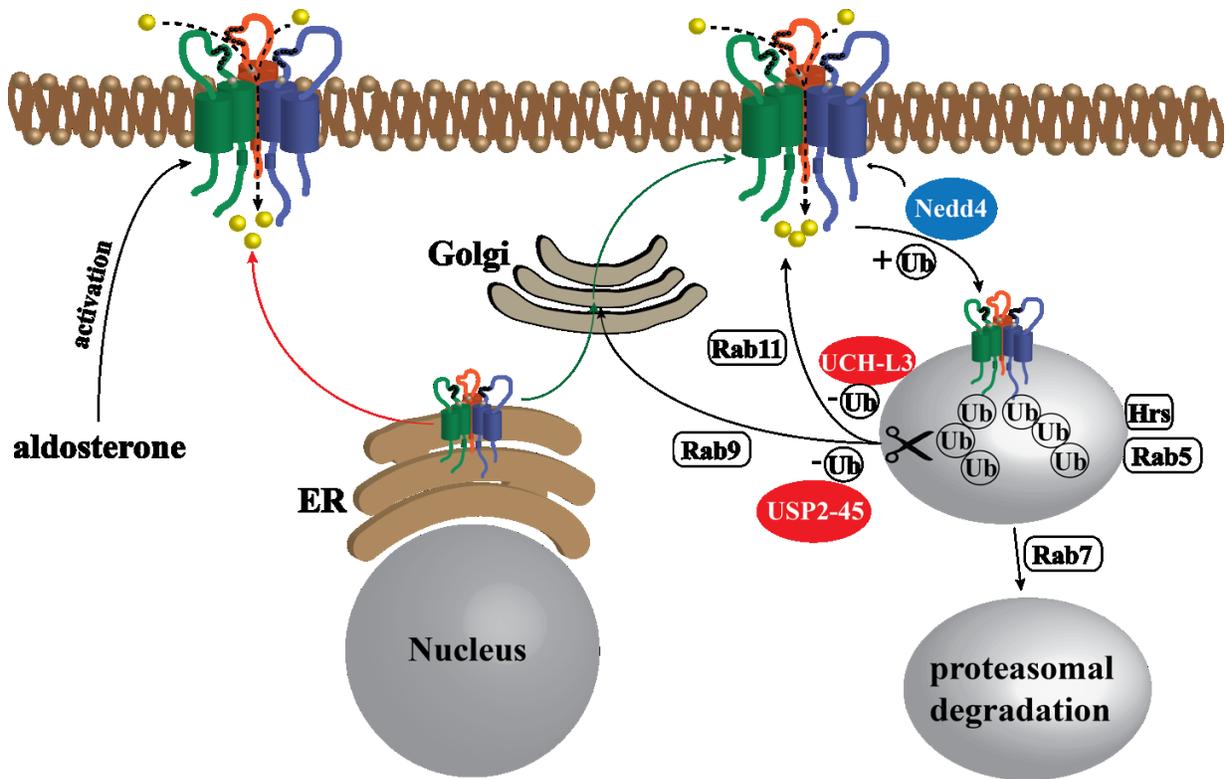
Figure 5.2 Schematic of ENaC trafficking in epithelial cells.

The ENaC chains appear to assemble together early on in the endoplasmic reticulum (ER). ENaC subunits are then passed from ER to Golgi, where they are processed by proteases. Matured ENaC is then delivered to cell membrane and perform their functions. This process is regulated by steroid hormones and proteins involved in endocytosis like Rab5 and Rab11.
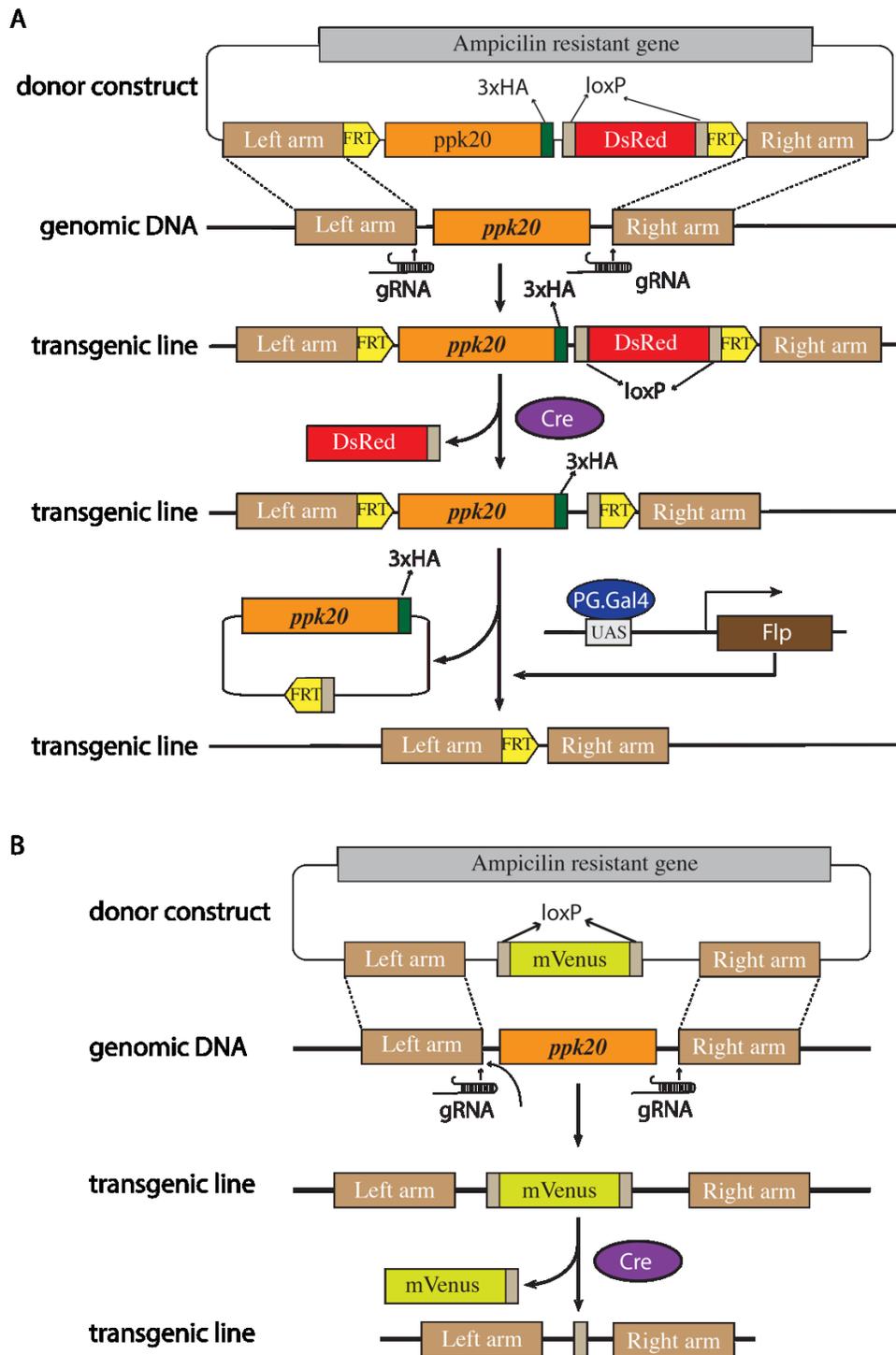
Figure 5.3 Diagram of *ppk20* CRISPR constructs.

A. Schematic of *ppk20^{FCH}* approach. In this approach, the endogenous *ppk20* is replaced with a FRT-flanked allele that carries FRT sites on either end of the gene. Using tissue-specific

173

expression of flippase (FLP), I can temporally and partially remove this allele. The $ppk20^{FCH}$ allele

also carries a 3xHA epitope tag at C-terminus for potential protein experiments **B.** Schematic of

classic $ppk20^{KO}$ deletion. The endogenous $ppk20$ allele is replaced with the mVenus marker, driven

by the 3xP3 promoter. In this approach, the $ppk20$ allele is removed in all the cells.
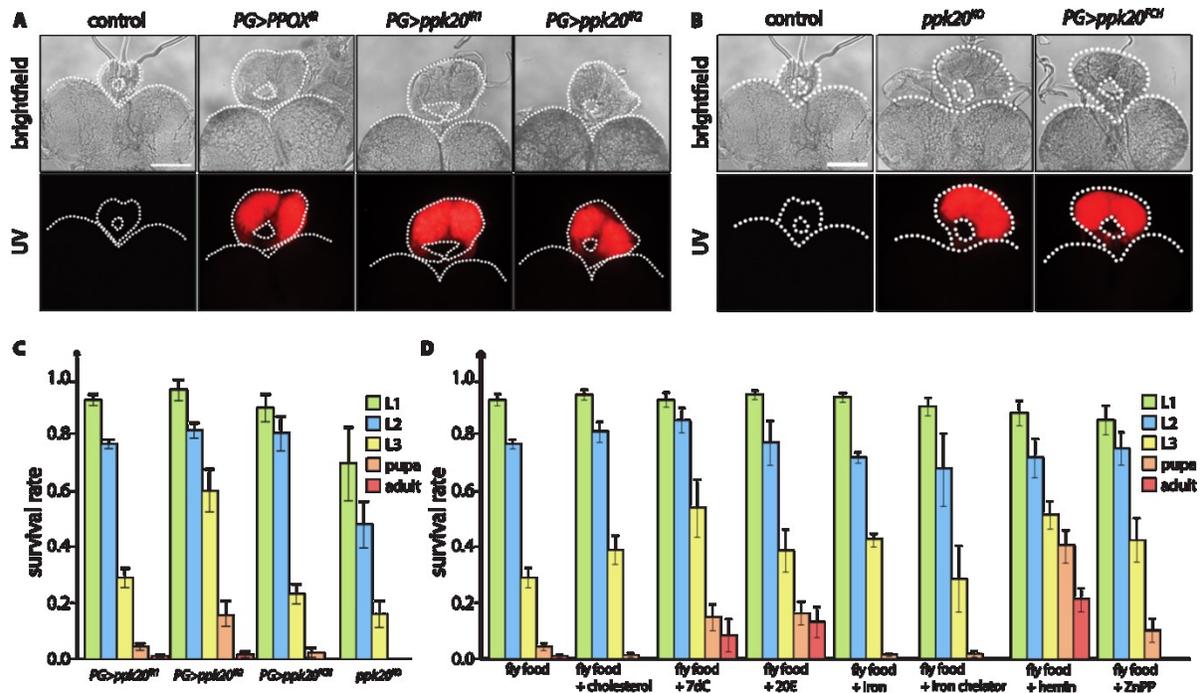
Figure 5.4 Developmental defects and supplementary rescue in *ppk20*-impaired animals.

A. Porphyria-like phenotype of *PG>ppk20*-RNAi animals beside wildtype and *PG>PPOX*-RNAi as controls. Scale bar = 250 μm. **B.** Porphyria phenotype in ring gland of *ppk20* mutants, *ppk20^KO* and *PG>ppk20^FCH*. Scale bar = 250 μm. **C.** Survival rate of *ppk20*-impaired animals, including PG-knock down, PG-knock out and classic deletion mutant. **D.** Survival rate of *PG>ppk20^IR1* animals on fly food with different supplements.
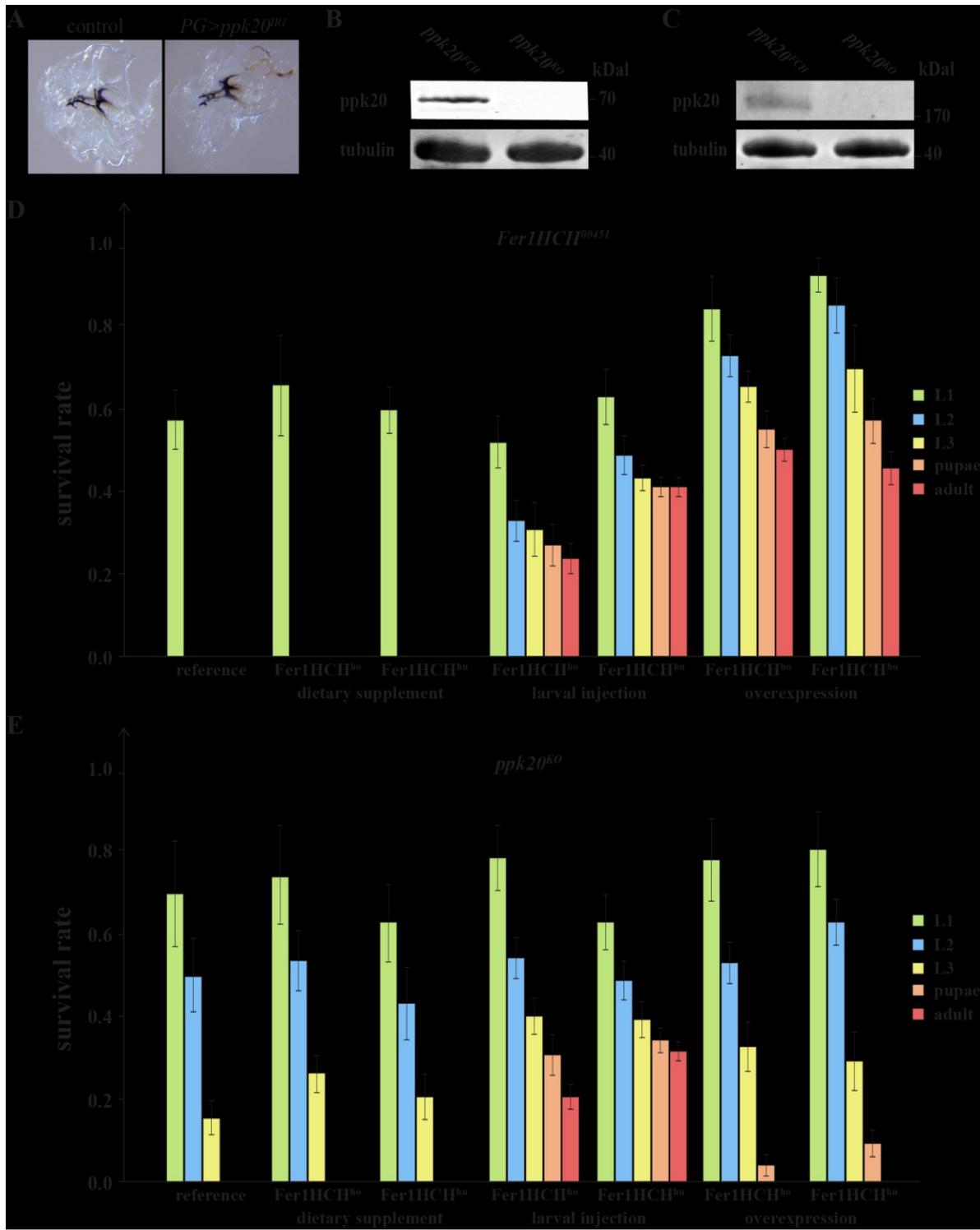
Figure 5.5 *ppk20* is expressed at low level in *Drosophila* larvae and has a link with iron.

A. Trachea necrosis in PG-knock down of *ppk20* animals. **B.** Western blot detection of ppk20 protein from whole body lysate using *ppk20^{FCH}* and *ppk20^{KO}* animals. **C.** Non-denatured western blot detection of ppk20 protein from whole body lysate using *ppk20^{FCH}* and *ppk20^{KO}* animals. **D.** Survival of *Fer1HCH* mutant (*Fer1HCH^{00451}*) in reference medium, fly food supplemented with either extracted horse-origin Fer1HCH (Fer1HCH^{ho}) or extracted human-origin Fer1HCH (Fer1HCH^{hu}), injected Fer1HCH^{ho} or Fer1HCH^{hu}, or overexpression of transgene carries Ferritin complex originated from either horse or human. Error bar represents standard deviation. **E.** Survival of *ppk20* mutant (*ppk20^{KO}*) in reference medium, fly food supplemented with either extracted Fer1HCH^{ho} or Fer1HCH^{hu}, injected Fer1HCH^{ho} or Fer1HCH^{hu}, or overexpression of transgene carries Ferritin complexes originated from either horse or human. Error bar represents standard deviation.
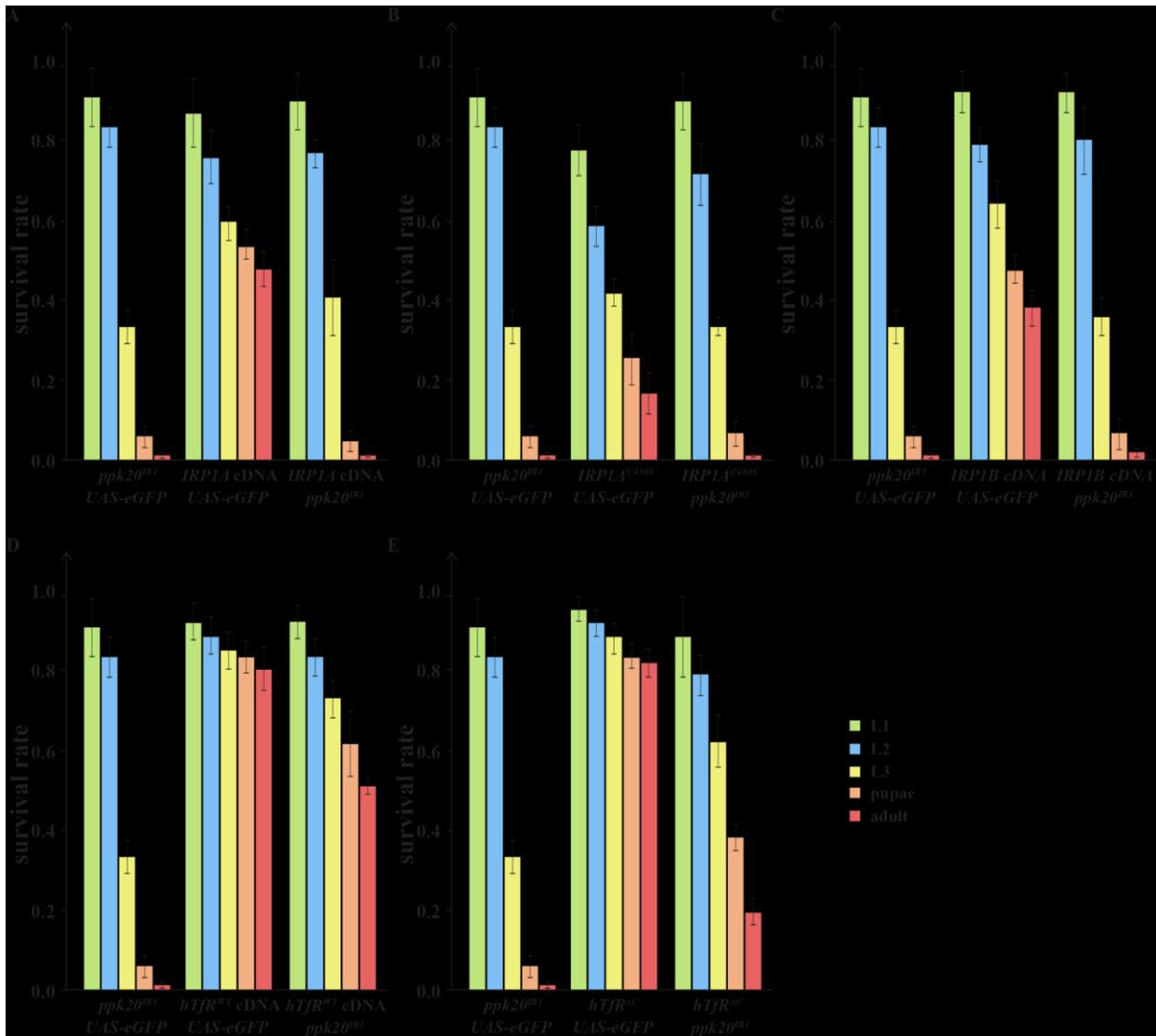
Figure 5.6 Genetic interactions reveal the link between ppk20 and human transferrin receptor (hTfR)

Survival of PG>*ppk20*-RNAi in animals with overexpression of wild-type *IRP1A* (*IRP1A^{WT}*) (**A**), apo-form *IRP1A* (*IRP1A^{C450S}*) (**B**), wild-type *IRP1B* (*IRP1B^{WT}*) (**C**), wild-type human Transferrin receptor (*hTfR^{WT}*) (**D**) and cell-membrane intact hTfR (*hTfR^{ΔC}*) (**E**). Error bar represents standard deviation.

5.6 Tables

Table 5.1 *ppk20* transgenic construct and properties

transgene	description
UAS-3xHA-ppk20^{PA}	expresses wild type N terminal 3xHA-tagged *ppk20 PA* cDNA under Gal4/UAS control
UAS-ppk20^{PA}-mVenus	expresses wild type C terminal mVenus-tagged *ppk20 PA* cDNA under Gal4/UAS control
UAS-ppk20^{PC}-3xHA	expresses wild type C terminal 3xHA-tagged *ppk20 PA* cDNA under Gal4/UAS control
UAS-mVenus-ppk20^{PC}	expresses wild type N terminal mVenus-tagged *ppk20 PA* cDNA under Gal4/UAS control
ppk20^{gRNA}	expresses double gRNA targeting coding sequence of *ppk20* under *Drosophila* U6:3 promoter

Table 5.2 Primers being used to generate *ppk20* constructs

primer name	sequence (5' – 3')
Ppk20 R XbaI FRT R	TTTCTAGAGAATAGGAACTTCGAATGTGTGGTAATG
Ppk20 R XbaI FRT F	TCTCTAGAAAGTATAGGAACTTCATATGCTGTATTTATTAAATG
Ppk20 gRNA Left F	CTTCGACGGCCAGGCCATCCATTA
Ppk20 gRNA Left R	AAACTAATGGATGGCCTGGCCGTC
Ppk20 gRNA Right F	CTTCGAATATACGAGATATATGAG
Ppk20 gRNA Right R	AAACCTCATATATCTCGTATATTC
Ppk20 Left NotI F	TAGCGGCCGCCGGACGAACTGCG
Ppk20 Left XbaI FRT R	TTTCTAGAGAATAGGAACTTCTTATGGCTCACAAGCTTTG
Ppk20 Mid XbaI FRT F	TCTCTAGAAAGTATAGGAACTTCTGGATGGCCTGGCCGTC
Ppk20 Mid HindIII R	ATAAGCTTGAGTGGTTCGCCATGCTG
Ppk20 R SpeI F	GCACTAGTATATATCTCGTATATTCG
Ppk20 R PstI R	AGCTGCAGCATTTAATAAATACAGCATATC
Ppk20 R XbaI FRT R	TTTCTAGAGAATAGGAACTTCCATTACCACACATTC
Ppk20 EcoRI Long FP	TAGCGGCCGCAGCCATTGAAGGTCTTGGTG
Ppk20 R PstI Long RP	AGCTGCAGAACCACTCCAGAGGGAAGAC
Ppk20 gRNA FP dual	AAGATATCCGGGTGAACTTCGAATATACGAGATATATGAGGTTTT AGAGCTAGAAATAG
Ppk20 gRNA RP dual	TAGAGCTAGAAATAGCAAGTAATGGATGGCCTGGCCGTCGTTTTA

primer name	sequence (5' – 3')
	GAGCTAGAAATAGC
ppk20 mut HindIII RP	ATAAGCTTTTATGGCTCACAAGCTTTG
ppk20 R FRT New RP	TTTCTAGAGAATAGGAACTTCGGGTCTTATTTAAGCACC
ppk20 R FRT New FP	TCTCTAGAAAGTATAGGAACTTCTGAAGGCTAAGTGGGTC
Ppk20 gRNA FP dual (2)	AAGATATCCGGGTGAACTTCGGTCTTATTTAAGCACCTGAGTTTTA GAGCTAGAAATAG
ppk20 Dual gRNA RP (2)	TGCTATTTCTAGCTCTAAAACTAATGGATGGCCTGGCCGTCGACG TTAAATTGAAAATAG
Ppk20 BamHI RP Left	AGAGGATCCGAGTGGTTCGCCATGCTG
Ppk20 PstI FP Right	GACCTGCAGATATATCTCGTATATTCG
Ppk20 HindIII RP Right	GCCAAGCTTAACCACTCCAGAGGGAAGAC
Ppk20 NotI long N FP	TAGCGGCCGCCAGCCATTGAAGGTCTTGGTG
Ppk20 NdeI left RP	GTGCATATGGAGTGGTTCGCCATGCTG
ppk20 NotI L G FP	CAGGTGGAATTCTTGCATGCTAGCTCTTATCGGACTCACCCCCAT TG
ppk20 ko G L RP	GCACTACGATCGCAGGTGTGCATTATGGCTCACAAGCTTTG
ppk20 ko G R FP	GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCATATATCGTA TATTCGCAG
ppk20 ko G R RP	CCTTATGCATGGAGATCTTTACTAGAACCACTCCAGAGGGAAGAC

primer name	sequence (5' – 3')
ppk20 FRT L Mut FP	GAAAGTATAGGAACTTCTGGATGGCCTGGCCGTCGGAACTTTG
ppk20 FRT L Mut RP	TAGAGAATAGGAACTTCTTATGGCTCACAAGCTTTG
NH_CG7577 F1 FP(1)	GTCGCCCTTCGCTGAAGCAGGTGGCCATTTTCAGGTTATGATTAAG
NH_CG7577 F1 RP(1)	AATTGACCCACTTAGCCTTCAGATGTTTGAATAGCACCCCATAGTACTGGTG
NH_CG7577 F2 FP(2)	GTACTATGGGGTGCTATTCAAACATCTGAAGGCTAAGTGGGTCAATTTAAAGGGCTAC
NH_CG7577 F2 RP(2)	GCACTACGATCGCAGGTGTGCATAAAATACAGCATATCATTAAGCGTAATCTGGAACATCGTATGGGTACCACACATTCTTTC
NH_CG7577 F3 FP(3)	CGAAGTTATAGAAGAGCATTAAATGTTTAATGTC
NH_CG7577 F3 RP(3)	GACGGAAGAGCCTCGAGCTATGGCGAGGGCTC
NH_CG7577 F3 Mut RP	ACTTATATATTCATGGCAAGAGAAAG
NH_CG7577 F3 Mut FP	ACAGCCCGATGGAGAAGTAGTAG
NH_CG7577 pCFD4 FP	TATATAGGAAAGATATCCGGGTGAACTTCGGGTCTTATTTAAGCACCTGAAGGGTTTTAGAGCTAGAAATAGCAAG
NH_CG7577 pCFD4 RP	ATTTTAACTTGCTATTTCTAGCTCTAAAACGGCCTCTATTTGTACATTCATGGCGACGTTAAATTGAAAATAGGTC
ppk20 MRG G FP	TACAAAAAGCTGAACGAGAACGATGGCCAAGGGAGATAATTC
ppk20 MRG G RP	TTCACAAGATCCTCTAGAGTTACCACACATTCTTTCTGTAG

primer name	sequence (5' – 3')
NH_CG7577 PCFD4 FP(2)	TATATAGGAAAGATATCCGGGTGAACTTCGGTCTTATTTAAGCACC TGAGTTTTAGAGCTAGAAATAGCAAG
NH_CG7577 PCFD4 RP(2)	ATTTTAACTTGCTATTTCTAGCTCTAAAACGGCCTCTATTTGTACAT TCACGACGTTAAATTGAAAATAGGTC
NH_ppk20 p1_2 attB1 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGCCAAGGGAGATAATTCG GTC
NH_ppk20 p1_con_p2 RP	GACATTGTTCAGTTGGTCGGGGGCCTGGCAGATACTTACGTAG
NH_ppk20 p1_con_p2 FP	CTACGTAAGTATCTGCCAGGCCCCCGACCAACTGAACAATGTC
NH_ppk20 p1_2 attB2 RP	CCCACTTTGTACAAGAAAGCTGGGTCTCTGTGGTTGTACAGGAAG
NH_ppk20 p456 attB1 FP	TCAAGTTTGTACAAAAAAGCAGGCTACATCTTCAGCTACGACAAG
NH_ppk20 p45 RP	ACAAAGGTGTGGTCTCCACTTGGTAGTGCACATCCACCCTGAAGA GTTTCTGATGCGGATCCTTGGGTAGTTTATGGGCTGACAAAG
NH_ppk20 p56 FP	TGCACTACCAAGTGGAGACCACACCTTTGTATCGCACAAGCCTGGA GTTTACCATTATCGATCTAATTGCCAATCTGGGTGGAATCTTTGGAC
NH_ppk20_p456 attB2 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACCACACATTCTTTCTGTAG
NH_ppk20_p123 RP	AGGCAGCGCAGGCTGACTGCGTTGCAATCCGGAACTCTGTGGTTGT ACAGGAAG
NH_ppk20_p3456 FP	ACGCAGTCAGCCTGCGCTGCCTTGCGCGTCACAATGACATCTTCA

primer name	sequence (5' – 3')
	GCTACGACAAG
NH_ppk20 attB2 NRP	CCCACTTTGTACAAGAAAGCTGGGTCCACACATTCTTTCTGTAG

Table 5.3 Developmental phenotype and ring gland morphology summary for PG-specific pickpocket genes lost-of-function

| symbol | RNAi stock(s) | | | | | | gRNA stock(s) | |
| | VDRC | | Bloomington | | NIG | | | |
	ID	defect	ID	defect	ID	defect	ID	defect
ppk	V108683	No	BL29571	4-6 hrs delay[a]	N3478R-1	No	BL82690	No
	V330392	No			N3478R-2	No		
rpk	V105463	No	BL25847	delay[a]	N1058R-1	No	BL82869	No
	V8549		BL39053	L3 arrested[a]	N1048R-3	No		
ppk3	V104593	No	BL61995	4 hrs delay[a]				
Nach	V106647	No	BL27262	No			BL68006	No
	V45921	No	BL62894	No				
ppk5	V101664	No	BL25816	6 hrs delay[a]				
	V48290	No						
	V48289	No						
ppk6	V101091	No	BL53010	1 day delay[a]			BL68010	No
			BL25880	No				
ppk7	V100643	No	BL25922	No			N 2LG-0352	No
	V7900	No						
ppk8	V47047	No	BL25814	N/A				
	V47048	No						
ppk9	V109685	No	BL25892	4 hrs delay[a]				
	V104952	No						
	V23391	No						
	V17213	No						
ppk10	V330625	No	BL27256	No				
ppk11	V107741	No	BL23781	No				
	V330319	No						

ppk12	V105131	No	BL27092	No				
ppk13	V110084	No	BL25817	No	N14398R-1	No		
	V9494	No			N14398R-2	No		
ppk14	V110258	No	BL27091	**1 day delay[a]**			N 2LG-0754	No
	V7903	No						
	V7904	No						
ppk15	V109855	**L2 arrested[2]**	BL28012	**5 days delay[b]**	N14239R-1	No		
	V42523	No			N14239R-4	No		
ppk16	V22990	No	BL25890	No			BL68009	No
	V22989	No						
ppk17	V109927	No	BL58557	No	N13278R-1	No		
	V8595	No			N13278R-2	No		
	V8596	No						
ppk18	V105199	No	BL61949	**1 day delay[a]**				
	V13211	No	BL27240	**1 day delay[a]**				
	V13209	No	BL25883	**1 day delay[a]**				
ppk19	V107638	No	BL58203	No				
	V36660	No	BL25887	No				
ppk20	V36659	**L2 arrested[c]**	V25897	**50% L2 arrested, L3 arrested[c]**				
ppk21	V107892	No	BL25849	No				
	V1345	No	BL62487	No				
	V1346	No						
ppk22	V106384	No	BL28706	No				
	V47946	No	BL61821	No				
ppk23	V106873	No	BL28350	No				
	V39581	No						
	V39580	No						
ppk24	V102923	No	BL26006	No				

	V30196	No						
ppk25	V101808	No	BL27088	No				
	V7343	No						
ppk26	V100834	No	BL25825	No			BL82844	No
	V5110	No						
	V5109	No						
ppk27	V330515	No	BL27239	**6 hrs delay[a]**				
ppk28	V100946	No	BL31878	No			BL68038	No
	V44412	No						
ppk29	V106888	No	BL27241	No				
	V330294	No						
ppk30	V105896	No	BL25810	No	N18110R-1			
	V1349	No			N18110R-3			
	V1351	No						
ppk31	V106385	No	BL27087	No				
	V1269	No	BL44013	No				

[a] No ring gland morphology phenotype

[b] Enlarged ring gland but no detectable red autofluorescence

[c] Enlarged ring gland with detectable red autofluorescence

Table 5.4 Phenotypes of PG-specific somatic CRISPR for pickpocket genes

CG (Flybase)	symbol	defect	RG phenotype	gRNA functional?[a]
CG3478	ppk	No	Normal	Yes
CG1058	rpk	No	Normal	Yes
CG30181	ppk3	No	Normal	Yes
CG8178	Nach	No	Normal	Yes
CG33289	ppk5	No	Normal	Yes
CG11209	ppk6	No	Normal	Yes
CG9499	ppk7	No	Normal	Yes
CG32792	ppk8	No	Normal	Yes
CG34369	ppk9	No	Normal	Yes
CG34042	ppk10	No	Normal	Yes
CG34058	ppk11	No	Normal	Yes
CG10972	ppk12	No	Normal	Yes
CG33508	ppk13	No	Normal	Yes
CG9501	ppk14	No	Normal	Yes
CG14239	ppk15	No	Normal	Yes
CG34059	ppk16	No	Normal	Yes
CG13278	ppk17	No	Normal	Yes
CG44152	ppk18	No	Normal	Yes
CG18287	ppk19	No	Normal	Yes
CG7577	ppk20	L2/L3 arrest	red RG	Yes
CG12048	ppk21	No	Normal	Yes
CG31105	ppk22	No	Normal	Yes
CG8527	ppk23	No	Normal	Yes
CG15555	ppk24	No	Normal	Yes
CG33349	ppk25	No	Normal	Yes
CG8546	ppk26	No	Normal	Yes
CG10858	ppk27	No	Normal	Yes
CG4805	ppk28	No	Normal	Yes
CG13568	ppk29	No	Normal	Yes
CG18110	ppk30	No	Normal	Yes
CG31065	ppk31	No	Normal	Yes

[a] gRNA efficiency was tested by crossing with transgenic line carries *act-Cas9* and *lig4* mutant.

Lig4 is responsible for repairing the double strand break (DSB). Animals with *lig4* mutation failed

to repair the DSB, resulted in the lethality of the male progeny. This phenotype was used to screen for gRNA efficiency.

Table 5.5 Transcription profile of pickpocket genes during larval development

CG number	symbol	RG_WT				whole body_WT			
		4hr	8hr	24hr	36hr	4hr	8hr	24hr	36hr
CG3478	ppk	19.351	20.01	26.504	43.162	140.83	131.83	118.11	70.581
CG1058	rpk	85.331	107.66	63.465	117.12	38.147	51.419	50.76	53.454
CG30181	ppk3	28.88	23.779	30.374	30.683	38.292	35.412	35.491	27.324
CG8178	Nach	17.188	16.08	16.67	20.053	15.714	14.543	13.842	14.108
CG33289	ppk5	17.415	23.214	20.414	20.107	19.726	17.577	16.585	15.085
CG11209	ppk6	29.519	42.387	73.366	58.91	2642	1550.7	2392	1103
CG9499	ppk7	20.492	22.811	24.328	22.782	22.076	20.947	19.362	16.849
CG32792	ppk8	27.85	26.352	26.642	31.251	21.281	23.89	23.983	24.191
CG34369	ppk9	22.332	23.64	26.011	31.058	21.194	18.757	19.083	15.817
CG34042	ppk10	36.601	67.539	47.219	51.056	27.441	23.369	22.844	18.352
CG34058	ppk11	18.997	19.742	16.658	19.559	14.316	14.09	14.028	13.149
CG10972	ppk12	19.867	21.357	24.298	25.642	27.68	31.529	23.418	25.165
CG33508	ppk13	358.11	547.58	813.59	995.3	151.77	198.83	291.31	233.68
CG9501	ppk14	23.879	24.689	30.091	32.991	23.416	23.322	21.025	19.741
CG14239	ppk15	17.594	20.936	18.68	19.654	17.594	16.131	17.218	14.398
CG34059	ppk16	58.042	31.509	32.806	36.746	22.728	20.094	19.85	19.497
CG13278	ppk17	47.476	40.076	32.228	22.598	198.77	110.89	112.39	80.596
CG13120	ppk18								
CG18287	ppk19	25.895	32.082	26.581	30.388	18.719	21.115	21.859	18.937
CG7577	ppk20	27.295	33.332	42.546	36.406	27.333	35.011	31.145	24.661
CG12048	ppk21	18.181	17.739	21.205	20.374	16.208	15.926	15.534	14.366
CG31105	ppk22	44.181	34.265	132.58	539.36	49.928	49.187	50.145	65.467
CG8527	ppk23	16.943	22.502	21.924	21.684	17.568	20.306	17.956	15.576
CG15555	ppk24	18.02	18.608	19.018	17.507	17.619	16.833	17.009	15.057
CG33349	ppk25	43.153	58.709	53.451	53.02	39.208	49.453	39.041	30.924
CG8546	ppk26	45.212	37.293	29.876	50.135	262.59	277.63	215.25	100.54
CG10858	ppk27	13.404	13.204	14.246	15.772	12.005	11.689	11.746	12.117
CG4805	ppk28	118.17	109.08	121.21	168.03	76.076	59.685	73.23	55.069
CG13568	ppk29	261.27	389.05	280.12	358.16	98.006	102.89	163.81	258.76
CG18110	ppk30	17.667	24.275	22.145	28.032	25.412	23.931	23.282	17.665
CG31065	ppk31	22.924	21.821	23.251	22.213	20.241	18.939	17.176	15.857

Table 5.6 Transcription profiles of pickpocket genes in porphyria phenotypes

symbol	ctrl vs *PPOX* mut	ctrl vs *FECH*[IR]	ctrl vs *spz5*[IR]	ctrl vs *NOS*[IR]	ctrl vs *DHR51*[IR]	ctrl vs *ZFRP8*[IR]	ctrl vs *CG8145*[IR]
ppk	1.456546426	-1.45786442	1.044266939	1.125317335	1.897830725	1.018172741	1.599416733
rpk	6.861706257	1.371869683	-1.35426436	1.591439009	-6.047974016	-6.047974016	1.739299417
ppk3	-2.61599444	-21.82327563	-3.362627454	-21.82327563	-21.82327563	-21.82327563	-21.82327563
Nach	7.326560974	-2.466301359	-14.87927353	3.510206223	-14.87927353	2.913243771	-2.423878559
ppk5	-1.013085341	-1.345123531	-10.98957267	-1.004189386	-10.98957267	-10.98957267	-1.591446123
ppk6	-3.726082184	1	1	8.87733078	9.240600586	11.35909748	5.947887897
ppk7	1.794116735	1	1	1	1	8.095867157	9.535419464
ppk8	-1.835366556	-6.16897864	2.341956139	1.488930583	3.36320734	1.264209747	2.752328873
ppk9	-1.091706385	1.252330065	-1.018240082	-6.637844226	-1.646524408	7.182887077	6.885770321
ppk10	136.1710358	1.563071251	4.57103014	1.985240817	-41.68780964	1.983455539	6.564687729
ppk11	-8.020067838	-202.3667571	-202.3667571	-6.371575173	-46.7532927	-2.125770046	-1.643215713
ppk12	1.928617597	-1.03086579	1.044266939	-6.026631138	-1.461974512	1.348054051	1.739299417
ppk13	-4.84079034	1.187773347	1.445486665	1.48332262	12.12083817	5.417603016	1.147320986
ppk14	2.362064362	1	1	1	1	1	58.51076889
ppk15	9.75215435	-5.225947115	-5.225947115	-1.004189386	-5.225947115	-1.026665546	10.73355103
ppk16	2.537264347	-32.39099796	-142.4943805	-18.71100151	-14.93469305	-2.667820264	1.20121479
ppk17	1.443811536	-1.168047798	-5.001124098	-3.475807998	-1.151574398	3.435445786	2.010016918
ppk18	1.651568413	-12.33418733	-73.14178593	-16.97557882	-73.14178593	-10.87043057	-5.289709217
ppk19	-2.949705851	-1.393251309	1.194758415	1.215706706	1.11725986	1.767122507	1.436005592
ppk20	-3.004917473	-4.251021393	-4.251021393	-4.251021393	2.042833567	8.251578331	13.45433998
ppk21	-4.098183914	-85.07137683	-85.07137683	-10.06077404	-13.66874376	-85.07137683	-85.07137683
ppk22	-7.636288494	-5.136659893	-5.090789482	-4.998097027	-1.261137208	2.154399395	-2.62169541
ppk23	-8.515487162	-1.960881978	5.164149284	1.670808554	-8.39846202	7.124034405	-8.39846202
ppk24	-13.79175313	-6.660081302	-6.660081302	-6.660081302	-6.660081302	2.922478676	-6.660081302
ppk25	10.37589931	-1.736714492	-1.565966407	-8.288738543	-1.207319412	7.584482193	1.066277862
ppk26	1.95851922	-47.9136369	-6.367083622	-47.9136369	-11.35244081	-10.46287491	1.059371948
ppk27	-1.563355719	-10.4862286	-6.333881725	-73.16331587	-73.16331587	-8.386571636	-5.662660547
ppk28	-1.565855037	-3.43863936	-2.494351245	-1.14865202	1.262998104	1.304704189	-1.035646024
ppk29	-1.261178542	-2.767418516	-2.695950598	-257.3175526	-1.349302135	2.82487154	-1.62914648
ppk30	91.38722992	1	1	1	1	7.400157452	59.27728653
ppk31	-1.876026828	-2.017685216	8.26654911	12.04202843	6.082243443	1.487850904	16.25641823

Chapter 6   **A *Drosophila* CRISPR toolkit for conditionally and temporally manipulating gene expression in the prothoracic gland as a test case for polytene tissues[3]**

6.1   Introduction

6.1.1   CRISPR allows precise genetic modification

It is crucial to have the ability to conditionally manipulate the activity of genes, be it to overcome embryonic lethality of null mutants to study later roles of a given gene, distinguish between cell-autonomous and non-autonomous mechanisms, or to study tissue-specific gene functions. In *Drosophila*, the standard techniques for conditionally altering gene function have been RNA interference (RNAi) to block or impair gene activity, the overexpression of cDNAs for gain-of-function studies as well as dominant negative constructs for induced inhibition of function. Most commonly, both RNAi and cDNA expression are temporally controlled via the Gal4-UAS system, resulting in a highly versatile set of tools. However, each of these commonly used components has its limitations and downsides. In particular, RNAi suffers from the frequent occurrence of off-targets, requiring rigorous validation, and often the expression of a target mRNA is only partially blocked. In addition, combining two or more RNAi transgenes to test for synthetic lethality or interaction of pathway components is cumbersome and exponentially increases the risk of non-specific effects. On the other hand, to achieve overexpression of a gene of interest, traditional cDNA overexpression requires the cloning of a full-length cDNA, which may be difficult and time-consuming. Further, in the case of alternatively spliced genes, one usually has to choose which isoform to use for the transgenic cDNA line, which may limit the conclusions that can be drawn from the experiment. It should also be noted that the use of Gal4 itself has its drawbacks. In particular, Gal4 may result in signal amplification due to its strong activation domain, and one has only limited control over how strongly a given cDNA is expressed. Further, certain UAS-regulated transgenes show some degree of leakiness, depending on the tissue and developmental time, potentially confounding experimental outcomes or the tools being used [345].

Similarly, the presence of a second unrelated UAS-transgene may alter phenotypes seen with a single UAS-transgene alone, as both compete for Gal4-binding, which may quench the expression of either transgene. Finally, Gal4 binds non-specifically to endogenous loci, resulting in the up- and down-regulation of hundreds of genes, which may complicate the interpretation of genome-wide gene expression studies [346].

The recent discovery of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and the generation of guide RNA- (gRNA-) dependent Cas9 endonucleases has been quickly adapted by *Drosophila* researchers [89–91] and we now possess a universal and powerful toolkit that can be used for both loss- and gain-of-function studies by using distinct versions of Cas9 [92,93]. As such, CRISPR-based techniques are ideal to replace, validate and complement traditional approaches relying on conditionally expressing RNAi or cDNAs. Recent advances in CRISPR-based approaches include codon-optimizations of Cas9, utilizing Cas9 variants as a RNA-guided transcription factors that block or increase target gene transcription, and large-scale transgenic *Drosophila* gRNA collections launched at Harvard Medical School (https://fgr.hms.harvard.edu/vivo-crispr-0), the German Cancer Research Center in Heidelberg (https://www.crisprflydesign.org/library/) and the National Institute of Genetics in Mishima, Japan (https://shigen.nig.ac.jp/fly/nigfly/).

6.1.2 Tissue specific and temporal gene manipulation via CRISPR

Our lab investigates signaling pathways that control ecdysone and heme biosynthesis in the larval prothoracic gland (PG), which is part of the larval ring gland (Figure 6.1). The PG is an endoreplicating tissue that reaches a C-value of 64 by the end of the 3rd instar (L3) [347] and represents a popular model for studying endocrine function, as multiple checkpoints converge on this gland that dictate whether an upcoming pulse of ecdysone can be produced [67]. In a recent

study, the King-Jones lab carried out a genome-wide PG-specific RNAi screen, resulting in the identification of ~1,906 genes that where critical for larval development [66]. However, a frequent issue in the follow-up studies was that we could not validate the RNAi-induced phenotypes by independent non-overlapping RNAi lines, either because no such lines existed or because independent lines did not replicate the phenotype. This prompted us to look into CRISPR-based methods that could be used to confirm the RNAi results. However, no studies have addressed whether somatic CRISPR is feasible in the PG, nor have there been any reports on the usage of tissue-specific CRISPR/CAS9 for other commonly studied polytene tissues such as the larval fat body and the salivary glands. Previous studies have established that somatic CRISPR/CAS9 is highly efficient in disrupting genes in a biallelic fashion, however endoreplicating tissues such as the salivary gland contain up to 1024 copies of a gene [348,349] raising the question as to whether CRISPR/CAS9 would be effective in polytene tissues. Furthermore, our initial attempts to express Cas9 via the most commonly used PG-specific Gal4 drivers resulted in substantial larval lethality, which rendered this approach impractical. We therefore developed several strategies that induced tissue-specific CRISPR/CAS9 without using Gal4.

Here we present two collections of CRISPR tools designed for tissue-specific as well as temporal genome modification in *Drosophila*, which we refer to as the first and second general Gateway Cas9 (gG-Cas9 I and gG-Cas9 II), the PG-Cas9 and the PG-gRNA vector collections. Most of these new tools have in common that they are not based on Gal4, but rather use enhancer regions to achieve tissue-specific expression of Cas9 or gRNAs. This greatly simplifies the genetics of tissue-specific CRISPR/Cas9, since one only requires a single cross to build the CRISPR/CAS9-gRNA combination, while the Gal4-UAS-based approach requires combining at least three transgenes. Importantly, we show that the lethality associated with Gal4-driven Cas9

can be prevented by several strategies, including tissue-specific expression of gRNAs coupled with ubiquitously expressed Cas9. To accomplish tissue-specific gRNA production, we took advantage of inserting ribozyme sequences, which demonstrates for the first time in *Drosophila* the that ribozymes can be used to effectively release gRNAs from mRNAs [350]. We also present the first drug-activated Cas9 collection (gG-Cas9 II) which include the fusion of Cas9 with the Ligand-binding domain of the human progesterone receptor and demonstrate that it is a highly effective tool for achieving both temporal and spatial control over Cas9-mediated gene manipulation. We also generated a split Cas9 collection that works under control of rapamycin supplemented from diet. We evaluated the efficiency of each tool by targeting two well-studied genes acting in ecdysone biosynthesis, *phantom* and *disembodied*. Finally, we provide a general, gateway-based vector collection (gG-Cas9 I and gG-Cas9 II) that allows the quick generation of seven different Cas9-based vectors. These tissue-specific and temporal specific vectors enable the user to i) disrupt target genes of interest, ii) block the assembly of the transcription apparatus near the transcription start site or iii) upregulate the activity of a given gene.

6.2   Modified materials and Methods

6.2.1   *Drosophila* stocks and husbandry

For list of flies generated in this study, see Appendix A.8. We obtained the following stocks from Bloomington stock center: *y¹v¹P(nos-PhiC31/int.NLS)X; P(carryP)attP40(II)* (#25709), *y¹v¹P(nos-PhiC31/int.NLS)X; P(carryP)attP2(III)* (#25710), *phm^{E7}/FM7c* (#2208), *dib²/TM3 Sb¹* (#2776), *UAS-Cas9.P2* (#58985), *UAS-Cas9.P* (#54594), *UAS-Cas9.P* (#54595), *act-Cas9* (#58590), *spz5 P(OE.gRNA)attP40* (#67547), *Alas P(OE.gRNA)attP40* (#68083), *Coprox P(OE.gRNA)attP40* (#68124), *FeCH P(OE.gRNA)attP40* (#78206), *IRP1A P(OE.gRNA)attP40*

(#68039), *Nach P(OE.gRNA)attP40* (#67562) were obtained from the Bloomington *Drosophila* Stock Center.

y²cho²v¹ (TBX-0004), *y²cho²v¹; sco/CyO* (TBX-0007), *y²cho²v¹/Yʰˢ⁻ʰⁱᵈ; Sp/CyO* (TBX-0008), *y²cho²v¹; Sp hs-hid/CyO* (TBX-0009), *y²cho²v¹; Pr Dr/TM6C, Sb Tb* (TBX-0010) were obtained from the National Institute of Genetics of Japan (NIG). *UAS-phm*-RNAi (#108359), *UAS-dib*-RNAi (#101117) were obtained from the Vienna *Drosophila* Resource Center.

y¹v¹/SM5,CyO, y¹w*P(nos-PhiC31/int.NLS)X; P(carryP)attP40(II) and *y¹w*P(nos-PhiC31/int.NLS)X; P(carryP)*[carryP]*attP2(III)* were gifts from the BestGene Inc. *phm22-Gal4*, *spok-Gal4/TM6 Tb* and *spok-Switch-Gal4* were kind gifts from Michael O'Connor's lab. Stocks were maintained on a cornmeal diet unless otherwise specified.

6.2.2 Generating prothoracic gland-specific gRNA plasmids (PG-gRNA)

PG-specific gRNA plasmids were generated based on the pCFD5 plasmids that utilize tRNA- flanked gRNAs [94]. We amplified the pCFD5 backbone via PCR and fused the 1.45kb *spok* regulatory region obtained from the CRII-TOPO Spok plasmid (a kind gift from Michael O'Connor). To ensure proper processing of the Pol II-derived transcript, we added either an HV or HDV ribozyme-containing region to the 3' end, which was then amplified together with tRNA-gRNA duplexes (Figure 6.4 and Tables

Table 6.1). These fragments were cloned together via Gibson reactions, transformed into DH5α, and validated by Sanger sequencing.

6.2.3 Other experiments

Other experiments, including survival studies, generation of gateway plasmids, surveyor nuclease assay, genome extraction, immunostaining, ring gland DNA extraction, *ex vivo* culturing, RNA extraction and qPCR were done as described in chapter 2.

6.3   Results and Discussions

6.3.1   PG-specific expression of Cas9 via the Gal4/UAS system is toxic

6.3.1.1   Classic Gal4/UAS-Cas 9

Current CRISPR/Cas9 tools that conditionally modify gene expression include UAS-Cas9 for GAL4-directed sequence cuts, as well as UAS-dCas9-VPR for gene overexpression, which utilizes a non-cutting version of Cas9 (dead Cas9 = dCas9) fused to the VPR co-activator domain [90,91,94]. Our initial attempts to block gene function in the prothoracic gland (PG) were based on expressing UAS-Cas9.C (the original Cas9 transgene) with *phm22-Gal4* (aka *phm22>Cas9.C* animals), a widely used PG-specific Gal4 driver [351]. However, this approach caused significant lethality, with only ~15% of animals reaching adulthood compared to ~85% in controls (Figure 6.5A). We then tried another PG-specific Gal4 driver, *spok-Gal4* (= *spok>*), which has overall lower expression levels compared to *phm22>* [352]. This combination resulted in only slightly improved survival rates, with 25% of the population reaching adulthood (Figure 6.5A). This observation is consistent with previous studies where high expression levels of Cas9 via Gal4/UAS caused toxicity that was independent of the endonuclease activity [91]. The lethality was also observed when we tried different Cas9 versions, namely Cas9.P (codon-optimized for *Drosophila*) and Cas9.P2 (codon-optimized for human cells) [91,94,353]. UAS-Cas9.P2 was considered to be safer for using the Gal4/UAS approach [94]. Unfortunately, in our hand, *phm22>Cas9.P2* animals, showed only moderately improved survival rates compared to *phm22>Cas9.P* and *phm22>Cas9.C* populations, with only ~50% reaching the third instar stage, and 35% surviving to adulthood (Figures 6.5A and 6.6A). Using *spok-Gal4* instead of *phm22-Gal4* as a PG-specific driver did not make a significant difference (Figure 6.6B). Interestingly, ubiquitous expression of *Cas9.P2*

198

(*act>Cas9.P2*) caused no obvious lethality and the majority of the population reaches adulthood, while *act>Cas9.C* and *act>Cas9.P* were completely and partially lethal, respectively (Figure 6.6C). We reasoned that the very high *Cas9* expression levels that result from expressing UAS-Cas9 in combination with a strong PG-specific Gal4 drivers causes substantial cytotoxicity. Since the PG is responsible for producing ecdysteroids, high levels of Cas9 may interfere with ecdysteroid production and thus disrupt larval and pupal development. Similarly, PG-specific expression of *dCas9.VPR* had only 45% surviving adults, indicating that the toxicity is not necessarily linked to chromosomal breaks, as dCas9 does not cut DNA (Figure 6.5A). Taken together, these data indicate that combining PG-specific Gal4 with *UAS-Cas9* is not an optimal approach to carry out conditional CRISPR in this tissue, and that other tissues may pose similar issues.

6.3.1.2   GeneSwitch Gal4/UAS-Cas9

To bypass the toxicity associated with high levels of Cas9, we tested whether temporally controlling Gal4 via the GeneSwitch (GS) system would resolve the problem. The GS system is based on a Gal4 DNA-binding domain that is fused to the human progesterone receptor ligand-binding domain and the activator domain from human p65 [354]. The chimeric Gal4 protein is only activated in the presence of the steroid mifepristone (RU486), which is provided in the diet. Using PG-specific *spok-Gal4GS*, we activated Gal4 during the first (L1), second (L2), third (L3) or mid-third instar larval stages by transferring larvae to a RU486-supplemented diet. Temporal activation of Cas9 as late as early second instar still caused substantial lethality, while later stages (early and mid L3) displayed 60-70% survival. This might provide a suitable approach for studying gene function at later stages but would likely not be ideal for most genes expressed throughout larval development, such as the Halloween genes, which encode ecdysteroid-producing enzymes

(Figure 6.5B). However, given that the PG is a polytene tissue, it remains unclear whether inducing Cas9 during the L3 stage can efficiently disrupt gene function [355].

6.3.2 The pG-Cas9 system to generate conditional CRISPR in a tissue of interest

We reasoned that omitting Gal4 altogether and instead opting for endogenous regulatory regions may result in lower but equally specific expression of Cas9, and thus reduce its toxicity. To accomplish this, we sought to generate a vector that would allow for quick insertion of tissue-specific enhancers. For this, we used the existing pBPGUw plasmid, a modular gateway-compatible Gal4 vector [95] and replaced the Gal4 sequence by a fragment encoding Cas9 or variants thereof (Figure 6.2). In brief, en̲hancer-D̲m̲Cas9 (en-DmC), en̲hancer-H̲s̲Cas9 (en-HsC) and en̲hancer-F̲okI-d̲Cas9 (en-dFC) function by generating double-strand breaks and deletions via site-directed C̲leavage (C = cleavage), while enhancer-dCas9 (en-dI) acts via transcriptional interference. On the other hand, enhancer-dCas9-VP64 (en-64bO), enhancer-dCas9 GeneSwitch (en-GSO) and enhancer-dCas9-VPR (en-VPRO) are designed to overexpress target genes (O = overexpression) (Figure 6.2). More specifically, en-DmC is codon-optimized for *Drosophila*, while en-HsC is codon-optimized for humans (and identical to the afore-mentioned Cas9.P and Cas9.P2, respectively). en-dFC is a fusion of dCas9 with the nuclease domain of FokI and designed to cut target DNA upon dimerization of the FokI nuclease domain (requires two gRNAs ~ 15-25 bp apart). en-dI is designed for transcription interference (CRISPRi), where promoter-bound dCas9 will not cut DNA but rather sterically inhibit the proper formation of the pre-initiation complex. Taken together, this set of modified Gateway plasmids can be easily adapted to generate specific enhancer/Cas9 combinations, followed by PhiC31-mediated locus-specific transformation. In order to examine the efficiency of this system, we generated PG-specific versions of these vectors.

6.3.3 Transgenic lines carrying PG-specific expression of Cas9

For each of the above-listed vectors, we generated corresponding versions that express Cas9 and its variants under the control of the *spookier* (*spok*) regulatory region (Figure 6.3), which mediates highly specific expression in the PG (Figure 6.7A) [356,357]. When we examined the survival rates of these transgenic lines, we noticed that populations heterozygous for any of the *spok-Cas9* constructs were healthy and showed no significant adulthood lethality compared to controls (Figures 6.5 and 6.7A, B). However, homozygous animals, with the exception of *spok_GSO*, displayed some lethality during larval development and substantial or complete lethality during late larval and early pupal stages (Figures 6.7A, B). In conclusion, heterozygous transgenic lines are viable and can be kept as balanced stocks.

6.3.4 Localization of Cas9 in the PG

Before examining whether our Cas9 transgenes caused PG-specific alterations in gene expression, we first examined the presence of Cas9 protein in PG nuclei. Previous studies have shown that epitope tags might affect the DNA-binding properties of Cas9, which prompted us to remove the 3xFLAG tags found in the original dCas9.VPR plasmid, which ensured that all Cas9 transgenes were untagged. As expected, immunostaining with anti-Cas9 antibodies showed robust presence in PG nuclei, while the expression in nearby tissues, including the CA and CC was negligible (Figure 6.8).

6.3.5 Mutation efficiency of PG-specific gene disruption via Cas9

The ability to generate somatic gene mutations is still limited in *Drosophila* and has not been reported for the PG. We therefore used three different strategies to generate transgenes with PG-specific Cas9 expression, comprising spok_DmC (fly codon-optimized), spok_HsC (human

codon-optimized) and spok_dFC (FokI nuclease domain fused to fly codon-optimized dCas9). In order to test the efficiency for each approach, we generated transgenic lines carrying gRNAs against either *phantom* (*phm*) or *disembodied* (*dib*), two well characterized genes involved in ecdysteroid synthesis [358,359]. Classic mutants of *phm* and *dib* display embryonic lethality, while PG-specific *phm*- and *dib*-RNAi cause L1 and L3 arrest, respectively. Both *phm*- and *dib*-RNAi populations can be rescued to adulthood when reared on 20E-supplemented media (Figures 6.9A, B) [358,360], and we reasoned that the specificity of *phm*- and *dib*-gRNAs could be easily assessed by 20E-feeding as well. To generate double-strand breaks (DSB) in the coding region of *phm* or *dib*, we generated transgenic lines that carried at least two gRNAs (dib^{gR1} and phm^{gR1}), where the distance between target sequences would not exceed 400 bp, although a distance of up to 750 bp has been reported to work as well (Table 6.2 and Figure 6.10) [358,360,361]. For spok_dFC, DSBs are not achieved by the endonuclease activity of Cas9 (which is missing in dCas9) but require dimerization of the FokI nuclease domain, a bacterial type IIs restriction enzyme [362]. Dimerization of FokI is dependent on the recruitment of two Cas9 molecules guided by two distinct gRNAs that are 15-25 bp apart. We therefore generated transgenic lines that carry two pairs of gRNAs to allow for Cas9-FokI-mediated deletions (Table 6.2 and Figure 6.10).

In summary, using either spok_DmC, spok_HsC or spok_dFC to induce PG-specific DSBs in the *phm* and *dib* genes yielded similar results, and all caused phenotypes that were similar to those seen in *phm*- and *dib*-RNAi animals. The Cas9 lines were less leaky than the RNAi approach, with very few pupal and no adult escapers [199]. Importantly, Cas9/gRNA animals were rescued to adulthood when reared on a diet supplemented with 20E, with typically 70-80% of the population developing into adults [199], suggesting that the phenotypes resulted specifically from gene disruptions in *phm* and *dib* (Figures 6.9A, B).

To further confirm the specificity of the *dib* and *phm* gRNAs, we crossed either line to *act-Cas9*, allowing us to target *dib* and *phm* in a ubiquitous manner, which should give rise to phenotypes that are similar to the corresponding classic mutants. In agreement with this, both *act-Cas9.P2>phm^{gR1}* and *act-Cas9.P2>dib^{gR2}* were embryonic lethal, and thus phenocopied the classic mutants (Figures 6.9A, B). To ensure that these phenotypes were indeed caused by a disruption of the intended target genes, we extracted genomic DNA from hand-dissected ring glands and sequenced the *phm* and *dib* gene regions. As a control, we isolated genomic DNA from the adjacent brain. Upon sequencing at least 10 clones per line, we found that both *Drosophila*- and human-optimized Cas9 (spok_DmC and spok_HsC), in combination with two gRNAs, were highly efficient in generating deletions in the predicted region (Figures 6.9C-D). Some of clones appeared to be wild type alleles (not more than three out of ≥10 per line), however, since the ring gland samples comprised two non-targeted tissues (the corpora cardiaca and the corpora allata, Figure 6.1), we cannot distinguish between loci that were not targeted in the PG and loci that originate from the other two Cas9-free cell types. In comparison to spok_DmC and spok_HsC, using spok_dFC in combination with two gRNA pairs resulted in fewer large deletions, suggesting this approach was less efficient in this regard. However, on a phenotypic level, spok_dFC was just as efficient as spok_DmC and spok_HsC, all of which were 100% lethal. All tested clones derived from brain samples were wild type, indicating that the *spok* regulatory region does promote little or no expression in brain cells (Figure 6.9E).

6.3.6 *In vivo* transcription interference via PG specific dCas9 (spok_dI)

CRISPR applications are not limited to ablating gene function via DSBs. An alternative strategy is to interfere with the transcription of a target gene (CRISPRi). This is a desirable approach for selectively targeting specific promoters of genes that harbor alternative promoters to

repress specific mRNA isoforms. Alternatively, one could interfere with the expression of a gene for a defined duration, and then revert back to normal expression, thus studying dynamically expressed genes. We wanted to test whether PG-specific CRISPRi would work as efficiently as the other tools at our disposal. As candidate genes, we again chose *phm* and *dib*. We generated lines expressing a single gRNA targeting either -423 or -174 bp upstream of the *phm* TSS, and for *dib* we selected -482 and -110 bp upstream (Figure 6.10). When we crossed these four gRNA lines to flies carrying *spok_dI* transgenes, we observed developmental arrest during the L3 stage, which, in the case of *dib*, was comparable to what we had observed in *phm22-Gal4>dib*-RNAi animals (Figures 6.11A, B) and the corresponding Cas9-driven gene knockouts (Figure 6.9A). In contrast, targeting *phm* via CRISPRi, while lethal, was less efficient compared to the other strategies, as larvae died at later stages. This may suggest that the chosen gRNA sites were too far away from the *phm* TSS, since no alternative promoters have been reported for this gene. However, based on qPCR analysis, the relative reduction of transcript levels in the CRISPRi lines were comparable between *phm* and *dib*, as we observed a 2- to 6-fold reduction for *phm* and a 2.5- to 4-fold reduction for *dib* (Figure 6.11B). It is possible that *phm* transcript levels need to be even more strongly reduced to elicit phenotypes that are comparable to the Cas9 knockouts. Finally, to ensure that these phenotypes arise not from DSBs, we sequenced these loci and found them to be wild type in all cases (Figure 6.11C).

6.3.7 Upregulating gene expression via PG-specific CRISPR/Cas9

Previous approaches aimed at overexpressing a specific gene were based on the generation of transgenic lines that carry a cDNA, either driven by heat-shock promoters, nearby enhancers or the Gal4-UAS system [363–365]. Later improvements included the use of the PhiC31 system to ensure locus-specific integration and consistent expression of the transgene [120]. However, these

approaches require the generation and cloning of a cDNA, which may be time-consuming and difficult. Using dCas9 variants that harbor activation domains, one can now direct dCas9 to specific endogenous promoters and activate any given target gene, referred to as CRISPRa (a= activation). We therefore generated PG-specific versions of dCas9, to which we fused the VP64 or VPR activation domains [90,366], named here spok_64bO and spok_VPRO (short for spok_dCas9-VP64b and spok_dCas9-VPR) (Figure 6.7A). A report by the Perrimon lab showed that dCas9-VP64 was not as efficient as dCas9-VPR to activate target genes [90]. However, the dCas9-VP64 construct only contained two nuclease-attenuating mutations D10A and H840A compared to the dCas9-VPR, which contained four (D10A, H839A, H840A and N863A). We therefore modified the original *dCas9-VP64* to *dCas9-VP64b*, so that it contained the same four nuclease-attenuating mutations as the *dCas9-VPR* construct.

We first examined the efficiency of the *spok_64aO*, *spok_64bO* and *spok_VPRO* constructs by transfecting cultured brain-ring gland complexes (BRGC) that carried gRNA transgenes targeting either the *phm* or *dib* promoters upstream of the TSS, which were the same lines as used for CRISPRi. Since the plasmid-encoded *Cas9* alleles were driven by the *spok* regulatory region, we reasoned that this approach should result in PG-specific *Cas9* expression. Indeed, when we used ring gland-specific qPCR, *spok_VPRO* resulted in a 10- to 30-fold induction, while *spok_64bO* ranged from 5- to 15-fold upregulation. In contrast, the *spok_64aO* plasmid showed essentially no increased gene expression, suggestion that the two additional point mutations in *spok_64bO* (839A and N863A) are critical for induction (Figure 6.12A). In order to ensure that *spok_64bO* and *spok_VPRO* worked similarly efficient *in vivo*, we generated corresponding transgenic lines, and crossed them to transgenic lines carrying gRNAs that target regions upstream of the *phm* and *dib* TSS. Similar to our BGRC transfection results, *spok_VPRO* resulted in a 9- to

28-fold induction, while *spok_64bO* upregulated expression ranging from 4- to 18-fold (Figure 6.12B).

Finally, we also generated a version of CRISPRa that allows for temporally-controlled gene induction. We wondered whether we could render Cas9 ligand-inducible, similar to the GeneSwitch (GS) system, where the Gal4 DNA-binding domain is fused to the human progesterone receptor-ligand-binding domain (hPR-LBD) and the p65 activation domain. The resulting chimeric Gal4 protein can only be activated in the presence of steroid mifepristone (RU486), typically provided in the diet. Therefore, we cloned a *spok*-driven chimeric cDNA encoding the catalytically inactive dCas9 fused to hPR-LBD and p65 (Figure 6.12C) and generated the corresponding transgenic line (aka spok_GSO). In order to assess the efficiency of RU486-mediated induction, we chose target genes we are actively studying in the lab (*Alas*, *Coprox*, *FeCH*, *IRP1A*, *spz5*, *Nach*) and that have comparatively flat expression profiles in the PG compared to *phm* and *dib* during larval development [67]. When we crossed the corresponding gRNA transgenic lines to *spok_GSO* and switched larvae to a RU486-containing diet, we observed PG-specific upregulation as early as two hours after exposure to RU486, similar to what has been reported in Gal4GS system (Figure 6.12D) [354]. After four hours, induction of target genes ranged from 4- to 15-fold compared to controls, indicating that the GeneSwitch system works well, and is a powerful tool to temporally control gene upregulation.

6.3.8 Using PG-specific gRNAs for modulating gene expression

To manipulate gene expression in a tissue-specific manner via CRISPR, one can utilize two main strategies: (i) restricting Cas9 expression to specific tissues or (ii) limiting the expression of gRNA to the tissue of interest. In the approaches outlined above, we employed PG-specific Cas9 expression. We therefore tested whether reversing gRNA and Cas9 expression patterns from

tissue-specific to ubiquitous (and vice versa) was a viable strategy, because an existing line with ubiquitous Cas9 expression was reported to be homozygous viable (Bloomington stock #58590). A previous study described using UAS-driven multiplexed gRNA cloned into pCFD6 (Addgene 73915) to mediate tissue-specific gRNA expression in *Drosophila* imaginal wing discs [94]. In contrast to pCFD6, other gRNA-generating vectors use U6-type promoters (pU6), which are RNA Polymerase III promoters that drive ubiquitous expression of gRNAs [91,94]. This pU6-based approach has the potential to cause non-specific mutagenesis in non-target tissues where Cas9 expression is leaky [94]. However, since pCFD6 requires an additional Gal4-expressing transgene, and building the corresponding fly lines to achieve tissue-specific lesions is thus more complex.

In an effort to improve available tools for tissue-specific gRNA production, we replaced the pU6:3 promoter in the commonly used pCFD5 plasmid [94] with the *spok* regulatory region and added sequences mediating hammerhead (HH) or Hepatitis delta virus (HDV) ribozyme function to induce self-cleavage and proper release of gRNAs [350,367] (Figure 6.4). The pCFD5 plasmid harbors two Gly-tRNA sequences that allow insertion of two gRNAs. However, additional tRNA-gRNA pairs can be added if one requires more than two gRNAs, which we recommend for targeting large genes or if one wishes to target multiple genes with a single construct. Unlike pU6 promoters, the *spok* enhancer recruits Polymerase II (Pol II), resulting in mRNAs that will be subjected to 5'-capping and 3'-polyadenylation, which have the potential to interfere with proper gRNA maturation [368,369]. Therefore, we added the HH and/or HDV ribozyme sequences to three of our vectors (PG2-4) to test whether this would result in more efficient phenotypes due to increased processing of gRNAs, while one vector (PG1) received no ribozyme sequence (Figure 6.4). As a consequence of this design, the resulting transgenic lines require only a single cross to combine Cas9 and gRNA in the F1 generation. The use of ribozyme sequences has been

successfully used in zebrafish and *Arabidopsis* [370,371], but has not been described in *Drosophila*. Compared to the original pCFD5 vector, all cloning steps for our PG-gRNA plasmids are exactly the same, requiring no additional adjustments in terms of cloning strategy [94].

In total, we generated four different PG-gRNA vectors, three of which are designed for generating DSBs (PG1-3), while one them (PG4) harbors an MS2 aptamer [372] to mediate gene upregulation (Figure 6.4). For PG1-gRNA, we did not add any ribozyme sequences, while PG2-gRNA has an HDV ribozyme sequence at the 3' end and PG3-gRNA harbors an HH ribozyme at the 5' end as well as HDV region at the 3' end of the multiplex. In order to evaluate the efficiency of this approach, we inserted the same two gRNAs targeting the *dib* gene that we used for ubiquitous pCFD5-driven gRNA expression, ensuring that any differences in phenotypes arise from ubiquitous vs. tissue-specific gRNA expression (Table 6.2, Figure 6.10).

When we crossed transgenic lines carrying either *PG1-dib^{gR1}*, *PG2-dib^{gR1}* or *PG3-dib^{gR1}* to Act-Cas9, we observed consistently 100% L3 arrest, similar to *spok9>pU6-dib^{gR1}* animals that produced ubiquitous gRNA and PG-specific Cas9. In addition, supplementation with 20E efficiently rescued the L3 lethality, resulting in 72-82% normal-looking adults, indicating the specific disruption of the *dib* gene (Figure 6.13A). However, when we sequenced genomic DNA from dissected ring glands, we noticed that the *PG1-dib^{gR1}* transgene, which lacked the ribozyme sequences, was less efficient compared to *PG2-dib^{gR1}* and *PG3-dib^{gR1}*. Specifically, while *PG2-dib^{gR1}* or *PG3-dib^{gR1}* consistently caused deletions resulting from mutations at both gRNA loci, the *PG1-dib^{gR1}* construct failed to produce mutations for the downstream gRNA and therefore lacked corresponding deletions (Figure 6.13C). This suggests that the addition of the 3' HDV ribozyme sequences was necessary to allow for effective processing of the downstream gRNA. Taken together, these data demonstrated that restricting the expression of gRNA to the PG is highly

effective and provides an alternative solution to the lethality issue caused by PG-specific Cas9 expression, provided one uses ribozyme sequences to ensure appropriate processing of gRNAs from Pol II-derived mRNAs.

We have not tested the functionality of the PG4 construct, which harbors the MS2 aptamer, but make it available for public testing. The addition of the MS2 sequence promotes the recruitment of the MS2 coat protein (MCP) [372] and MCP fusion proteins such as MCP_p65_hsf1 and MCP_VP64, which are available as transgenic fly lines ("flySAM") [373]. The binding of MCP-coactivator fusions is aimed at causing stronger gene upregulation compared to VP64 and VPR alone.

6.3.9 Overview over new CRISPR/Cas9 constructs (gG-Cas9 II) and their properties

In total, we generated three new all-purpose vectors (Figure 6.14). Based on these, we then generated three vectors with a PG-specific enhancer, three vectors under control of the actin promoter and three vectors that are based on UAS (Figure 6.15). In this section, we will provide an overview over the purpose of these different vectors. In the subsequent paragraphs, we will i) discuss the experiments that helped us to identify the specific Cas9 variants with the desired properties and ii) provide proof-of-principle that these Cas9 variants work *in vivo* and in cell cultures.

The new vectors are based on our previously reported pG_Cas9 plasmids [199], which we used to replace the original Cas9 gene with our newly developed Cas9 variants (Figure 6.14). The new vector series falls into three groups, GSD (GeneSwitch-dependent Disruption), RDX (Rapamycin-dependent Disruption) and RDO (Rapamycin-dependent Overexpression). In brief, GSD and RDX are inducible by RU486 (aka Mifepristone) and rapamycin, respectively, and

function by generating double-strand breaks (DSBs) and deletions via site-directed cleavage. In contrast, RDO can be used to temporally overexpress target genes (Figure 6.14).

The GSD construct is based on a fusion of nuclease-dead Cas9 (dCas9), human Progesterone-Receptor ligand-binding domain (hPR-LBD) and the nuclease domain of FokI. This chimeric Cas9 enzyme can be activated by RU486, upon which it forms homodimers and cuts target DNA via the FokI nuclease domain. This approach requires two gRNAs that are ~15-25 bp apart [374]. In the RDX construct, the Cas9 gene is split and was separated at the recognition lobe. The C-terminal part of Cas9 is fused to the 12-kDa FK506 binding protein (FKBP12 or FKBP) while the N-terminal part is fused to the FKBP-rapamycin binding domain of mTOR (FRB). These two Cas9 halves are initially expressed as a single polypeptide, since the vectors use a single regulatory region. However, the protein is then cut by a P2A self-cleavage peptide, which allows for equivalent expression of the two Cas9 fragments [98,375]. Upon binding to (exogenously provided) rapamycin, FRB will bind to FKBP and form a heterodimer and cause the reconstitution of Cas9 [376,377].

Finally, Cas9-RDO is similar to Cas9-RDX, however, the Cas9 is now replaced with dCas9, and thus unable to cut DNA. In addition, the C-terminal dCas9 fragment, besides having FKBP and an NLS, is also fused to three transactivation domains, VP64, p65 and Rta (collectively referred to as VPR). Upon full assembly in the presence of rapamycin, the promoter-bound dCas9 will not cut DNA but rather act as a potent transcriptional activator by recruiting general transcription factors. Taken together, this set of modified gateway-compatible plasmids can be easily adopted to generate specific enhancer/Cas9 combinations, followed by PhiC31-mediated locus-specific transformation. The PhiC31 integrase system has been widely used in different insects species like silkworm *Bombyx mori*, yellow fever mosquito *Aedes aegypti* or the moth

Plutella xylostella [378–380]. Thus, the gateway-compatible plasmids presented here should be – after appropriate modifications - suitable for insect species other than the fruit fly. In order to examine the efficiency of this system, we generated transgenic lines that carry these constructs and tested their *in vivo* activity.

6.3.10 Mutation efficiency and tissue specificity of gene disruption via GSD constructs

As outlined above, certain limits still exist with current somatic CRISPR/Cas9 approaches, in particular because Gal4/UAS-driven Cas9 may cause increased lethality that is independent of Cas9 nuclease activity [91,199,381]. Even when bypassing Gal4 through the use of tailored regulatory regions to express Cas9 in a tissue-specific manner, one may trigger developmental defects that prevents the analysis of later stages [199]. We therefore attempted to solve this issue by designing Cas9 variants that can be controlled by the addition or removal of a dietary compound.

We previously reported the development of a RU486-dependent dCas9 variant as a tool to ectopically activate genes. In this approach, the hPR-LBD was fused with dCas9 as well as the p65 activation domain [96,199]. We showed that the resulting chimeric Cas9 protein activated target genes only in the presence of RU486, effectively allowing us to overexpress target genes with spatial and temporal control. This succesful approach prompted us to examine whether we could design a RU486-dependent Cas9 variant for gene disruption as well. We reasoned that one challenge with this approach was that one cannot simply add the LBD to the C-terminus of Cas9, and expect that this would render the folding of the nuclease domain ligand-dependent. Likewise, it was unclear whether the unliganded hPR-LBD would retain the protein in the cytoplasm, since the proteins that mediate nucleocytoplasmic shuttling of hPR are not necessarily conserved in *Drosophila* [324]. Another issue was that it is not straightforward to insert the LBD within the

Cas9 nuclease domain (to render the domain ligand-dependent) without disrupting the nuclease activity.

Cas9 is composed of several domains involved in gRNA recognition, PAM recognition, and nuclease activity [382,383] (Figure 6.16). There are two nuclease domains (NUC) present in Cas9. Upon binding to gRNA, the Cas9 protein undergoes a conformational change that re-aligns the two nuclease domains so that target DNA can be cut. Each domain is responsible for cutting a single target strand [382,383]. Based on this, one can modify the Cas9 protein structure and still ensure its cleavage efficiency as long as the protein can fold properly. Based on the crystal structure of Cas9 in complex with gRNA and appropriate target DNA, there are 11 sites which can be potentially targeted to modify the Cas9 sequence without affecting its function (Figure 6.16A) [377,383]. To examine whether insertion of the hPR-LBD would yield ligand-dependent Cas9 nuclease activity, we inserted the hPR-LBD at all 11 positions and tested whether this would trigger DNA cleavage in the presence of RU486 (Figure 6.16A). Specifically, we generated plasmids that harboured the Cas9_hPR-LBD construct as well as multiplexed tRNA-gRNAs to simultaneously express the chimeric Cas9 protein and two gRNAs in S2 cells. The gRNAs were based on previously validated target sites of the *disembodied* gene (*dib*) dU6-dib[gR1] (Figure 6.16B) [199]. We then used the surveyor mutation detection assay (Integrated DNA Technologies), which uses the surveyor nuclease to detect and cleave mismatches between the reference DNA and Cas9-treated DNA, which may be caused by single nucleotide polymorphisms (SNPs), small insertions or deletions. By measuring the intensity of cleaved fragments, one can evaluate the cleavage efficiency by which Cas9 variants cut target DNA in the presence and absence of RU486 [384]. Among the 11 inserts we examined, none resulted in detectable cleavage (Figure 6.16C). We also sequenced 10-15 clones for each construct but found no evidence of gene editing at the target sites.

We concluded that the large size of the hPR-LBD (47.4 kDa) likely interfered with the proper folding of Cas9 regardless of which insertion site was used, thus abolishing Cas9 nuclease activity.

We then took a different approach where we fused dCas9 with the hPR-LBD and FokI nuclease domains. In this approach the hPR-LBD was attached to the C-terminus of the dCas9 protein and as such, should not affect the folding of dCas9 protein. Since dCas9 cannot cut DNA, the nuclease activity is provided by the FokI nuclease domain, which requires dimerization in order to cut DNA (Figure 6.17) [362]. Dimerization and successful DNA cleavage via FokI requires two Cas9 molecules each bound to a different gRNAs that correspond to target sites that are 15-25 bp apart. This vastly reduces potential off-target effects, since one requires two distinct gRNAs for the FokI nuclease domain to function [199,374]. A key challenge with this approach was to ensure that the two FokI nuclease domains only dimerize in the presence of RU486. We reasoned that a critical factor of this design was the linker length between hPR-LBD and the FokI nuclease domain. A linker that is too short might inhibit FokI dimerization altogether because it is too close the LBD, while a linker that is too long may cause FokI to dimerize independent of RU486. We hypothesized that the FokI domain requires a certain degree of movement for dimerization. This prompted us to test the widely used flexible linker $(GGGGS)_n$, which is composed of four glycines and a single serine per unit. Glycine is a non-polar amino acid, and its small size provides flexibility and mobility for the FokI domain while the presence of a serine residue is thought to maintain the stability of the linker and reduce the unfavorable interactions between the linker and protein moieties [385] We experimented with the number of linker repeats to find the optimal length for proper dimerization of FokI. To test for cleavage efficiency, we generated co-transfection plasmids for expression in S2 cells, each of which expressed the GSD construct with different linker lengths as well as two gRNAs targeting *dib* (dU6-dib^{gR2}) (Figure 6.18A) [199]. We

then used the surveyor mutation detection assay to evaluate Cas9 activity. We tried 5 different linkers length ranging from 5-25 amino acids (in increments of 5 amino acids) in the absence or presence of RU486. Remarkably, we found that the length of the linker played a decisive role. When we used a linker length of 15 amino acids, we measured that 25% of the target DNA was cleaved in the presence of RU486, with no detectable cleavage in the absence of the ligand. In contrast, Cas9-GSD constructs with shorter linkers (5 or 10 amino acids) failed to cleave DNA even in the presence of RU486, while longer linkers (20 and 25 amino acids) caused DNA cleavage both in the presence and absence of RU486 (Figure 6.18B). This suggests that a linker length of ≥20 amino acids is too long and allows FokI domains to dimerize independently of the conformation in the LBD. Taken together, this approach determined that a linker comprising 15 amino acids ((GGGGS)$_3$) allows for sufficient nuclease activity in a RU486-dependent manner. As such, the GSD constructs we then generated for our *in vivo* experiments were based on a single linker with a length of 15 amino acids to separate the hPR-LBD and FokI domains.

Based on the above *in vitro* results, we then generated transgenic lines carrying either *spok_GSD*, which expresses Cas9-GSD specifically in PG, *act_GSD*, which expresses the construct ubiquitously, as well as *UAS-GSD*, which requires a Gal4 driver transgene for expression. In order to ensure these constructs worked efficiently *in vivo*, we crossed them to transgenic lines carrying two pairs of validated double gRNAs that target coding sequences of the *dib* gene (*dU6-dib^{gR2}*) [199]. To test the *UAS-GSD* line, we generated flies that also harboured *dU6-dib^{gR2}* and *phm22-Gal4* (aka *phm22>*), a widely used PG-specific Gal4 driver [67,351].

In the presence of RU486, both *spok_GSD* or *phm22>UAS-GSD* potently interfered with *dib* function and caused uniform arrest of larval development in the third instar. Compared to the PG-specific *dib*-RNAi line, the effects appeared to be stronger, as we observed only one out of

300 larvae that reached (and died during) the pupal stage, while the RNAi line yielded 6-10% pupae, with a total of 2% reaching adulthood. Animals from the same cross showed no developmental defects and were fully viable when being reared on media that lacked RU486 (Figure 6.19A). Since dib^2 mutants are embryonic lethal, we wanted to know whether ubiquitous expression of Cas9-GSD (in the form of *act-GSD*) in the presence of RU486 caused a similar result. To bypass the need for oral uptake of RU486, we stripped embryos of the chorion and immersed them into a solution that either contained RU486 or was absent of the drug, allowing uptake across the vitelline and egg cell membranes. This approach caused widespread embryonic lethality (89%), similar to what we found for ubiquitous *dib*-RNAi (92%). Compared to dib^2 mutants, the phenotypes seen in two conditional lines were not quite as strong, since dib^2 mutants displayed 100% embryonic lethality (Figure 6.19A). Taken together, the Cas9-GSD lines caused very similar phenotypes to those observed with established *dib* mutants and *dib*-RNAi lines, and were absent of any obvious phenotypes when RU486 was omitted.

To further assess the specificity of these results, we examined whether we could rescue the developmental arrest of the Cas9-GSD lines by providing embryos and larvae with 20-hydroxyecdysone (20E), a biologically active insect steroid hormone. The *dib* gene is required for the production of 20E and therefore both *dib* mutants and *dib*-RNAi lines can be effectively rescued by exogenously providing this hormone, which allows ~40-75% of the population to reach adulthood [386]. If our Cas9-GSD line was specifically targeting the *dib* gene, and was not causing additional off-target effects, we would expect a similar rescue when providing 20E. Indeed, when we added 20E to the diet or to the solution we used for embryo immersion, RU486-treated Cas9-GSD animals were rescued to adulthood. For the PG-specific lines targeting *dib*, we observed that 65-75% of the population developed into adults, comparable to *phm>dib*-RNAi animals (75%).

When we ubiquitously interrupted *dib* function in the presence of 20E, both the RU486-treated actin-*dib*-GSD and the actin>*dib*-RNAi displayed 41% survival to adulthood, suggesting that the phenotypes resulted specifically from disrupting *dib* function (Figure 6.19A). Finally, for all crosses mentioned above, we conducted equivalent crosses on media without RU486. In all cases, the resulting animals developed normally and displayed not obvious defects, suggesting that any residual activity Cas9-GSD may have is negligible or non-existent.

To ensure that these phenotypes were indeed caused by a disruption of the intended target gene, we extracted genomic DNA from hand-dissected ring glands (for PG-specific GSD constructs) or embryos (for ubiquitous GSD) and sequenced the *dib* locus. As a control, we either isolated genomic DNA from the adjacent brain or from animals raised on media without RU486. After examining the sequences of at least 10 clones per line, we found that all PG-specific GSD constructs, in combination with two gRNAs, were highly efficient in generating deletions in the predicted region (Figure 6.19B). Some of clones appeared to be wild type alleles (not more than three out of ≥10 per line), however, since the ring gland samples comprised two non-targeted tissues (the corpora cardiaca and the corpora allata) [387] we cannot distinguish between loci that were not targeted in the PG and loci that originate from the other two Cas9-free cell types. Importantly, all tested clones derived from either brain samples or hand-dissected ring gland from animals raised on media without RU486 represented wild type sequences, indicating that the *spok* regulatory region causes negligible or no expression in brain cells.

When we tested clones isolated from RU486-treated act-GSD animals in combination with gRNAs targeting *dib*, we found that 90% of the clones showed mutations at the predicted target sites. In contrast, animals reared on media without RU486 results only yielded *dib* wild-type sequences. Taken together, these data strongly suggest that the GSD constructs can only be

activated in the presence of RU486, and thus can be used for temporally controlled gene interference.

6.3.11 Rapamycin-dependent mutation efficiency and tissue specificity of CRISPR/Cas9 gene disruption

In addition to the GSD constructs, we also attempted to develop a second approach for temporal gene disruption via CRISPR/Cas9. We reasoned that it would be useful to have a system that uses a different compound in case researchers want to combine the commonly used GeneSwitch approach in *Drosophila* Gal4/UAS system (which already uses RU486) [388,389] with CRISPR/Cas9.

A rapamycin-inducible split Cas9 system was demonstrated to work in human embryonic kidney 293FT cells. In this approach, Cas9 is split into two halves, where each half is fused with a specific rapamycin dimerization domain, namely FK506 binding protein 12 (FKBP) for the C-terminal Cas9 fragment and the FKBP rapamycin binding domain (FRB) of the mammalian target of rapamycin (mTOR) for the N-terminal Cas9 fragment. Upon binding to rapamycin, the FKBP and FRB domains will dimerize and result in the formation of fully reconstituted Cas9 for gene editing activity [376,377,390,391]. The C-terminal fragment also encodes two nuclear localization signals (NLS), while the N-terminal Cas9 part was fused with FRB and a nuclear export signal (NES), which effectively separates the two Cas9 polypeptides into the nuclear and cytoplasmic compartment.

We wondered if a similar approach would work for *Drosophila in vivo*. We used the same construct from the mammalian study [377], and inserted it into the phiC31 vector to make it suitable for transformation (Figure 6.20A). In the absence of rapamycin, the NES sequence ensures that the N-terminal fragment will stay in the cytoplasm while the NLS sequences will result in the

nuclear localization of the C-terminal Cas9 fragment. This separation will reduce random

dimerization the two Cas9 fragments. The presence of rapamycin, however, will trigger

dimerization of FKBP and FRB (presumably newly translated protein), causing the assembly of

functional Cas9. In addition, the two NLS sequences on C-terminal fragment will ensure the proper

nuclear presence of Cas9 (Figure 6.20B). In our approach, these two Cas9 fragments were driven

by the same promoter and were only separated by a P2A self-cleavaging sequence. Even though

both fragments are initially translated together as a single polypeptide, the self-cleaving P2A

peptide will trigger the split at the weak peptide bond between Pro$^{(2B)}$ and Gly$^{(2A)}$ located near the

end of the P2A sequence, ensuring that the two Cas9 fragments will be separated into two

independent polypeptides (Figure 6.20B) [392]. We refer to this construct as CRISPR/Cas9

Rapamycin-dependent gene Disruption, or RDX.

We generated transgenic lines carrying *spok_RDX*, which expresses the Cas9 RDX variant

specifically in PG, *act_RDX*, which expresses the construct ubiquitously, and *UAS-RDX* for the

Gal4/UAS approach (Figure 6.15). To ensure that these constructs work efficiently *in vivo*, we

crossed them to transgenic lines carrying validated gRNAs that target coding sequences of *dib*

(*dib^{gR1}*). In the presence of rapamycin, using either *spok_RDX* or *phm22>UAS-RDX* to induce PG-

specific disruption of the *dib* gene yielded very similar results, and both caused 100% lethality in

the third instar, somewhat stronger to what we observed for PG>*dib*-RNAi animals. In contrast,

animals from the same crosses showed no developmental defects when being raised on media that

lacked rapamycin (Figure 6.21A). Similar to the results for *act-GSD*, rapamycin-treated *act-RDX*

animals displayed predominantly embryonic lethality, again comparable to *dib^2* mutants or

ubiquitous expression of *dib*-RNAi (Figure 6.21A). Importantly, Cas9-RDX/gRNA animals were

rescued to adulthood when reared on a diet supplemented with 20E, with typically 70-75% of the

population developing into adults for PG-specific deletion and 37% for the ubiquitous deletion of *dib*, suggesting that the phenotypes resulted specific mutations in the *dib* gene (Figure 6.21A). The lower 20E rescue rates are likely caused by inefficient uptake of 20E, since we had to immerse the embryos in a buffer containing 20E. This rescue protocol is time-sensitive, since longer immersion times cause asphyxiation (Chapter 2).

We also tested the gene editing efficiency via sequencing the *dib* locus to ensure the developmental defects in these animals were caused by mutations in the target gene. Upon sequencing at least 10 clones from each sample, we found the PG-specific expression of *Cas9-RDX* generated about 50-60% of mutations in the presence of rapamycin while DNA isolated from the brain from ring glands of rapamycin-free animals did not show any sequence alteration at the *dib* locus. Again, we did not expect to see 100% efficiency, because the ring gland is composed of three glands, two of which do not express the Cas9 construct. In addition, DNA extracted from embryos of ubiquitous *act-RDX* animals showed a high mutation rate (approximately 80-95%) (Figure 6.21B). In summary, these data indicate that the Cas9-RDX approach can efficiently edit target genes and represents a viable strategy to temporally control the onset of Cas9-mediated gene editing.

We also investigated how long it takes before exposure to dietary rapamycin results in gene editing events. For this, we reared *spok_RDX*; *dU6-dib^{gR1}* animals on a rapamycin-free diet until they reached 24 hours after L2/L3 molt. We then switched larvae to a diet supplemented with rapamycin, and isolated ring glands for DNA extraction at 0, 2, 4, 6 and 8 hours after the switch. For each sample we analyzed at least 10 clones via Sanger sequencing. Using this strategy, we detected the first targeting event at two hours, and by four hours we observed a significant mutation rate of 50% in the tested samples. In animals with 6 and 8-hour treatments, we found a stable 70%

mutation rate (Figure 6.21C). These data suggest that significant Cas9 activity occurs in as little as four hours of rapamycin exposure, with potentially faster rates in gut cells. Taken together, our second approach based on rapamycin is equally effective as the one based on RU486, indicating that either approach can be used to conditionally edit target sequences.

6.3.12 Temporal upregulating gene expression via split dCas9

We previously reported a tool for temporal gene upregulation using the GeneSwitch activation (GSO) [199]. In this approach, the dCas9 was fused with hPR-LBD and the activation domains (VPR) which includes VP64, p65 and Rta activation domains. The resulting chimeric Cas9 protein can only be activated in the presence of RU486. Upon binding to this ligand, the hPR-LBD will change its conformation and bring the VPR domain closer to the DNA and help recruit the transcription machinery. As a result, target gene expression was induced 10-15-fold [199]. Given that the split Cas9 system described above (Cas9-RDX) worked effectively, we wondered whether the same strategy could be used to induce gene expression. This approach has been successfully applied in mammalian cell culture systems where the dCas9 was fused with just the VP64 activation domain. To adopt this for *Drosophila*, we used dCas9 and split the gene into two halves. Each half was fused with the same domains and signal sequences we used for the RDX constructs (Figures 6.14 and 6.15), with the exception that the C-terminal dCas9 fragment was also fused with the VPR activation domains, which differs from the construct used in mammalian cells, which only used the VP64 domain. Favoring VPR over VP64 is based on previous studies in *Drosophila,* which showed that VPR works more efficiently compared to VP64 alone (Figures 6.15 and 6.22A) [90,199,377]. We refer this Cas9 variant as RDO (RDO = Rapamycin-dependent Overexpression).

We then generated transgenic lines carrying spok_RDO for PG-specific expression, act_RDO for ubiquitous expression and UAS-RDO for Gal4/UAS-driven expression (Figure 6.15). In order to assess the efficiency of rapamycin-mediated induction, we chose previously validated gRNAs to target genes we are actively studying in the lab (*Alas*, *Coprox*, *FeCH*). In addition, these genes have comparatively flat expression profiles in the PG compared to *dib* during larval development [67,166,199]. When we crossed the corresponding gRNA transgenic lines to *spok_RDO*, *phm22-Gal4>UAS-RDO* as well as *act_RDO* and switched 24-hours old third-instar larvae to a rapamycin-containing diet, we observed significant PG-specific upregulation of the three tested genes, ranging from 7- to 15-fold induction. The effects were detected as early as 4 hours after drug administration, indicating that the rapamycin-dependent induction system works well and fast, and as such represents a powerful tool to temporally control gene upregulation (Figure 6.22B).

6.3.13 Viability of transgenically expressed Cas9 variants

Finally, we used survival charts to assess whether any of the Cas9 variants caused lethality on their own. When we examined the survival rates of these transgenic lines, we noticed that populations homozygous for any of the *act-Cas9* and *UAS-Cas9* transgenes were healthy and can develop without any significant issues to adulthood (Figure 6.23), which is in agreement with previously established Cas9 stocks [91]. In contrast, animals homozygous for the PG-specific Cas9-GSD variant showed some lethality during development, which resulted in only about 50% of animals reaching adulthood. In comparison, the heterozygous PG-Cas9-GSD animals appeared to be healthy with about 75% of the population reaching adulthood (Figure 6.23). Animals that carried other PG-specific Cas9 variants, including *spok-RDX* and *spok-RDO* are also homozygous healthy (Figure 6.23). In conclusion, all transgenic lines were sufficiently healthy for experiments.

6.4   Conclusions and Significance

We showed here that PG-specific expression of Cas9 causes lethality that is independent of its nuclease activity. We have generated a series of strategies to solve this problem, which now allows somatic generation of DSBs, CRISPRi and CRISPRa. Generating tissue-specific gRNAs is also a viable strategy, provided ubiquitously expressed Cas9 levels are sufficiently low to avoid lethality. Since endoreplicating tissues harbor multiple copies of the same locus, Cas9 should be activated early enough to ensure efficient gene disruption. In our hands, somatic gRNAs work even better than RNAi lines, and appear to be highly specific, indicating that polytene tissues pose no issue for somatic CRISPR approaches. Even though we generated tools for PG-specific CRISPR, our tools can be used for any tissue of interest, polytene or not.

6.5   Figures
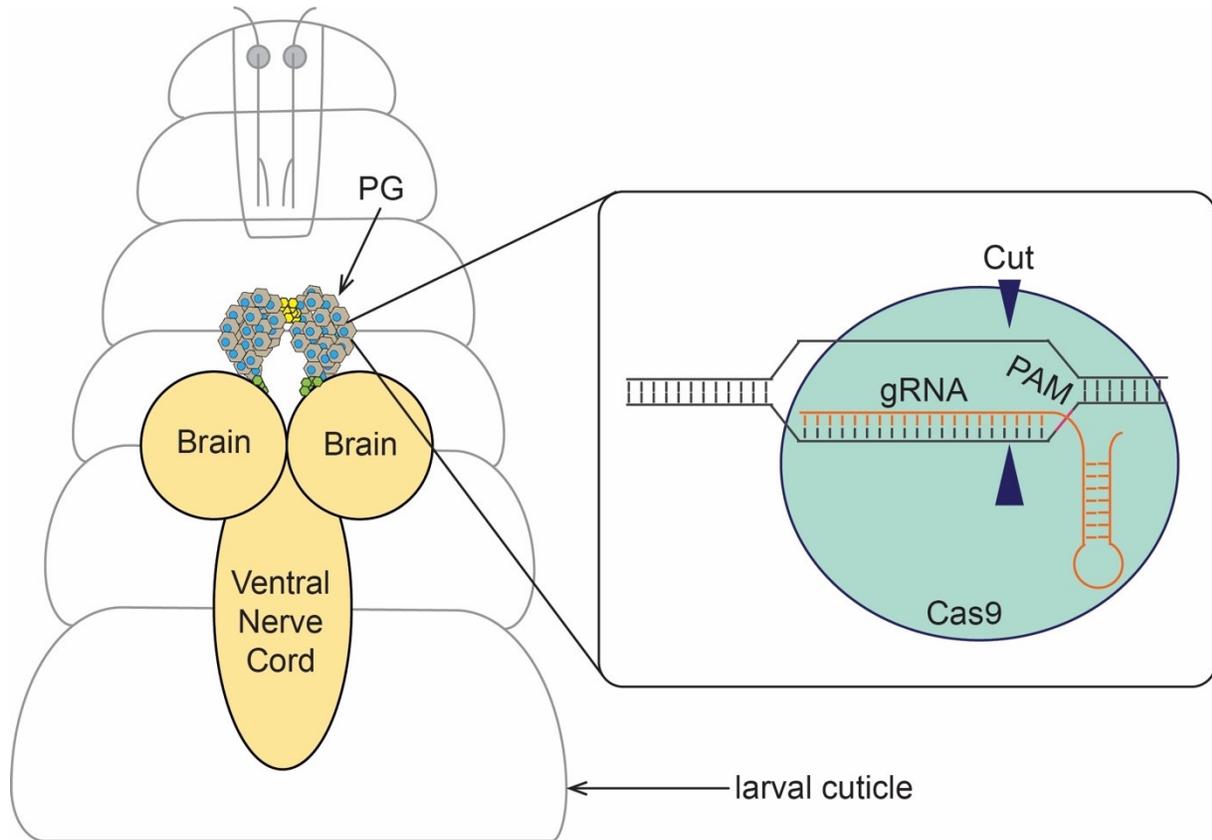
Figure 6.1. Somatic CRISPR in the Drosophila prothoracic gland.

In Drosophila larvae, the prothoracic gland (PG) is the principal source for ecdysteroid production. The PG is a part of the ring gland, which also harbours the corpora allata (yellow) and corpora cardiaca (green). PG-specific genome editing via Clustered Regularly Interspaced Short Palidromic Repeats (CRISPR) requires the recruitment of CRISPR-associated protein 9 (Cas9: blue) to the target site recognized by the guide RNA (gRNA: orange). Target site cleavage by Cas9 is ensured by the presence of the protospacer adjacent motif (PAM: purple) sequence immediately following the target site. This sequence will direct the cut site of Cas9 to a region of about three nucleotides upstream of the PAM.
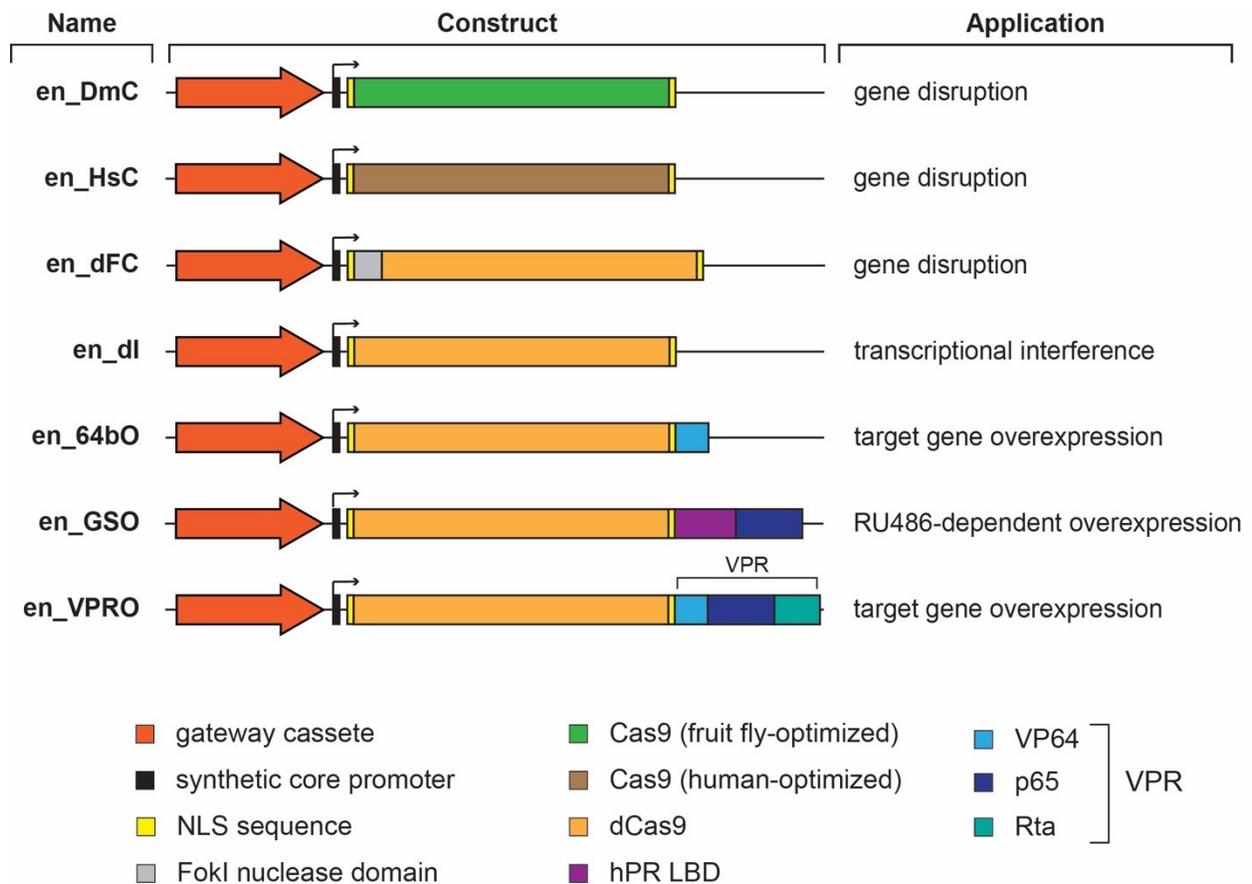
Figure 6.2. The general Gateway Cas9 (gG-Cas9) vector collection.

Each gG-Cas9 vector backbone is composed of a mini-*white* gene as a marker, a *PhiC31* integrase-compatible *attB* site, and the *bla* coding sequence to mediate ampicillin resistance. Shown here are the gateway cassette, the Cas9 variant, the regions encoding Nuclear Localization Sequences (NLS), activations domains (VP64, p65 and Rta), the human Progesterone Receptor ligand-binding domain (hPR LBD) and the FokI nuclease domain. The Gateway cassette allows to use LR recombination to insert enhancer/promoter regions to drive tissue-specific Cas9 expression. en_DmC, en_HsC and en_dFC can be used to generate somatic mutations. En_DmC uses a fruit fly codon-optimized Cas9 version, while en_HsC is optimized for human cells. En_dFC cuts DNA upon FokI-mediated dimerization followed by FokI cleavage, since dCas9 (= dead Cas9) is unable

to cut DNA [362,374]. However, the dCas9 vector can be used to interfere with transcription by guiding Cas9 into the vicinity of transcriptional start sites where it may block the assembly of the pre-initiation complex. en_64bO, en_GSO and en_VPRO (O = overexpression) were designed achieve upregulation of target genes. en_GSO (GS = GeneSwitch) encodes a protein where Cas9 is fused to the hPR LBD and p65 domain, allowing activation via RU486
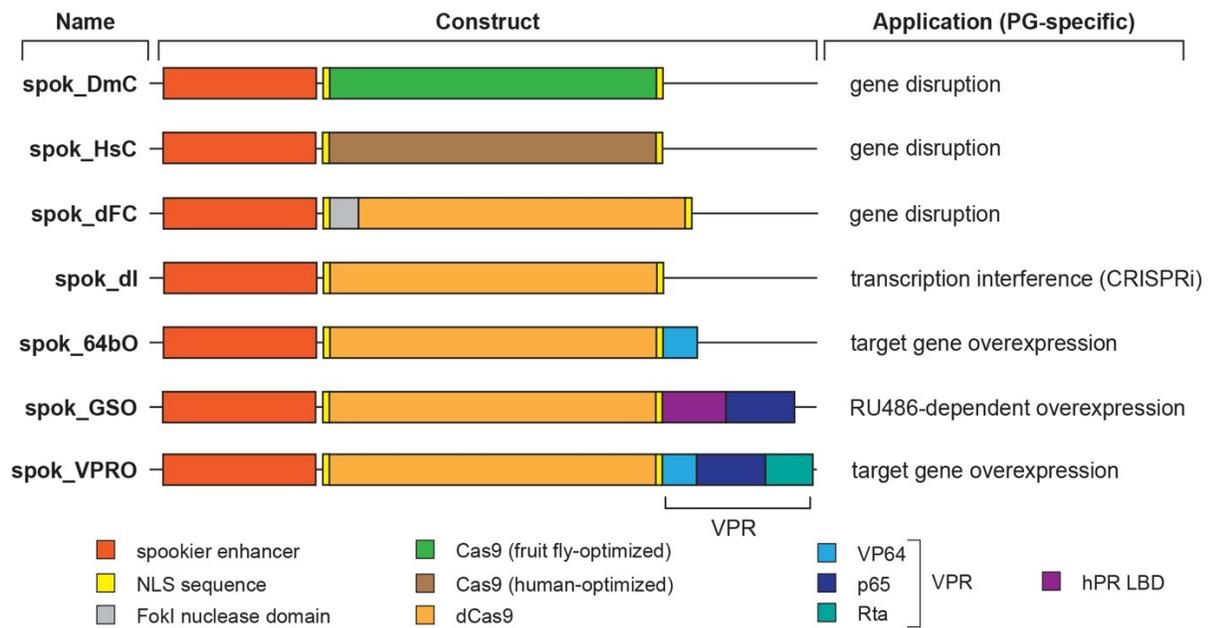
Figure 6.3. The PG-specific Cas9 (PG-Cas9) vector collection.

All vectors are based on the general Gateway Cas9 vector collection (Figure 6.2). Each PG-Cas9 vector backbone is composed of a mini-*white* gene as a marker, a *PhiC31* integrase-compatible *attB* site, the *bla* coding sequence to mediate ampicillin resistance, and a synthetic core promoter. Shown here are the *spookier* (*spok*) regulatory region, the Cas9 variant, the regions encoding Nuclear Localization Sequences (NLS), activations domains (VP64, p65 and Rta), the human Progesterone Receptor Ligand-Binding Domain (hPR LBD) and the FokI nuclease domain. spok_DmC, spok_HsC and spok_dFC can be used to generate somatic mutations (C = cleavage). spok_DmC uses a fruit fly codon-optimized Cas9 version, while spok_HsC is optimized for human cells. spok_dFC cuts DNA upon FokI-mediated dimerization followed by FokI cleavage, since dCas (= dead Cas9) is unable to cut DNA. However, the spok_dI (I = interference) vector harbors dCas9 and can be used to interfere with transcription (CRISPRi) by guiding Cas9 into the vicinity of transcriptional start sites where it may block the assembly of the pre-initiation complex. spok_64bO, spok_GSO and spok_VPRO (O = overexpression)were designed to achieve

226

upregulation of target genes. spok_GSO (GeneSwitch activation) encodes a protein where Cas9 is

fused to the hPR LBD and p65 domain, allowing activation via RU486.
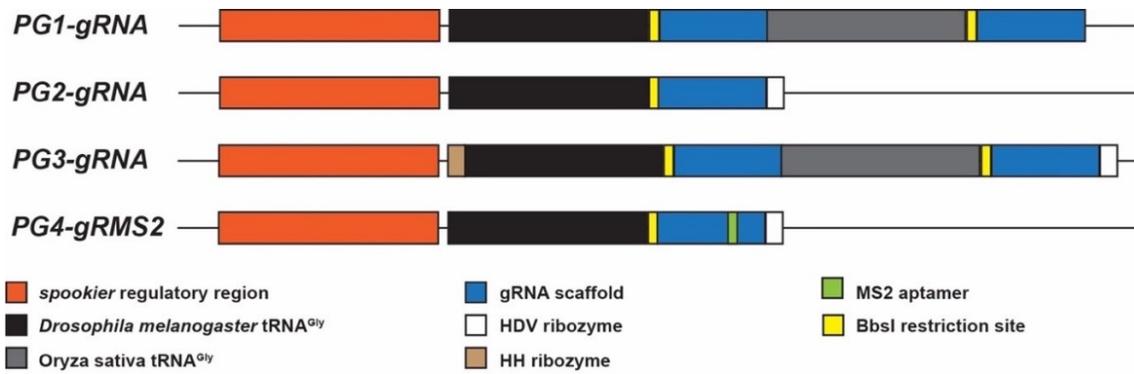
Figure 6.4. Schematic illustration of PG-gRNA constructs.

These constructs allow for prothoracic gland-specific expression of multiple gRNAs in a single vector, which are based on the commonly used pCFD5 plasmid where we replaced the pU6:3 promoter with the *spok* regulatory region. We also added sequences mediating hammerhead (HH) or Hepatitis delta virus (HDV) ribozyme function to promote proper processing of gRNAs from Pol II-derived mRNAs [350,370]. Like the original pCFD5 plasmid, this vector series harbors two tRNA^{Gly} sequences that natively allow the insertion of two gRNAs, but additional gRNA-tRNA fragments can be added to target larger regions of DNA. PG1-3 are used for gene disruption, while PG4 (not tested in this study), which harbors an MS2 aptamer, is intended for gene activation in combination with fly lines that carry MCP_p65_hsf1 or MCP_VP64 transgenes ("flySAM").
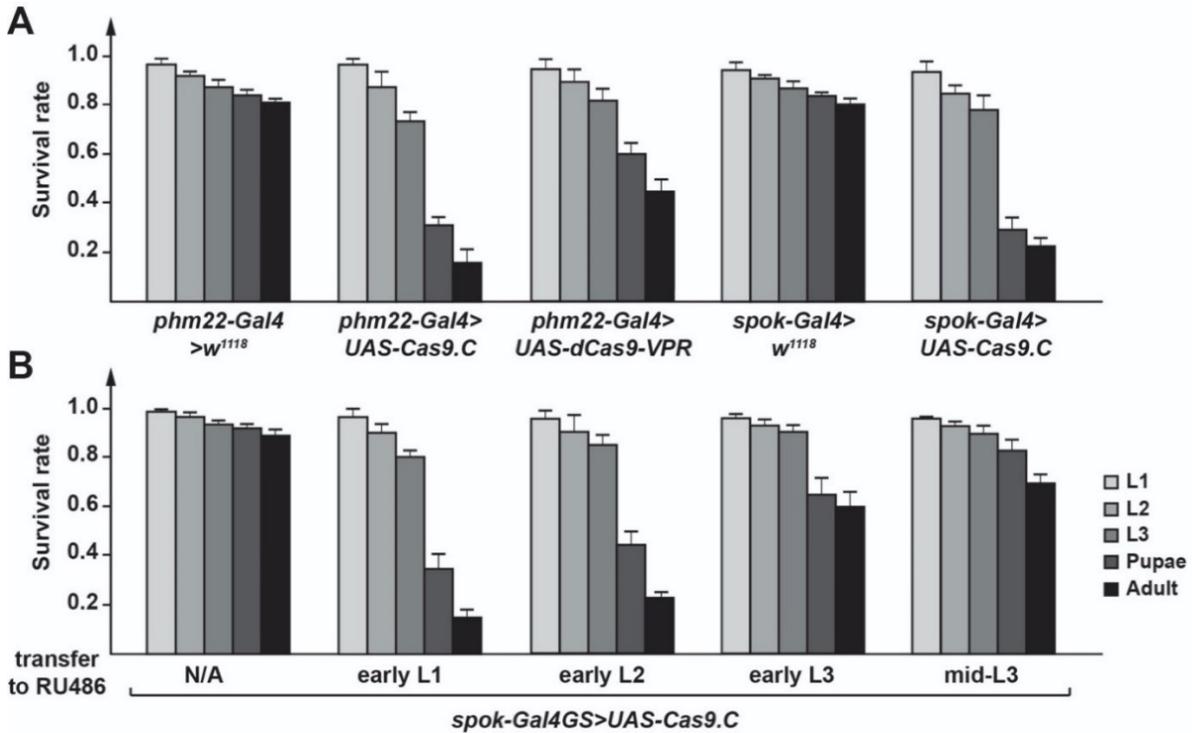
Figure 6.5. PG-specific Gal4-driven expression of Cas9 causes lethality.

A. The survival rates of flies harboring a single copy of *UAS-Cas9* or *UAS-Cas9-VPR* in combination with a single copy of a PG-specific Gal4 driver (*phm22>* or *spok>*). Error bars represent standard error. The Cas9 cDNA used here is the original allele that is not codon-optimized. UAS-dCas9-VPR is a transgene encoding nuclease-dead Cas9 (dCas9) that is fused to a chimeric co-activator domain (comprising VP64, p65 and Rta). **B.** The survival rates of flies harboring a single copy of *UAS-Cas9* in combination with a single copy of a PG-specific Gal4GS (Gal4-GeneSwitch) driver. GAL4GS was activated at different developmental time points by transferring larvae to RU486-supplemented media. Survival rates were quantified for each larval stage and represent surviving animals relative to the number of embryos used per condition (50 embryos for each replicate, three replicates in total). Error bars represent standard error.
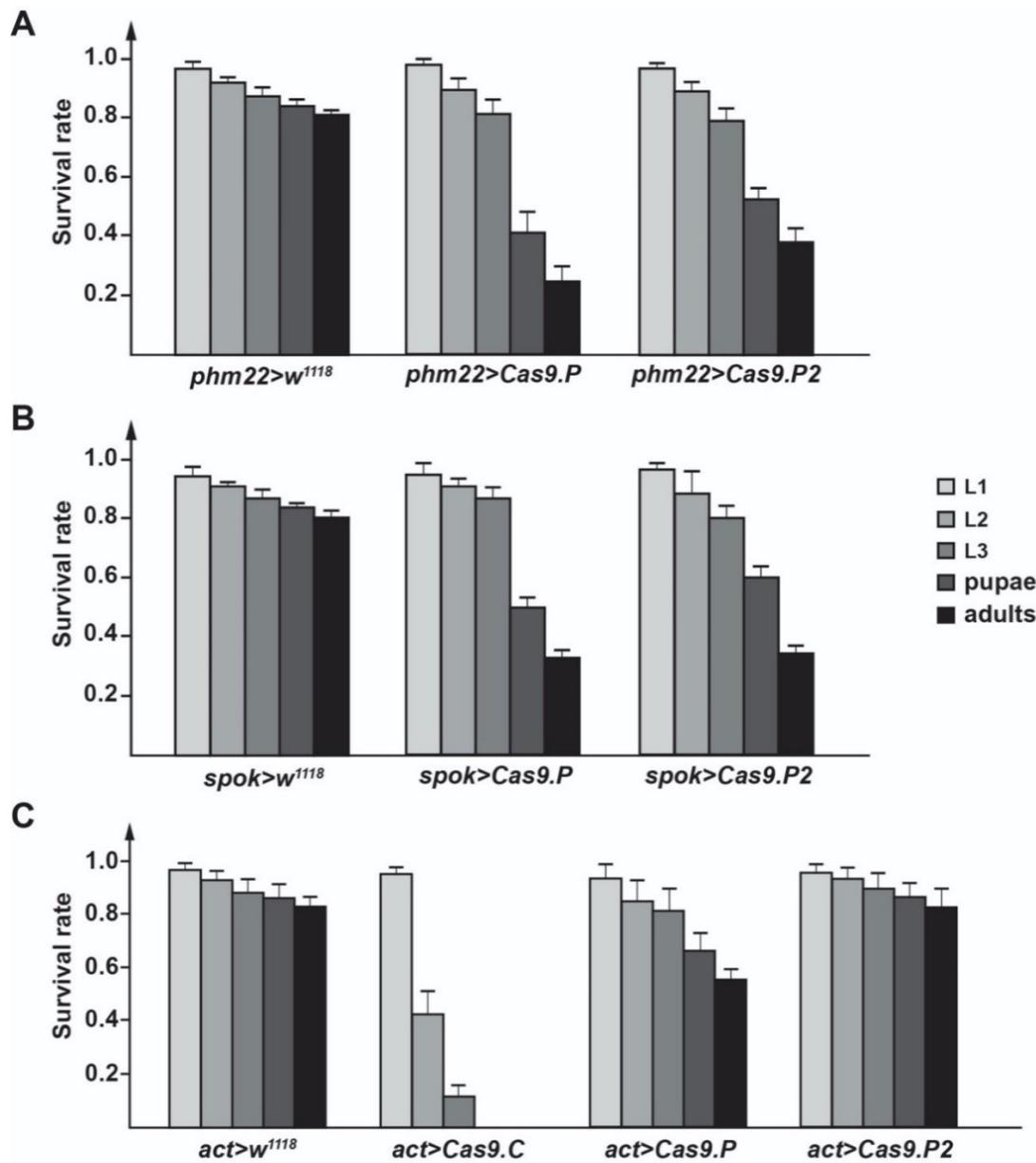
Figure 6.6. PG-specific and ubiquitous Gal4-driven expression of Cas9 causes lethality.

A. The survival rates of *phm22>Cas9.P* and *phm22>Cas9.P2* animals. Cas9.P is codon-optimized for *Drosophila* and Cas9.P2 is codon-optimized for human cells. **B.** The survival rates of *spok>Cas9.P* and *spok>Cas9.P2* animals. **C.** The survival rates of *act>Cas9.C*, *act> Cas9.P* and *act>Cas9.P2* animals. **A-C.** Data was normalized to the starting number of embryos and error bars

represent standard error. *phm22>* is short for *phm22-Gal4*; *spok>* represents *spok-Gal4*, and *act>* stands for *act-Gal4*.
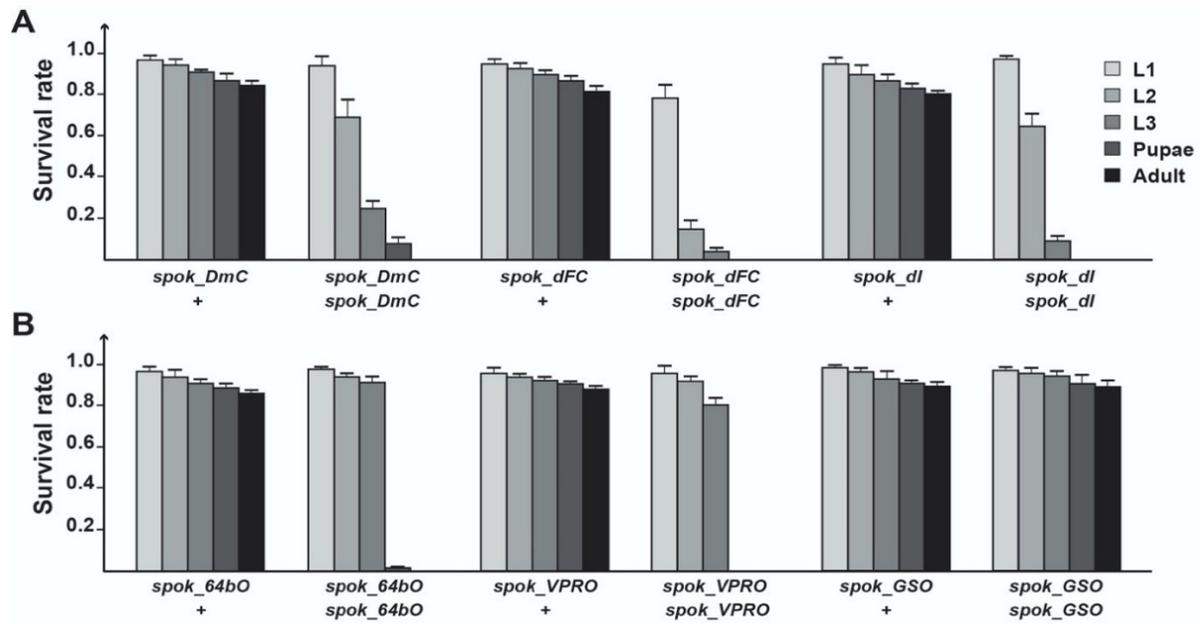
Figure 6.7. PG-specific expression of Cas9 without Gal4.

A. The survival rates of flies harboring a single copy or two copies of *spok_DmC*, *Spok_dFC* or *spok_dI*. Error bars represent standard error. **B.** The survival rates of flies harboring a single copy or two copies of *spok_64bO*, *spok_VPRO* or *spok_GSO*. Error bars represent standard error. For A and B data was normalized to the number of embryos in the starting population.
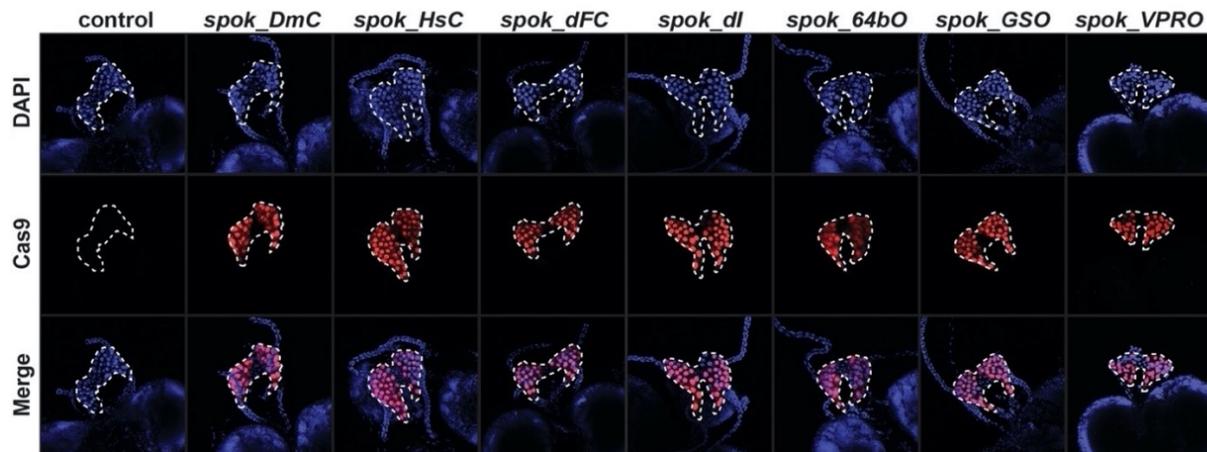
Figure 6.8. Nuclear localization of Cas9 in the *Drosophila* prothoracic gland (PG).

Nuclear localization sequences were added to the 5' and 3' ends of the *Cas9* cDNA to ensure transport of Cas9 into nuclei. The *spok* regulatory region drives the expression of Cas9 specifically in PG cells with no detectable signal in the adjacent corpora allata and the corpora cardiaca. DAPI (blue) was used to stain DNA while anti-Cas9 antibodies (red) was used to detect Cas9.

Figure 6.9. Efficiency of tissue-specific CRISPR/CAS9 in the *Drosophila* prothoracic gland (PG).

A. Comparing phenotypes of a classic *disembodied* mutant (*dib²*) and PG-specific RNAi (*dibIR*) with PG-specific (*spok_DmC*, *spok_HsC* and *spok_dFC*) or ubiquitous CRISPR/Cas9 (*act-*

Cas9.P2). **B.** Comparing phenotypes of a classic *phantom* mutant (*phm^{E7}*) and PG-specific RNAi (*phm^{IR}*) with PG-specific (*spok_DmC*, *spok_HsC* and *spok_dFC*) or ubiquitous CRISPR/Cas9 (*act-Cas9.P2*) in combination with *dU6-dib^{gR1}* or *dU6-phm^{gR2}*. **C and D.** Sequences of *dib* (C) and *phm* (D) loci from brain and PG nuclei, using either *spok_DmC* or *spok_HsC* in combination with *dU6-dib^{gR1}* or *dU6-phm^{gR1}*. **E.** Sequences of *dib* locus from brain and PG nuclei using *spok_dFC*.
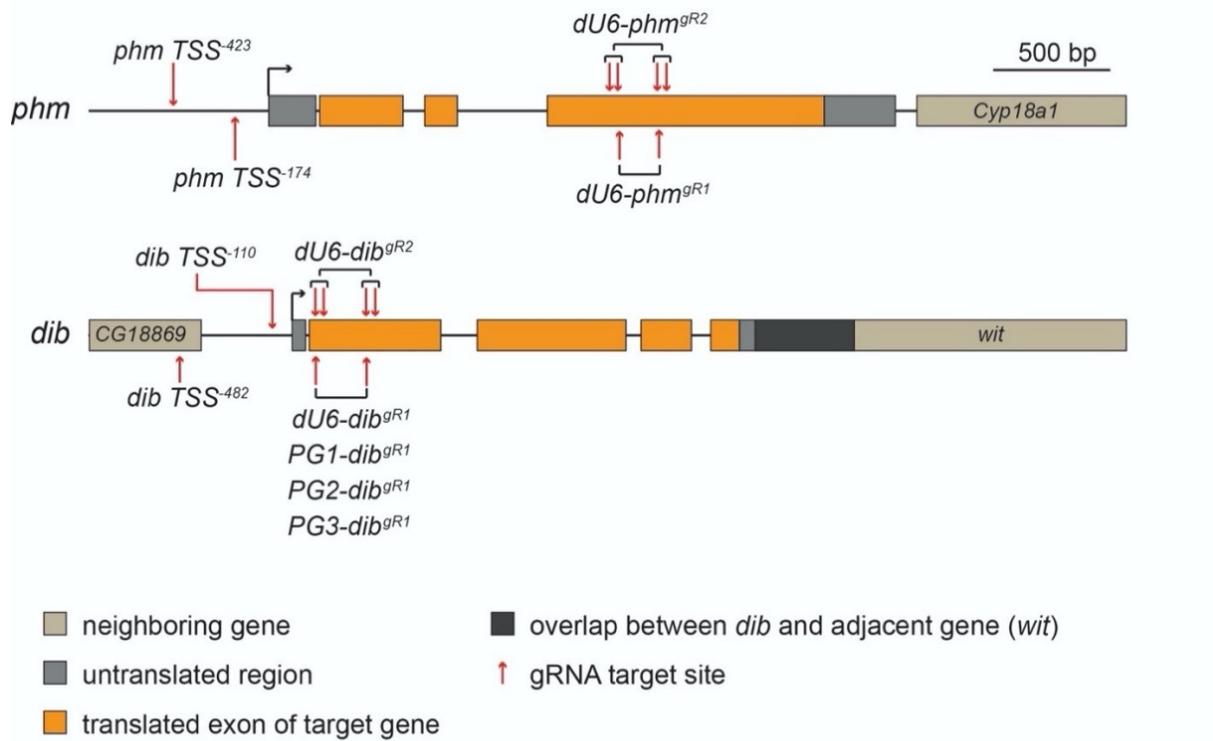
Figure 6.10. Target sites of gRNAs.

Genes that encode enzymes acting as ecdysteroid-synthesizing enzymes in the *Drosophila* prothoracic gland (PG), *phantom* (*phm*) and *disembodied* (*dib*). gRNAs targeting coding sequence (CDS, orange) were used for somatic disruption via CRISPR. gRNAs that target the upstream region of the of transcription start site (TSS, black arrow) were used either to activate (CRISPRa) or to interfere with (CRISPRi) target gene transcription.
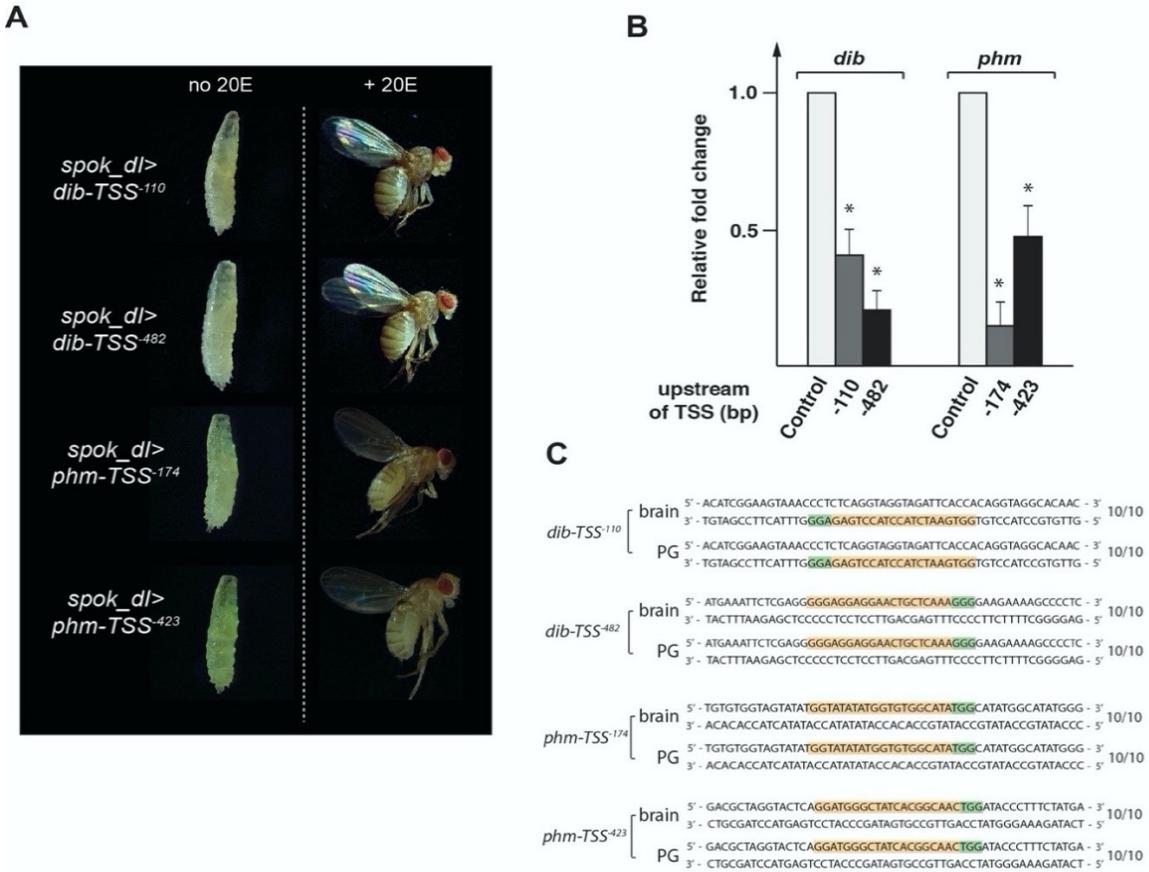
Figure 6.11 Efficiency of PG-specific CRISPR interference (CRISPRi) in *Drosophila*.

A. *spok_dI* (= *spok-dCas9*) is used to ubiquitously express gRNAs targeting -110 and -482 bp upstream of *dib* as well as -174 and -423 bp upstream of *phm* (relative to transcription start site = TSS), respectively. In each case, this resulted in L3 arrest (left) and rescue to adulthood when the diet was supplemented with 20E (right). For comparison to PG-specific *dib*- and *phm*-RNAi and classic mutant phenotypes, see Figures 6.9A, B. **B.** RG-specific qPCR for *dib*- and *phm*-CRISPRi. Ring glands were dissected at 42 hrs after the L2/L3 molt, three replicates per condition. * => p-value < 0.05. **C.** Sequences of *dib* and *phm* loci obtained from DNA of CRISPRi-treated PG nuclei show no alterations. For each condition, we sequenced 10 clones, all of which were wild type.
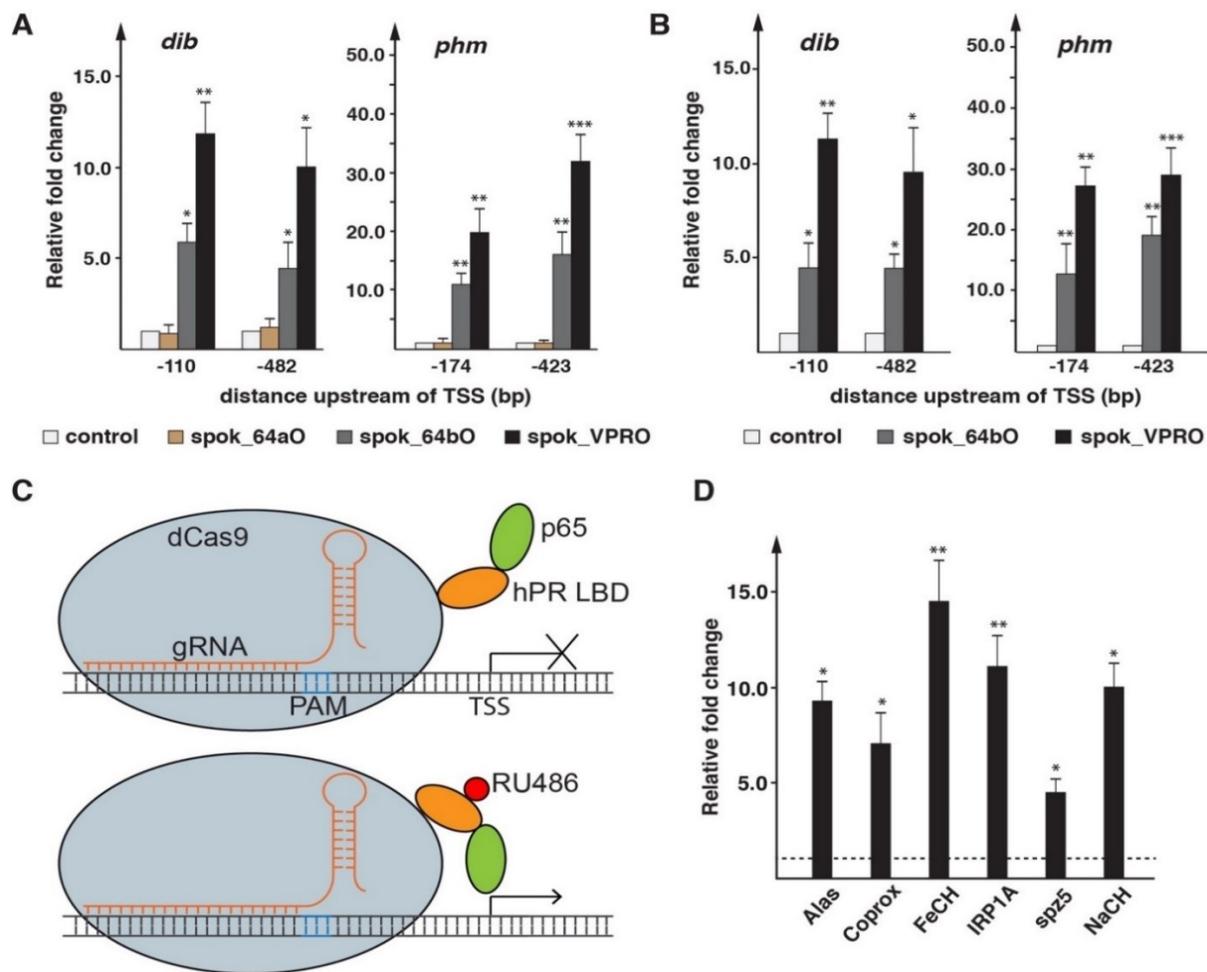
Figure 6.12 Efficiency of PG-specific CRISPRa.

A. qPCR of *ex vivo*-cultured ring glands transfected with *spok_64aO*, *spok_64bO* and *spok_VPRO* plasmids. Transfected glands were isolated from transgenic larvae expressing gRNAs targeting -110 and -482 bp upstream of *dib* as well as -174 and -423 bp upstream of *phm* (relative to TSS), respectively. All results normalized to controls (= no plasmid added). spok_64aO differs from spok_64bO, as it encodes different amino acids at positions 839 and 863, which are important for attenuation of the endonuclease, as previously reported [90]. *spok_VPRO* has the same changes as *spok_64bO*, but in addition harbors p65 and Rta domains (Figure 6.3). **B.** Same as A, however ring

glands were transgenic for both gRNA and Cas9 constructs. *spok_64aO* was not used to make transgenics, due to the lack of activity shown in A. **C.** Schematic of dCas9 fused to the human progesterone ligand-binding domain (hPR LBD) and the p65 activation domain, resulting in dCas9GSO (= dead Cas9-GeneSwitch for activation). This approach allows for temporal control over the activation via dCas9, by switching animals to a diet supplemented with RU486. **D.** qPCR analysis of six target genes. Lines were obtained from Bloomington stock centre. Shown are the fold-changes relative to the same gene in samples of the same genotype, but raised on RU-486-free medium (dotted line = 1). Ring glands were dissected from larvae that were reared for four hours on media supplemented with RU486. * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. Error bars represent standard error.
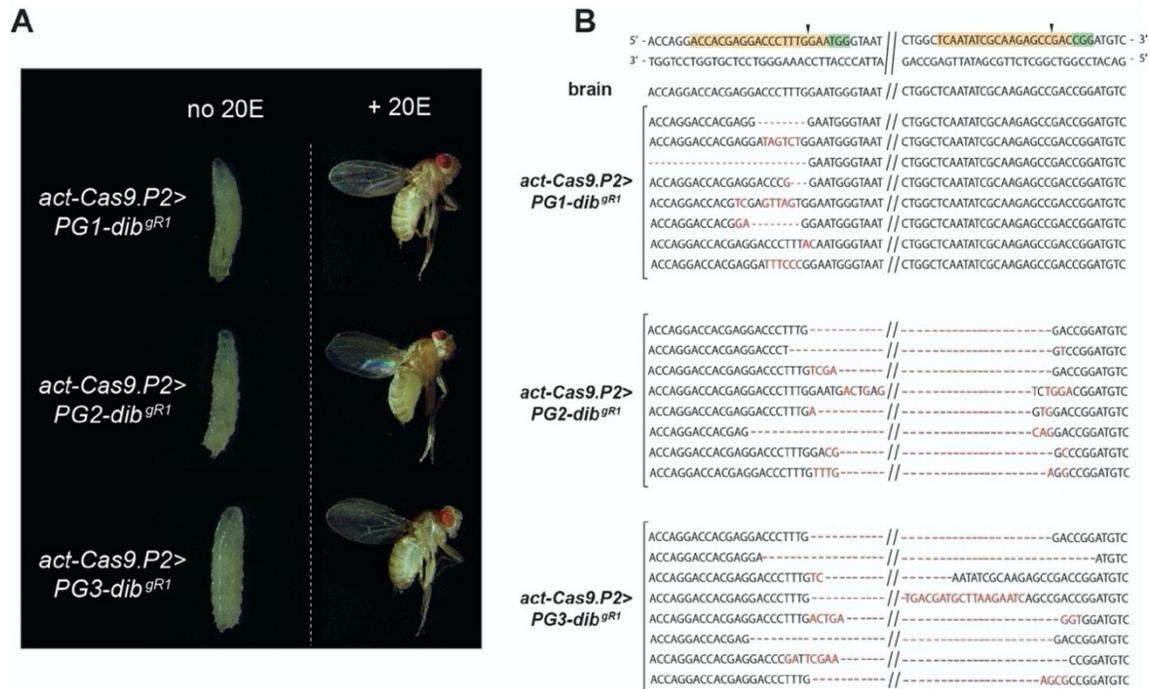
Figure 6.13 PG-specific gRNA in combination with ubiquitous Cas9 expression.

A. Phenotypes associated with PG-specific gRNA- and ubiquitous Cas9-expression targeting the *dib* gene, in the absence (left) or presence (right) of dietary 20OH-ecdysone (20E). **B.** Sequences of *dib* locus resulting from using the same *dib* gRNA pair (gR1), but different PG-gRNA vectors (PG1-3). Red letters and dashes indicate altered or missing nucleotides.

Figure 6.14 The general Gateway Cas9 (gG-Cas9) vector collection II.

Each gG-Cas9 vector backbone is composed of a *mini-white* gene as a marker, a PhiC31 integrase-compatible *attB* site, and the *bla* coding sequence to mediate ampicillin resistance. Shown here are the gateway cassette for an enhancer of interest, the Cas9 variants, the regions encoding Nuclear Localization Signals (NLS), Nuclear Export Signals (NES), activation domains (VPR = VP64, p65 and Rta), the human Progesterone Receptor ligand-binding domain (hPR LBD), the FokI nuclease domain, the P2A self-cleaving peptide, the FK506 binding protein (FKBP) rapamycin binding domain, FKBP-rapamycin binding domain of mTOR (FRB). The gateway cassette allows using LR Clonase-based recombination (ThermoFisher) to insert enhancer/promoter regions to drive tissue-specific Cas9 expression. en_GSD (GSD = GeneSwitch-dependent Disruption) and RDX (RDX = Rapamycin-dependent Disruption) can be used to generate temporal somatic mutations. GSD encodes a protein where nuclease-dead Cas9 (=dCas9) is fused to the hPR LBD and FokI nuclease domain, separated by a linker. Upon binding to RU486, the hPR LBD will change its conformation and allow the dimerization of FokI nuclease domains for DNA cutting. RDX encodes two proteins where the C-terminal Cas9 fragment is fused with FKBP and the other fragment contains N-terminal Cas9 fragment fused with FRB rapamycin binding domain. Upon

241

binding to rapamycin, the FKBP and FRB domains will dimerize and allow the full assemble of Cas9 which triggers cleavage of target DNA. RDO (RDO = Rapamycin-dependent Overexpression) is a similar approach as RDX, however, the dCas9 carries the D10A and N863A mutation, which abolishes the nuclease activity and the C-terminus fragment of dCas9 is also fused with the activation domains (VPR). RDO is activated by rapamycin and can trigger gene overexpression using gRNA target the upstream region of transcription start site (TSS).
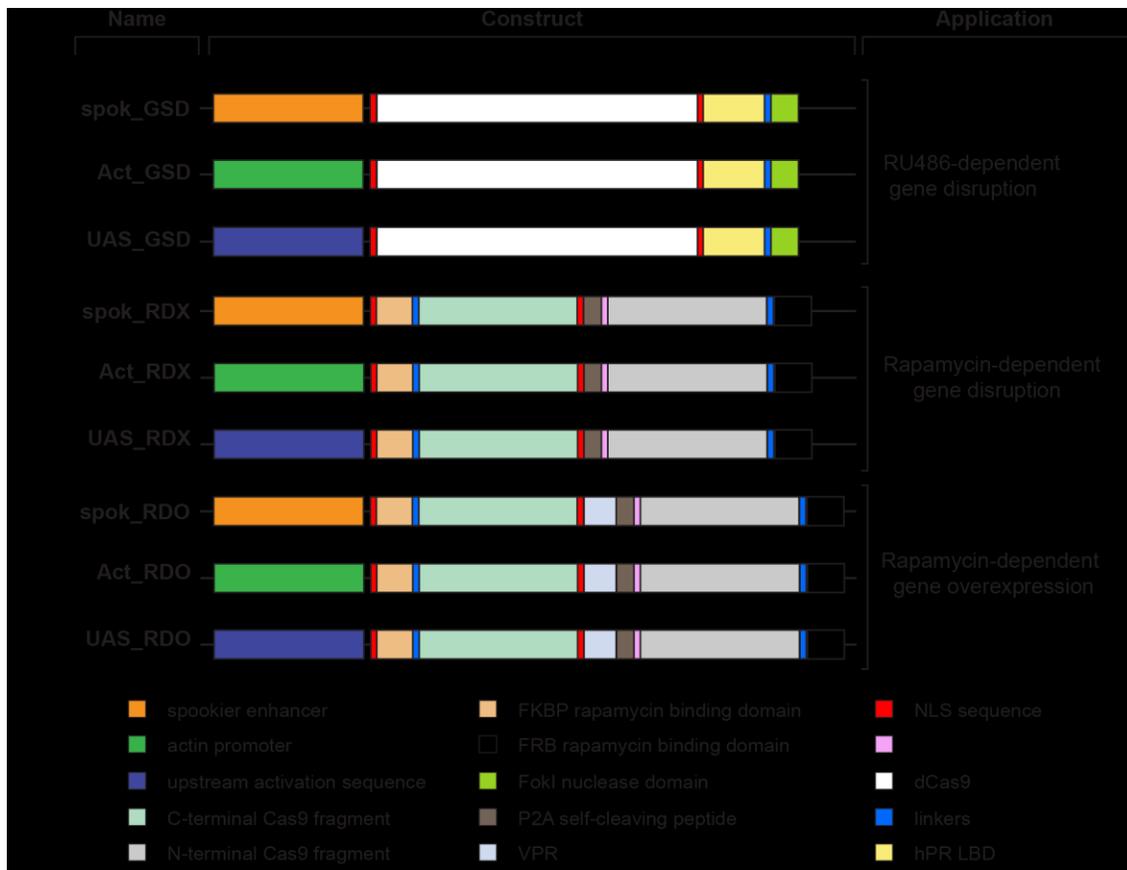
Figure 6.15 Overview of GeneSwitch-dependent gene Disruption (GSD).

In this approach, the dead Cas9 (dCas9) is fused with <u>h</u>uman <u>P</u>rogesterone <u>R</u>eceptor <u>l</u>igand-<u>b</u>inding <u>d</u>omain (hPR LBD) and FokI nuclease domain, resulting in dCas9 GSD (= dead Cas9-GeneSwitch dependent gene Disruption). This approach allows for temporal control over gene disruption via dCas9, by exposing animals to exogneous RU486.

Figure 6.16 *In vitro* evaluation of Cas9_hPR LBD efficiency in genome editing.

A. Cas9 domain structure. The arrows indicate the insertion sites of human Progesterone Receptor ligand-binding domain (hPR LBD) into Cas9 for RU486-dependent gene editing. Among these sites, 5, 6, 7, 8, 9, 11 are inserted in Cas9 loops while 1, 2, 3, 4, 10 are inserts in uncharacterized regions. **B.** Schematic of co-transfection construct of Cas9_hPR LBD and multiplex tRNA-gRNA expressing previously validated gRNA target *disembodied* (*dib*) gene. **C.** Genome editing

efficiency of Cas9_hPR_LBD in the presence of RU486 using surveyor nuclease assay, which allows detection of small deletions or insertions.
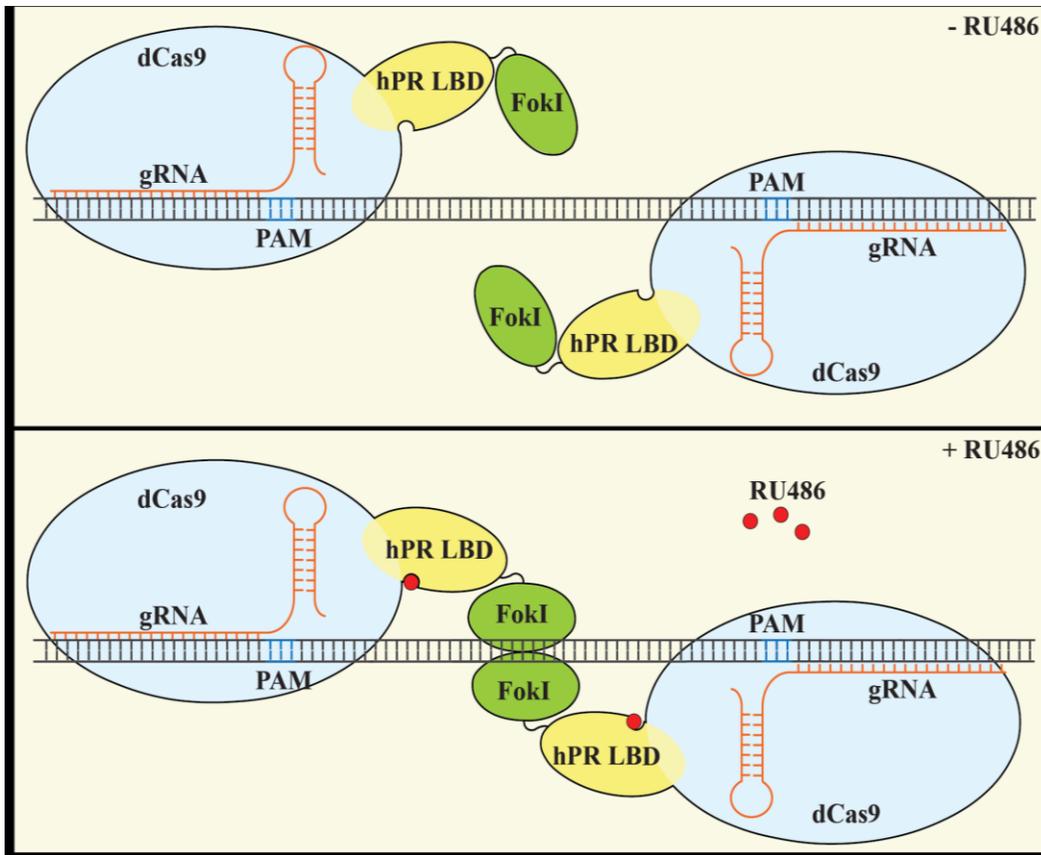
Figure 6.17 Overview of GeneSwitch-dependent gene Disruption (GSD).

In this approach, the dead Cas9 (dCas9) is fused with human Progesterone Receptor ligand-binding domain (hPR LBD) and FokI nuclease domain, resulting in dCas9 GSD (= dead Cas9-GeneSwitch dependent gene Disruption). This approach allows for temporal control over gene disruption via dCas9, by exposing animals to exogenous RU486.

Figure 6.18 *In vitro* evaluation of dCas9_GSD efficiency in genome editing.

A. Schematic of a co-transfection construct that carries both Cas9_hPR LBD and multiplex tRNA-gRNA expressing previously validated gRNAs target *disembodied* (*dib*) gene. **B.** Evaluation of Cas9_hPR_LBD cleavage efficiency using surveyor nuclease assay, which detects mismatches in DNA, such as small deletions or insertions. Indel % was calculated following manufacturer's directions.
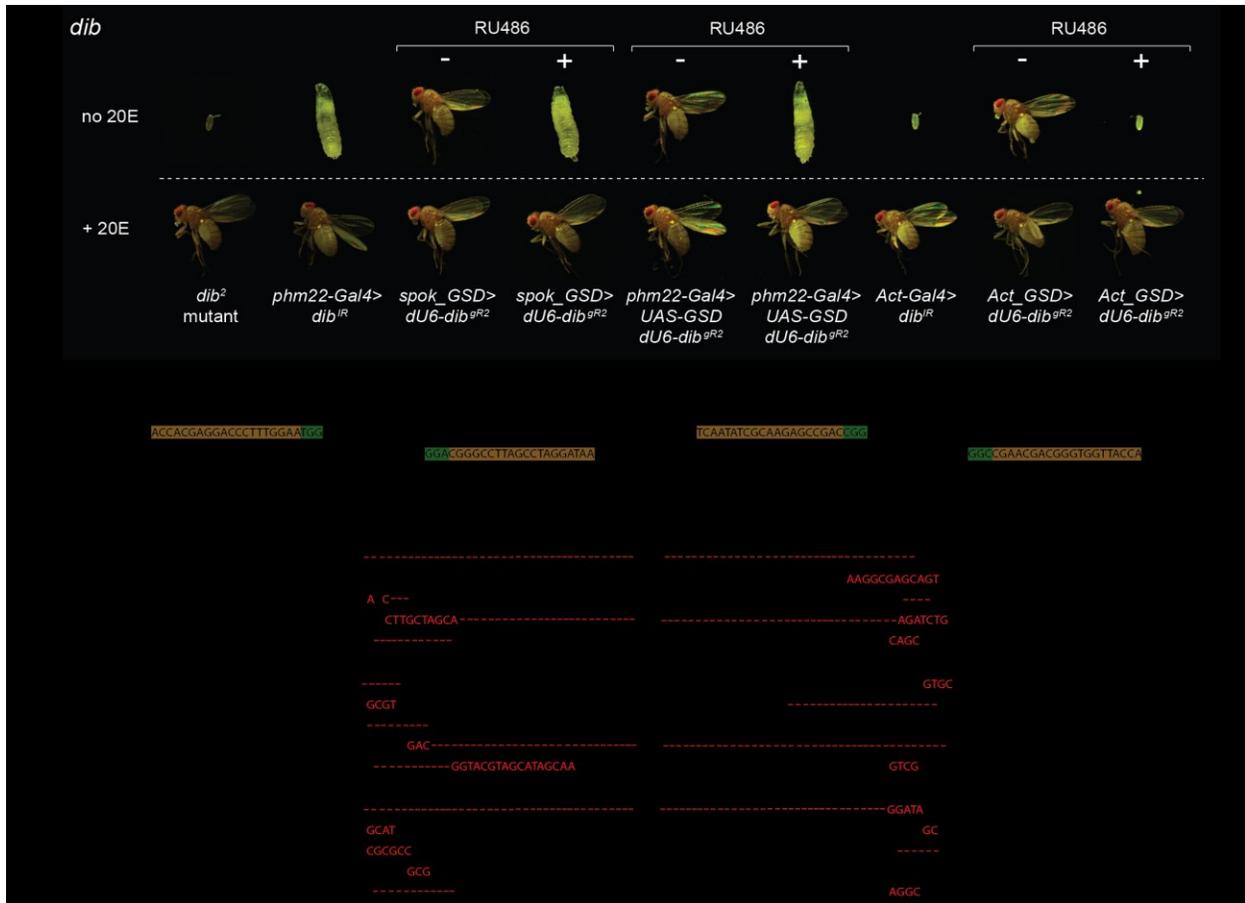
Figure 6.19 Efficiency of temporal RU486-dependent gene editing CRISPR/CAS9 in the Drosophila prothoracic gland (PG).

A. Comparison of phenotypes from a classic *disembodied* mutant (*dib²*), PG-specific *dib*-RNAi (*phm22-Gal4>dib^IR*), ubiquitous *dib*-RNAi (*act-Gal4>dib^IR*) with PG-specific (*spok_GSD, phm22-Gal4>UAS-GSD*) or ubiquitous CRISPR/Cas9 (*act-GSD*) in the presence or absence of RU486. **B.** Sequences of the *dib* locus from brain and PG DNA (for PG-specific GSD in combination with *dU6-dib^gR2*) or embryonic DNA (for ubiquitous GSD in combination with *dU6-dib^gR2*). Left: presence or absence of RU486 is indicated by "+" and "-", respectively. Ratios on the right indicate the frequency of mutated sequences.
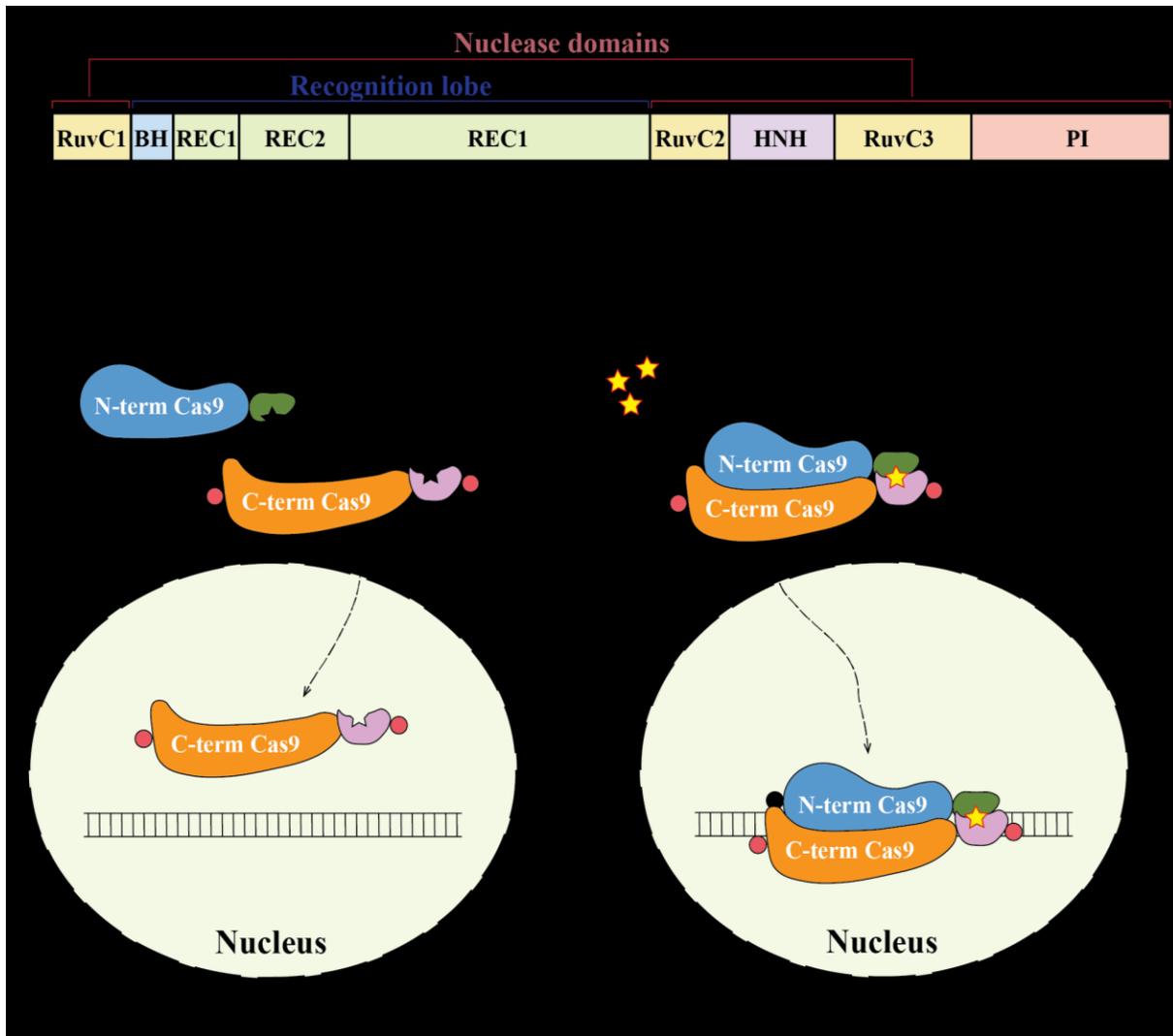
Figure 6.20 Schematic of <u>R</u>apamycin-dependent gene <u>D</u>isruption (RDX).

In this approach, the human Cas9 is split into two halves in the middle of the recognition lobe (REC) (A). The C-terminal Cas9 fragment is fused with the FK506 binding protein 12 (FKBP) and two Nuclear Localization Signals (NLS), while the N-terminal Cas9 fragment is fused with the FKBP rapamycin binding domain (FRB) and a Nuclear Export Signal (NES). In the absence of rapamycin, two Cas9 fragments remain disassociated. However, in the presence of rapamycin, dimerization of FKBP and FRB will promote the association of the two Cas9 polypeptides into a functional complex.

Figure 6.21 Efficiency of rapamycin-dependent gene editing CRISPR/CAS9 in the *Drosophila* prothoracic gland (PG).

A. Comparison of phenotypes from a classic *disembodied* mutant (*dib²*), PG-specific *dib*-RNAi (*phm22-Gal4>dib^{IR}*) and ubiquitous *dib*-RNAi (*act-Gal4>dib^{IR}*) with PG-specific (*spok_RDX, phm22-Gal4>dib^{gR1}*) or ubiquitous CRISPR/Cas9 (*act-RDX>dib^{gR1}*). **B.** Sequences of the *dib* locus from brain and PG DNA, using either *spok_RDX*, or *phm22>UAS-RDX* (for PG-specific RDX in combination with *dU6-dib^{gR1}*) or embryo DNA (for ubiquitous RDX) in combination with *dU6-dib^{gR1}*) in the presence or absence of rapamycin in diet. **C.** Sequences of the *dib* locus from brain and PG DNA using *spok_RDX* in combination with *dU6-db^{gR1}* upon feeding on a rapamycin-diet. Time after treatment indicates how long larvae were reared on rapamycin-supplemented media. Ratios (far right) indicate the frequencies of observed mutations.

Figure 6.22 Schematic of rapamycin-dependent gene Overexpression (RDO).

A. In this approach, the nuclease-dead Cas9 is split into two halves in the middle of the recognition lobe. The C-terminal Cas9 fragment is fused with the FK506 binding protein 12 (FKBP), VPR activation region (VP64, p65, Rta) and two Nuclear Localization Signals (NLS), while the N-terminal Cas9 fragment is fused with the FKBP rapamycin binding domain (FRB) and a Nuclear Export Signal (NES). In the absence of rapamycin, two Cas9 fragments remain separated from each other. However, in the presence of rapamycin, dimerization of FKBP and FRB will promote the association of the two Cas9 polypeptides into a functional complex. **B.** qPCR analysis of three target genes. Stocks were obtained from Bloomington stock center. Shown are fold changes relative to the same gene in samples of the same genotype that were reared on rapamycin-free medium (dotted line = 1). Ring glands were dissected from larvae that were reared for four hours on media supplemented with rapamycin. * = p-value < 0.05, ** = p-value < 0.01, *** = p - value < 0.001. Error bars represent standard error.

251
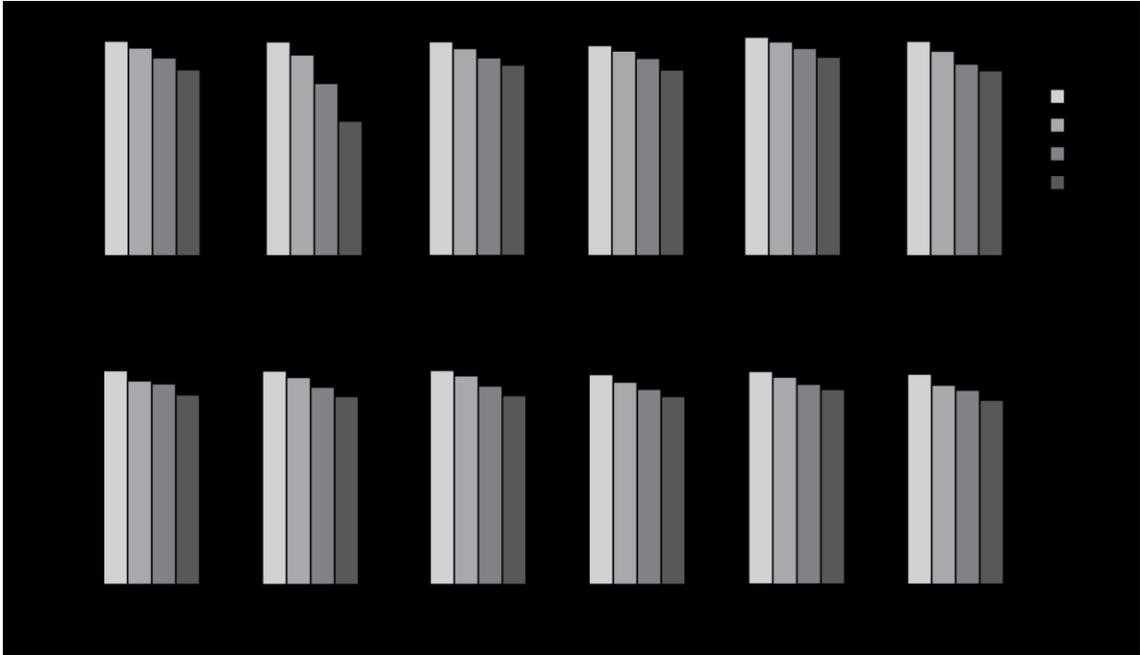
Figure 6.23 Survival rates of PG-specific and ubiquitously expressed drug-inducible Cas9 transgenes.

The survival rates of flies harbouring a single copy or two copies of spok_GSD, act_GSD, spok_RDX, act_RDX, spok_RDO or act_RDO. Error bars represent standard deviation. Data was normalized to the number of embryos in the starting population.

6.6 Tables

Table 6.1 Primers used for transgenic lines and vectors

primer name	sequence (5' – 3')
gG-Cas9 vector collection	
gG-Cas9 backbone	
SV40 pBPGUw RP	CACCTTTCTCTTCTTCTTGGGCTTTCAGGAGGCTTGCTTCAAG
pBPGUw BB FP	CTGCTTACCCACCCAAAACCAATC
pBPGUw BB RP	GATTGGTTTTGGGTGGGTAAGCAG
PCas9 BB FP	AATGAATCGTAGATACTGAAAAACCCCGCAAG
gG-en_DmC	
SV DmCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTAC
DmCas9 RP	TCAGTATCTACGATTCATTTTATCACACCTTCCTCTTCTTG
gG- en_HsC	
SV HsCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCATGGACAAGAAGTAC
HsCas9 RP	TCAGTATCTACGATTCATTTCACACCTTCCTCTTCTTCTTG
gG- en_dFC	
SV FokI FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGC CAACTTGTGAAGTCTGAAC
dCas9 RP	TCAGTATCTACGATTCATTTTATCACACCTTCCTCTTCTTG
gG- en_dI	
SV dCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTAC
dCas9 RP	TCAGTATCTACGATTCATTTTATCACACCTTCCTCTTCTTG
gG- en_64bO	
SV dCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTAC
VP64 RP	TCAGTATCTACGATTCATTTTACAGCATGTCCAGGTC
H839A Mut FP	ACGATGTGGCTGCTATCGTGCCTCAG
H839A Mut RP	AGTCGGACAGCCGGTTGATGTC
N863A Mut FP	AGCGACAAGGCCCGGGGCAAGAGC
N863A Mut RP	TCTGGTCAGCACCTTGTTGTC
gG- en_VPRO	
SV dhsCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTACTCCATTG
VPR RP	TCAGTATCTACGATTCATTTCAAAACAGAGATGTGTCGAAGATG

primer name	sequence (5' – 3')
gG- en_GSO	
SV40 dCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTACAGCATC
SV40 dCas9 RP	GAGGAAGCGGAGGAGGAGGTAGCGGACCTAAGAAAAAGAGGAAGGTG
SV40 hPR FP	CCTAAGAAAAAGAGGAAGGTGGCGGCCGCTGGATCCGGACGGGCTAAAA AGTTCAATAAAGTCAGAG
XbaI p65 RP	TCAGTATCTACGATTCATTCTAGTAGGAGCTGATCTGAC
gG-Cas9 backbone	
SV40 pBPGUw RP	CACCTTTCTCTTCTTCTTGGGCTTTCAGGAGGCTTGCTTCAAG
pBPGUw BB FP	CTGCTTACCCACCCAAAACCAATC
pBPGUw BB RP	GATTGGTTTTGGGTGGGTAAGCAG
PCas9 BB FP	AATGAATCGTAGATACTGAAAAACCCCGCAAG
gG-en_GSD	
SV dCas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGACAAGAAGTAC
hPR dCas9 RP	TCTGACTTTATTGAACTTTTTAGCCCGTCCGGATCCAG
dCas9 hPR FP	CTGGATCCGGACGGGCTAAAAAGTTCAATAAAGTCAGAG
FokI hPR RP	GAGTTCAGACTTCACAAGTTGCGAGCCACCGCCACCCGAGCCAC
hPR FokI FP	GTGGCTCGGGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
BB FokI RP	TCAGTATCTACGATTCATTCACACCTTCCTCTTCTTCTTGGGGTCAG
gG- en_RD	
SV FKBP FP	CCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAG
NES P2A RP	GTCTCTCAAGCGGTGGTAGGAGGATTAAGCTAGCTAAATC
P2A NES FP	GATTTAGCTAGCTTAATCCTCCTACCACCGCTTGAGAGAC
BB FRB RP	TCAGTATCTACGATTCATTTACTGCTTGCTGATTCTTC
gG- en_RO	
SV FKBP FP	CCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAG
NLS C.dCas9 RP	CACCTTTCTCTTCTTCTTGGGGCTGTCGCCTCCCAGCTGAGAC
NLS VPR FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGGACGGGC
P2A VPR RP	GCAGAGAGAAGTTTGTTGCGCCGGATCCAAACAGAGATG
VPR P2A FP	CATCTCTGTTTGGATCCGGCGCAACAAACTTCTCTCTGC
BB FRB RP	TCAGTATCTACGATTCATTTACTGCTTGCTGATTCTTC
S2 cell transfection construct	
Generation of dual Cas9-hPR LBD gRNA	

primer name	sequence (5' – 3')
SV BB RP	CACCTTTCTCTTCTTCTTGGGGGCCATGGTGGCACCGGTCGTCTCTG
SV Cas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAG
BB SVC RP	GTGGCACTTTTCGGGGAAATGTGTTACACCTTCCTCTTTTTCTTAG
SVC BB FP	CTAAGAAAAAGAGGAAGGTGTAACACATTTCCCCGAAAAGTGCCAC
BB dU6 FP	GATACTTCTAAAAAAATTTTTTTGCTCACCTGTGATTG
dU6 BB RP	CAATCACAGGTGAGCAAAAAAATTTTTTTTAGAAGTATC
BB gRNA RP	CATGTCTGGATCCCTCGAGGCACCGACTCGGTGCCAC
gRNA BB FP	GTGGCACCGAGTCGGTGCCTCGAGGGATCCAGACATG
hPR LBD FP	AGTTCAATAAAGTCAGAGTTGTGAGAG
hPR LBD RP	CTTGGGGGCAGGTGGGGCCAC
hPR Cas9 RP1	CAACTCTGACTTTATTGAACTTTTTGGCGTTGATGGGGTTTTCCTC
hPR Cas9 FP1	TGGCCCCACCTGCCCCCAAGAGCGGCGTGGACGCCAAGGCCATC
hPR Cas9 RP2	CTGACTTTATTGAACTTTTTGTTGCTCTTGAAGTTGGGGGTC
hPR Cas9 FP2	TGGCCCCACCTGCCCCCAAGTTCGACCTGGCCGAGGATG
hPR Cas9 RP3	CTGACTTTATTGAACTTTTTGGTGTTCACTCTCAGGATG
hPR Cas9 FP3	TGGCCCCACCTGCCCCCAAGGAGATCACCAAGGCCCCCCTGAG
hPR Cas9 RP4	CTGACTTTATTGAACTTTTTCATTCCCTCGGTCACGTATTTC
hPR Cas9 FP4	GTGGCCCCACCTGCCCCCAAGAGAAAGCCCGCCTTCCTG
hPR Cas9 RP5	CTGACTTTATTGAACTTTTTCTCGATTTTCTTGAAGTAGTC
hPR Cas9 FP5	GTGGCCCCACCTGCCCCCAAGTGCTTCGACTCCGTGGAAATC
hPR Cas9 RP6	CTGACTTTATTGAACTTTTTGGACACCTGGGCTTTCTGGATG
hPR Cas9 FP6	GTGGCCCCACCTGCCCCCAAGGGCCAGGGCGATAGCCTGCAC
hPR Cas9 RP7	CTGACTTTATTGAACTTTTTCAGCTTAGGGTACTTTTTGATC
hPR Cas9 FP7	GTGGCCCCACCTGCCCCCAAGGAAAGCGAGTTCGTGTACGGCGAC
hPR Cas9 RP8	CTGACTTTATTGAACTTTTTGCCGTTGGCCAGGGTAATC
hPR Cas9 FP8	GTGGCCCCACCTGCCCCCAAGGAGATCCGGAAGCGGCCTCTG
hPR Cas9 RP9	CTGACTTTATTGAACTTTTTGTTCCTCTTGGGCAGGATAG
hPR Cas9 FP9	GTGGCCCCACCTGCCCCCAAGAGCGATAAGCTGATCGCCAG
hPR Cas9 RP10	CTGACTTTATTGAACTTTTTCTTGCCCTTTTCCACTTTGGCCAC
hPR Cas9 FP10	GTGGCCCCACCTGCCCCCAAGTCCAAGAAACTGAAGAGTGTG
hPR Cas9 RP11	CTGACTTTATTGAACTTTTTCTTCAGCTTCTCATAGTGGCTG
hPR Cas9 FP11	GTGGCCCCACCTGCCCCCAAGGGCTCCCCCGAGGATAATGAG

primer name	sequence (5' – 3')
Generation of dual dCas9-hPR LBD-FokI gRNA	
SV BB RP	CACCTTTCTCTTCTTCTTGGGGGCCATGGTGGCACCGGTCGTCTCTG
SV Cas9 FP	CCCAAGAAGAAGAGAAAGGTGGAGGCCAG
Cas9 Middle RP	GATGTCCAGTTCCTGGTCCAC
Cas9 Middle FP	GTGGACCAGGAACTGGACATC
Linker1 GSD RP	CGAGCCACCGCCACCCTTGGGGGCAGGTGGGGCCAC
Linker1 GSD FP	GGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
Linker2 GSD RP	CGAGCCACCGCCACCCGAGCCACCGCCACCCTTGGGGGCAGGTGGGGCCAC
Linker2 GSD FP	GGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
Linker3 GSD RP	CGAGCCACCGCCACCCGAGCCACCGCCACCCGAGCCACCGCCACCCTTGGGGGCAGGTGGGGCCAC
Linker3 GSD FP	GGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
Linker4 GSD RP	CGAGCCACCGCCACCCGAGCCACCGCCACCCGAGCCACCGCCACCCTTGGGGGCAGGTGGGGCCAC
Linker4 GSD FP	GGTGGCGGTGGCTCGGGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
Linker5 GSD RP	CGAGCCACCGCCACCCGAGCCACCGCCACCCGAGCCACCGCCACCCTTGGGGGCAGGTGGGGCCAC
Linker5 GSD FP	GGTGGCGGTGGCTCGGGTGGCGGTGGCTCGGGTGGCGGTGGCTCGCAACTTGTGAAGTCTGAACTC
FokI RP	CACTTTTCGGGGAAATGTGTCACACCTTCCTCTTCTTCTTGGGGTC
SVC BB FP	CTAAGAAAAGAGGAAGGTGTAACACATTTCCCCGAAAAGTGCCAC
BB dU6 FP	GATACTTCTAAAAAAATTTTTTTGCTCACCTGTGATTG
dU6 BB RP	CAATCACAGGTGAGCAAAAAAATTTTTTTTAGAAGTATC
BB gRNA RP	CATGTCTGGATCCCTCGAGGCACCGACTCGGTGCCAC
gRNA BB FP	GTGGCACCGAGTCGGTGCCTCGAGGGATCCAGACATG
Generation of dib gRNA for cell transfection plasmid	
dib gR1 5KO sFP	ACCACGAGGACCCTTTGGAA
dib gR1 5KO sRP	TTCCAAAGGGTCCTCGTGGT
dib gR1 5KO dFP	GCGGCCCGGGTTCGATTCCCGGCCGATGCACCACGAGGACCCTTTGGAAGTTTTAGAGCTAGAAATAGCAAG

primer name	sequence (5' – 3')
dib gR1 5KO dRP	ATTTTAACTTGCTATTTCTAGCTCTAAAACGTCGGCTCTTGCGATATTGATGCACCAGCCGGGAATCGAACCC
dib screening FP	CTATGTATACGATGATTCAC
dib screening RP	ACCAATCATCCTAGCGTGTAG
PG-Cas9 vector collection	
attB1 Spok FP	CCAAGTTTGTACAAAAAAGCAGGCTCATTTCGGTGGAAGGTCCTG
attB2 Spok RP	CCCACTTTGTACAAGAAAGCTGGGTTTTCAGCCTTAGTAAATAG
attB2 BB FP	ACCCAGCTTTCTTGTACAAAGTGGGATAAACGGCCGGCCGAGCTCGCCCGGGGATC
Miniwhite RP	GACGCAAGGAGTAGCCGACATATATC
Miniwhite FP	GATATATGTCGGCTACTCCTTGCGTC
attB1 BB RP	AGCCTGCTTTTTTGTACAAACTTGGATAAACGAATTCTTGAAGACGAAAGGGCCTC
PG1-gRNA	
pCFD5 Spok FP	GATCAATTGAGATCTGAATTCCATTTCGGTGGAAGGTCCTG
tRNA Spok RP	GTCTACACACACTCAAAGCCCTTTCAGCCTTAGTAAATAG
Spok tRNA FP	CTATTTACTAAGGCTGAAAGGGCTTTGAGTGTGTGTAGAC
pCFD5 Middle RP	CTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTAC
pCFD5 Middle FP	GTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAG
Spok pCFD5 RP	CAGGACCTTCCACCGAAATGGAATTCAGATCTCAATTGATC
PG2-gRNA	
pCFD5 Spok FP	GATCAATTGAGATCTGAATTCCATTTCGGTGGAAGGTCCTG
HDV gRNA RP	GTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGCCGCACCGACTCGGTGCCAC
gRNA HDV FP	GCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTTTTTGCCTACCTGGAGCCTG
pCFD5 Middle RP	CTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTAC
pCFD5 Middle FP	GTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAG
Spok pCFD5 RP	CAGGACCTTCCACCGAAATGGAATTCAGATCTCAATTGATC
PG3-gRNA	
pCFD5 Spok FP	GATCAATTGAGATCTGAATTCCATTTCGGTGGAAGGTCCTG

primer name	sequence (5' – 3')
HH Spok RP	CGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGGGGCTTTTTCAGCCTTAGTAAATAG
HH tRNA FP	TGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCGGGCTTTGAGTGTGTGTAGACATCAAG
HDV gRNA RP	GTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGCCGCACCGACTCGGTGCCAC
gRNA HDV FP	GCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTTTTTGCCTACCTGGAGCCTG
pCFD5 Middle RP	CTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTAC
pCFD5 Middle FP	GTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAG
Spok pCFD5 RP	CAGGACCTTCCACCGAAATGGAATTCAGATCTCAATTGATC
Generation of gRNAs targeting *phantom* (*phm*) and *disembodied* (*dib*) genes	
phm gR1 FP	CGGCCCGGGTTCGATTCCCGGCCGATGCACGGCGCCTGGTAGGCTCCATGTTTTAGAGCTAGAAATAGCAAG
phm gR1 gRNA RP	ATTTTAACTTGCTATTTCTAGCTCTAAAACGTAAAGAGCGTGAGTATCATTGCACCAGCCGGGAATCGAACCC
phm TSS -174 FP	TGCA GGTATATATGGTGTGGCATA
phm TSS -174 RP	AAAC TATGCCACACCATATATACC
phm TSS -423 FP	TGCA GGATGGGCTATCACGGCAAC
phm TSS -423 RP	AAAC GTTGCCGTGATAGCCCATCC
phm gR2 FPa	TTCGATTCCCGGCCGATGCGGAACCGGAGGAGTTCCGTCGTTTTAGAGCTAGAAATAGC
phm gR2 RPa	ATGGAGCCTACCAGGCGCCG TGCACCAGCCGGGAATCGAACCC
phm gR2 FPb	CGGCGCCTGGTAGGCTCCAT GTTTTAGAGCTAGAAATAGCAAG
phm gR2 RPb	GTAAAGAGCGTGAGTATCAT TGCACCAGCCGGGAATCGAACCC
phm gR2 FPc	ATGATACTCACGCTCTTTAC GTTTTAGAGCTAGAAATAGCAAG
phm gR2 RPc	CTATTTCTAGCTCTAAAACGCTTCCACTTGGAACTGCCCTGCACCAGCCGGGAATCGAAC
dib 5KO gRNA FP	GCGGCCCGGGTTCGATTCCCGGCCGATGCACCACGAGGACCCTTTGGAAGTTTTAGAGCTAGAAATAGCAAG
dib 5KO gRNA RP	ATTTTAACTTGCTATTTCTAGCTCTAAAACGTCGGCTCTTGCGATATTGATGCACCAGCCGGGAATCGAACCC

primer name	sequence (5' – 3')
dib TSS -110 FP	TGCA GGAAGTAAACCCTCTCAGGT
dib TSS -110 RP	AAAC ACCTGAGAGGGTTTACTTCC
dib TSS -482 FP	TGCA GGGAGGAGGAACTGCTCAAA
dib TSS -482 RP	AAAC TTTGAGCAGTTCCTCCTCCC
dib gR2 FPa	TTCGATTCCCGGCCGATGCAACCACGAGGACCCTTTGGAAGTTTTAGAGC TAGAAATAGC
dib gR2 RPa	GCCCGGAATCGGATCCTATT TGCACCAGCCGGGAATCGAACCC
dib gR2 FPb	AATAGGATCCGATTCCGGGC GTTTTAGAGCTAGAAATAGCAAG
dib gR2 RPb	GTCGGCTCTTGCGATATTGA TGCACCAGCCGGGAATCGAACCC
dib gR2 FPc	TCAATATCGCAAGAGCCGAC GTTTTAGAGCTAGAAATAGCAAG
dib gR2 RPc	CTATTTCTAGCTCTAAAACGCTTGCTGCCCACCAATGGTTGCACCAGCCGG GAATCGAAC
qPCR primers	
phm qPCR FP	GGCATCATGGGTGGATTT
phm qPCR RP	CAAGGCCTTTAGCCAATCG
dib qPCR FP	GTGACCAAGGAGTTCATTAGATTTC
dib qPCR RP	CCAAAGGTAAGCAAACAGGTTAAT
rp49 qPCR FP	CGGATCGATATGCTAAGCTGT
rp49 qPCR RP	CGACGCACTCTGTTGTCG
alas qPCR FP	CCTGCTGAAGCGAGAAGG
alas qPCR RP	GAGGGTCTCCGATCTTAATGG
Coprox qPCR FP	CCAAGTGAAACAGGAGTGAGG
Coprox qPCR RP	AGTCGGGATCCACTTGAGAA
FeCH qPCR FP	AACACAAAGTTTTGCAGACTGG
FeCH qPCR RP	ATCGCGGTCTTCGGTTTT
IRP1A qPCR FP	TCCATCGACAGCAAATATGAGT
IRP1A qPCR RP	CCAGCACATGAAAGTTGTCAC
spz5 qPCR FP	CAAGTCGACTCCCTACAATGC
spz5 qPCR RP	CGACTGAGATCCCTGACCA
Nach qPCR FP	CGAGGCCTTTCTGAACACTC
Nach qPCR RP	GATGTCCTCCGCCGAATA

Table 6.2 Transgenic gRNA constructs and properties

construct	gRNA	vector	characteristics
disembodied (*dib*)			
dU6-dib^{gR1}	2 (CDS)	pCFD5	ubiquitous expression, targets CDS
dU6-dib^{gR2}	4 (CDS)	pCFD5	ubiquitous expression, for FokI-based approach (en_dFC)
PG1-dib^{gR1}	2 (CDS)	PG1-gRNA	PG-specific expression, targets same sites as *dU6-dib^{gR1}*
PG2-dib^{gR1}	2 (CDS)	PG2-gRNA	PG-specific expression, targets same sites as *dU6-dib^{gR1}*
PG3-dib^{gR1}	2 (CDS)	PG3-gRNA	PG-specific expression, targets same sites as *dU6-dib^{gR1}*
dib-TSS^{-110}	1 (TSS)	pCFD5	ubiquitous expression, targets upstream TSS
dib-TSS^{-482}	1 (TSS)	pCFD5	ubiquitous expression, targets upstream TSS
phantom (*phm*)			
dU6-phm^{gR1}	2 (CDS)	pCFD5	ubiquitous expression, targets CDS
dU6-phm^{gR2}	4 (CDS)	pCFD5	ubiquitous expression, for FokI-based approach
phm-TSS^{-174}	1 (TSS)	pCFD5	ubiquitous expression, targets upstream of TSS
phm-TSS^{-423}	1 (TSS)	pCFD5	ubiquitous expression, targets upstream of TSS

Table 6.3 Plasmids used for generating S2 constructs and transgenic lines

constructs	plasmids
dual Cas9-hPR LBD gRNA	Ac5-Stable2-neo (Addgene 32426), pCFD5 (73914), pAct:Cas9 (62209), pP(ELAV-Geneswitch) (83957)
dual dCas9-hPR LBD-FokI gRNA	Ac5-Stable2-neo (Addgene 32426), pCFD5 (73914), pAct:FokI-dCas9 (62211), pP(ELAV-Geneswitch) (83957)
gG-Cas9 GSD	pAct:FokI-dCas9 (62211), pP(ELAV-Geneswitch) (83957), pBPGUw (17575)
gG-Cas9 RD	pLSC-5 (62889), pBPGUw (17575)
gG-Cas9 RO	PX855 (62887), PX856 (62888), pWalium20-10xUAS-3xFlag-dCas9-VPR (78897)
y^1v^1;P(pCFD5 FG TSS gRNA)attP40	pCFD5 (73914)
spok_GSD	pAct:FokI-dCas9 (62211), pP(ELAV-Geneswitch) (83957), pBPGUw (17575), pCRII-TOPO Spok (gift from Michael O'Connor)
act_GSD	pAct:FokI-dCas9 (62211), pP(ELAV-Geneswitch) (83957), pBPGUw (17575),
spok_RD	pLSC-5 (62889), pBPGUw (17575), pCRII-TOPO Spok (gift from Michael O'Connor)
act_RD	pAct:FokI-dCas9 (62211), pLSC-5 (62889), pBPGUw (17575)
spok_RO	PX855 (62887), PX856 (62888), pWalium20-10xUAS-3xFlag-dCas9-VPR (78897), pCRII-TOPO Spok (gift from Michael O'Connor)
act_RO	PX855 (62887), PX856 (62888), pWalium20-10xUAS-3xFlag-dCas9-VPR (78897), pAct:FokI-dCas9 (62211)
UAS-GSD	pBID-UASc (35200), pAct:FokI-dCas9 (62211), pP(ELAV-Geneswitch) (83957)
UAS-RD	pBID-UASc (35200), pLSC-5 (62889)
UAS-RO	pBID-UASc (35200), PX855 (62887), PX856 (62888), pWalium20-10xUAS-3xFlag-dCas9-VPR (78897)

Chapter 7   CRISPR-Cas13 RNA manipulation in *Drosophila* [4]

---

[4] A version of this chapter is in preparation for publication

7.1 Introduction

7.1.1 Overview of bacterial CRISPR systems and recent discoveries

Most bacterial and archaeal genomes harbor Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and encode CRISPR-associated proteins (Cas) as a defense system against bacteriophages and other invading nucleic acids [393–395]. The immune response of all CRISPR/Cas systems characterized to date includes three steps: i) adaptation and spacer acquisition, where a piece of the invading genome is incorporated into the CRISPR array; ii) the expression of mature CRISPR RNAs (gRNAs) from the processed CRISPR array and iii) interference, where Cas enzymes are guided by the gRNAs to the corresponding region of the invading genome for cleavage and degradation [396,397]. The CRISPR/Cas class II systems use a single, multidomain Cas effector protein [398]. Because of its simplicity, the single multidomain effector found in class II organisms is used in current CRISPR methods.

7.1.2 Development of CRISPR-based genome editing techniques in eukaryotes

Class II type II CRISPR Cas9 was one of the first Cas proteins studied in detail, which led to its widespread use for genomic engineering (Figure 7.1A) [91,199,374,381,398,399]. Currently, CRISPR/Cas9 approaches allow scientists to precisely alter gene function via i) classic CRISPR to introduce short INDELs, ii) HR-based CRISPR for homology-based gene replacements or deletions, iii) somatic CRISPR for conditional gene disruption, iv) CRISPRi, (i = interference) to interfere with gene transcription, and v) CRISPRa (a = activation) to upregulate gene activity. Studies have shown that it is possible to conditionally target genes of interest by exerting spatial and temporal control over Cas9 expression or using ligand-activated Cas9 variants [91,94,199,400]. The rapid advances in CRISPR technologies have made it a popular choice over earlier nuclease-based gene editing approaches like meganucleases (MNs)

263

[401,402], zinc finger nucleases (ZFNs) [403–405], and transcription activator-like effector nucleases (TALENs) [406,407].

7.1.3   CRISPR/Cas13: Firing the RNA editing race

The recent introduction of the class II type VI CRISPR/Cas13 system further expands the existing technology in significant ways. Like Cas9, Cas13 uses a guide RNA (CRISPR-RNA, aka crRNA) to identify its substrate, which is RNA rather than DNA (Figure 7.1B). Cas13 enzymes have two distinct catalytic activities: i) an RNAse activity that is mediated by two higher eukaryotic and prokaryotic nucleotide (HEPN) binding domains and ii) a gRNA maturation activity, possibly a combination of activities located in the HEPN2 and Helical-1 domains [408,409]. There are currently four subtypes identified in the Cas13 family, including Cas13a (aka C2c2), Cas13b, Cas13c, and Cas13d. All Cas13 family members are smaller than Cas9, with Cas13d being the smallest protein. The small size of Cas13 proteins makes them suitable for molecular genetics (Figure 7.1C). All Cas13 enzymes require a 60-66 nucleotide long crRNA to ensure target specificity [110,394,395]. Similar to the gRNA in the CRISPR/Cas9 system, the crRNA used by Cas13 forms a short hairpin structure next to a short spacer sequence (28-30 nucleotides) that is specific to the target transcript (Figure 7.1D). Since CRISPR/Cas13 mediates RNA degradation, it holds the promise to replace or complement RNA interference (RNAi) approaches or other systems that interfere with transcript levels, such as CRISPRi. Despite being a powerful tool, RNAi often suffers from low efficiencies or off-target effects, whereas Cas9-based CRISPRi requires a protospacer adjacent motif (PAM), thus limiting the flexibility by which target sequences can be selected [372,410–413]. It is desirable to examine whether CRISPR/Cas13 can offer better specificity and efficiency than these other interference techniques.

7.1.4 The necessity of adopting CRISPR/Cas13 in *Drosophila* research

Drosophila melanogaster is a versatile genetic model organism that is used to study a wide variety of biological processes. Traditional techniques to analyze gene function in *Drosophila* include the generation of mutations via chemical mutagens and transposable P-elements, or the use of transgenes to trigger RNAi and to express cDNAs for gain-of-function studies via the Gal4/UAS system [339,414–417]. Like other model organisms, the CRISPR/Cas9 endonucleases have been quickly adopted by *Drosophila* researchers [89–91,410,411,418–421]. CRISPR-based techniques are remarkably precise and, therefore, ideal for replacing, validating, and complementing traditional approaches, in particular procedures relying on the expression of RNAi or cDNA transgenes [93,366]. Also, the large worldwide collection of gRNAs stocks has ensured the quick adaptation of CRISPR/Cas9 into mainstream *Drosophila* research [398,422,423]. Given the potential of CRISPR/Cas13-based methods to replace current techniques, we explored its feasibility and reliability in *Drosophila*.

Our lab studies signaling pathways that control ecdysone and heme biosynthesis in the larval prothoracic gland (PG), which is part of a larger structure called the ring gland. The PG is a popular model for investigating fundamental aspects of insect endocrinology and allows for the study of external cues that control the timing of ecdysone pulses [67]. Recently, we carried out a genome-wide PG-specific RNAi screen that identified 1,906 genes with critical roles in larval development [66]. In follow-up experiments, however, we often were unable to validate the RNAi-induced phenotypes by independent RNAi lines, either because no such lines existed or because other RNAi lines did not replicate the phenotype. This prompted us to develop CRISPR-based methods that could validate the RNAi results by an unrelated methodology. We previously generated two CRISPR/Cas9 toolkit collections and could use them to validate some RNAi

phenotypes. However, specific issues still exist, including inconsistent gRNA efficiency and early lethality. We sought to investigate the possibility of adapting the CRISPR/Cas13 system for interference and other potential applications of this system in *Drosophila melanogaster*.

We generated and evaluated the catalytic activity of *Drosophila* codon-optimized Cas13 (a-d) variants in a cell line derived from Sg4 embryonic cells. We refer to these Cas13 variants as CasFA[n], CasFB[n], CasFC[n] and CasFX[n], respectively (F = Fruit fly, A-C indicates the Cas13 subfamily, CasFX is the fly version of CasRX, and [n] indicates variant number)(Figure 2A-D). "CasRX" was coined by Konermann et al. for the Cas13d ortholog isolated from *Ruminococcus flavefaciens* XPD3002 to distinguish it from other Cas13d variants [112]. Since we generated fly-optimized versions of CasRX, we refer to these versions as CasFX. Once we had identified a fly-optimized Cas13 variant, we used this variant to adapt existing Cas13 mammalian cell culture applications for *Drosophila* cells, such as transcript tracking and RNA modification [108,110,111,113,424–426]. These *ex vivo* procedures formed the basis for generating a collection of transgenic CRISPR/Cas13 tools designed for *in vivo* RNA targeting. In particular, we generated four Cas13 transgenic lines, namely two that either ubiquitously express CasFB or CasFX, and two that express either CasFB or CasFX under UAS control. The UAS lines allow tissue-specific expression of CasFX and CasFB by crossing them to Gal4-expressing flies. As proof-of-principle that these Cas13 transgenes work effectively *in vivo*, we generated seven crRNA transgenes to target three genes we are studying in our lab.

7.2   Modified materials and Methods

7.2.1   Nuclease-dead dCasFX-IRP1A^{C450S} co-immunoprecipitation

The dCasFX/crRNA complex and IRP1A^{C450S}/Fer1HCH RA cDNA was transfected independently. In one sample, dCasFX and the crRNA corresponding for Fer1HCH RA transcript

were cloned into the same plasmid pC13cr01 (Figure 7.2D) while in another approach, IRP1A^{C450S} and Fer1HCH RA cDNA were cloned into the same plasmid with the similar principle as pC13cr01. IRP1A^{C450S}/Fer1HCH RA co-transfection was done at 10x higher amount in comparing with each dCasFX/crRNA transfection alone. Transfected samples were lysed using 200 μl lysis buffer (1x PBS, 1% Triton, 1x proteinase K inhibitor) by vortexing for 15 seconds every 10 minutes for up to 1 hour. Lysate of IRP1A^{C450S}/Fer1HCH RA transfected samples were combined together and evenly distributed to 10 groups of dCasFX/crRNA lystate. This will ensure each lysate has the similar amount of IRP1A^{C450S}/Fer1HCH RA complex. The mixed lysate was incubated with pre-crosslinked HA Dynabeads protein G (Invitrogen 10004D) following manufacturer's protocol. At the elution step, samples were eluted in 4x Laemli buffer (0.25M Tris pH 6.8, 8% SDS, 40% Glycerol, 25% β-Mercaptoethanol, 0.2% bromophenol blue). Later steps, including running gel and western blot were done as described in chapter 2.

Other experiments were done as described in chapter 2.

7.3   Results

7.3.1   Generation and characterization of *Drosophila* optimized Cas13s

We generated ten Cas13 variants for each of the four Cas13 family members (a-d) by optimizing different codon subsets for codon usage in *Drosophila*. Specifically, we made ten constructs based on the *Leptotrichia wadei Cas13a* gene (*LwaCas13a*), ten variants based on the *Prevotella* sp. P5-125 *Cas13b* gene (*PspCas13b*), ten versions based on the *Fusobacterium perfoetens Cas13c* gene (*FpeCas13c*) and ten forms of the *Ruminococcus flavefaciens* XPD3002 Cas13d gene (aka CasRX). We chose these Cas13 orthologs for the following reasons: i) based on studies in mammalian and plant cells, LwaCas13a, PspCas13b and CasRX showed improved and robust catalytic efficiency when compared to other Cas13 orthologs [108–110,112], ii) unlike

some Cas13 orthologs, the Cas13 genes we chose for our studies do not require a specific

protospacer flanking sequence (PFS) for efficient target RNA identification [108–110,112]. In

the case of PspCas13b, the original study, which was performed in *Escherichia coli,* showed that

the PFS is necessary for RNA cleavage activity. However, when the same enzyme was tested in

mammalian cells and plants, the PFS was no longer required [111,113,425]. Finally, iii) we also

selected Cas13c, since only a few studies have examined this Cas13 subtype [110].

    To evaluate the RNA degradation efficiency of these fruit fly-optimized Cas13
enzymes, we needed to establish a stable reporter gene cell line. For this, we used the PhiC31
integrase system to generate a dual-reporter transgene in the *Drosophila* embryo cell line Sg4-PP-
27F [427] that simultaneously expressed eCFP (enhanced Cyan Fluorescent Protein) and DsRed
(*Discosoma* Red fluorescent protein) (Figures 7.2A, C). Sg4 is one of four embryonic cell lines
isolated from the original Schneider's line 2 (S2) and differs from the popular S2 cells in a range
of transcriptional properties [428]. Importantly, Sg4-PP-27F cells were modified from the original
Sg4 cells by adding a PhiC31 docking site to the second chromosome [427]. The inserted eCFP
and DsRed transgenes are each controlled by the ubiquitous *actin 5C* promoter (*act5C*). To ensure
this transgene's stability, we added a Neo^R gene cassette, which encodes aminoglycoside kinase,
and ensures cell survival in the presence of G418 antibiotics [429]. We refer to this new transgenic
cell line as Sg4_CD (C = eCFP, D = DsRed), and our subsequent cell culture experiments were
based on this line. To transform the Sg4_CD cell line with appropriate vectors, we generated
plasmids that harbored a single copy of a given Cas13 variant and a single crRNA (the vector
allows for adding multiple crRNAs). These constructs, here referred to as pC13cr01, allowed us
to simultaneously express Cas13 as well as its crRNA in transfected cells (Figure 7.2D and
Tables

   Table 7.1). To ensure stable transfection, we also included the *PURO* gene in the
pC13cr01 vector. The *PURO* gene encodes the puromycin N-acetyltransferase, which allows
cells to survive in medium supplemented with puromycin [430,431] (Figure 7.2D and Tables

Table **7.1**). Thus, the presence of two resistance markers allowed for dual selection during the transfection experiments. Besides testing the *Drosophila*-optimized Cas13 variants, we also examined the efficiency of the original Cas13 orthologs in the Sg4_CD cell line (Tables Table **7.1**).

We measured the efficiency of the Cas13 variants by targeting one of the two reporter gene mRNAs and quantifying mRNA levels via qPCR. To accomplish this, for each Cas13 variant, we used two independent single crRNAs targeting eCFP mRNA (crRNA1 and crRNA2) while the DsRed mRNA was not targeted and served as a control (Figure 7.12 and Table 7.2). To ensure that any observed differences derived only from the catalytic activity of the Cas13/crRNA complex, and not from either Cas13 or the crRNA itself, we also tested the eCFP expression level in the presence of a non-targeting (NT) Cas13/crRNA complex. In our hands, the different Cas13 variants showed a wide range of RNA-targeting efficiency, with some of the variants failing to trigger RNA degradation. The original Cas13a, (aka LwaCas13a) showed roughly 35-40% eCFP knock-down efficiency, while the best-performing *Drosophila* variant, CasFA5, was only slightly better and exhibited 47% efficiency (Figure 7.3A). For the Cas13b (aka PspCas13b) variants, we measured 45-51% efficiency for the original Cas13b enzyme, while the best-performing *Drosophila* variants were CasFB5 and CasFB8, both of which were 65-70% efficient (Figure 7.3B). The Cas13c group was the least efficient in knocking down eCFP, with the best line, CasFC4, only accomplishing a 25% knock-down (Figure 7.3C). In contrast, the Cas13d group performed best, displaying 82% efficiency for the original Cas13d (CasRX) enzyme, whereas the CasFX4 variant was even better and reached a 90% knock-down (Figure 7.3D).

To validate these qPCR data, we quantified the protein levels of eCFP and DsRed based on their fluorescence and Western blotting. We selected the best-performing enzyme variants from all four groups, namely three CasFA variants, four CasFB versions, one CasFC enzyme, and six CasFX forms. We then assessed the efficiency of the eCFP knock-down via immunofluorescence

269

and Western blotting (Figures 7.3E and 7.4A-D). Both approaches showed comparable results and confirmed that CasFX4 was the overall most efficient Cas13 enzyme of the entire cohort, showing ~90% and ~95% efficiency on the mRNA and protein levels, respectively.

Next, we sought to investigate whether the subcellular localization of Cas13 would affect the enzyme's catalytic activity. Since mRNAs mature in the nucleus but are translated in the cytoplasm, we wondered if Cas13 performance could be improved by identifying which cellular compartment is optimal for Cas13 activity. To test this, we selected the original Cas13 variants and their corresponding best-performing *Drosophila* counterparts (CasFA5, CasFB5, CasFC4, and CasFX4), and fused them either with a nuclear localization signal (NLS) or a nuclear export signal (NES) (Figure 7.4E). These constructs were based on similar designs from other studies and our approaches (Figure 7.4F). [89,91,94,199,373,382,421,432,433]. Then, as described above, we again examined how efficiently eCFP was knocked down. Overall, we observed similar efficiencies when the same Cas13 variant was tested in the nucleus or cytoplasm, indicating that the catalytic activity of these Cas13 variants was independent of the subcellular localization (Figure 7.4G). For LwaCas13a, PspCas13b and CasRX, this result is consistent with a previous study in plants [113]. Since we found no significant differences, we decided to use Cas13 variants without any localization signal for experiments that followed.

Together, these data suggested that the Cas13 variants retain their RNA-cleaving activity in *Drosophila* Sg4_CD cells, but efficiencies varied considerably. Among the *Drosophila* codon-optimized Cas13 enzymes we generated, we noticed consistent and robust efficiency of two CasFB versions (namely CasFB5 and CasFB8) and the overall best Cas13 variant, CasFX4.

270

7.3.2   Evaluating the collateral activity of *Drosophila*-optimized Cas13 variants

Studies in *Escherichia coli* showed that once the Cas13/crRNA complex is bound to its target RNA, the HEPN-nuclease domains become active and are capable of cleaving not just the intended target, but also RNA molecules that are in the vicinity of the Cas13/RNA complex, resulting in the non-specific RNA degradation referred to as "collateral activity" (Figure 7.6A) [110,408,409,434]. Subsequent studies reported that the collateral activity of Cas13 varied from system to system. While non-specific RNA degradation was detected in human U87 glioblastoma cells [434], no collateral activity was detected in human embryonic kidney 293FT cells or in the plant *Nicotiana benthamiana* [108,110,111]. To test for collateral activity in our hands, we examined the best-performing Cas13 variants using the same transgenic cell line Sg4_CD. Specifically, we co-expressed eCFP, DsRed, and Neo[R] independently, each with an *act5C* promoter. Since eCFP, DsRed, and aminoglycoside kinase (encoded by *Neo[R]* gene) are foreign genetic components, we reasoned that manipulating their expression via Cas13 would not have a significant impact on the physiology of SG4_CD cells. The idea was to target eCFP with specific crRNAs in the presence of Cas13 and monitor the expression of DsRed as a readout for collateral activity. Both eCFP and DsRed were presumed to be highly expressed in a coordinate fashion, since the act5C promoter controlled each transgene. As such, if the interference activity of Cas13 was not specific to eCFP, we expected to detect differences in DsRed expression via qPCR. Using this approach, our data showed that the selected Cas13/crRNA complexes only affected target-eCFP expression, while DsRed expression appeared unperturbed (Figure 7.6B). These data suggest that the tested Cas13 enzymes did not have any detectable collateral activity, at least not in the *Drosophila* Sg4_CD cell line.

271

7.3.3 Testing the fidelity of *Drosophila* Cas13 variants

Our efforts identified several Cas13 versions that efficiently degraded target RNAs in *Drosophila* cells while exhibiting no detectable collateral activity. Next, we wanted to assess how mismatches between crRNAs and their cognate target RNA would affect RNA degradation as a means to define Cas13 fidelity. In particular, we were curious as to whether Cas13 would display higher fidelity - and as such, lower off-target rates - than RNA interference (RNAi), which is widely used in a variety of research models, ranging from cell culture to whole organisms [435–437]. While RNAi is an attractive and powerful tool, its usability is often hampered by its off-target activity, which can make it challenging to interpret phenotypes, and validation strategies involving codon-modified genes/cDNAs are cumbersome and harbor pitfalls [412,438,439]. Other validation strategies include non-overlapping RNAi constructs targeting distinct regions on the mRNA, classic mutants, or conditional CRISPR/Cas9 approaches. To test the propensity of our Cas13 enzymes to degrade off-target RNAs due to small sequence differences, we selected the six top-performing variants for which we had not detected any collateral activity (CasFA5, CasFB5, CasFB8, CasFC4, CasFX4, and CasFX8). Specifically, we generated mismatches in the crRNA-2 spacer sequence and measured the ability to degrade its target RNA, eCFP. To indicate the mismatch location, we referenced the position of the altered nucleotide relative to the stem loop-forming direct repeat of the crRNA. The nucleotide at position 1 represents the one closest to the DR, and the highest number corresponds to the nucleotide farthest away from the DR.

Among all variants that we tested, all had a central region that appeared to be intolerant to single mismatches. The CasFA5, CasFB5, CasFB8, and CasFC4 variants showed some tolerance to single mismatches outside the core region, namely nucleotides #1-6 at the 5'-end and nucleotides #27 and higher at the 3'-end. In contrast, the core region showed no tolerance to mismatches

(Figures 7.5A-D). Remarkably, CasFX4 and CasFX8 variants showed no tolerance for mismatches throughout the entire range, including the extreme 5' and 3' ends. To examine this further, we tested the outermost nucleotides for both CasFX variants (position #1 and #30). Even single mismatches at either end of the spacer region abrogated interference activity, indicating that these two variants are highly specific and have the lowest off-target potential (Figures 7.5 E, F). Since four of the variants had some tolerance towards a single mismatch, we further examined mismatch tolerance by introducing more than one mutation per crRNA. Specifically, we generated constructs encoding two, three, or four mismatches in the eCFP-crRNA. In all tested conditions, we included at least one mismatch from the extreme 5' or 3' end of the spacer. In our hands, none of the *Drosophila* Cas13 variants exhibited tolerance to crRNAs with mismatches of more than one nucleotide (Figures 7.6 C-H). These data are in agreement with other studies using similar approaches [426,440,441]. Taken together, this suggests that the *Drosophila* Cas13 variants tested here are highly specific and display no tolerance to a single mismatch in the core region of the spacer, and none of the enzymes were functional with two mismatches in the crRNA. The CasFA, CasFB, and CasFC variants did tolerate a single mismatch located at either end outside the core region. In contrast, the CasFX4 and CasFX8 variants appeared to require a perfect match of the entire spacer region to mediate interference. We conclude that the CasFX4 and CasFX8 variants will likely have the lowest off-target rate of the Cas13 enzymes tested here.

7.3.4 Nuclease-dead CasFX for applications involving transcript detection

The CRISPR/Cas9 system has been modified to allow for non-nuclease activities, such as for transcription interference (CRISPRi) as well as transcriptional activation (CRISPRa) [91,94,199,421]. Similarly, the Cas13 system can also be adapted for other purposes and may be more suitable for certain applications than CRISPR/Cas9-based methods. For instance, the ability

to target RNA instead of DNA has the advantage that it is reversible. Also, Cas13 may allow for the development of techniques that cannot be accomplished by corresponding CRISPR/Cas9 approaches: By abolishing the nuclease activity of Cas13 while retaining its RNA binding capability, one could use the enzyme to specifically target RNAs to track these transcripts in the cell. Another option would be to fuse Cas13 with different protein domains to affect post-transcriptional processing of target mRNAs, e.g., altering transcript splicing or stability. Specific efforts have been made to investigate these applications with promising results [108,111,113,425,442].

We were particularly interested in a nuclease-deficient Cas13 variant as a tool to validate specific RNA-protein interactions. For our proof-of-principle approach, we selected the Cas13 variant with the most consistent, robust, and specific interference activity, CasFX4 (hereafter referred to as simply CasFX), and introduced quadruple mutations in the catalytic HEPN domains (R239A/H244A/R858A/H863A). These mutations abolish the nuclease activity but not RNA binding activity in the CasRX variant [108,113,442] (Figure 7.7A). We first tested whether the mutant CasFX still retained nuclease activity by testing our validated crRNAs against eCFP in the Sg4_CD cell line. As expected, the mutant CasFX failed to interfere with the expression level of eCFP, whereas the wild-type variant worked efficiently (Figures 7.7B, C). We conclude that this mutant CasFX variant, similar to the corresponding variants in other species, lost its nuclease activity. We hereafter refer this variant as dCasFX (d = dead).

Introducing mutations into Cas13 led to the development of the nuclease-dead dLwaCas13a and dCasRX variants, which abolished their RNA degradation ability while maintaining their RNA-binding activity [425,442]. Similarly, we tested whether introducing equivalent mutations would allow us to generate a nuclease-dead dCasFX that still retained RNA

binding capability. To assess whether crRNA-guided dCasFX would specifically interact in a non-destructive manner with its intended target mRNA, we tested its ability to co-IP a known mRNA-binding protein, which we considered a much more stringent test than mere RNA-immunoprecipitation since this would be a better indication that the mRNA had remained intact. For this, we used an isoform of the ferritin heavy chain 1 mRNA (Fer1HCH-RA), which carries a canonical iron-responsive element (IRE) at its 5' end. This IRE allows iron regulatory protein 1A (IRP1A), the *Drosophila* ortholog of human iron regulatory protein 1 (IRP1), to bind to the Fer1HCH-RA mRNA [124,166,198,443,444]. Specifically, we used the IRP1A^{C450S} form , which is constitutively RNA-binding. We then designed a series of crRNAs that directs dCasFX to its target, Fer1HCH-RA, and tested whether immunoprecipitation of dCasFX would also pull down IRP1A. We transfected and lysed cells containing the dCasFX and crRNA components, and mixed this lysate with a second sample obtained by lysing cells containing transfected Fer1HCH-RA mRNA and IRP1A^{C450S}. By combining the two lysates together, the dCasFX/crRNA enzyme should bind to the Fer1HCH-RA mRNA/IRP1A^{C450S} complex. If the interaction occurs, immunoprecipitation of dCasFX could also pull down IRP1A^{C450S} (Figure 7.7D).

A key question for this strategy was how far the recognition site for dCasFX/crRNA needed to be away from the IRE to allow binding of both proteins, dCas13 and IRP1A, to the Fer1HCH-RA mRNA. To this end, we generated nine different crRNAs, representing binding sites spaced ~150 bases apart to roughly cover the entire 1.7 kb Fer1HCH-RA mRNA. One of the sites (crRNA #3) partially overlapped with the IRE site, which served as a control to disrupt IRP1A binding. Using this strategy, we found that immunoprecipitation of dCasFX successfully pulled down IRP1A, as long as the cRNA binding site was sufficiently removed from the IRE. As expected, this interaction appeared to be dependent on the distance between the crRNA target

site and IRE sequence, since an insufficient distance should cause steric hindrance between the two proteins (Figure 7.7EAs a control, we used a non-targeting (NT) crRNA to ensure the interactions we observed were specific. The control showed that immunoprecipitation of dCasFX with a non-*Fer1HCH-RA* mRNA-targeting cRNA was not able to pull down IRP1A.

We also tested whether we can simply detect immunoprecipitated *Fer1HCH-RA* mRNA via real-time PCR (qPCR). In the absence of IRP1A, dCasFX appears to bind to the *Fer1HCH-RA* mRNA efficiently, and we found no significant differences between the nine different crRNAs (Figure 7.8A). Interestingly, when we repeated the experiment in the presence of IRP1A, we noticed a ~4-fold reduction of immunoprecipitated *Fer1HCH-RA* mRNA when we used cRNAs #1-4 (Figure 7.8B). This is consistent with the results for co-immunoprecipitated IRP1A (Figure 4E), suggesting that competition between IRP1A and dCasFX (bound to crRNAs #1-4) affected the RNA-binding ability of both proteins. We conclude that dCasFX is a reliable tool to validate interactions between a protein and its candidate target RNA. In addition to RNA-immunoprecipitation, dCasFX could potentially also used for other *in vivo* studies, such as locating a transcript of interest to elucidate its subcellular localization or for co-localization studies, or to determine whether a given protein is bound to its target RNA or unbound.

7.3.5   Targeting mitochondrial RNAs via Cas13

Like CRISPR/Cas9, Cas13 needs to form a complex with a crRNA before it can identify and cleave its target transcript [110,409]. Since the Cas13/crRNA complex harbors a single protein, it can be easily tagged with a mitochondrial targeting sequence to cleave RNA in mitochondria, which is not feasible with RNAi. *Drosophila* mitochondria contain multiple copies of circular DNA (mtDNA), which encode tRNAs, rRNAs, and polypeptides important for oxidative phosphorylation. The study of mitochondrial genes is important, because mutations in

mtDNA can cause devastating human disorders, such as Leber's hereditary optic neuropathy, which causes blindness [445,445,446,446,447,447]. To modify CRISPR/Cas13 applications for mitochondrial-encoded transcripts, we added a sequence encoding an N-terminal mitochondrial targeting peptide derived from the nuclear-encoded *translocase of the inner mitochondrial membrane 23* (*tim23*) gene. For this approach, we generated a modified version of our highly efficient CasFX variant, which we termed CasFX^mt. The CasFX^mt/crRNA complex is predicted to be imported into the mitochondrial matrix, where it should bind to and cleave the target transcripts (Figures 4F, G).

To test the functionality and efficiency of the CasFX^mt variant, we co-transfected CasFX^mt with constructs encoding a crRNAs against either *mitochondrial cytochrome c oxidase subunit I* (*mt:CoI*, aka *COXI*) or *mitochondrial cytochrome c oxidase subunit II* (*mt:CoII*, aka *COXII*). Both *COXI* and *COXII* are highly expressed mitochondrial-encoded genes critical for oxidative phosphorylation [448–450]. We analyzed the expression levels of *COXI* and *COXII* via qPCR as well as western blots. To put these results into context, we generated RNAi samples against each of these targets, and used the original CasFX (CasFX^O, O = original) variant, which lacks the mitochondrial sequence, as a control. In our hands, RNAi targeting either *COXI* or *COXII* had no significant effect on the expression of these two transcripts. Similarly, CasFX^O/crRNA produced no significant effects (Figures 7.7H, I). In stark contrast, CasFX^mt caused a 4-5-fold reduction of the COX transcripts and resulted in a substantial drop in protein levels as well (Figures 7.7 H,I). To ensure that this result was reproducible, we tested additional RNAi as well as crRNAs sequences, all of which target *COXI* or *COXII* transcripts (Figure 7.12). In all cases, the observed results were comparable (Figure 7.8), suggesting that CasFX^mt is a useful tool to target mitochondrial-encoded transcripts.

7.3.6 Cas13-ADAR2 for RNA modification

One intriguing aspect of CRISPR/Cas13 has focused on the modification of RNA, which led to two approaches, namely "<u>R</u>NA <u>e</u>diting for <u>p</u>rogrammable <u>A</u> to <u>I</u> replacement" (REPAIR) and "<u>R</u>NA <u>e</u>diting for <u>s</u>pecific <u>C</u> to <u>U</u> exchange" (RESCUE) [108,424]. These methods allow for programmable adenosine-to-inosine editing as well as cytosine-to-uridine editing, respectively. The ability to modify genetic information at the RNA level may be advantageous, because, unlike Cas9 which causes a permanent change in the genome, RNA modifications via Cas13 are reversible due to RNA turnover [94,166,199,421]. As such, Cas13-based approaches may be suitable for future therapies, where Cas13 could be used to repair missense mutations in transcripts without affecting a patient's genome.

In the REPAIR systems used in mammalian cells, the nuclease-dead PspCas13b was fused to the RNA-modifying domain of Adenosine Deaminase Acting on RNA 2 (ADAR2). In their original approach, Cox et al. found that the first REPAIR version (REPAIRv1) had substantial off-target activity. Subsequently, they generated REPAIRv2, which harbored two point mutations in the ADAR2 domain (T375G and E488Q). This version showed high specificity and robustness in mammalian cells [108].

Given its success in mammalian cell systems, we wondered whether a Cas13-ADAR fusion would be functional in *Drosophila*. The insect ADAR protein appears to function similarly to its human counterpart [451], suggesting that constructs based on mammalian ADAR2 would work in *Drosophila*. We first fused the above-described dCasFX to the mutant human ADAR2 domain that carries equivalent mutations as the REPAIRv2 we mentioned earlier. We refer to this construct as FREPAIRv2 (F = fruit fly), and tested for its editing efficiency (Figure 7.9A). To test for Cas13-ADAR2 activity, we generated a system that uses a dual reporter transgene in the *Drosophila*

embryo cell line Sg4-PP-27F. Similar to the earlier described Sg4_CD line; this cell line carries the independently expressed *eCFP* and *DsRed* transcription units in the genome, each with their own *actin5* promoters. However, unlike the Sg4_CD line, we introduced a point mutation into the eCFP coding region that converts a tryptophan residue 57 (W57*) TGG into an early stop codon (TGA), which we refer to as eCFP*. Also, we termed this new cell line "Sg4*" line to distinguish it from the original Sg4_CD (Figure 7.2B). Next, we co-expressed FREPAIRv2 and an eCFP-crRNA, which carries a single mismatch A to C at the position that corresponds to the introduced stop codon (Figures 7.9A, B). If the FREPAIRv2 is capable of editing its target RNA encoded by *eCFP**, the stop codon should be reverted to the wild-type tryptophan residue (W57), and the resulting full-length eCFP should be detectable via Western blotting and, if efficiency is sufficiently high, via fluorescence from the restored CFP. Using this strategy, we found that we were able to detect fluorescence at a wavelength of 405 nm as early as 36 hours after transfection, indicating that detectable levels of eCFP had been produced. eCFP fluorescence continued to increase, with substantially higher levels at the 60-hour time point (Figure 7.9D). When we conducted Western Blots to validate these data, we saw corresponding results, with detectable eCFP protein at 36 hours and progressively higher levels from 42 to 60 hrs after transfection (Figure 7.9C). We conclude that Cas13-ADAR2 works effectively in *Drosophila* and can be used to modify target mRNAs, such as reverting transcripts carrying missense mutations without altering the genome.

For the above approach, we followed a similar path that was used in the original study [108] where the mismatch (C→A) was placed in the center of the crRNA spacer, measured at the 26th nucleotide of 50 nucleotides (nt) spacer, relative to the stem loop-forming direct repeat of the crRNA. To evaluate the editing efficiency in correlation to mismatch position and spacer

length, we tested a series of crRNA constructs with the same spacer length of 50 nt; however, we changed the relative mismatch distance to the hairpin by increments of 8 nt (Figure 7.9E). We then performed reverse transcription and sequenced a minimum of ten randomly selected eCFP cDNAs per construct. This was followed by sequencing to assess the fraction of clones that harbored the repaired codon for tryptophan #57, expressed as editing rate (Figure 7.9F). Based on our findings, the crRNA that carried the mismatch at position 26 relative to the hairpin ("mismatch distance", Figure 7.9E) resulted in the highest efficiency (Figure 7.9F), consistent with other studies. We then tested the effect of varying spacer length while keeping the mismatch distance at 26 nt. We tested spacer lengths from 30 nt to 80 nt, and in all cases, we observed similar efficiencies, all of which were comparable to a 50 nt spacer (Figure 7.9G). Based on these findings, we conclude that FREPAIRv2 works best when using a mismatch distance of 26 nt, whereas the spacer length did not appear to affect the editing efficiency [108].

To evaluate the off-target tendencies of FREPAIRv2 in *Drosophila* cells, we examined the cDNA sequences for additional A→I modifications, which is straightforward since inosine is recognized as guanosine by the reverse transcriptase. However, we scored any unpredicted sequence deviations as potential off-target events and plotted them relative to the mismatch distances and spacer lengths (Figures 7.9F, G). This strategy revealed that some off-target effects persisted, albeit at a low level across all crRNAs that we tested. Given that these effects are random, and distributed across multiple RNA molecules, it appears likely that this off-target activity has no or inconsequential impact on phenotypes. However, future studies are needed to improve the specificity of this editing system further.

7.3.7 Generation and characterization of Cas13 flies

Our data demonstrated that Cas13 works well in *Drosophila* Sg4 cells and can be used for purposes beyond RNA cleavage. We next sought to generate transgenic fly lines carrying Cas13 variants and characterize their efficacy *in vivo*. To this date, no study has analyzed the usability Cas13 in live organisms to the best of our knowledge. As such, it is critical to establish whether Cas13-based technology is suitable for *in vivo* studies. Furthermore, we were interested in creating a system that allows for temporal and spatial control over Cas13 expression. To this end, we have previously created a *Drosophila* toolkit for CRISPR/Cas9 based on Gateway-compatible cassettes that allow researchers to insert specific enhancers that drive the expression of the Cas transgene in a tissue of interest [199,199,432,432]. While this generates more upfront work compared to Gal4/UAS-based systems driving the expression of Cas9 [90,94], it does simplify the downstream workflow. Also, it reduces unspecific effects since one requires fewer transgenes to build the necessary fly genotype. We, therefore, decided to create a similar Cas13 toolkit. In total, we manufactured two general Cas13 vectors, one based on CasFB and one that uses CasFX, both of which displayed the highest catalytic efficiency in Sg4_CD cells. For our *in vivo* strategy, we limited our efforts to constructs that would interfere with RNA expression (Figure 7.10A). Based on these all-purpose vectors, we then generated four transgenic lines for further characterization, named here *act_CasFB*, *UAS-CasFB*, *act_CasFX*, *UAS-CasFX* (Figure 7.10A). For the generation of crRNAs, we used the previously described multiplexed pCFD5 vector and implemented changes suitable for Cas13 crRNA processing [94]. We refer to the new plasmids as i) pC13B, which expresses CasFB-compatible crRNAs under control of the U6:3 promoter and ii) pC13X, which expresses CasFX-compatible crRNAs under control of the U6:3 promoter (Figure 7.10B, C). Both plasmids will ubiquitously express the tRNA:crRNA units. As the tRNA is processed, its cleavage

will result in the release of mature crRNAs that form complexes with Cas13 enzymes. The cloning procedures for these new crRNA plasmids are overall similar to those for the pCFD5 vector, but, since some differences exist, we include a detailed protocol in the supplementary material (see Appendix A.3).

To evaluate the efficiency of our transgenic Cas13 constructs *in vivo*, we generated seven transgenic crRNAs targeting three genes that we study in our lab. This includes *phantom* (*phm*) and *disembodied* (*dib*), two well-characterized genes involved in ecdysone synthesis in *Drosophila* [358,359] as well as the third gene, *Iron Regulatory Protein 1A* (*IRP1A*), a gene critical for cellular iron homeostasis [166,248]. Classic mutants of *phm* and *dib* display embryonic lethality while *IRP1A* mutant animals die as first instar larvae (L1) [166,199,358,360,432]. In contrast, using PG-specific somatic CRISPR/Cas9 strategies, *phm*^{gR} (gRNA for CRISPR Cas9) caused L1 arrest, while *dib*^{gR} and *IRP1A*^{gR} both caused third instar (L3) larval arrest (Figures 7.11A-C) [166,199,432]. In addition, PG-specific disruption of *IRP1A* via somatic CRISPR/Cas9 caused a porphyria-like phenotype due to iron deficiency (Figure 7.11D) [166].

When we crossed the Cas13-compatible crRNAs (referred to as 13B for CasFB-compatible cRNAs and 13X for CasFX-compatible crRNAs) targeting either *phm*, *dib* or *IRP1A* with either ubiquitously expressed or PG-specific Cas13 variants, we observed the same developmental defects we found with our previous strategies (Figures 7.11 A-C, E and 7.12), indicating that Cas13 worked effectively in *Drosophila*. The fact that *phm*^{13B}, *phm*^{13X}, *dib*^{13B}, *and dib*^{13X} individuals were rescued to adulthood when reared on 20E-supplemented media [199,432], and that *IRP1A*^{13B}, as well as *IRP1A*^{13X} animals, reached adulthood when dietary iron was provided [166], strongly suggested that the activity Cas13 was highly specific (Figures 7.11A-C).

In addition to the above phenotypic analysis, we evaluated *dib* expression levels via qPCR. We compared the results to other tissue-specific loss-of-function techniques, including samples from two independent RNAi lines and samples from one line where we used transcriptional interference via dead Cas9 (dCas9) to target *dib*. We found that the two RNAi lines reduced dib expression by 30-40%, whereas the CRISPRi approach via dCas9 lowered *dib* expression by 50-60%. Concerning the new Cas13 lines, CasFB reduced *dib* expression by 55-65%, equivalent to the dCas9 data. Remarkably, CasFX showed the strongest knock-down, and robustly reduced *dib* expression by 80-90% (Figure 7.11E). Taken together, these data indicated that Cas13 transgenes work *in vivo* and may exceed the efficacy of other techniques.

We also tested the ability to target multiple transcripts with a single transgene. For this, we used the pC13X vector and generated a dual-crRNA transgenic line (termed dI^{13X}) that ubiquitously expressed a crRNA targeting *dib* mRNA as well as a crRNA targeting the *IRP1A* transcript (Figure 7.11F). Target sites for either of these transcripts were the same as before (Figures 7.11A, C and 7.12). As expected, the animals arrested development at the L3 stage, similar to targeting the *dib* and *IRP1A* transcripts individually. Consistent with this, neither 20E-nor iron-supplementation alone could rescue these double knock-downs, however, a diet supplemented with both 20E and iron caused a significant rescue to adulthood (Figure 7.11G). This makes sense since the two cRNAs interfered with ecdysone production and the regulation of cellular iron homeostasis. To assess whether the simultaneous knock-down of two genes was as efficient as targeting these genes individually, we evaluated *dib* and *IRP1A* expression levels via qPCR. We found no significant difference in any of these approaches suggesting that there is no penalty when targeting two genes at the same time (Figure 7.11H).

7.4   Discussion

7.4.1   RNA targeting efficiency of Cas13 in *Drosophila*

We evaluated eleven variants of each reported Cas13 ortholog in *Drosophila* Sg4 cells, including the well-characterized variant from the original studies and ten *Drosophila*-optimized variants. Among all Cas13 enzymes tested, we observed a wide range of efficiencies, even between the versions from the same ortholog. Among them, CasRX and its *Drosophila*-optimized variants CasFX appeared to have the highest efficiency. For the Cas13a and Cas13b variants, we also identified the optimized variants with reliable efficiency. Even though they were less efficient than CasFX, these variants may still prove useful in circumstances where only a moderate knock-down is desired. On the other hand, Cas13c variants did not significantly alter the expression of target transcripts. We hypothesize that this was caused by several factors: (i) Cas13c is the least characterized Cas13 enzyme, and it might use a mechanism that differs from the other Cas13 enzymes. (ii) Even though the low efficiency of Cas13c was in agreement with previous studies conducted in other species, we cannot rule out the possibility that the Cas13c variants we used were not ideally suited for *Drosophila*, and (iii) Cas13c might still require a PFS for optimal activity in the fruit fly. Future studies will need to address this.

We noticed that the expression of the PspCas13b and CasRX variants resulted in considerable toxicity when animals were homozygous for these transgenes, causing lethality during the first (L1) or second (L2) instar larvae (Figure S8). Interestingly, animals heterozygous for PspCas13b and CasRX transgenes showed no significant lethality. In contrast, animals homozygous for our *Drosophila*-optimized Cas13 transgenes, namely CasFB and CasFX, showed only moderate lethality, with 51% to 58% reaching adulthood, respectively (80-85% is expected in wild type populations). As expected, animals heterozygous for these transgenes appeared

normal (Figure S8). The lethality of Cas13 transgenic animals was also reported in a recent study [452], similar to the results of early versions of Cas9 in *Drosophila* [91,199]. Since we observed a wide range of efficiencies between the variants, it is possible that each variant also exhibits different levels of toxicity. While the reasons for the relatively high lethality of the original PspCas13b and CasRX constructs (in a homozygous setting) remain unclear, our data suggest that each variant is unique and that perhaps using codon-optimized versions help to reduce the toxicity associated with Cas13.

7.4.2 Beyond RNA interference

A few studies have shown that Cas13 may be useful in a broad range of applications, and not just RNA cleavage. In this study, we have demonstrated that dCasFX can validate RNA-protein interactions by using an appropriately designed crRNA. We also showed that by adding a mitochondrial localization sequence, one could recruit the CasFX^{mt}/crRNA complex into mitochondria and target mitochondrial-encoded transcripts. We also adopted the REPAIRv2 system from mammalian cell culture into *Drosophila* Sg4 cells and showed that this system, FREPAIRv2, can efficiently modify target transcripts with an overall low off-target rate. We have not tested other potential applications; however, in theory, Cas13 can be modified for many approaches to study RNA, including splicing, transcript stabilization, or RNA localization. Cas13 may have far-reaching implications for simplifying diagnostics. Recently, the outbreak COVID-19 caused by SARS-CoV-2 has resulted in a global health threat. To develop a fast test for COVID-19, the specific high-sensitivity enzymatic reporter unlocking (SHERLOCK) protocol, a recently developed Cas13-based diagnostic test for infectious diseases, can detect the virus in 50 min [453,454] (https://mcgovern.mit.edu/2020/02/14/enabling-coronavirus-detection-using-crispr-cas13-an-open-access-sherlock-research-protocol/). In an independent study,

CRISPR/Cas13 was also used to detect SARS-CoV-2 [455]. Together, these studies demonstrate the enormous potential of Cas13 as a diagnostic and therapeutic tool.

7.4.3   From in vitro to in vivo

A significant part of the work presented here was based on cell culture experiments. These approaches were ideal to economically evaluate the efficiencies of multiple Cas13 versions in *Drosophila*. However, our ultimate goal is to establish CRISPR/Cas13 approaches for *in vivo* studies in model organisms, which has not been accomplished yet. Based on our results of transgenic CRISPR/Cas13 flies, CasFX and CasFB can efficiently target and cleave transcripts of interest *in vivo*, and as such, represent a compelling alternative to existing methods. This study may also help scientists working with other model organisms to optimize their approach for implementing Cas13 *in vivo*.

7.4.4   The CRISPR/Cas13-based toolkit

This study has generated two collections of Cas13/crRNA toolkits to study in either cell culture or organisms. For the cell culture toolkit, we have produced the pC13cr01 vectors, which allow the co-transfection of Cas13 variants and the crRNA corresponding to the target transcript. With this vector, one only needs to digest the crRNA backbone with the BbsI enzyme and clone the target site for the crRNA, similar to the generation of the Cas9-compatible gRNA system in pCFD5 or pCFD6 plasmids. For *in vivo* work, we also established a similar system with Cas13 transgenes already available from our study. Researchers will need to generate their crRNAs against the target transcript. For this, we provide the pC13B and pC13D vectors with the same cloning procedure as pCFD5. We also provided a supplemental method section with a detailed description of the cloning procedures. On the other hand, the UAS-based versions of Cas13

transgenes will also allow scientists to spatially and temporally manipulate Cas13 activity and study transcript of interest at desired tissues.

7.4.5   Conclusions and future directions

Just like CRISPR/Cas9 allows for the manipulation of DNA, Cas13 enables us to target any transcript of interest. This is beneficial for approaches where researchers do not want to alter the DNA of the gene of interest, since Cas13 controls gene expression on the RNA level, similar to RNAi. Furthermore, current evidence suggests that Cas13, especially variants from the Cas13d family, display minimal off-target tendencies, and this might help quell concerns regarding RNA targeting. Even though it might be too early to make conclusions about the off-target activity of Cas13, we believe that its high specificity holds excellent promise for future applications. Also, the ability to modify Cas13, such as targeting Cas13 to mitochondria, further expands the range of future applications for this methodology.

7.5   Figures
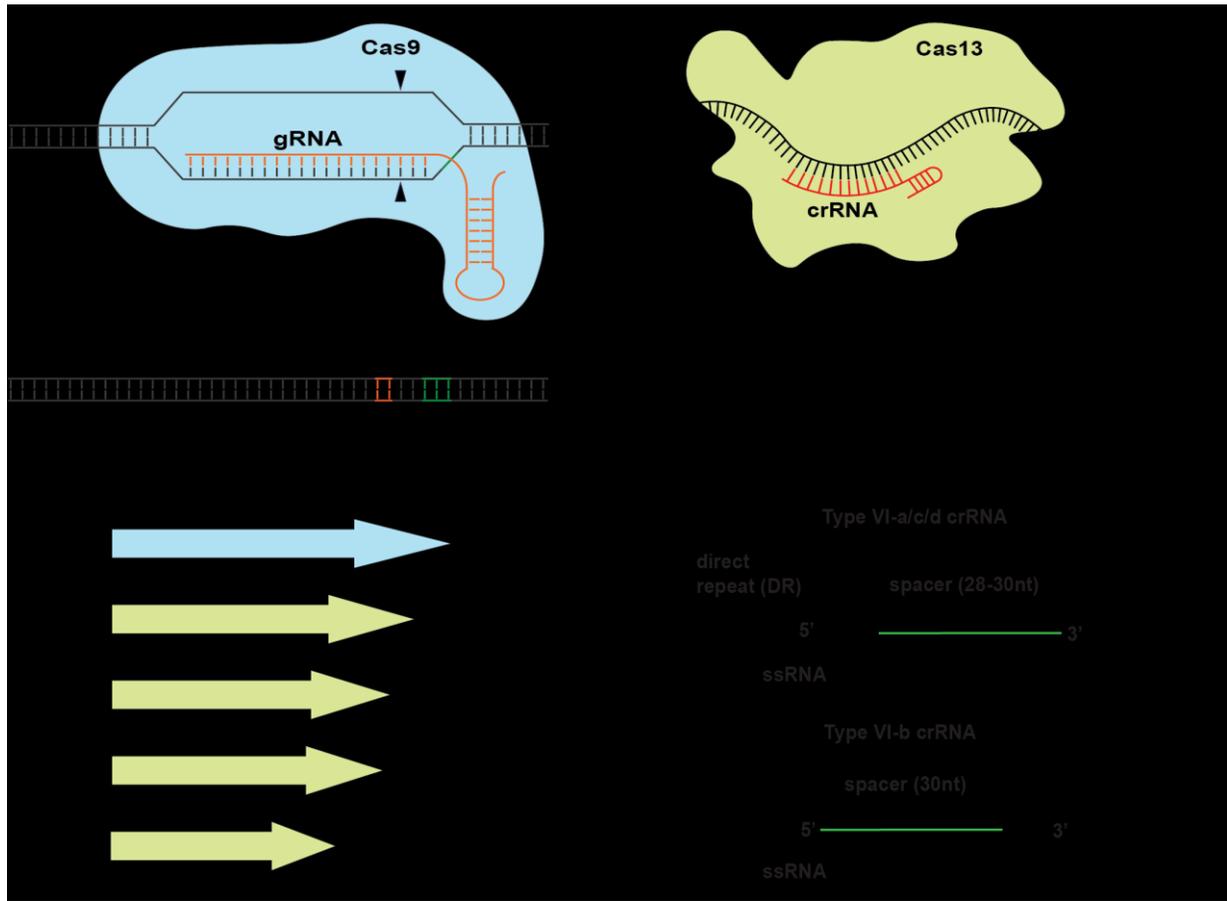
Figure 7.1 Functional overview of CRISPR/Cas9 and CRISPR/Cas13 systems.

(A) Schematic of Cas9 mechanism in genome editing. This system requires the recruitment of CRISPR-associated protein Cas9 (blue) to the target site recognized by the guide RNA (gRNA: orange). Target site cleavage by Cas9 is ensured by the presence of the protospacer adjacent motif (PAM) (green), a sequence that immediately follows the target site. The PAM will direct cleavage site Cas9 to a region of about three nucleotides upstream of PAM. **(B)** Schematic of Cas13 mechanism in RNA cleavage. This system requires the pre-assembly of Cas13 (green) with the CRISPR RNA (crRNA: red). The complex will then be recruited to target RNAs. Upon RNA-binding, Cas13 will undergo a conformational change and induce the catalytic activity of its

nuclease domains, resulting in the cleavage of target transcripts. **(C)** Comparisons of Cas9 size with different Cas13 subtypes (a-d). Polypeptide sizes are indicated as the number of amino acids. **(D)** Relative structural representation of different Cas13 subtype-compatible crRNAs. All four subtype crRNAs carry a direct repeat (DR) to facilitate the binding with their corresponding Cas13 enzyme, as well as a spacer sequence specific for the target transcript. Cas13b-compatible crRNAs carry a DR at the 3' end while compatible crRNAs for Cas13a,c, and d carry the DR at the 5' end.
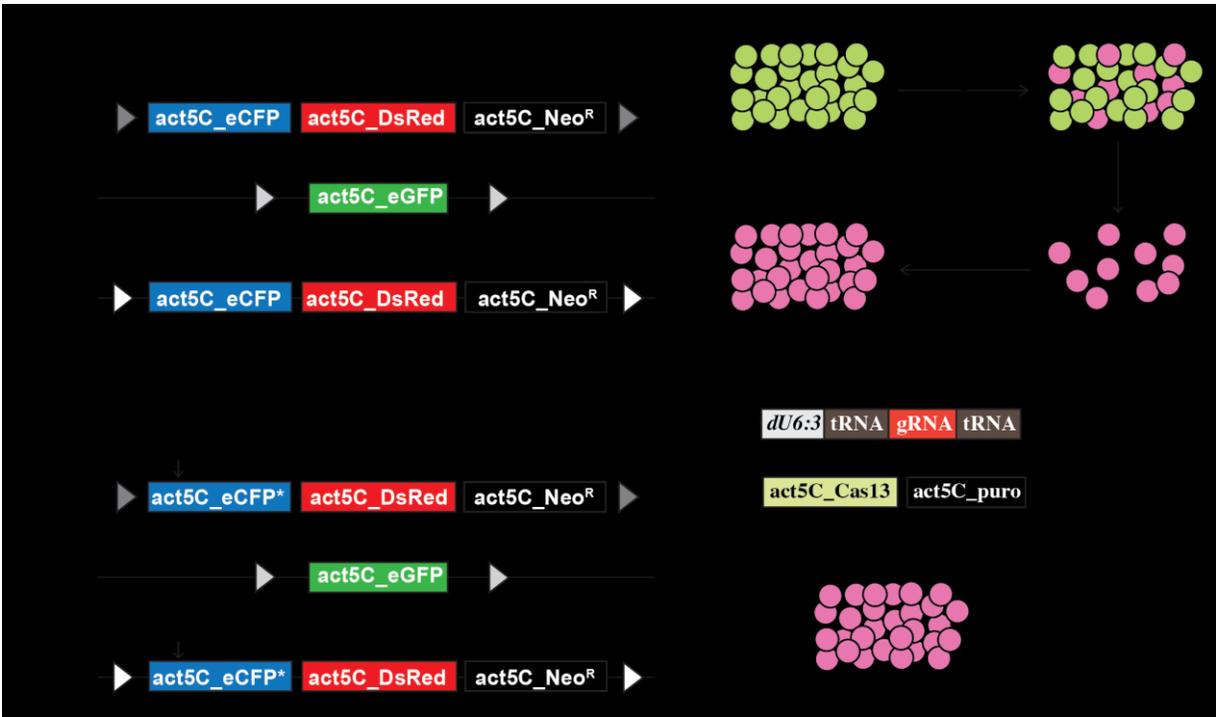
Figure 7.2 Schematic of transgenic cell culture and *in vitro* study.

(A) Generation of Sg4_CD cell line that expresses *eCFP*, *DsRed*, and *Neo^R* genes under independent *actin5C* (*ac5*) promoters. **(B)** Generation of Sg4_cD cell line that expresses mutant *eCFP**, *DsRed*, and *Neo^R* genes under independent *actin5C* (*ac5*) promoters. **(C)** Establishment of the transgenic cell line. Two days after transfection, cells were supplemented with geneticin (G418). Cells without transfected plasmid will be eliminated eventually, leaving cells with successful integration. Cells were passaged for at least four rounds, and integration was confirmed via sequencing. **(D)** Schematic of pC13cr01 vector activity. Upon transfection with the pC13cr01 vector, the cells were selected with geneticin as well as puromycin to eliminate untransfected cells. Seven days after transfection, cells were collected for later study.
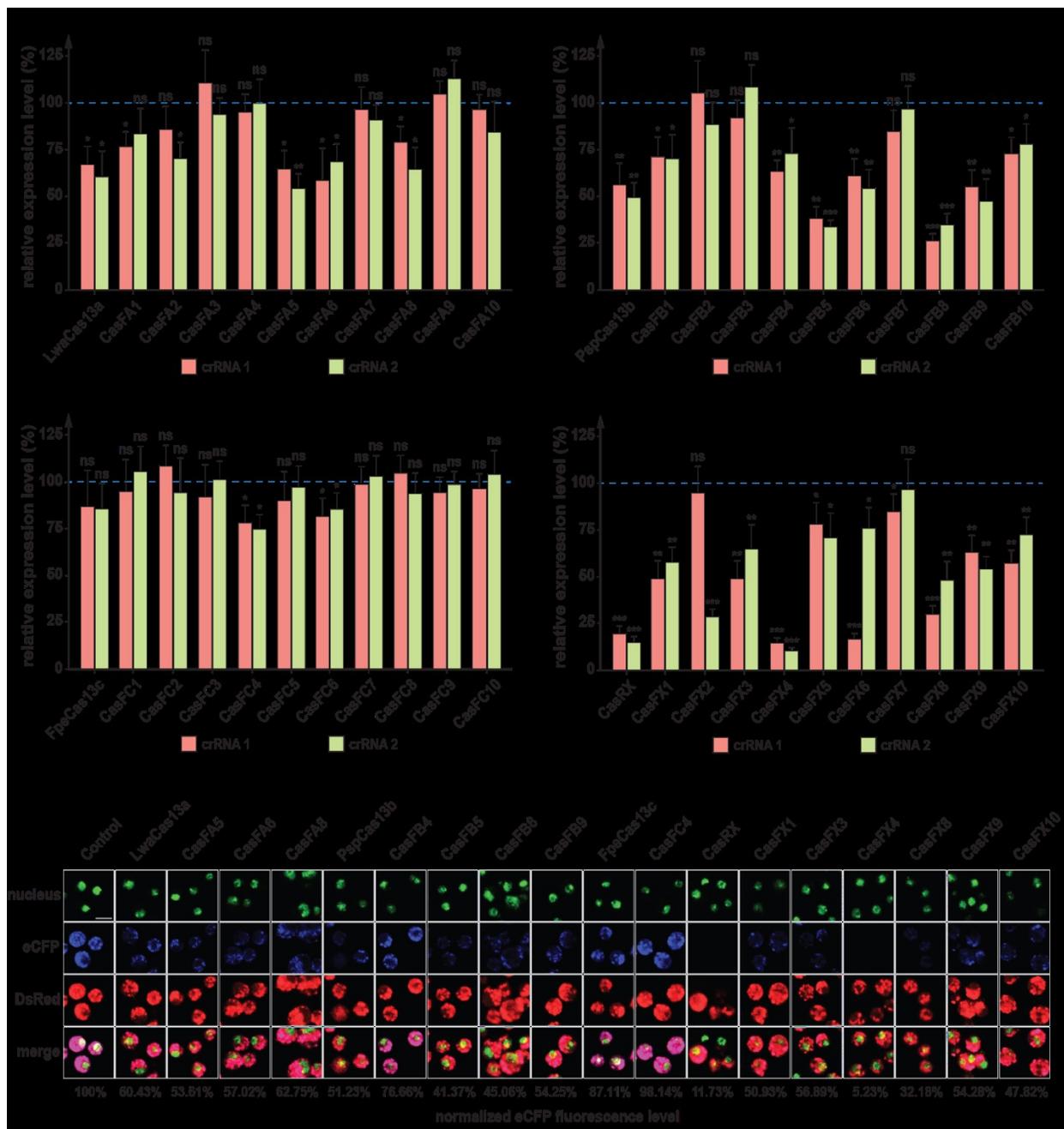
Figure 7.3 Efficiency evaluation of *Drosophila* codon-optimized Cas13 variants.

(A-D) qPCR analysis showing eCFP transcript levels in Sg4 cells, as a function of the different Cas13 variants that were expressed in these cells (a-d, respectively). Shown are relative fold changes of eCFP transcript being targeted by two independent crRNAs, crRNA 1 (red) and crRNA 2 (green). Data were normalized to eCFP expression levels when using a blank crRNA as a control

(blue dotted line = 1). * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. Error bars represent standard errors. **(E)** Fluorescence changes of eCFP across samples targeted by the Cas13/crRNA 2 complex. Fluorescence levels were measured using ImageJ and normalized to signals obtained with a blank crRNA (control). Nuclei were stained with nuclear green. eCFP and DsRed fluorescence were measured using their native fluorescence properties (no antibody staining). Scale bar = 50 μm.
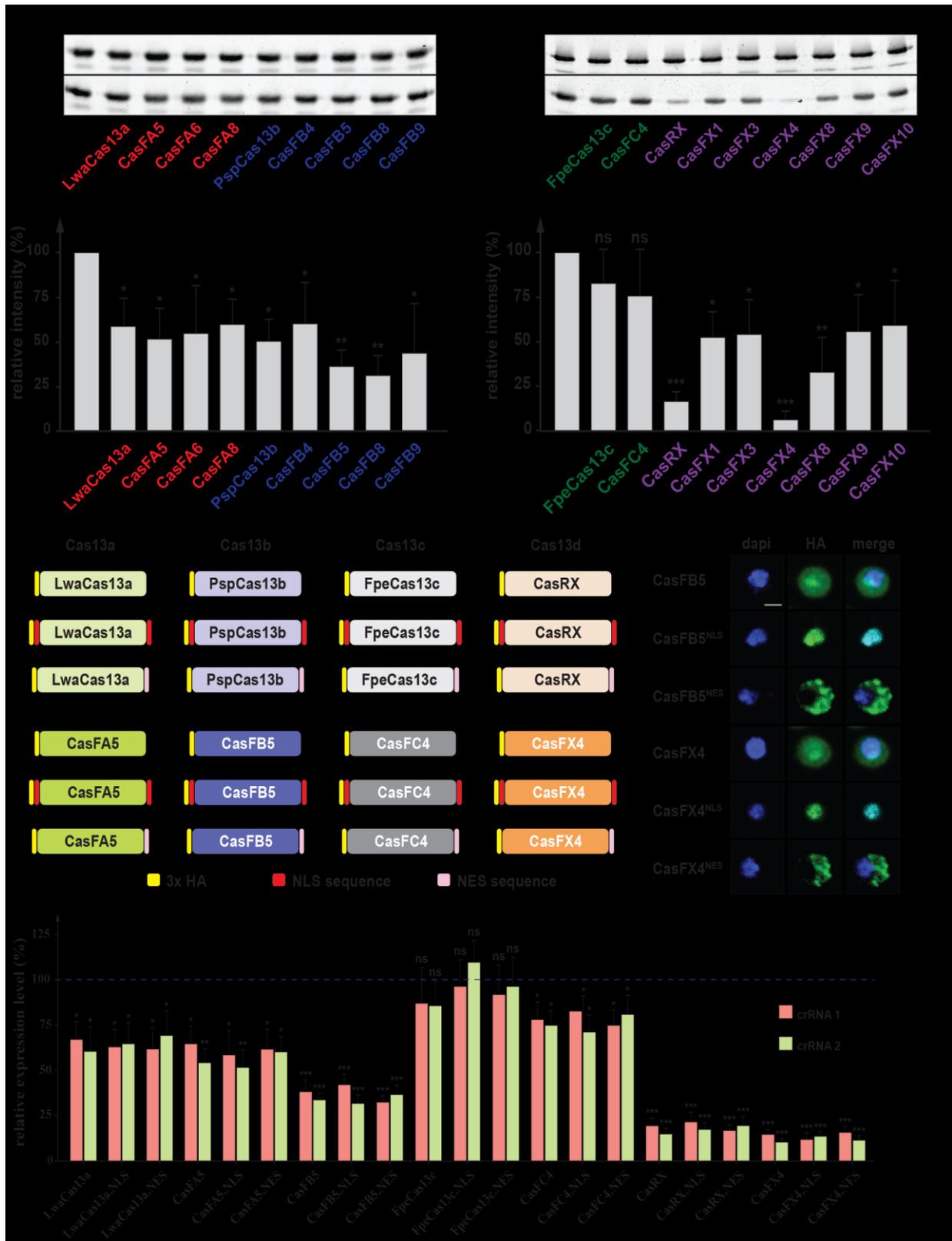
Figure 7.4 Evaluation of *Drosophila* codon-optimized Cas13.

(A-D) Western blotting of eCFP in samples that were treated with Cas13 variants that showed the highest efficiency in qPCR experiments. Band intensities were quantified with ImageJ and normalized to samples treated with blank crRNA. * = p-value < 0.05, ** = p-value < 0.01, *** = p - value < 0.001. Error bars represent standard error. **(E)** Schematic of Cas13 variants with different signaling sequences, including nuclear localization signal (NLS), nuclear export signal (NES). **(F)** Evaluation of NLS and NES on Cas13 efficiency. Each Cas13 variant being tested were fused with either NLS or NES and test for their interference efficiency on eCFP expression. Data were normalized to samples treated with blank crRNA (blue dotted line = 1). * = p-value < 0.05, ** = p-value < 0.01, *** = p - value < 0.001. Error bars represent standard error.
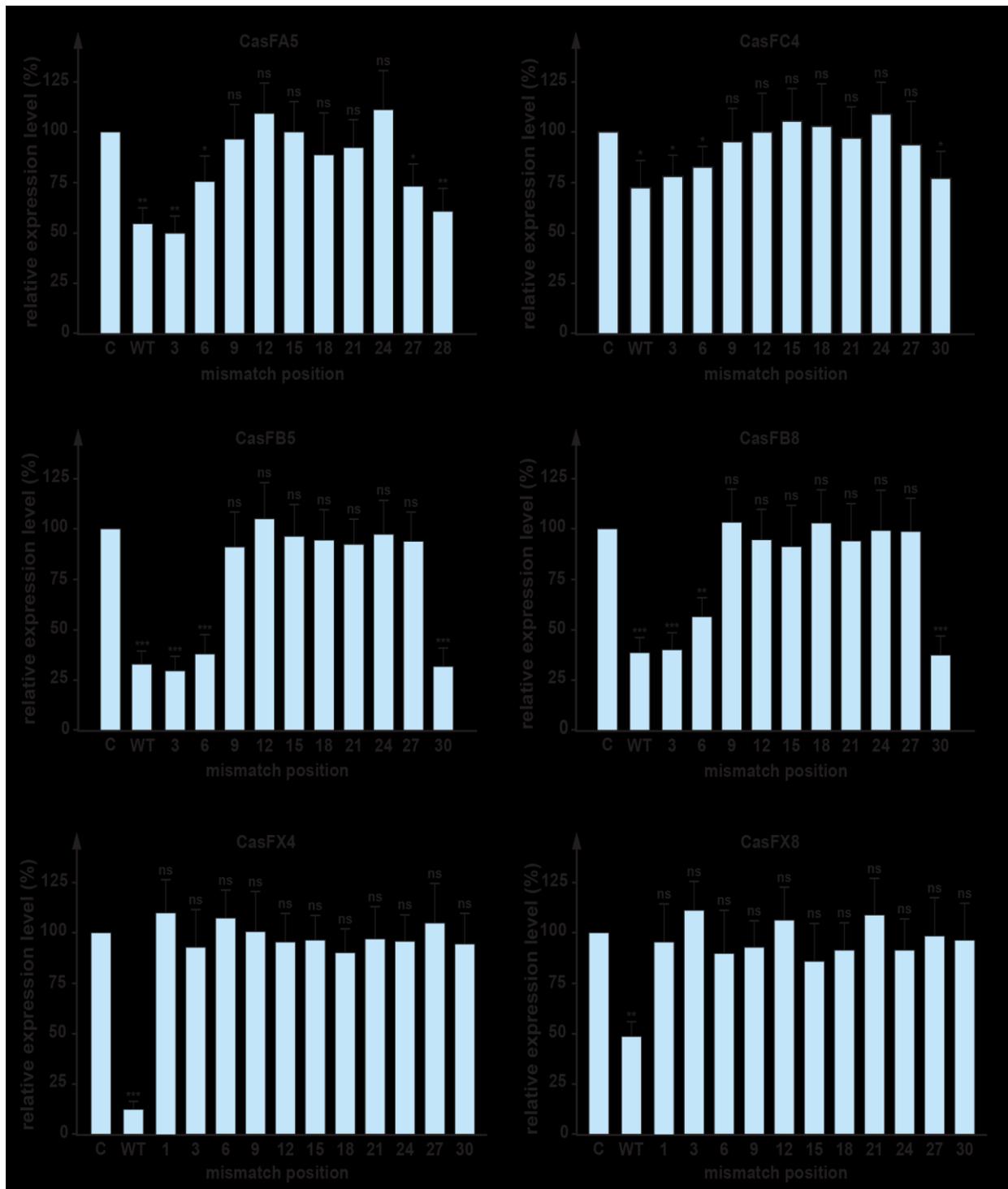
Figure 7.5 Specificity evaluation of *Drosophila* codon-optimized Cas13 variants in Sg4 cells.

(A-F) Relative expression level of eCFP when using different Cas13 variants and crRNAs that carry a range of single mismatches along the eCFP crRNA-2. Data were normalized to samples

treated with blank crRNA (control = C). eCFP expression level in Cas13/ wild-type (WT) crRNA

samples were also included as a reference of changes. * = p-value < 0.05, ** = p-value < 0.01, ***

= p-value < 0.001. Error bars represent standard errors.
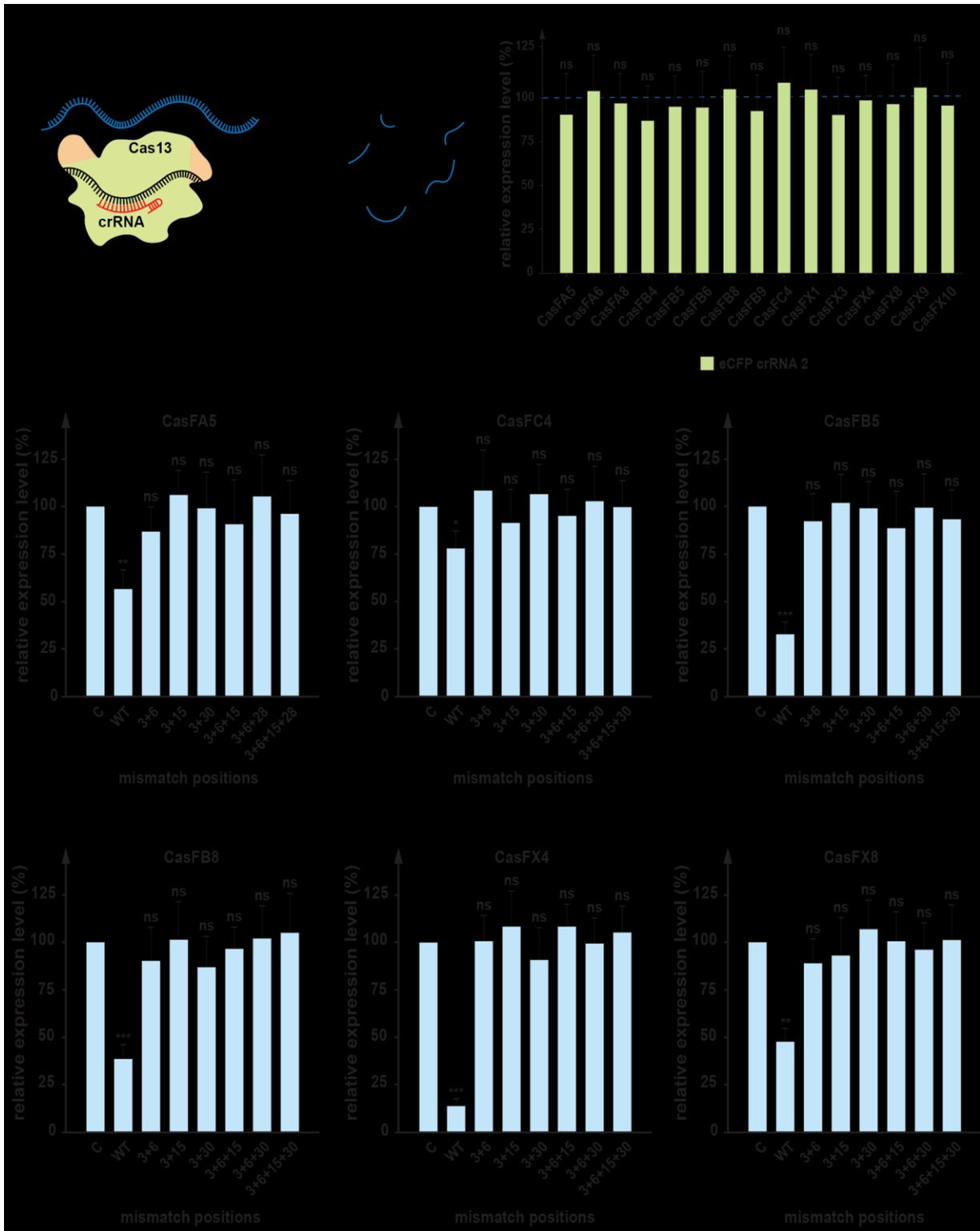
Figure 7.6 Collateral activity and specificity evaluation of Cas13 variants.

(A) Schematic of collateral activity in Cas13. Overall, once a complex is formed with its crRNA, and upon binding to target transcripts, Cas13 will undergo a conformational change, which results in the exposure of two nuclease domains (HEPN). This exposure allows the domains to interact with nearby non-specific transcripts and also results in the degradation of them. **(B)** Relative expression of *DsRed* in samples treated with Cas13/crRNA against eCFP. DsRed is driven by the *act5* promoter. It is believed that the DsRed transcript is present in high amounts, and is more likely to interact with Cas13. Therefore, if the collateral activity is an issue, we would be expected that DsRed transcript levels are affected. Data were normalized to samples treated with blank crRNA (blue dotted line = 1). * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. Error bars represent standard error. **(C-H)** Relative expression levels of eCFP that were exposed to different Cas13 variants and crRNAs carrying different combinations of mismatches along the eCFP crRNA 2. Data were normalized to samples treated with blank crRNA (control = C). eCFP expression level in Cas13/wild-type (WT) crRNA samples were also included as a reference for changes. * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. Error bars represent standard errors.
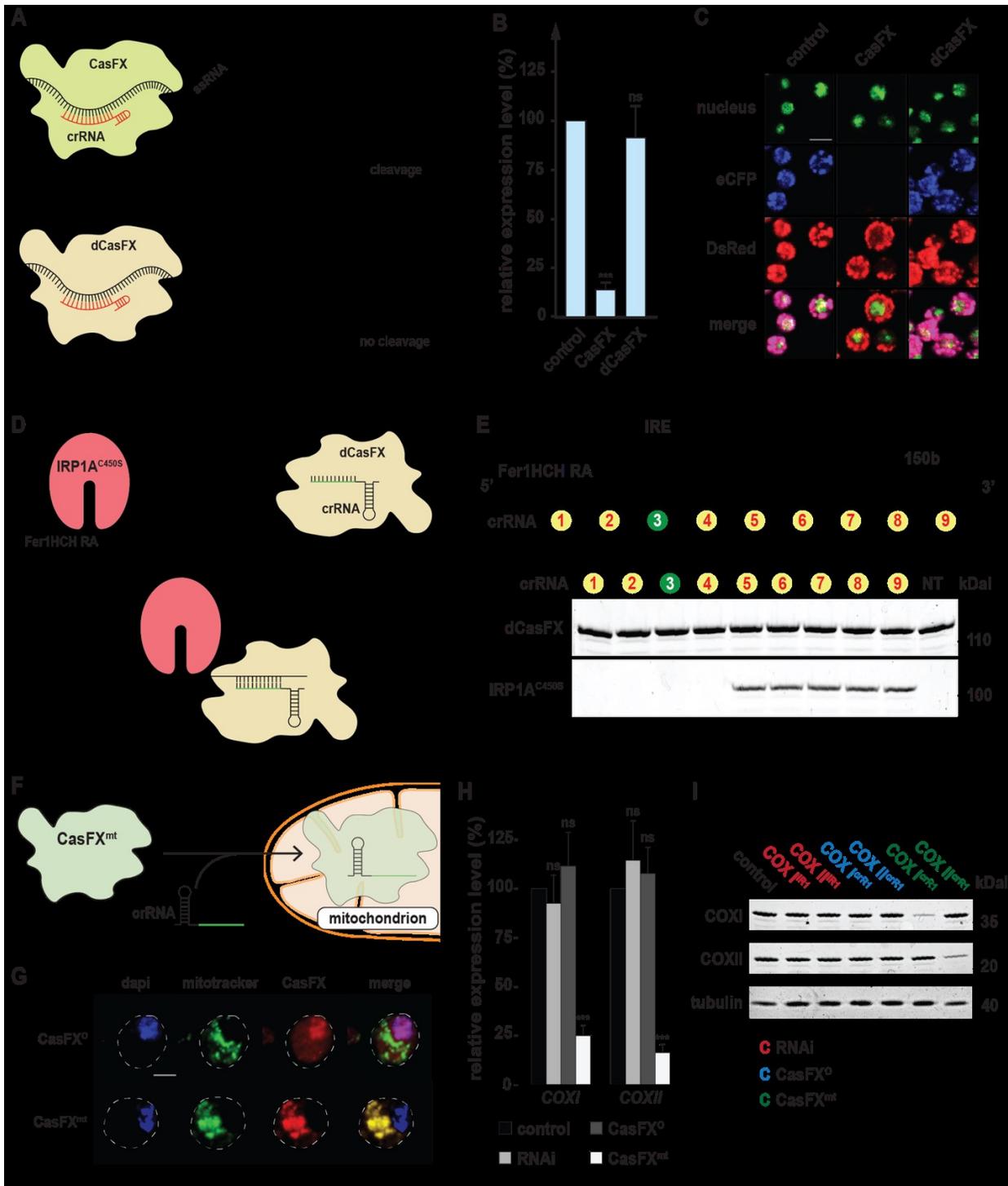
Figure 7.7 Investigation of modified Cas13 for different purposes

(A) Schematic of nuclease-dead CasFX (dCasFX) activity. dCasFX carries quadruple point mutations that abolish its nuclease activity. As a result, the dCasFX/crRNA complex can be

recruited and bind to target transcripts, but it cannot cleave the RNA. **(B)** Evaluation of Cas13 cleavage efficiency of dCasFX compared to wild-type CasFX. qPCR data represent expression levels of eCFP. Data were normalized to samples treated with blank crRNA (control). $*$ = p-value < 0.05, $**$ = p-value < 0.01, $***$ = p-value < 0.001. Error bars represent standard errors. **(C)** eCFP fluorescence when targeted by either CasFX or dCasFX. Nuclei were stained with nuclear green DCS1 (Abcam ab138904). eCFP and DsRed fluorescence were measured using their native fluorescence property without using antibody staining. Scale bar = 50 µm. **(D)** Schematic of dCasFX for the validation of RNA-protein interactions. dCasFX and crRNA targeting Fer1HCH-RA mRNA were transfected together in one sample. On the other hand, Fer1HCH-RA and IRP1A^{C450S}, the constitutively RNA-binding form of IRP1A that interacts with the iron-responsive element (IRE) in the Fer1HCH-RA mRNA, were transformed together in another sample. Two samples were lysed and combined together, followed by immunoprecipitation (IP) of dCasFX to test for the presence of IRP1A in the pull-down assay. **(E)** Western blot showing the IP of dCasFX in combination with different crRNAs along Fer1HCH-RA mRNA and the detection of IRP1A in corresponding samples. **(F)** Functional schematic of CasFX that carries a mitochondrial localization signal (CasFX^{mt}). At the N terminus, CasFX^{mt} is fused with the tim23 mitochondrial signal sequence. Upon binding with crRNA, the complex will localize into mitochondria and target mitochondrial-encoded transcripts. **(G)** Mitochondrial localization of CasFX^{mt}. Nuclei were stained with DAPI (blue) while mitochondria were stained with mitotracker green and CasFX polypeptide was stained with anti-HA antibody (red). Scale bar = 25 µm. **(H)** The relative expression level of mitochondrial-encoded transcripts, *COXI* and *COXII*, targeted by RNAi, CasFX^O, and CasFX^{mt}. Data were normalized to samples treated with no transfected plasmid

(control). * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. **(I)** Western blotting of

COXI and COXII when being targeted by RNAi, CasFX^{O}, and CasFX^{mt}.

A Fer1HCH RA mRNA IP in the absence of IRP1A

B Fer1HCH RA mRNA IP in the presence of IRP1A

C RNAi / CasFX° / CasFX^mt

D RNAi / CasFX° / CasFX^mt

E

F post transfection time (hr)

FREPAIRv2 + crRNA in Sg4_CD

FREPAIRv2 + blank crRNA in Sg4°
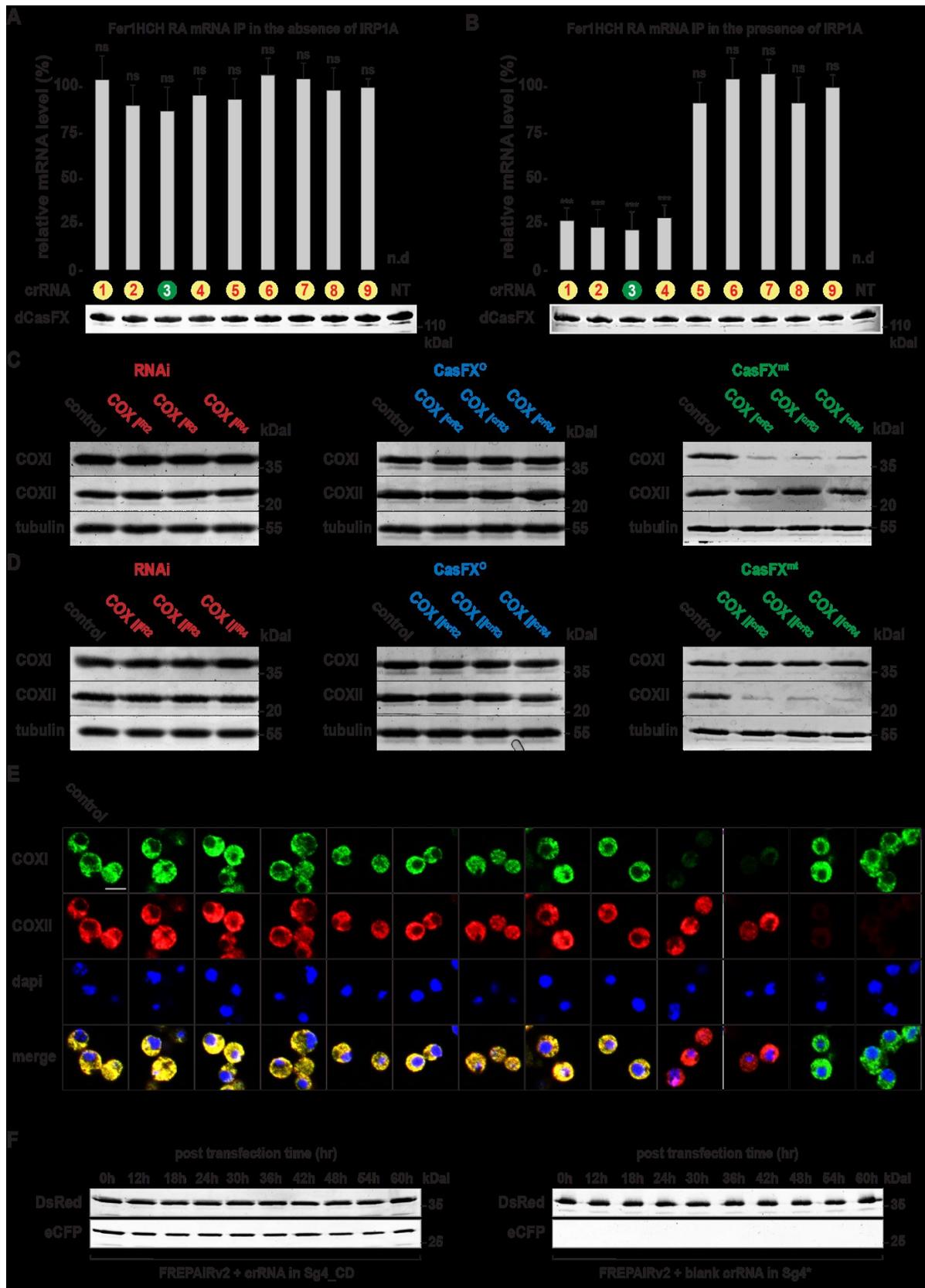
Figure 7.8 Evaluation of modified CasFX for different approaches.

(A-B) Western blotting of COXI and COXII that were targeted by independent RNAi (IR), CasFX^O, and CasFX^{mt}. **(C)** Immunofluorescence of COXI and COXII when being targeted by two independent RNAi, CasFX^O and CasFX^{mt}. Nuclei were stained with DAPI (blue), COXI was stained with anti-COXI antibody (green), and COXII was stained with anti-COXII antibody (red). Scale bar = 50 μm. **(D)** Western blotting of wild-type eCFP or mutant eCFP* with blank crRNA under the same condition as FREPAIRv2.
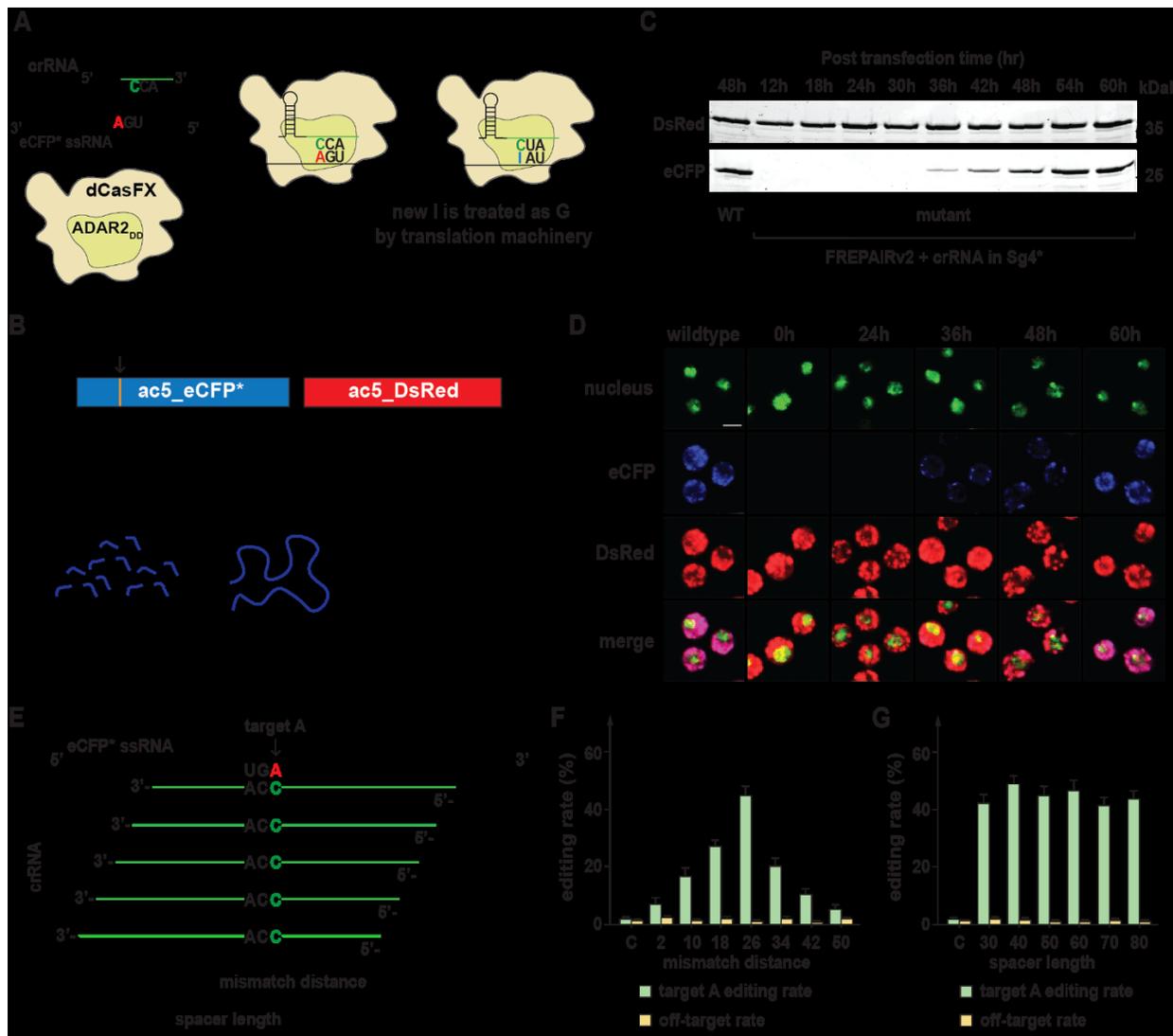
Figure 7.9 Adaptation of the REPAIRv2 system to modify RNA in *Drosophila* Sg4 cell culture.

Schematic for the Drosophila-modified REPAIRv2 system (FREPAIRv2), to modify a mutant eCFP transcript. Mutant eCFP carries an early stop codon that normally encodes Tryptophan at residue 57 (W57*). By generating an A to C mismatch in the crRNA spacer that corresponds to the stop codon, the ADAR2$_{DD}$ domain will change the equivalent adenosine (A) to inosine (I). Inosine will be treated as guanosine by the translation machinery. **(B)** Schematic of FREPAIRv2 outcome. Originally, the mutant eCFP transcript harbors a stop codon at position 57, which will generate a short polypeptide with 56 amino acids. However, once modified by FREPAIRv2, codon

304

57 will be reverted back to wild-type tryptophan, and restore the production of a full-length polypeptide. **(C)** Western blotting monitoring eCFP productions relative to transfection time. **(D)** Fluorescence emitted by eCFP relative to transfection time. Nuclei were stained with nuclear green DCS1 (Abcam ab138905). eCFP and DsRed fluorescence were measured based on their natively emitted fluorescence. Scale bar = 50 μm. **(E)** Schematic of crRNAs that we used for FREPAIRv2. Two criteria we considered for the crRNA design were i) mismatch distance from the first nucleotide and ii) spacer length. **(F)** Editing rate and off-target rate of FREPAIRv2 in relation to mismatch distance when spacer length was kept at a constant 50 nucleotides. **(G)** Editing rate and off-target rates of FREPAIRv2 in relation to spacer lengths when the mismatch distance was kept at the constant position 26.
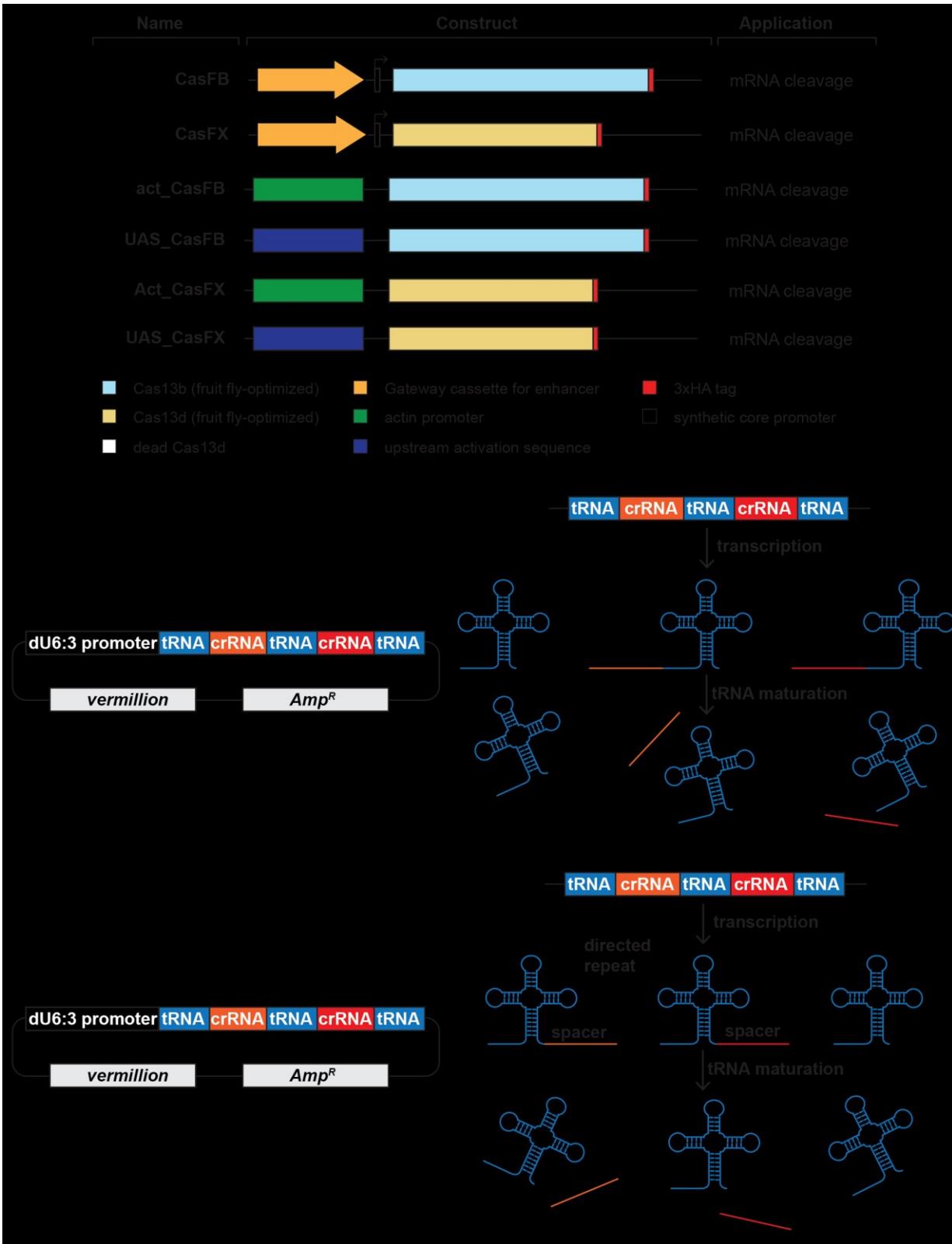
Figure 7.10 CRISPR/Cas13 transgenes and crRNA vector for *in vivo* **RNA targeting.**

(A) Collection of Cas13 transgenes. The general Cas13 collection is composed of a *mini-white* gene as a marker, a PhiC31 integrase-compatible *attB* site, and the *bla* coding sequence to mediate ampicillin resistance and a synthetic core promoter. Shown here are the gateway cassette for an enhancer of interest, and the Cas13 variants. The gateway cassette allows using LR Clonase-based recombination (ThermoFisher) to insert enhancer/promoter regions to drive tissue-specific *Cas9* expression. The *act-Cas13* transgenes drive the expression of *Cas13* via *actin 5C* (*ac5*) promoter while the *UAS-Cas13* transgenes allow tissue-specific expression of *Cas13* via the Gal4/UAS system. In all cases, Cas13 variants were fused with a 3xHA epitope tag at the C-terminal end. **(B)** Collection of Cas13-compatible crRNA vectors. pC13X is compatible with CasF, whereas pC13B is designed for CasFB. Both vectors carry a *vermillion* marker, a PhiC31 integrase-compatible *attB* site, and the *bla* coding sequence to mediate ampicillin resistance. Each vector holds a multiplex tRNA:crRNA cassette to facilitate the cloning of corresponding crRNA via BbsI digestion. The cassette is driven by the ubiquitous *Drosophila* U6:3 promoter (dU6:3) and will be transcribed as a single transcript. Upon tRNA maturation, crRNA will be released and ready to form a complex with Cas13 nuclease.
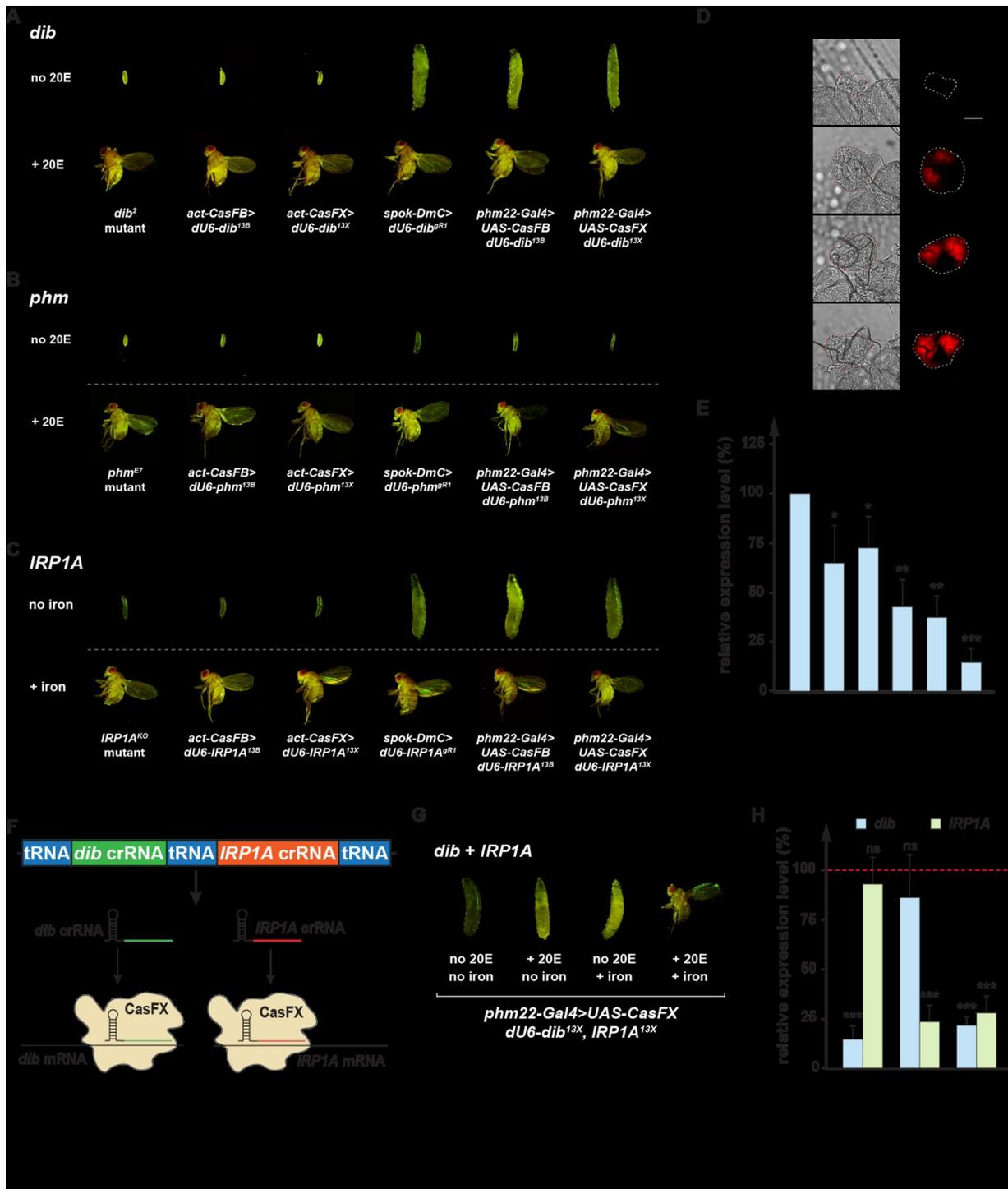
Figure 7.11 Efficiency of *Drosophila* codon-optimized CRISPR/Cas13 *in vivo*.

(A) Comparison of phenotypes from a classic *disembodied* mutant (*dib²*), ubiquitous knock-down

of *dib* via CasFB/*dib*^13B, CasFX/*dib*^13X, prothoracic gland (PG)-specific manipulation via

308

CRISPR/Cas9, or Cas13 of *dib* in the presence or absence of 20OH-ecdysone (20E). **(B)** Comparison of phenotypes from a classic *phantom* mutant (*phm^{E7}*), ubiquitous knock-down of *phm* via CasFB/*phm*^{13B}, CasFX/*phm*^{13X}, PG-specific manipulation via CRISPR/Cas9, or Cas13 of *phm* in the presence or absence of 20OH-ecdysone (20E). **(C)** Comparison of phenotypes from a classic *iron regulatory protein 1* mutant (*IRP1A^{KO}*), ubiquitous knock-down of *IRP1A* via CasFB/*IRP1A*^{13B}, CasFX/*IRP1A*^{13X}, PG-specific manipulation via CRISPR/Cas9, or Cas13 of *IRP1A* in the presence or absence of iron in the diet. **(D)** Porphyria phenotype in PG-specific *IRP1A* knock-down. Scale bar = 250 μm. **(E)** Relative *dib* expression levels in samples representing different PG-specific loss-of-function strategies, including RNAi (IR), dCas9-mediated transcriptional interference, and Cas13 cleavage. Ring glands were dissected from larvae at 42-hour after the L2/L3 molt. * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001. Error bars represent standard error. **(F)** Schematic of dual crRNA targeting both *dib* and *IRP1A* mRNA. **(G)** Comparison of phenotypes from PG-CasFX/*dI*^{13X} in the presence or absence of either 20E, iron, or both.
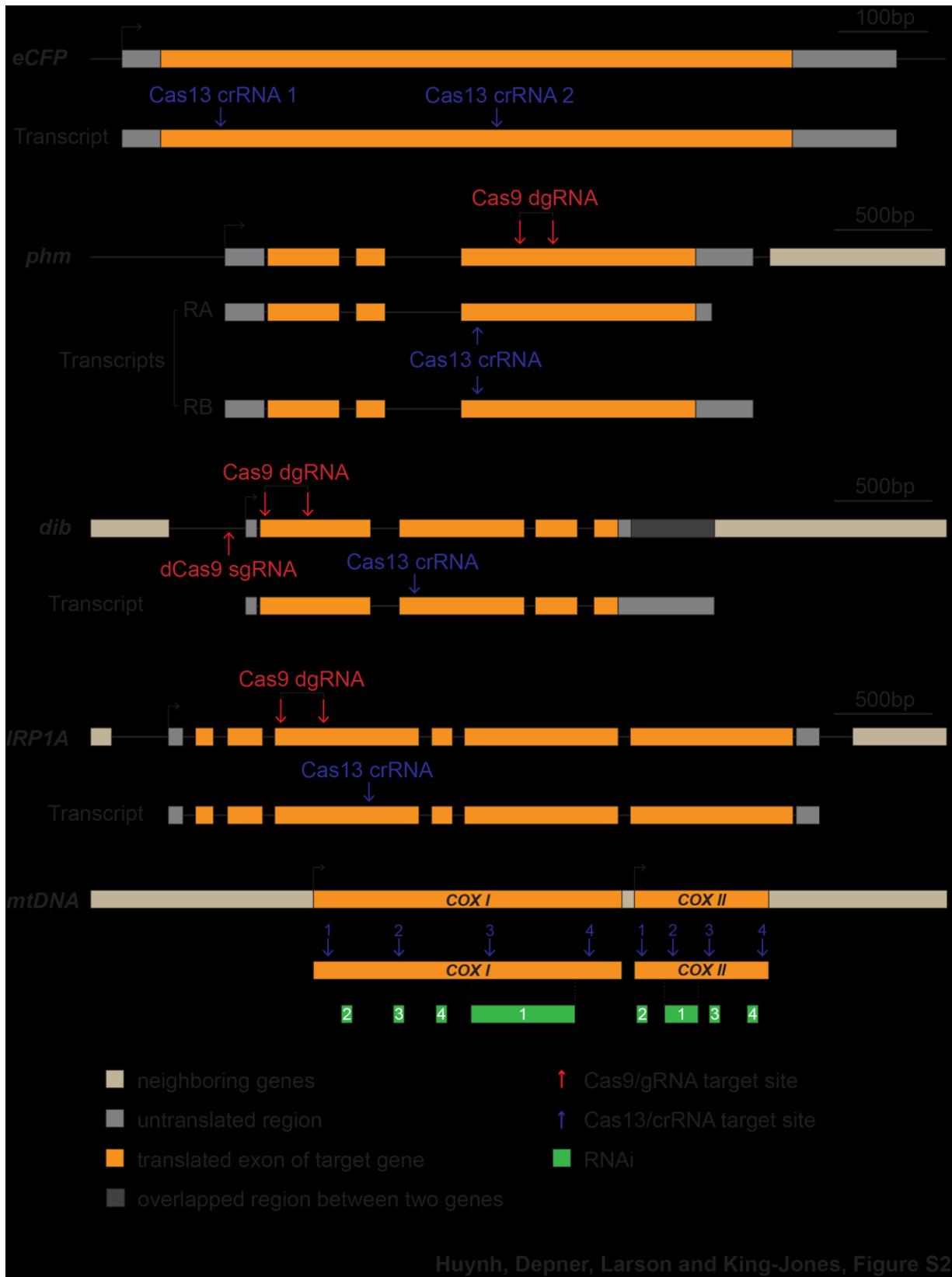
Figure 7.12 Target sites of crRNAs.

For *in vitro* evaluation, we tested *eCFP* expression as well as two mitochondrial-encoded transcripts, *COXI* and *COXII*. For the *in vivo* approach, we tested two genes that encode enzymes acting as ecdysteroid-synthesizing enzymes in the *Drosophila* prothoracic gland (PG), *phantom* (*phm*) and *disembodied* (*dib*) and a gene involved in cellular iron homeostasis, namely *iron regulatory protein 1A (IRP1A).* Shown here are the target sites for crRNA (Cas13-compatible, blue), gRNA (Cas9-compatible, red), and RNAi (green) for transcripts we tested
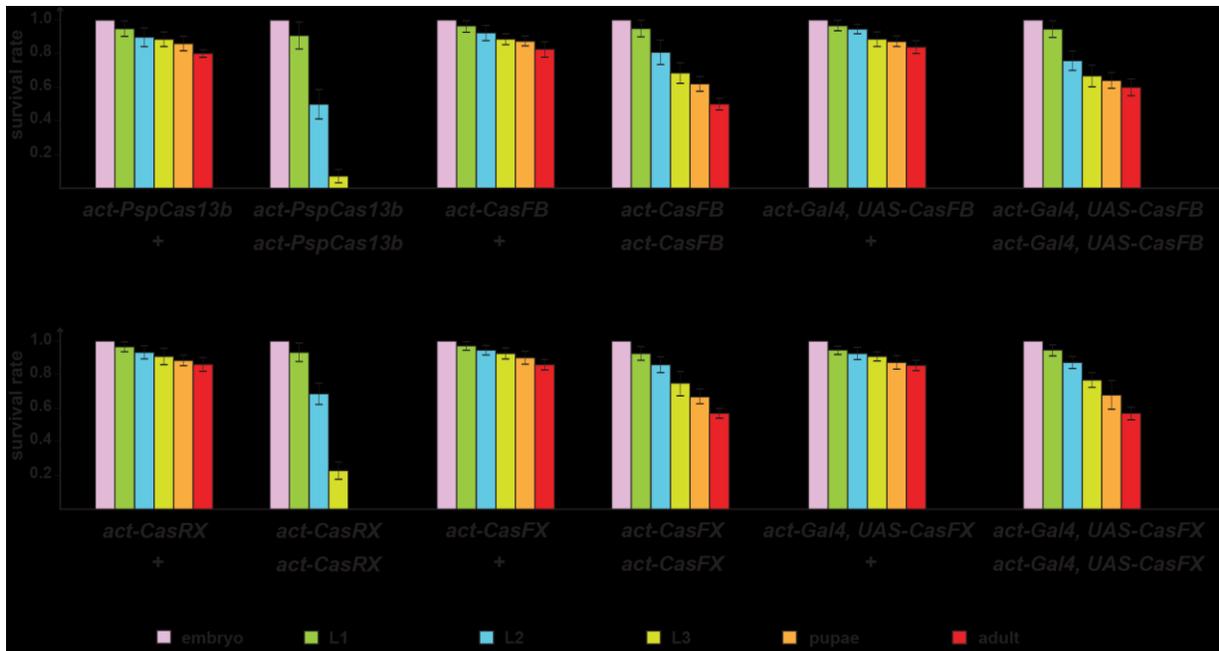
Figure 7.13 Survival rates of transgenic *Drosophila* lines carrying codon-optimized CRISPR/Cas13.

Survival rates of populations heterozygous or homozygous for Cas13 transgenes, including *act-PspCas13b*, *act-CasFB*, *act-Gal4>UAS-CasFB*, *act-CasRX*, *act-CasFX*, and *act-Gal4>UAS-CasFX*. Survival rates of the w^{1118} strain were used as a control. Data were normalized to the number of embryos used in the starting population. Error bars represent standard deviation.

7.6 Tables

Table 7.1 List of plasmids

name	expression system	characteristics
pAc5-CD	Cell culture	Modified from pAFW (DGRC #1111). Expresses dual-reporters eCFP and DsRed under distinct *ac5* promoter, also carries *Neo^R* gene and *attB* integration site.
pAc5-cD	Cell culture	Modified from pAFW (DGRC #1111). Expresses dual-reporters eCFP* and DsRed under distinct *ac5* promoter. eCFP* is eCFP mutant with early stop codon W57*, also carries *Neo^R* gene and *attB* integration site.
pAc5-PhiC31	Cell culture	Modified from pAFW (DGRC #1111). Expresses PhiC31 under *ac5* promoter for integration reaction, also carries *Neo^R* gene.
pC13gR01 collection		
pLwaC13a1cr1_1	Cell culture	Expresses *Leptotrichia wadei* Cas13a (LwaCas13a) as well as crRNA against eCFP at target site 1.
pCasFA1cr1_1	Cell culture	Expresses CasFA-01 as well as crRNA against eCFP at target site 1.
pCasFA2cr1_1	Cell culture	Expresses CasFA-02 as well as crRNA against eCFP at target site 1.
pCasFA3cr1_1	Cell culture	Expresses CasFA-03 as well as crRNA against eCFP at target site 1.
pCasFA4cr1_1	Cell culture	Expresses CasFA-04 as well as crRNA against eCFP at target site 1.
pCasFA5cr1_1	Cell culture	Expresses CasFA-05 as well as crRNA against eCFP at target site 1.
pCasFA6cr1_1	Cell culture	Expresses CasFA-06 as well as crRNA against eCFP at target site 1.
pCasFA7cr1_1	Cell culture	Expresses CasFA-07 as well as crRNA against eCFP at target site 1.
pCasFA8cr1_1	Cell culture	Expresses CasFA-08 as well as crRNA against eCFP at target site 1.
pCasFA9cr1_1	Cell culture	Expresses CasFA-09 as well as crRNA against eCFP at target site 1.
pCasFA10cr1_1	Cell culture	Expresses CasFA-10 as well as crRNA against eCFP at target site 1.
pPspC13bcr1_1	Cell culture	Expresses *Prevotella* sp. P25-125 Cas13b (PspCas13b) as well as crRNA against eCFP at target site 1.
pCasFB1cr1_1	Cell culture	Expresses CasFB-01 as well as crRNA against eCFP at target site 1.
pCasFB2cr1_1	Cell culture	Expresses CasFB-02 as well as crRNA against eCFP at target site 1.
pCasFB3cr1_1	Cell culture	Expresses CasFB-03 as well as crRNA against eCFP at target site 1.
pCasFB4cr1_1	Cell culture	Expresses CasFB-04 as well as crRNA against eCFP at target site 1.
pCasFB5cr1_1	Cell culture	Expresses CasFB-05 as well as crRNA against eCFP at target site 1.

name	expression system	characteristics
pCasFB6cr1_1	Cell culture	Expresses CasFB-06 as well as crRNA against eCFP at target site 1.
pCasFB7cr1_1	Cell culture	Expresses CasFB-07 as well as crRNA against eCFP at target site 1.
pCasFB8cr1_1	Cell culture	Expresses CasFB-08 as well as crRNA against eCFP at target site 1.
pCasFB9cr1_1	Cell culture	Expresses CasFB-09 as well as crRNA against eCFP at target site 1.
pCasFB10cr1_1	Cell culture	Expresses CasFB-10 as well as crRNA against eCFP at target site 1.
pFneC13ccr1_1	Cell culture	Expresses *Fusobacterium perfoetens* Cas13c (FneCas13c) as well as crRNA against eCFP at target site 1.
pCasFC1cr1_1	Cell culture	Expresses CasFC-01 as well as crRNA against eCFP at target site 1.
pCasFC2cr1_1	Cell culture	Expresses CasFC-02 as well as crRNA against eCFP at target site 1.
pCasFC3cr1_1	Cell culture	Expresses CasFC-03 as well as crRNA against eCFP at target site 1.
pCasFC4cr1_1	Cell culture	Expresses CasFC-04 as well as crRNA against eCFP at target site 1.
pCasFC5cr1_1	Cell culture	Expresses CasFC-05 as well as crRNA against eCFP at target site 1.
pCasFC6cr1_1	Cell culture	Expresses CasFC-06 as well as crRNA against eCFP at target site 1.
pCasFC7cr1_1	Cell culture	Expresses CasFC-07 as well as crRNA against eCFP at target site 1.
pCasFC8cr1_1	Cell culture	Expresses CasFC-08 as well as crRNA against eCFP at target site 1.
pCasFC9cr1_1	Cell culture	Expresses CasFC-09 as well as crRNA against eCFP at target site 1.
pCasFC10cr1_1	Cell culture	Expresses CasFC-10 as well as crRNA against eCFP at target site 1.
pCasRXcr1_1	Cell culture	Expresses *Ruminococcus flavefaciens* XPD3002 Cas13d (CasRX) as well as crRNA against eCFP at target site 1.
pCasFX1cr1_1	Cell culture	Expresses CasFX-01 as well as crRNA against eCFP at target site 1.
pCasFX2cr1_1	Cell culture	Expresses CasFX-02 as well as crRNA against eCFP at target site 1.
pCasFX3cr1_1	Cell culture	Expresses CasFX-03 as well as crRNA against eCFP at target site 1.
pCasFX4cr1_1	Cell culture	Expresses CasFX-04 as well as crRNA against eCFP at target site 1.
pCasFX5cr1_1	Cell culture	Expresses CasFX-05 as well as crRNA against eCFP at target site 1.
pCasFX6cr1_1	Cell culture	Expresses CasFX-06 as well as crRNA against eCFP at target site 1.
pCasFX7cr1_1	Cell culture	Expresses CasFX-07 as well as crRNA against eCFP at target site 1.
pCasFX8cr1_1	Cell culture	Expresses CasFX-08 as well as crRNA against eCFP at target site 1.
pCasFX9cr1_1	Cell culture	Expresses CasFX-09 as well as crRNA against eCFP at target site 1.
pCasFX10cr1_1	Cell culture	Expresses CasFX-10 as well as crRNA against eCFP at target site 1.

name	expression system	characteristics
pLwaC13a1cr1_2	Cell culture	Expresses *Leptotrichia wadei* Cas13a (LwaCas13a) as well as crRNA against eCFP at target site 2.
pCasFA1cr1_2	Cell culture	Expresses CasFA-01 as well as crRNA against eCFP at target site 2.
pCasFA2cr1_2	Cell culture	Expresses CasFA-02 as well as crRNA against eCFP at target site 2.
pCasFA3cr1_2	Cell culture	Expresses CasFA-03 as well as crRNA against eCFP at target site 2.
pCasFA4cr1_2	Cell culture	Expresses CasFA-04 as well as crRNA against eCFP at target site 2.
pCasFA5cr1_2	Cell culture	Expresses CasFA-05 as well as crRNA against eCFP at target site 2.
pCasFA6cr1_2	Cell culture	Expresses CasFA-06 as well as crRNA against eCFP at target site 2.
pCasFA7cr1_2	Cell culture	Expresses CasFA-07 as well as crRNA against eCFP at target site 2.
pCasFA8cr1_2	Cell culture	Expresses CasFA-08 as well as crRNA against eCFP at target site 2.
pCasFA9cr1_2	Cell culture	Expresses CasFA-09 as well as crRNA against eCFP at target site 2.
pCasFA10cr1_2	Cell culture	Expresses CasFA-10 as well as crRNA against eCFP at target site 2.
pPspC13bcr1_1	Cell culture	Expresses *Prevotella* sp. P25-125 Cas13b (PspCas13b) as well as crRNA against eCFP at target site 2.
pCasFB1cr1_2	Cell culture	Expresses CasFB-01 as well as crRNA against eCFP at target site 2.
pCasFB2cr1_2	Cell culture	Expresses CasFB-02 as well as crRNA against eCFP at target site 2.
pCasFB3cr1_2	Cell culture	Expresses CasFB-03 as well as crRNA against eCFP at target site 2.
pCasFB4cr1_2	Cell culture	Expresses CasFB-04 as well as crRNA against eCFP at target site 2.
pCasFB5cr1_2	Cell culture	Expresses CasFB-05 as well as crRNA against eCFP at target site 2.
pCasFB6cr1_2	Cell culture	Expresses CasFB-06 as well as crRNA against eCFP at target site 2.
pCasFB7cr1_2	Cell culture	Expresses CasFB-07 as well as crRNA against eCFP at target site 2.
pCasFB8cr1_2	Cell culture	Expresses CasFB-08 as well as crRNA against eCFP at target site 2.
pCasFB9cr1_2	Cell culture	Expresses CasFB-09 as well as crRNA against eCFP at target site 2.
pCasFB10cr1_2	Cell culture	Expresses CasFB-10 as well as crRNA against eCFP at target site 2.
pFneC13ccr1_1	Cell culture	Expresses *Fusobacterium perfoetens* Cas13c (FneCas13c) as well as crRNA against eCFP at target site 2.
pCasFC1cr1_2	Cell culture	Expresses CasFC-01 as well as crRNA against eCFP at target site 2.
pCasFC2cr1_2	Cell culture	Expresses CasFC-02 as well as crRNA against eCFP at target site 2.
pCasFC3cr1_2	Cell culture	Expresses CasFC-03 as well as crRNA against eCFP at target site 2.
pCasFC4cr1_2	Cell culture	Expresses CasFC-04 as well as crRNA against eCFP at target site 2.
pCasFC5cr1_2	Cell culture	Expresses CasFC-05 as well as crRNA against eCFP at target site 2.

name	expression system	characteristics
pCasFC6cr1_2	Cell culture	Expresses CasFC-06 as well as crRNA against eCFP at target site 2.
pCasFC7cr1_2	Cell culture	Expresses CasFC-07 as well as crRNA against eCFP at target site 2.
pCasFC8cr1_2	Cell culture	Expresses CasFC-08 as well as crRNA against eCFP at target site 2.
pCasFC9cr1_2	Cell culture	Expresses CasFC-09 as well as crRNA against eCFP at target site 2.
pCasFC10cr1_2	Cell culture	Expresses CasFC-10 as well as crRNA against eCFP at target site 2.
pCasRXcr1_2	Cell culture	Expresses *Ruminococcus flavefaciens* XPD3002 Cas13d (CasRX) as well as crRNA against eCFP at target site 2.
pCasFX1cr1_2	Cell culture	Expresses CasFX-01 as well as crRNA against eCFP at target site 2.
pCasFX2cr1_2	Cell culture	Expresses CasFX-02 as well as crRNA against eCFP at target site 2.
pCasFX3cr1_2	Cell culture	Expresses CasFX-03 as well as crRNA against eCFP at target site 2.
pCasFX4cr1_2	Cell culture	Expresses CasFX-04 as well as crRNA against eCFP at target site 2.
pCasFX5cr1_2	Cell culture	Expresses CasFX-05 as well as crRNA against eCFP at target site 2.
pCasFX6cr1_2	Cell culture	Expresses CasFX-06 as well as crRNA against eCFP at target site 2.
pCasFX7cr1_2	Cell culture	Expresses CasFX-07 as well as crRNA against eCFP at target site 2.
pCasFX8cr1_2	Cell culture	Expresses CasFX-08 as well as crRNA against eCFP at target site 2.
pCasFX9cr1_2	Cell culture	Expresses CasFX-09 as well as crRNA against eCFP at target site 2.
pCasFX10cr1_2	Cell culture	Expresses CasFX-10 as well as crRNA against eCFP at target site 2.
pLwaC13a1cr1_0	Cell culture	Expresses *Leptotrichia wadei* Cas13a (LwaCas13a) as well as no-targeting crRNA.
pCasFA1cr1_0	Cell culture	Expresses CasFA-01 as well as no-targeting crRNA.
pCasFA2cr1_0	Cell culture	Expresses CasFA-02 as well as no-targeting crRNA.
pCasFA3cr1_0	Cell culture	Expresses CasFA-03 as well as no-targeting crRNA.
pCasFA4cr1_0	Cell culture	Expresses CasFA-04 as well as no-targeting crRNA.
pCasFA5cr1_0	Cell culture	Expresses CasFA-05 as well as no-targeting crRNA.
pCasFA6cr1_0	Cell culture	Expresses CasFA-06 as well as no-targeting crRNA.
pCasFA7cr1_0	Cell culture	Expresses CasFA-07 as well as no-targeting crRNA.
pCasFA8cr1_0	Cell culture	Expresses CasFA-08 as well as no-targeting crRNA.
pCasFA9cr1_0	Cell culture	Expresses CasFA-09 as well as no-targeting crRNA.
pCasFA10cr1_0	Cell culture	Expresses CasFA-10 as well as no-targeting crRNA.
pPspC13bcr1_0	Cell culture	Expresses *Prevotella* sp. P25-125 Cas13b (PspCas13b) as well as no-targeting crRNA.

name	expression system	characteristics
pCasFB1cr1_0	Cell culture	Expresses CasFB-01 as well as no-targeting crRNA.
pCasFB2cr1_0	Cell culture	Expresses CasFB-02 as well as no-targeting crRNA.
pCasFB3cr1_0	Cell culture	Expresses CasFB-03 as well as no-targeting crRNA.
pCasFB4cr1_0	Cell culture	Expresses CasFB-04 as well as no-targeting crRNA.
pCasFB5cr1_0	Cell culture	Expresses CasFB-05 as well as no-targeting crRNA.
pCasFB6cr1_0	Cell culture	Expresses CasFB-06 as well as no-targeting crRNA.
pCasFB7cr1_0	Cell culture	Expresses CasFB-07 as well as no-targeting crRNA.
pCasFB8cr1_0	Cell culture	Expresses CasFB-08 as well as no-targeting crRNA.
pCasFB9cr1_0	Cell culture	Expresses CasFB-09 as well as no-targeting crRNA.
pCasFB10cr1_0	Cell culture	Expresses CasFB-10 as well as no-targeting crRNA.
pFpeC13ccr1_0	Cell culture	Expresses *Fusobacterium perfoetens* Cas13c (FpeCas13c) as well as no-targeting crRNA.
pCasFC1cr1_0	Cell culture	Expresses CasFC-01 as well as no-targeting crRNA.
pCasFC2cr1_0	Cell culture	Expresses CasFC-02 as well as no-targeting crRNA.
pCasFC3cr1_0	Cell culture	Expresses CasFC-03 as well as no-targeting crRNA.
pCasFC4cr1_0	Cell culture	Expresses CasFC-04 as well as no-targeting crRNA.
pCasFC5cr1_0	Cell culture	Expresses CasFC-05 as well as no-targeting crRNA.
pCasFC6cr1_0	Cell culture	Expresses CasFC-06 as well as no-targeting crRNA.
pCasFC7cr1_0	Cell culture	Expresses CasFC-07 as well as no-targeting crRNA.
pCasFC8cr1_0	Cell culture	Expresses CasFC-08 as well as no-targeting crRNA.
pCasFC9cr1_0	Cell culture	Expresses CasFC-09 as well as no-targeting crRNA.
pCasFC10cr1_0	Cell culture	Expresses CasFC-10 as well as no-targeting crRNA.
pCasRXcr1_0	Cell culture	Expresses *Ruminococcus flavefaciens* XPD3002 Cas13d (CasRX) as well as no-targeting crRNA.
pCasFX1cr1_0	Cell culture	Expresses CasFX-01 as well as no-targeting crRNA.
pCasFX2cr1_0	Cell culture	Expresses CasFX-02 as well as no-targeting crRNA.
pCasFX3cr1_0	Cell culture	Expresses CasFX-03 as well as no-targeting crRNA.
pCasFX4cr1_0	Cell culture	Expresses CasFX-04 as well as no-targeting crRNA.
pCasFX5cr1_0	Cell culture	Expresses CasFX-05 as well as no-targeting crRNA.
pCasFX6cr1_0	Cell culture	Expresses CasFX-06 as well as no-targeting crRNA.
pCasFX7cr1_0	Cell culture	Expresses CasFX-07 as well as no-targeting crRNA.

name	expression system	characteristics
pCasFX8cr1_0	Cell culture	Expresses CasFX-08 as well as no-targeting crRNA.
pCasFX9cr1_0	Cell culture	Expresses CasFX-09 as well as no-targeting crRNA.
pCasFX10cr1_0	Cell culture	Expresses CasFX-10 as well as no-targeting crRNA.
Plasmids for specificity evaluation		
pCasFA5_cr1-a3	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFA5_cr1-a6	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFA5_cr1-a9	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 9th nucleotide.
pCasFA5_cr1-a12	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 12th nucleotide.
pCasFA5_cr1-a15	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFA5_cr1-a18	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFA5_cr1-a21	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFA5_cr1-a24	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFA5_cr1-a27	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 27th nucleotide.
pCasFA5_cr1-a28	Cell culture	Expresses CasFA-05 as well as crRNA with single mismatch at the 28th nucleotide.
pCasFC4_cr1-a3	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFC4_cr1-a6	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFC4_cr1-a9	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 9th nucleotide.

name	expression system	characteristics
pCasFC4_cr1-a12	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 12th nucleotide.
pCasFC4_cr1-a15	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFC4_cr1-a18	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFC4_cr1-a21	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFC4_cr1-a24	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFC4_cr1-a27	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 27th nucleotide.
pCasFC4_cr1-a30	Cell culture	Expresses CasFC-04 as well as crRNA with single mismatch at the 30th nucleotide.
pCasFB5_cr1-a3	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFB5_cr1-a6	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFB5_cr1-a9	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 9th nucleotide.
pCasFB5_cr1-a12	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 12th nucleotide.
pCasFB5_cr1-a15	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFB5_cr1-a18	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFB5_cr1-a21	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFB5_cr1-a24	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFB5_cr1-a27	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 27th nucleotide.

name	expression system	characteristics
pCasFB5_cr1-a30	Cell culture	Expresses CasFB-05 as well as crRNA with single mismatch at the 30th nucleotide.
pCasFB8_cr1-a3	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFB8_cr1-a6	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFB8_cr1-a9	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 9th nucleotide.
pCasFB8_cr1-a12	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 12th nucleotide.
pCasFB8_cr1-a15	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFB8_cr1-a18	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFB8_cr1-a21	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFB8_cr1-a24	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFB8_cr1-a27	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 27th nucleotide.
pCasFB8_cr1-a30	Cell culture	Expresses CasFB-08 as well as crRNA with single mismatch at the 30th nucleotide.
pCasFX4_cr1-a1	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 1st nucleotide.
pCasFX4_cr1-a3	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFX4_cr1-a6	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFX4_cr1-a9	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 9th nucleotide.
pCasFX4_cr1-a12	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 12th nucleotide.

name	expression system	characteristics
pCasFX4_cr1-a15	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFX4_cr1-a18	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFX4_cr1-a21	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFX4_cr1-a24	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFX4_cr1-a27	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 27th nucleotide.
pCasFX4_cr1-a30	Cell culture	Expresses CasFX-04 as well as crRNA with single mismatch at the 30th nucleotide.
pCasFX_cr1-a1	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 1st nucleotide.
pCasFX8_cr1-a3	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 3rd nucleotide.
pCasFX8_cr1-a6	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 6th nucleotide.
pCasFX8_cr1-a9	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 9th nucleotide.
pCasFX8_cr1-a12	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 12th nucleotide.
pCasFX8_cr1-a15	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 15th nucleotide.
pCasFX8_cr1-a18	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 18th nucleotide.
pCasFX8_cr1-a21	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 21st nucleotide.
pCasFX8_cr1-a24	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 24th nucleotide.
pCasFX8_cr1-a27	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 27th nucleotide.

name	expression system	characteristics
pCasFX8_cr1-a30	Cell culture	Expresses CasFX-08 as well as crRNA with single mismatch at the 30th nucleotide.
pCasFA5_cr1-a3+6	Cell culture	Expresses CasFA-05 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.
pCasFA5_cr1-a3+15	Cell culture	Expresses CasFA-05 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFA5_cr1-a3+28	Cell culture	Expresses CasFA-05 as well as crRNA with dual mismatches at the 3rd and 28th nucleotides.
pCasFA5_cr1-a3+6+15	Cell culture	Expresses CasFA-05 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFA5_cr1-a3+6+28	Cell culture	Expresses CasFA-05 as well as crRNA with triple mismatches at the 3rd, 6th and 28th nucleotides.
pCasFA5_cr1-a3+6+15+28	Cell culture	Expresses CasFA-05 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 28th nucleotides.
pCasFC4_cr1-a3+6	Cell culture	Expresses CasFC-04 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.
pCasFC4_cr1-a3+15	Cell culture	Expresses CasFC-04 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFC4_cr1-a3+30	Cell culture	Expresses CasFC-04 as well as crRNA with dual mismatches at the 3rd and 30th nucleotides.
pCasFC4_cr1-a3+6+15	Cell culture	Expresses CasFC-04 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFC4_cr1-a3+6+30	Cell culture	Expresses CasFC-04 as well as crRNA with triple mismatches at the 3rd, 6th and 30th nucleotides.
pCasFC4_cr1-a3+6+15+30	Cell culture	Expresses CasFC-04 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 30th nucleotides.
pCasFB5_cr1-a3+6	Cell culture	Expresses CasFB-05 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.
pCasFB5_cr1-a3+15	Cell culture	Expresses CasFB-05 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFB5_cr1-a3+30	Cell culture	Expresses CasFB-05 as well as crRNA with dual mismatches at the 3rd and 30th nucleotides.

name	expression system	characteristics
pCasFB5_cr1-a3+6+15	Cell culture	Expresses CasFB-05 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFB5_cr1-a3+6+30	Cell culture	Expresses CasFB-05 as well as crRNA with triple mismatches at the 3rd, 6th and 30th nucleotides.
pCasFB5_cr1-a3+6+15+30	Cell culture	Expresses CasFB-05 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 30th nucleotides.
pCasFB8_cr1-a3+6	Cell culture	Expresses CasFB-08 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.
pCasFB8_cr1-a3+15	Cell culture	Expresses CasFB-08 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFB8_cr1-a3+30	Cell culture	Expresses CasFB-08 as well as crRNA with dual mismatches at the 3rd and 30th nucleotides.
pCasFB8_cr1-a3+6+15	Cell culture	Expresses CasFB-08 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFB8_cr1-a3+6+30	Cell culture	Expresses CasFB-08 as well as crRNA with triple mismatches at the 3rd, 6th and 30th nucleotides.
pCasFB8_cr1-a3+6+15+30	Cell culture	Expresses CasFB-08 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 30th nucleotides.
pCasFX4_cr1-a3+6	Cell culture	Expresses CasFX-04 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.
pCasFX4_cr1-a3+15	Cell culture	Expresses CasFX-04 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFX4_cr1-a3+30	Cell culture	Expresses CasFX-04 as well as crRNA with dual mismatches at the 3rd and 30th nucleotides.
pCasFX4_cr1-a3+6+15	Cell culture	Expresses CasFX-04 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFX4_cr1-a3+6+30	Cell culture	Expresses CasFX-04 as well as crRNA with triple mismatches at the 3rd, 6th and 30th nucleotides.
pCasFX4_cr1-a3+6+15+30	Cell culture	Expresses CasFX-04 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 30th nucleotides.
pCasFX8_cr1-a3+6	Cell culture	Expresses CasFX-08 as well as crRNA with dual mismatches at the 3rd and 6th nucleotides.

name	expression system	characteristics
pCasFX8_cr1-a3+15	Cell culture	Expresses CasFX-08 as well as crRNA with dual mismatches at the 3rd and 15th nucleotides.
pCasFX8_cr1-a3+30	Cell culture	Expresses CasFX-08 as well as crRNA with dual mismatches at the 3rd and 30th nucleotides.
pCasFX8_cr1-a3+6+15	Cell culture	Expresses CasFX-08 as well as crRNA with triple mismatches at the 3rd, 6th and 15th nucleotides.
pCasFX8_cr1-a3+6+30	Cell culture	Expresses CasFX-08 as well as crRNA with triple mismatches at the 3rd, 6th and 30th nucleotides.
pCasFX8_cr1-a3+6+15+30	Cell culture	Expresses CasFX-08 as well as crRNA with quadruple mismatches at the 3rd, 6th, 15th and 30th nucleotides.
Plasmids for nuclease dead CasFX		
pdCasFX_cr1	Cell culture	Expresses nuclease-dead CasFX_04 (dCasFX) as well as crRNA against eCFP at target site 1
pdCasFX_cr1C	Cell culture	Expresses dCasFX as well as crRNA cassette with no target site.
Plasmids for transcript tracking		
pdCasFX_crA1	Cell culture	Expresses dCasFX as well as crRNA targeting site 300 bp upstream of IRE on Fer1HCH RA mRNA.
pdCasFX_crA2	Cell culture	Expresses dCasFX as well as crRNA targeting site 150 bp upstream of IRE on Fer1HCH RA mRNA.
pdCasFX_crA3	Cell culture	Expresses dCasFX as well as crRNA targeting IRE on Fer1HCH RA mRNA.
pdCasFX_crA4	Cell culture	Expresses dCasFX as well as crRNA targeting site 150 bp downstream of IRE on Fer1HCH RA.
pdCasFX_crA5	Cell culture	Expresses dCasFX as well as crRNA targeting site 300bp downstream of IRE on Fer1HCH RA.
pdCasFX_crA6	Cell culture	Expresses dCasFX as well as crRNA targeting site 450bp downstream of IRE on Fer1HCH RA.
pdCasFX_crA7	Cell culture	Expresses dCasFX as well as crRNA targeting site 600bp downstream of IRE on Fer1HCH RA.
pdCasFX_crA8	Cell culture	Expresses dCasFX as well as crRNA targeting site 750bp downstream of IRE on Fer1HCH RA.

name	expression system	characteristics
pdCasFX_crA9	Cell culture	Expresses dCasFX as well as crRNA targeting site 900 bp downstream of IRE on Fer1HCH RA.
pdCasFX_crAC	Cell culture	Expresses dCasFX as well as crRNA cassette with no target site.
pAFW-IRP1A^{C450S}-Fer1HCH RA	Cell culture	Expresses 3xFlag-tagged IRP1A^{C450S} as well as Fer1HCH RA.
Plasmids for mitochondrial-encoded RNA target		
pCasFX^{MT}_cr1C	Cell culture	Expresses CasFX fused with mitochondrial signaling sequence (CasFX^{MT}) and crRNA cassette with no target site.
pCasFX^{MT}_cr1COXI-1	Cell culture	Expresses CasFX^{MT} and crRNA1 targeting *COXI* transcript
pCasFX^{MT}_cr1COXI-2	Cell culture	Expresses CasFX^{MT} and crRNA2 targeting *COXI* transcript
pCasFX^{MT}_cr1COXI-3	Cell culture	Expresses CasFX^{MT} and crRNA3 targeting *COXI* transcript
pCasFX^{MT}_cr1COXI-4	Cell culture	Expresses CasFX^{MT} and crRNA4 targeting *COXI* transcript
pCasFX^{MT}_cr1COXII-1	Cell culture	Expresses CasFX^{MT} and crRNA1 targeting *COXII* transcript
pCasFX^{MT}_cr1COXII-2	Cell culture	Expresses CasFX^{MT} and crRNA2 targeting *COXII* transcript
pCasFX^{MT}_cr1COXII-3	Cell culture	Expresses CasFX^{MT} and crRNA3 targeting *COXII* transcript
pCasFX^{MT}_cr1COXII-3	Cell culture	Expresses CasFX^{MT} and crRNA3 targeting *COXII* transcript
Plasmids for RNA editing application		
pFREPAIRv2-cr50*2	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 2 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr50*10	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 10 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.

name	expression system	characteristics
pFREPAIRv2-cr50*18	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 18 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr50*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr50*34	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 34 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr50*42	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 42 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv50-cr*50	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 50-nt long spacer and a mismatch at location 50 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr30*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 30-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr40*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 40-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr60*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 60-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr70*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 70-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
pFREPAIRv2-cr80*26	Cell culture	Expresses the dCasFX fused with ADAR2$_{DD}$ (FREPAIRv2) together with crRNA carrying 80-nt long spacer and a mismatch at location 26 from the first nucleotide against W57* on eCFP* in Sg4-cD cell line.
General gateway-Cas13 variants plasmids for transgenic fly		

name	expression system	characteristics
CasFB	Transgenic fly	Expresses the 3xHA tagged CasFB-05 for RNA cleavage, carries embedded attB integration site.
CasFX	Transgenic fly	Expresses the 3xHA tagged CasFX-04 for RNA cleavage, carries embedded attB integration site.
pC13X	Transgenic fly	Expresses the CasFX-compatible crRNA under control of dU6:3 promoter, carries embedded attB integration site.
pC13B	Transgenic fly	Expresses the CasFB-compatible crRNA under control of dU6:3 promoter, carries embedded attB integration site.

Table 7.2 Primers used to generate cell culture constructs and transgenes

primer name	primer sequences (5' - 3')
pC13acr01	
attB1 Lwa FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAAGTGACCAAGGTCGAC
attB2 Lwa RP	CCCACTTTGTACAAGAAAGCTGGGTTTATTCCAGGGCCTTGTACTCGAACATGAC
attB1 CasFA-01 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTGGATGGCATC
attB2 CasFA-01 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATC
attB1 CasFA-02 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATC
attB2 CasFA-02 RP	CCCACTTTGTACAAGAAAGCTGGGTTTATTCCAGGGCCTTGTACTCGAACATGAC
attB1 CasFA-03 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAAGTGACCAAGGTCGACGGCATCAG
attB2 CasFA-03 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGAC
attB1 CasFA-04 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATCAG
attB2 CasFA-04 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTC
attB1 CasFA-05 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATCAGCCAC
attB2 CasFA-05 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGAC
attB1 CasFA-06 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATC
attB2 CasFA-06 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTCAC
attB1 CasFA-07 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATCAGCCAC
attB2 CasFA-07 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTC
attB1 CasFA-08 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATC
attB2 CasFA-08 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTC
attB1 CasFA-09 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATCAGCCAC
attB2 CasFA-09 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTC

primer name	primer sequences (5' - 3')
attB1 CasFA-10 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAAGGTGACCAAGGTCGACGGCATCAGCCAC
attB2 CasFA-10 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTCCAGGGCCTTGTACTCGAACATGACCTTC
pC13bcr01	
attB1 Psp FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAAAAC
attB2 Psp RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCCCCAAAG
attB1 CasFB-01 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCCCTGGTGGAGAAC
attB2 CasFB-01 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCGAAG
attB1 CasFB-02 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAAAAC
attB2 CasFB-02 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-03 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-03 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-04 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-04 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-05 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-05 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-06 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-06 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCCCCAAAG
attB1 CasFB-07 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-07 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-08 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-08 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-09 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-09 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
attB1 CasFB-10 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAACATCCCCGCTCTGGTGGAGAAC
attB2 CasFB-10 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACTTCATGATGGCGTACTCGCCAAAG
pC13ccr01	
attB1 Ppe FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGAAAACCAAATAGAAGTTC
attB2 Ppe RP	CCCACTTTGTACAAGAAAGCTGGGTTTAAAGGGTTATTTTTAAGTTAAAGC
attB1 CasFC-01 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCCAACCGCAGCAGCATCATC

primer name	primer sequences (5' - 3')
attB2 CasFC-01 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTGAAGCAG
attB1 CasFC-02 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAACCCAACCGCAGCAGCATC
attB2 CasFC-02 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATTTTCAGGTTGAAGCAG
attB1 CasFC-03 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAACCCAACCGCAGCAG
attB2 CasFC-03 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATTTTCAGGTTGAAGCAGTAG
attB1 CasFC-04 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGAAAGCCAAATCGCAGTAGCATC
attB2 CasFC-04 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTAAAGCAATAATG
attB1 CasFC-05 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGAAAGCCAAATCGCAGTAGCATC
attB2 CasFC-05 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTAAAG
attB1 CasFC-06 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCAAATCGCAGTAGCATCATC
attB2 CasFC-06 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTAAAGCAATAATG
attB1 CasFC-07 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCAAACCGCAGTAGCATC
attB2 CasFC-07 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTAAAGCAATAATG
attB1 CasFC-08 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCAAACCGCAGTAGCATC
attB2 CasFC-08 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTGAAGCAATAATG
attB1 CasFC-09 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCAAACCGCAGTAG
attB2 CasFC-09 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTGAAGCAATAATG
attB1 CasFC-10 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGGGCAAGCCCAACCGCAGCAGCATCATC
attB2 CasFC-10 RP	CCCACTTTGTACAAGAAAGCTGGGTTTACAGGGTGATCTTCAGGTTGAAGCAG
pC13dcr01	
attB1 RX FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAAAAAAAAAAGTCCTTCGCCAAG
attB2 RX RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTCTTTTTC
attB1 CasFX-01 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGAGCTTC
attB2 CasFX-01 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGCTGTTGCCGCTCACCTTCTTC
attB1 CasFX-02 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTCCTTCGCCAAG
attB2 CasFX-02 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTCTTCTTC

330

primer name	primer sequences (5' - 3')
attB1 CasFX-03 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAAAAAAGTCCTTCGCCAAG
attB2 CasFX-03 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTCTTTTTC
attB1 CasFX-04 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAAAAGAAGAAGTCCTTCGCCAAG
attB2 CasFX-04 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTCTTC
attB1 CasFX-05 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTCCTTCGCCAAG
attB2 CasFX-05 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
attB1 CasFX-06 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTC
attB2 CasFX-06 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
attB1 CasFX-07 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTC
attB2 CasFX-07 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
attB1 CasFX-08 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTC
attB2 CasFX-08 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
attB1 CasFX-09 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTCCTTC
attB2 CasFX-09 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
attB1 CasFX-10 FP	TCAAGTTTGTACAAAAAAGCAGGCTATGATCGAGAAGAAGAAGTC
attB2 CasFX-10 RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTC
pC13cr01 eCFP crRNAs	
C13A eCFP 01 FP	ACTAAGGATGGGCACCACCCCGGTGAACAGCT
C13A eCFP 01 RP	TTTTAGCTGTTCACCGGGGTGGTGCCCATCCT
C13A eCFP 02 FP	ACTAGTTCACCAGGGTGTCGCCCTCGAACTTC
C13A eCFP 02 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAAC
C13B eCFP 01 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC
C13B eCFP 01 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
C13B eCFP 02 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC
C13B eCFP 02 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAACCG
C13C eCFP 01 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC
C13C eCFP 01 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
C13C eCFP 02 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC
C13C eCFP 02 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAACCG
C13D eCFP 01 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC

primer name	primer sequences (5' - 3')
C13D eCFP 01 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCCTG
C13D eCFP 02 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTC
C13D eCFP 02 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAACCG
pC13cr eCFP2 mismatch crRNA	
pC13A eCFP2 3 FP	ACTAGTGCACCAGGGTGTCGCCCTCGAACTTC
pC13A eCFP2 3 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGCAC
pC13A eCFP2 6 FP	ACTAGTTCATCAGGGTGTCGCCCTCGAACTTC
pC13A eCFP2 6 RP	TTTTGAAGTTCGAGGGCGACACCCTGATGAAC
pC13A eCFP2 9 FP	ACTAGTTCACCATGGTGTCGCCCTCGAACTTC
pC13A eCFP2 9 RP	TTTTGAAGTTCGAGGGCGACACCATGGTGAAC
pC13A eCFP2 12 FP	ACTAGTTCACCAGGGCGTCGCCCTCGAACTTC
pC13A eCFP2 12 RP	TTTTGAAGTTCGAGGGCGACGCCCTGGTGAAC
pC13A eCFP2 15 FP	ACTAGTTCACCAGGGTGTAGCCCTCGAACTTC
pC13A eCFP2 15 RP	TTTTGAAGTTCGAGGGCTACACCCTGGTGAAC
pC13A eCFP2 18 FP	ACTAGTTCACCAGGGTGTCGCACTCGAACTTC
pC13A eCFP2 18 RP	TTTTGAAGTTCGAGTGCGACACCCTGGTGAAC
pC13A eCFP2 21 FP	ACTAGTTCACCAGGGTGTCGCCCTAGAACTTC
pC13A eCFP2 21 RP	TTTTGTTCACCAGGGTGTCGCACTCGAACTTC
pC13A eCFP2 24 FP	ACTAGTTCACCAGGGTGTCGCCCTCGAGCTTC
pC13A eCFP2 24 RP	TTTTGAAGCTCGAGGGCGACACCCTGGTGAAC
pC13A eCFP2 27 FP	ACTAGTTCACCAGGGTGTCGCCCTCGAACTGC
pC13A eCFP2 27 RP	TTTTGCAGTTCGAGGGCGACACCCTGGTGAAC
pC13A eCFP2 28 FP	ACTAGTTCACCAGGGTGTCGCCCTCGAACTTA
pC13A eCFP2 28 RP	TTTTTAAGTTCGAGGGCGACACCCTGGTGAAC
pC13B eCFP2 3 FP	ACTACGGTTCACCAGGGTGTCGCCCTCGAACGTC
pC13B eCFP2 3 RP	TTTTGACGTTCGAGGGCGACACCCTGGTGAACCG
pC13B eCFP2 6 FP	ACTACGGTTCACCAGGGTGTCGCCCTCGGACTTC
pC13B eCFP2 6 RP	TTTTGAAGTCCGAGGGCGACACCCTGGTGAACCG
pC13B eCFP2 9 FP	ACTACGGTTCACCAGGGTGTCGCCCGCGAACTTC
pC13B eCFP2 9 RP	TTTTGAAGTTCGCGGGCGACACCCTGGTGAACCG
pC13B eCFP2 12 FP	ACTACGGTTCACCAGGGTGTCGTCCTCGAACTTC

primer name	primer sequences (5' - 3')
pC13B eCFP2 12 RP	TTTTGAAGTTCGAGGACGACACCCTGGTGAACCG
pC13B eCFP2 15 FP	ACTACGGTTCACCAGGGTGGCGCCCTCGAACTTC
pC13B eCFP2 15 RP	TTTTGAAGTTCGAGGGCGCCACCCTGGTGAACCG
pC13B eCFP2 18 FP	ACTACGGTTCACCAGGATGTCGCCCTCGAACTTC
pC13B eCFP2 18 RP	TTTTGAAGTTCGAGGGCGACATCCTGGTGAACCG
pC13B eCFP2 21 FP	ACTACGGTTCACCCGGGTGTCGCCCTCGAACTTC
pC13B eCFP2 21 RP	TTTTGAAGTTCGAGGGCGACACCCGGGTGAACCG
pC13B eCFP2 24 FP	ACTACGGTTCGCCAGGGTGTCGCCCTCGAACTTC
pC13B eCFP2 24 RP	TTTTGAAGTTCGAGGGCGACACCCTGGCGAACCG
pC13B eCFP2 27 FP	ACTACGGCTCACCAGGGTGTCGCCCTCGAACTTC
pC13B eCFP2 27 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAGCCG
pC13B eCFP2 30 FP	ACTATGGTTCACCAGGGTGTCGCCCTCGAACTTC
pC13B eCFP2 30 RP	TTTTGAAGTTCGAGGGCGACACCCTGGTGAACCA
pC13C eCFP2 3 FP	ACTACATGATGGGCACCACCCCGGTGAACAGCTC
pC13C eCFP2 3 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCATG
pC13C eCFP2 6 FP	ACTACAGGAGGGGCACCACCCCGGTGAACAGCTC
pC13C eCFP2 6 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCCTCCTG
pC13C eCFP2 9 FP	ACTACAGGATGGACACCACCCCGGTGAACAGCTC
pC13C eCFP2 9 RP	TTTTGAGCTGTTCACCGGGGTGGTGTCCATCCTG
pC13C eCFP2 12 FP	ACTACAGGATGGGCATCACCCCGGTGAACAGCTC
pC13C eCFP2 12 RP	TTTTGAGCTGTTCACCGGGGTGATGCCCATCCTG
pC13C eCFP2 15 FP	ACTACAGGATGGGCACCATCCCGGTGAACAGCTC
pC13C eCFP2 15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCATCCTG
pC13C eCFP2 18 FP	ACTACAGGATGGGCACCACCCAGGTGAACAGCTC
pC13C eCFP2 18 RP	TTTTGAGCTGTTCACCTGGGTGGTGCCCATCCTG
pC13C eCFP2 21 FP	ACTACAGGATGGGCACCACCCCGGCGAACAGCTC
pC13C eCFP2 21 RP	TTTTGAGCTGTTCGCCGGGGTGGTGCCCATCCTG
pC13C eCFP2 24 FP	ACTACAGGATGGGCACCACCCCGGTGAGCAGCTC
pC13C eCFP2 24 RP	TTTTGAGCTGCTCACCGGGGTGGTGCCCATCCTG
pC13C eCFP2 27 FP	ACTACAGGATGGGCACCACCCCGGTGAACATCTC
pC13C eCFP2 27 RP	TTTTGAGATGTTCACCGGGGTGGTGCCCATCCTG

primer name	primer sequences (5' - 3')
pC13C eCFP2 30 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTA
pC13C eCFP2 30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCATCCTG
pC13D eCFP2 1 FP	ACTATAGGATGGGCACCACCCCGGTGAACAGCTC
pC13D eCFP2 1 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCCTA
pC13D eCFP2 3 FP	ACTACATGATGGGCACCACCCCGGTGAACAGCTC
pC13D eCFP2 3 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCATCATG
pC13D eCFP2 6 FP	ACTACAGGACGGGCACCACCCCGGTGAACAGCTC
pC13D eCFP2 6 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCGTCCTG
pC13D eCFP2 9 FP	ACTACAGGATGGACACCACCCCGGTGAACAGCTC
pC13D eCFP2 9 RP	TTTTGAGCTGTTCACCGGGGTGGTGTCCATCCTG
pC13D eCFP2 12 FP	ACTACAGGATGGGCATCACCCCGGTGAACAGCTC
pC13D eCFP2 12 RP	TTTTGAGCTGTTCACCGGGGTGATGCCCATCCTG
pC13D eCFP2 15 FP	ACTACAGGATGGGCACCATCCCGGTGAACAGCTC
pC13D eCFP2 15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCATCCTG
pC13D eCFP2 18 FP	ACTACAGGATGGGCACCACCCAGGTGAACAGCTC
pC13D eCFP2 18 RP	TTTTGAGCTGTTCACCTGGGTGGTGCCCATCCTG
pC13D eCFP2 21 FP	ACTACAGGATGGGCACCACCCCGGCGAACAGCTC
pC13D eCFP2 21 RP	TTTTGAGCTGTTCGCCGGGGTGGTGCCCATCCTG
pC13D eCFP2 24 FP	ACTACAGGATGGGCACCACCCCGGTGAGCAGCTC
pC13D eCFP2 24 RP	TTTTGAGCTGCTCACCGGGGTGGTGCCCATCCTG
pC13D eCFP2 27 FP	ACTACAGGATGGGCACCACCCCGGTGAACATCTC
pC13D eCFP2 27 RP	TTTTGAGATGTTCACCGGGGTGGTGCCCATCCTG
pC13D eCFP2 30 FP	ACTACAGGATGGGCACCACCCCGGTGAACAGCTA
pC13D eCFP2 30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCATCCTG
pC13A 3+6 FP	ACTAGTGCATCAGGGTGTCGCCCTCGAACTTC
pC13A 3+6 RP	TTTTGAAGTTCGAGGGCGACACCCTGATGCAC
pC13A 3+15 FP	ACTAGTGCACCAGGGTGTAGCCCTCGAACTTC
pC13A 3+15 RP	TTTTGAAGTTCGAGGGCTACACCCTGGTGCAC
pC13A 3+28 FP	ACTAGTGCACCAGGGTGTCGCCCTCGAACTTA
pC13A 3+28 RP	TTTTTAAGTTCGAGGGCGACACCCTGGTGCAC
pC13A 3+6+15 FP	ACTAGTGCATCAGGGTGTAGCCCTCGAACTTC

334

primer name	primer sequences (5' - 3')
pC13A 3+6+15 RP	TTTTGAAGTTCGAGGGCTACACCCTGATGCAC
pC13A 3+6+28 FP	ACTAGTGCATCAGGGTGTCGCCCTCGAACTTA
pC13A 3+6+28 RP	TTTTTAAGTTCGAGGGCGACACCCTGATGCAC
pC13A 3+6+15+28 FP	ACTAGTGCATCAGGGTGTAGCCCTCGAACTTA
pC13A 3+6+15+28 RP	TTTTTAAGTTCGAGGGCTACACCCTGATGCAC
pC13B 3+6 FP	ACTACGGTTCACCAGGGTGTCGCCCTCGGACGTC
pC13B 3+6 RP	TTTTGACGTCCGAGGGCGACACCCTGGTGAACCG
pC13B 3+15 FP	ACTACGGTTCACCAGGGTGGCGCCCTCGAACGTC
pC13B 3+15 RP	TTTTGACGTTCGAGGGCGCCACCCTGGTGAACCG
pC13B 3+30 FP	ACTATGGTTCACCAGGGTGTCGCCCTCGAACGTC
pC13B 3+30 RP	TTTTGACGTTCGAGGGCGACACCCTGGTGAACCA
pC13B 3+6+15 FP	ACTACGGTTCACCAGGGTGGCGCCCTCGGACGTC
pC13B 3+6+15 RP	TTTTGACGTCCGAGGGCGCCACCCTGGTGAACCG
pC13B 3+6+30 FP	ACTATGGTTCACCAGGGTGTCGCCCTCGGACGTC
pC13B 3+6+30 RP	TTTTGACGTCCGAGGGCGACACCCTGGTGAACCA
pC13B 3+6+15+30 FP	ACTATGGTTCACCAGGGTGGCGCCCTCGGACGTC
pC13B 3+6+15+30 RP	TTTTGACGTCCGAGGGCGCCACCCTGGTGAACCA
pC13C 3+6 FP	ACTACATGAGGGGCACCACCCCGGTGAACAGCTC
pC13C 3+6 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCCTCATG
pC13C 3+15 FP	ACTACATGATGGGCACCATCCCGGTGAACAGCTC
pC13C 3+15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCATCATG
pC13C 3+30 FP	ACTACATGATGGGCACCACCCCGGTGAACAGCTA
pC13C 3+30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCATCATG
pC13C 3+6+15 FP	ACTACATGAGGGGCACCATCCCGGTGAACAGCTC
pC13C 3+6+15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCCTCATG
pC13C 3+6+30 FP	ACTACATGAGGGGCACCACCCCGGTGAACAGCTA
pC13C 3+6+30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCCTCATG
pC13C 3+6+15+30 FP	ACTACATGAGGGGCACCATCCCGGTGAACAGCTA
pC13C 3+6+15+30 RP	TTTTTAGCTGTTCACCGGGATGGTGCCCCTCATG
pC13D 3+6 FP	ACTACATGACGGGCACCACCCCGGTGAACAGCTC
pC13D 3+6 RP	TTTTGAGCTGTTCACCGGGGTGGTGCCCGTCATG

primer name	primer sequences (5' - 3')
pC13D 3+15 FP	ACTACATGATGGGCACCATCCCGGTGAACAGCTC
pC13D 3+15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCATCATG
pC13D 3+30 FP	ACTACATGATGGGCACCACCCCGGTGAACAGCTA
pC13D 3+30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCATCATG
pC13D 3+6+15 FP	ACTACATGACGGGCACCATCCCGGTGAACAGCTC
pC13D 3+6+15 RP	TTTTGAGCTGTTCACCGGGATGGTGCCCGTCATG
pC13D 3+6+30 FP	ACTACATGACGGGCACCACCCCGGTGAACAGCTA
pC13D 3+6+30 RP	TTTTTAGCTGTTCACCGGGGTGGTGCCCGTCATG
pC13D 3+6+15+30 FP	ACTACATGACGGGCACCATCCCGGTGAACAGCTA
pC13D 3+6+15+30 RP	TTTTTAGCTGTTCACCGGGATGGTGCCCGTCATG
dCasFX Fer1HCH RA crRNA	
dCasFX RA cr01 FP	ACTATGTGATACACACGTCCTCAATATGGGTATA
dCasFX RA cr01 RP	TTTTTATACCCATATTGAGGACGTGTGTATCACA
dCasFX RA cr02 FP	ACTAGCTCCGAACGGCGCACAAAACACTTTTAAG
dCasFX RA cr02 RP	TTTTCTTAAAAGTGTTTTGTGCGCCGTTCGGAGC
dCasFX RA cr03 FP	ACTAAGAAGGCGTCGCCTGCTTCAATTTGATGGG
dCasFX RA cr03 RP	TTTTCCCATCAAATTGAAGCAGGCGACGCCTTCT
dCasFX RA cr04 FP	ACTACATCTTTGATCGTCGAACGTAGTCTTTACA
dCasFX RA cr04 RP	TTTTTGTAAAGACTACGTTCGACGATCAAAGATG
dCasFX RA cr05 FP	ACTACTGGTAGGAGGCGTTGATCTCCTCCTGGAT
dCasFX RA cr05 RP	TTTTATCCAGGAGGAGATCAACGCCTCCTACCAG
dCasFX RA cr06 FP	ACTATCGGCACATTGATCAGATCGCTGACTCCCT
dCasFX RA cr06 RP	TTTTAGGGAGTCAGCGATCTGATCAATGTGCCGA
dCasFX RA cr07 FP	ACTAGTGGAGCTGCTCCTCCAGATAGACACCGGT
dCasFX RA cr07 RP	TTTTACCGGTGTCTATCTGGAGGAGCAGCTCCAC
dCasFX RA cr08 FP	ACTAGGGCTGACAGATAACAGATAACTCGACTGG
dCasFX RA cr08 RP	TTTTCCAGTCGAGTTATCTGTTATCTGTCAGCCC
dCasFX RA cr09 FP	ACTAGCATAATGTGGAGCCCCCTCCCGAGGTGTA
dCasFX RA cr09 RP	TTTTTACACCTCGGGAGGGGGCTCCACATTATGC
Mitochondrial encoded genes	
COXI cr01 FP	ACTAAAGATGTTCCAACTATTCCAGCTCAAGCTC

336

primer name	primer sequences (5' - 3')
COXI cr01 RP	TTTTGAGCTTGAGCTGGAATAGTTGGAACATCTT
COXI cr02 FP	ACTAGAGGTGGATAAACAGTTCATCCTGTCCCAG
COXI cr02 RP	TTTTCTGGGACAGGATGAACTGTTTATCCACCTC
COXI cr03 FP	ACTAGAGGATTAACAGGAGTTGTTTTAGCTAATT
COXI cr03 RP	TTTTAATTAGCTAAAACAACTCCTGTTAATCCTC
COXI cr04 FP	ACTAATAATGAAATAGTTGATCCAATAGTTGATA
COXI cr04 RP	TTTTTATCAACTATTGGATCAACTATTTCATTAT
COXII cr01 FP	ACTAGAAGCTCTATCTTGTAAACCTAAATTAGCT
COXII cr01 RP	TTTTAGCTAATTTAGGTTTACAAGATAGAGCTTC
COXII cr02 FP	ACTAATAATATAAATATTAAATATCCCACCAATA
COXII cr02 RP	TTTTTATTGGTGGGATATTTAATATTTATATTAT
COXII cr03 FP	ACTACATTTGTTGGAATTATATATGAATCAAATT
COXII cr03 RP	TTTTAATTTGATTCATATATAATTCCAACAAATG
COXII cr04 FP	ACTAAATTAGTTTGATTTAATCGTCCAGGTGTAC
COXII cr04 RP	TTTTGTACACCTGGACGATTAAATCAAACTAATT
COXI RNAi 2 FP	CTAGCAGTCGAGCTGAATTAGGACATCCTAGTTATATTCAAGCATAGGATGTCCTAATTCAGCTCGGCG
COXI RNAi 2 RP	AATTCGCCGAGCTGAATTAGGACATCCTATGCTTGAATATAACTAGGATGTCCTAATTCAGCTCGACTG
COXI RNAi 3 FP	CTAGCAGTAGGTGCTCCTGATATAGCATTAGTTATATTCAAGCATAATGCTATATCAGGAGCACCTGCG
COXI RNAi 3 RP	AATTCGCAGGTGCTCCTGATATAGCATTATGCTTGAATATAACTAATGCTATATCAGGAGCACCTACTG
COXI RNAi 4 FP	CTAGCAGTAGCTGGGACAGGATGAACTGTAGTTATATTCAAGCATACAGTTCATCCTGTCCCAGCTGCG
COXI RNAi 4 RP	AATTCGCAGCTGGGACAGGATGAACTGTATGCTTGAATATAACTACAGTTCATCCTGTCCCAGCTACTG
COXII RNAi 2 FP	CTAGCAGTCATGATCATGCATTATTAATTAGTTATATTCAAGCATAATTAATAATGCATGATCATGGCG
COXII RNAi 2 RP	AATTCGCCATGATCATGCATTATTAATTATGCTTGAATATAACTAATTAATAATGCATGATCATGACTG

primer name	primer sequences (5' - 3')
COXII RNAi 3 FP	CTAGCAGTAGCTGCTGATGTTATTCATTTAGTTATATTCAAGCATAAATGAATAACATCAGCAGCTGCG
COXII RNAi 3 RP	AATTCGCAGCTGCTGATGTTATTCATTTATGCTTGAATATAACTAAATGAATAACATCAGCAGCTACTG
COXII RNAi 4 FP	CTAGCAGTTCTGTGGAGCTAATCATAGATAGTTATATTCAAGCATATCTATGATTAGCTCCACAGAGCG
COXII RNAi 4 RP	AATTCGCTCTGTGGAGCTAATCATAGATATGCTTGAATATAACTATCTATGATTAGCTCCACAGAACTG
FREPAIRv2 crRNAs	
eCFP* 50cr02 FP	ACTAGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGT
eCFP* 50cr02 RP	TTTTACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCC
eCFP* 50cr10 FP	ACTAGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACT
eCFP* 50cr10 RP	TTTTAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTC
eCFP* 50cr18 FP	ACTAGTGGTCACGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCA
eCFP* 50cr18 RP	TTTTTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTCGTGACCAC
eCFP* 50cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCG
eCFP* 50cr26 RP	TTTTCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTCGTGACCACCCTGACCT
eCFP* 50cr34 FP	ACTACACGCCCCAGGTCAGGGTGGTCACGAGGGTGGGCTAGGGCACGGGCAGCT
eCFP* 50cr34 RP	TTTTAGCTGCCCGTGCCCTAGCCCACCCTCGTGACCACCCTGACCTGGGGCGTG
eCFP* 50cr42 FP	ACTAAAGCACTGCACGCCCCAGGTCAGGGTGGTCACGAGGGTGGGCTAGGGCAC
eCFP* 50cr42 RP	TTTTGTGCCCTAGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTT
eCFP* 50cr50 FP	ACTAAGCGGCTGAAGCACTGCACGCCCCAGGTCAGGGTGGTCACGAGGGTGGGC
eCFP* 50cr50 RP	TTTTGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTGCTTCAGCCGCT
eCFP* 30cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACC
eCFP* 30cr26 RP	TTTTGGTAGCCCACCCTCGTGACCACCCTGACCT
eCFP* 40cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACCAGGGCACGGG
eCFP* 40cr26 RP	TTTTCCCGTGCCCTGGTAGCCCACCCTCGTGACCACCCTGACCT
eCFP* 60cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCAGA
eCFP* 60cr26 RP	TTTTTCTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTCGTGACCACCCTGACCT

primer name	primer sequences (5' - 3')
eCFP* 70cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAG
eCFP* 70cr26 RP	TTTTCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTCGTGACCACCCTGACCT
eCFP* 80cr26 FP	ACTAAGGTCAGGGTGGTCACGAGGGTGGGCTACCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGCAAGCTGACC
eCFP* 80cr26 RP	TTTTGGTCAGCTTGCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGTAGCCCACCCTCGTGACCACCCTGACCT
Mutagenesis primers	
dCasFX RH5 mut FP	TCCAGGATCTCCAGGACCTGGCTCTACAACCTCGATAAGAACCTC
dCasFX RH5 Mut RP	CTCTTCTTCGTTGTTTGCGACCACCCAGTGTGCCAGTCCGCTCAG
dCasFX RH3 Mut FP	CCACGCCTATATCAACGACATTGCCGAGGTCAATTCCTACTTCCAACTGTAC
dCasFX RH3 Mut RP	ACATACCTGGCCACTTCCAGTGCGACGGCCTTGTTTGCGAACAGGGTACATG
mtCasFX Mut FP	TCAAGTTTGTACAAAAAAGCAGGCTATGAGTGACAATTTTTCAAGAACACCATATATCGAAAAGAAGAAGTCCTTC
mtCasFX Mut RP	CCCACTTTGTACAAGAAAGCTGGGTTTAGGAATTGCCGGACACCTTCTTC
FREPAIRv2 F1 FP	CTCCGCGGCCGCCCCCTTCACCATGATCGAAAAAAAAAAGTCCTTC
FREPAIRv2 F1 RP	CAGTCTTTCAAGTGGAGGCAGCTGCAGGCTTCCGGAATTGCCGGACACCTTC
FREPAIRv2 F2 FP	TGCCTCCACTTGAAAGACTGACACTGGGATCCGGAGGAGGTGGAAGCCAGCTGCATTTAC
FREPAIRv2 F2 RP	TGGGTCGGCGCGCCCACCCTTTTACGTGAGTGAGAACTGGTCCTGCTC
eCFP W57* Mut FP	AGCTGCCCGTGCCCTAACCCACCCTCGTGACCAC
eCFP W57* Mut RP	TGCCGGTGGTGCAGATGAACTTCAGGGTCAG
crRNA transgenes	
pC13B F1 RP	CTTCGTCCCAGGAAGACATGGTGGCATCGGCCGGGAATCGAAC
pC13B F1 FP	GTCAGCGGTTTCGTGACGAAGCTCCAAG
pC13B F2 RP	CTTGGAGCTTCGTCACGAAACCGCTGAC
pC13B F2 FP	CACTTTGAAGGGTATTCACAACTTTTTTGCCTACCTGGAGCCTGAGAG
pC13B Middle FP	CCATGTCTTCCTGGGACGAAGACAAGTTGGGACTGCTCTCACTTTGAAGGGTATTCACAACTAACAAAGCACCAGTGGTCTAGTGGTAGAATAGTACCCTGCCACGGTACAG

primer name	primer sequences (5' - 3')
	ACCCGGGTTCGATTCCCGGCTGGTGCACACCATGTCTTCCTGGGACGAAGACAAGT TGGGACTGCTCTCACTTTGAAGGGTATTCACA
pC13B Middle RP	TGTGAATACCCTTCAAAGTGAGAGCAGTCCCAACTTGTCTTCGTCCCAGGAAGACA TGGTGTGCACCAGCCGGGAATCGAACCCGGGTCTGTACCGTGGCAGGGTACTATTC TACCACTAGACCACTGGTGCTTTGTTAGTTGTGAATACCCTTCAAAGTGAGAGCAGT CCCAACTTGTCTTCGTCCCAGGAAGACATGG
pC13X F1 RP	CCGACCAGTTGGTAGGGGTTGCATCGGCCGGGAATCGAAC
pC13X F1 FP	GTCAGCGGTTTCGTGACGAAGCTCCAAG
pC13X F2 RP	CTTGGAGCTTCGTCACGAAACCGCTGAC
pC13X F2 FP	ACGGGTCTTCGAGAAGACCTTTTTTTGCCTACCTGGAGCCTGAGAGTTG
pC13X Middle FP	AACCCCTACCAACTGGTCGGGGTTTGAAACGGGTCTTCGAGAAGACCTTAACAAAG CACCAGTGGTCTAGTGGTAGAATAGTACCCTGCCACGGTACAGACCCGGGTTCGAT TCCCGGCTGGTGCAACCCCTACCAACTGGTCGGGGTTTGAAACGGGTCTTCGAGAA GACCT
pC13X Middle RP	AGGTCTTCTCGAAGACCCGTTTCAAACCCCGACCAGTTGGTAGGGGTTGCACCAGC CGGGAATCGAACCCGGGTCTGTACCGTGGCAGGGTACTATTCTACCACTAGACCAC TGGTGCTTTGTTAAGGTCTTCTCGAAGACCCGTTTCAAACCCCGACCAGTTGGTAGG GGTT
dib crRNA pC13X FP	AAACTGGGCAGTAAAAGACTGCAGACGAGCTCCA
dib crRNA pC13X RP	AAAATGGAGCTCGTCTGCAGTCTTTTACTGCCCA
dib crRNA pC13B FP	CACCTGGGCAGTAAAAGACTGCAGACGAGCTCCA
dib crRNA pC13B RP	CAACTGGAGCTCGTCTGCAGTCTTTTACTGCCCA
phm crRNA pC13X FP	AAACTGGAGCGCCGGCAGCGGATTCACTTCCGAC
phm crRNA pC13X RP	AAAAGTCGGAAGTGAATCCGCTGCCGGCGCTCCA
phm crRNA pC13B FP	CACCTGGAGCGCCGGCAGCGGATTCACTTCCGAC
phm crRNA pC13B RP	CAACGTCGGAAGTGAATCCGCTGCCGGCGCTCCA
IRP1A crRNA C13X FP	AAACTGGAGCGCCGGCAGCGGATTCACTTCCGAC
IRP1A crRNA C13X RP	AAAAGTCGGAAGTGAATCCGCTGCCGGCGCTCCA
IRP1A crRNA C13B FP	CACCTGGAGCGCCGGCAGCGGATTCACTTCCGAC
IRP1A crRNA C13B RP	CAACGTCGGAAGTGAATCCGCTGCCGGCGCTCCA

primer name	primer sequences (5' - 3')
dib cr IRP1A cr C13X FP	AACCCCTACCAACTGGTCGGGGTTTGAAACTGGGCAGTAAAAGACTGCAGACGAGCTCCATAACAAAGCACCAGTGGTCTAG
dib cr IRP1A cr C13X RP	AGGCTCCAGGTAGGCAAAAAAGTCGGAAGTGAATCCGCTGCCGGCGCTCCAGTTTCAAACCCCGACCAGTTG
qPCR primers	
eCFP qPCR FP	GAAGCGCGATCACATGGT
eCFP qPCR RP	CCATGCCGAGAGTGATCC
DsRed qPCR FP	GAAGGGCGAGATCCACAAG
DsRed qPCR RP	GGACTTGAACTCCACCAGGTA
rp49 qPCR FP	CGGATCGATATGCTAAGCTGT
rp49 qPCR RP	CGACGCACTCTGTTGTCG
COXI qPCR FP	TGACTTCTACCTCCTGCTCTTTC
COXI qPCR RP	GCGGATAGAGGTGGATAAACAG
COXII qPCR FP	CCGAGTAGTTTTACCCATAAACTCA
COXII qPCR RP	AAGCAGGTACTGTTCAAGAATGAAT
dib qPCR FP	GTGACCAAGGAGTTCATTAGATTTC
dib qPCR RP	CCAAAGGTAAGCAAACAGGTTAAT

References

1. Abbaspour N, Hurrell R, Kelishadi R. Review on iron and its importance for human health. J Res Med Sci. 2014;19:164-174.

2. Muckenthaler MU, Rivella S, Hentze MW, Galy B. A Red Carpet for Iron Metabolism. Cell. 2017;168:344-361.

3. Boulos S, Nyström L. UPLC-MS/MS investigation of β-glucan oligosaccharide oxidation. Analyst. 2016;141:6533-6548.

4. Saran M, Michel C, Stettmaier K, Bors W. Arguments against the significance of the Fenton reaction contributing to signal pathways under in vivo conditions. Free Radic Res. 2000;33:567-579.

5. Zimmermann MB, Hurrell RF. Nutritional iron deficiency. Lancet. 2007;370:511-520.

6. Camaschella C. Iron-deficiency anemia. N Engl J Med. 2015;372:1832-1843.

7. Anderson GJ. Mechanisms of iron loading and toxicity. Am J Hematol. 2007;82:1128-1131.

8. Anderson GJ, Frazer DM. Current understanding of iron homeostasis. Am J Clin Nutr. 2017;106:1559S-1566S.

9. Chua AC, Graham RM, Trinder D, Olynyk JK. The regulation of cellular iron metabolism. Crit Rev Clin Lab Sci. 2007;44:413-459.

10. Collard KJ. Iron homeostasis in the neonate. Pediatrics. 2009;123:1208-1216.

11. Fuqua BK, Vulpe CD, Anderson GJ. Intestinal iron absorption. J Trace Elem Med Biol. 2012;26:115-119.

12. McKie AT et al. An iron-regulated ferric reductase associated with the absorption of dietary iron. Science. 2001;291:1755-1759.

13. Bradley JM, Le Brun NE, Moore GR. Ferritins: furnishing proteins with iron. J Biol Inorg Chem. 2016;21:13-28.

14. Chiou B, Connor JR. Emerging and Dynamic Biomedical Uses of Ferritin. Pharmaceuticals (Basel). 2018;11

15. Alkhateeb AA, Connor JR. Nuclear ferritin: A new role for ferritin in cell biology. Biochim Biophys Acta. 2010;1800:793-797.

16. Corsi B et al. Human mitochondrial ferritin expressed in HeLa cells incorporates iron and affects cellular iron metabolism. J Biol Chem. 2002;277:22430-22437.

17. Knovich MA, Storey JA, Coffman LG, Torti SV, Torti FM. Ferritin for the clinician. Blood Rev. 2009;23:95-104.

18. Kuiper MA, Mulder C, van Kamp GJ, Scheltens P, Wolters EC. Cerebrospinal fluid ferritin levels of patients with Parkinson's disease, Alzheimer's disease, and multiple system atrophy. J Neural Transm Park Dis Dement Sect. 1994;7:109-114.

19. Surguladze N, Thompson KM, Beard JL, Connor JR, Fried MG. Interactions and reactions of ferritin with DNA. J Biol Chem. 2004;279:14694-14702.

20. Surguladze N, Patton S, Cozzi A, Fried MG, Connor JR. Characterization of nuclear ferritin and mechanism of translocation. Biochem J. 2005;388:731-740.

21. Waldo GS, Wright E, Whang ZH, Briat JF, Theil EC, Sayers DE. Formation of the ferritin iron mineral occurs in plastids. Plant Physiol. 1995;109:797-802.

22. Frazer DM, Anderson GJ. The regulation of iron transport. Biofactors. 2014;40:206-214.

23. Fleming MD, Romano MA, Su MA, Garrick LM, Garrick MD, Andrews NC. Nramp2 is mutated in the anemic Belgrade (b) rat: evidence of a role for Nramp2 in endosomal iron transport. Proc Natl Acad Sci U S A. 1998;95:1148-1153.

24. Sharp P. The molecular basis of copper and iron interactions. Proc Nutr Soc. 2004;63:563-569.

25. Briley-Saebo K, Bjørnerud A, Grant D, Ahlstrom H, Berg T, Kindberg GM. Hepatic cellular distribution and degradation of iron oxide nanoparticles following single intravenous injection in rats: implications for magnetic resonance imaging. Cell Tissue Res. 2004;316:315-323.

26. Cohen LA et al. Serum ferritin is derived primarily from macrophages through a nonclassical secretory pathway. Blood. 2010;116:1574-1584.

27. Connor JR, Zhang X, Nixon AM, Webb B, Perno JR. Comparative evaluation of nephrotoxicity and management by macrophages of intravenous pharmaceutical iron formulations. PLoS One. 2015;10:e0125272.

28. Kidane TZ, Sauble E, Linder MC. Release of iron from ferritin requires lysosomal activity. Am J Physiol Cell Physiol. 2006;291:C445-55.

29. Leimberg MJ, Prus E, Konijn AM, Fibach E. Macrophages function as a ferritin iron source for cultured human erythroid precursors. J Cell Biochem. 2008;103:1211-1218.

30. Nixon AM, Neely E, Simpson IA, Connor JR. The role of HFE genotype in macrophage phenotype. J Neuroinflammation. 2018;15:30.

31. Truman-Rosentsvit M et al. Ferritin is secreted via 2 distinct nonclassical vesicular pathways. Blood. 2018;131:342-352.

32. Wilkinson N, Pantopoulos K. The IRP/IRE system in vivo: insights from mouse models. Front Pharmacol. 2014;5:176.

33. Anderson GJ, Vulpe CD. Mammalian iron transport. Cell Mol Life Sci. 2009;66:3241-3261.

34. Guo W et al. Testosterone administration inhibits hepcidin transcription and is associated with increased iron incorporation into red blood cells. Aging Cell. 2013;12:280-291.

35. Hou Y et al. Estrogen regulates iron homeostasis through governing hepatic hepcidin expression via an estrogen response element. Gene. 2012;511:398-403.

36. Peyssonnaux C, Nizet V, Johnson RS. Role of the hypoxia inducible factors HIF in iron metabolism. Cell Cycle. 2008;7:28-32.

37. Folwell JL, Barton CH, Shepherd D. Immunolocalisation of the D. melanogaster Nramp homologue Malvolio to gut and Malpighian tubules provides evidence that Malvolio and Nramp2 are orthologous. J Exp Biol. 2006;209:1988-1995.

38. D'Souza J, Cheah PY, Gros P, Chia W, Rodrigues V. Functional complementation of the malvolio mutation in the taste pathway of Drosophila melanogaster by the human natural resistance-associated macrophage protein 1 (Nramp-1). J Exp Biol. 1999;202:1909-1915.

39. Orgad S, Nelson H, Segal D, Nelson N. Metal ions suppress the abnormal taste behavior of the Drosophila mutant malvolio. J Exp Biol. 1998;201:115-120.

40. Rodrigues V, Cheah PY, Ray K, Chia W. malvolio, the Drosophila homologue of mouse NRAMP-1 (Bcg), is expressed in macrophages and in the nervous system and is required for normal taste behaviour. EMBO J. 1995;14:3007-3020.

41. Iliadi KG et al. nemy encodes a cytochrome b561 that is required for Drosophila learning and memory. Proc Natl Acad Sci U S A. 2008;105:19986-19991.

42. Kamyshev NG et al. Novel memory mutants in Drosophila: behavioral characteristics of the mutant nemyP153. BMC Neurosci. 2002;3:9.

43. Mandilaras K, Pathmanathan T, Missirlis F. Iron absorption in Drosophila melanogaster. Nutrients. 2013;5:1622-1647.

44. Petrak J, Vyoral D. Hephaestin--a ferroxidase of cellular iron export. Int J Biochem Cell Biol. 2005;37:1173-1178.

45. Tang X, Zhou B. Ferritin is the key to dietary iron absorption and tissue iron detoxification in Drosophila melanogaster. FASEB J. 2013;27:288-298.

46. McKie AT, Barlow DJ. The SLC40 basolateral iron transporter family (IREG1/ferroportin/MTP1). Pflugers Arch. 2004;447:801-806.

47. Winzerling JJ, Pham DQ. Iron metabolism in insect disease vectors: mining the Anopheles gambiae translated protein database. Insect Biochem Mol Biol. 2006;36:310-321.

48. Dittmer NT, Kanost MR. Insect multicopper oxidases: diversity, properties, and physiological roles. Insect Biochem Mol Biol. 2010;40:179-188.

49. Lang M, Braun CL, Kanost MR, Gorman MJ. Multicopper oxidase-1 is a ferroxidase essential for iron homeostasis in Drosophila melanogaster. Proc Natl Acad Sci U S A. 2012;109:13337-13342.

50. Bettedi L, Aslam MF, Szular J, Mandilaras K, Missirlis F. Iron depletion in the intestines of Malvolio mutant flies does not occur in the absence of a multicopper oxidase. J Exp Biol. 2011;214:971-978.

51. Wang X, Yin S, Yang Z, Zhou B. Drosophila multicopper oxidase 3 is a potential ferroxidase involved in iron homeostasis. Biochim Biophys Acta Gen Subj. 2018;1862:1826-1834.

52. Kaplan J, O'Halloran TV. Iron metabolism in eukaryotes: Mars and Venus at it again. Science. 1996;271:1510-1512.

53. Geiser DL, Winzerling JJ. Insect transferrins: multifunctional proteins. Biochim Biophys Acta. 2012;1820:437-451.

54. Lambert LA, Perri H, Halbrooks PJ, Mason AB. Evolution of the transferrin family: conservation of residues associated with iron and anion binding. Comp Biochem Physiol B Biochem Mol Biol. 2005;142:129-141.

55. Dunkov B, Georgieva T. Insect iron binding proteins: insights from the genomes. Insect Biochem Mol Biol. 2006;36:300-309.

56. Geiser DL, Zhang D, Winzerling JJ. Secreted ferritin: mosquito defense against iron overload. Insect Biochem Mol Biol. 2006;36:177-187.

57. Geiser DL, Shen MC, Mayo JJ, Winzerling JJ. Iron loaded ferritin secretion and inhibition by CI-976 in Aedes aegypti larval cells. Comp Biochem Physiol B Biochem Mol Biol. 2009;152:352-363.

58. Geiser DL, Thai TN, Love MB, Winzerling JJ. Iron and Ferritin Deposition in the Ovarian Tissues of the Yellow Fever Mosquito (Diptera: Culicidae). J Insect Sci. 2019;19

59. Gourley BL, Parker SB, Jones BJ, Zumbrennen KB, Leibold EA. Cytosolic aconitase and ferritin are regulated by iron in Caenorhabditis elegans. J Biol Chem. 2003;278:3227-3234.

60. Kim YI, Cho JH, Yoo OJ, Ahnn J. Transcriptional regulation and life-span modulation of cytosolic aconitase and ferritin genes in C.elegans. J Mol Biol. 2004;342:421-433.

61. Lind MI et al. Of two cytosolic aconitases expressed in Drosophila, only one functions as an iron-regulatory protein. J Biol Chem. 2006;281:18707-18714.

62. Tsuboi D, Qadota H, Kasuya K, Amano M, Kaibuchi K. Isolation of the interacting molecules with GEX-3 by a novel functional screening. Biochem Biophys Res Commun. 2002;292:697-701.

63. Moeder W, Del Pozo O, Navarre DA, Martin GB, Klessig DF. Aconitase plays a role in regulating resistance to oxidative stress and cell death in Arabidopsis and Nicotiana benthamiana. Plant Mol Biol. 2007;63:273-287.

64. Wang J, Chen G, Filebeen C, Pantopoulos K. Insights on regulation and function of the iron regulatory protein 1 (IRP1). Hemoglobin. 2008;32:109-115.

65. Papanikolaou G, Pantopoulos K. Systemic iron homeostasis and erythropoiesis. IUBMB Life. 2017;69:399-413.

66. Danielsen ET et al. A Drosophila Genome-Wide Screen Identifies Regulators of Steroid Hormone Production and Developmental Timing. Dev Cell. 2016;37:558-570.

67. Ou Q, Zeng J, Yamanaka N, Brakken-Thal C, O'Connor MB, King-Jones K. The Insect Prothoracic Gland as a Model for Steroid Hormone Biosynthesis and Regulation. Cell Rep. 2016;16:247-262.

68. Hamza I, Dailey HA. One ring to rule them all: trafficking of heme and heme synthesis intermediates in the metazoans. Biochim Biophys Acta. 2012;1823:1617-1632.

69. Wagener FA et al. Different faces of the heme-heme oxygenase system in inflammation. Pharmacol Rev. 2003;55:551-571.

70. Dailey TA, Woodruff JH, Dailey HA. Examination of mitochondrial protein targeting of haem synthetic enzymes: in vivo identification of three functional haem-responsive motifs in 5-aminolaevulinate synthase. Biochem J. 2005;386:381-386.

71. Lathrop JT, Timko MP. Regulation by heme of mitochondrial protein transport through a conserved amino acid motif. Science. 1993;259:522-525.

72. Yamauchi K, Hayashi N, Kikuchi G. Translocation of delta-aminolevulinate synthase from the cytosol to the mitochondria and its regulation by hemin in the rat liver. J Biol Chem. 1980;255:1746-1751.

73. Hibino A, Petri R, Büchs J, Ohtake H. Production of uroporphyrinogen III, which is the common precursor of all tetrapyrrole cofactors, from 5-aminolevulinic acid by Escherichia coli expressing thermostable enzymes. Appl Microbiol Biotechnol. 2013;97:7337-7344.

74. Hamza I, Dailey HA. One ring to rule them all: trafficking of heme and heme synthesis intermediates in the metazoans. Biochim Biophys Acta. 2012;1823:1617-1632.

75. Lash TD. The enigma of coproporphyrinogen oxidase: how does this unusual enzyme carry out oxidative decarboxylations to afford vinyl groups. Bioorg Med Chem Lett. 2005;15:4506-4509.

76. Silva PJ, Ramos MJ. Computational characterization of the substrate-binding mode in coproporphyrinogen III oxidase. J Phys Chem B. 2011;115:1903-1910.

77. Chiabrando D, Vinchi F, Fiorito V, Mercurio S, Tolosano E. Heme in pathophysiology: a matter of scavenging, metabolism and trafficking across cell membranes. Front Pharmacol. 2014;5:61.

78. Balla J, Jacob HS, Balla G, Nath K, Vercellotti GM. Endothelial cell heme oxygenase and ferritin induction by heme proteins: a possible mechanism limiting shock damage. Trans Assoc Am Physicians. 1992;105:1-6.

79. Kumar S, Bandyopadhyay U. Free heme toxicity and its detoxification systems in human. Toxicol Lett. 2005;157:175-188.

80. Ryter SW, Tyrrell RM. The heme synthesis and degradation pathways: role in oxidant sensitivity. Heme oxygenase has both pro- and antioxidant properties. Free Radic Biol Med. 2000;28:289-309.

81. Tolosano E, Fagoonee S, Morello N, Vinchi F, Fiorito V. Heme scavenging and the other facets of hemopexin. Antioxid Redox Signal. 2010;12:305-320.

82. Bari AU. Congenital erythropoietic porphyria in three siblings. Indian J Dermatol Venereol Leprol. 2007;73:340-342.

83. Ciftci V, Kılavuz S, Bulut FD, Mungan HN, Bisgin A, Dogan MC. Congenital erythropoietic porphyria with erythrodontia: A case report. Int J Paediatr Dent. 2019;29:542-548.

84. Koley S, Saoji V. Congenital erythropoietic porphyria: two case reports. Indian J Dermatol. 2011;56:94-97.

85. Ascenzi P et al. Hemoglobin and heme scavenging. IUBMB Life. 2005;57:749-759.

86. Hvidberg V, Maniecki MB, Jacobsen C, Højrup P, Møller HJ, Moestrup SK. Identification of the receptor scavenging hemopexin-heme complexes. Blood. 2005;106:2572-2579.

87. Hsiung CS, Andrade JD, Costa R, Ash KO. Minimizing interferences in the quantitative multielement analysis of trace elements in biological fluids by inductively coupled plasma mass spectrometry. Clin Chem. 1997;43:2303-2311.

88. Dupuy J, Volbeda A, Carpentier P, Darnault C, Moulis JM, Fontecilla-Camps JC. Crystal structure of human iron regulatory protein 1 as cytosolic aconitase. Structure. 2006;14:129-139.

89. Gratz SJ et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 2014;196:961-971.

90. Lin S, Ewen-Campen B, Ni X, Housden BE, Perrimon N. In Vivo Transcriptional Activation Using CRISPR/Cas9 in Drosophila. Genetics. 2015;201:433-442.

91. Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc Natl Acad Sci U S A. 2014;111:E2967-76.

92. Bier E, Harrison MM, O'Connor-Giles KM, Wildonger J. Advances in Engineering the Fly Genome with the CRISPR-Cas System. Genetics. 2018;208:1-18.

93. Gupta RM, Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. J Clin Invest. 2014;124:4154-4161.

94. Port F, Bullock SL. Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. Nat Methods. 2016;13:852-854.

95. Pfeiffer BD et al. Tools for neuroanatomy and neurogenetics in Drosophila. Proc Natl Acad Sci U S A. 2008;105:9715-9720.

96. Pfeiffer BD et al. Refinement of tools for targeted gene expression in Drosophila. Genetics. 2010;186:735-755.

97. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6:343-345.

98. Wang JW, Beck ES, McCabe BD. A modular toolset for recombination transgenesis and neurogenetic analysis of Drosophila. PLoS One. 2012;7:e42102.

99. Daniel E, Onwukwe GU, Wierenga RK, Quaggin SE, Vainio SJ, Krause M. ATGme: Open-source web application for rare codon identification and custom DNA sequence optimization. BMC Bioinformatics. 2015;16:303.

100. Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res. 2007;35:W126-31.

101. Behura SK, Severson DW. Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. Biol Rev Camb Philos Soc. 2013;88:49-61.

102. Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 2000;28:292.

103. Vicario S, Moriyama EN, Powell JR. Codon usage in twelve species of Drosophila. BMC Evol Biol. 2007;7:226.

104. Sharp PM, Li WH. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 1987;15:1281-1295.

105. Goetz RM, Fuglsang A. Correlation of codon bias measures with mRNA levels: analysis of transcriptome data from Escherichia coli. Biochem Biophys Res Commun. 2005;327:4-7.

106. Henry I, Sharp PM. Predicting gene expression level from codon usage bias. Mol Biol Evol. 2007;24:10-12.

107. Wright F. The 'effective number of codons' used in a gene. Gene. 1990;87:23-29.

108. Cox DBT et al. RNA editing with CRISPR-Cas13. Science. 2017;358:1019-1027.

109. Smargon AA et al. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. Mol Cell. 2017;65:618-630.e7.

110. Abudayyeh OO et al. RNA targeting with CRISPR-Cas13. Nature. 2017;550:280-284.

111. Aman R et al. RNA virus interference via CRISPR/Cas13a system in plants. Genome Biol. 2018;19:1.

112. Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell. 2018;173:665-676.e14.

113. Mahas A, Aman R, Mahfouz M. CRISPR-Cas13d mediates robust RNA virus interference in plants. Genome Biol. 2019;20:263.

114. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008;36:W70-4.

115. Mathews DH. RNA Secondary Structure Analysis Using RNAstructure. Curr Protoc Bioinformatics. 2014;46:12.6.1-25.

116. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 2010;11:129.

117. Tan Z, Fu Y, Sharma G, Mathews DH. TurboFold II: RNA structural alignment and secondary structure prediction informed by multiple homologs. Nucleic Acids Res. 2017;45:11570-11581.

118. Tafer H et al. The impact of target site accessibility on the design of effective siRNAs. Nat Biotechnol. 2008;26:578-583.

119. Zhu H, Richmond E, Liang C. CRISPR-RT: a web application for designing CRISPR-C2c2 crRNA with improved target specificity. Bioinformatics. 2018;34:117-119.

120. Fish MP, Groth AC, Calos MP, Nusse R. Creating transgenic Drosophila by microinjecting the site-specific phiC31 integrase mRNA and a transgene-containing donor plasmid. Nat Protoc. 2007;2:2325-2331.

121. Tomoyasu Y, Denell RE. Larval RNAi in Tribolium (Coleoptera) for analyzing adult development. Dev Genes Evol. 2004;214:575-578.

122. Huang DW et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8:R183.

123. Prithviraj R, Trunova S, Giniger E. Ex vivo culturing of whole, developing Drosophila brains. J Vis Exp. 2012

124. González-Morales N, Mendoza-Ortíz MÁ, Blowes LM, Missirlis F, Riesgo-Escovar JR. Ferritin Is Required in Multiple Tissues during Drosophila melanogaster Development. PLoS One. 2015;10:e0133499.

125. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. 2007;4:207-214.

126. Elias JE, Gygi SP. Target-decoy search strategy for mass spectrometry-based proteomics. Methods Mol Biol. 2010;604:55-71.

127. Khan SJ, Abidi SN, Tian Y, Skinner A, Smith-Bolton RK. A rapid, gentle and scalable method for dissociation and fluorescent sorting of imaginal disc cells for mRNA sequencing. Fly (Austin). 2016;10:73-80.

128. Deutsch EW et al. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res. 2017;45:D1100-D1106.

129. Hendrickson DG et al. Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol. 2009;7:e1000238.

130. Khalil AM et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009;106:11667-11672.

131. Rinn JL et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129:1311-1323.

132. Sentmanat MF, Peters ST, Florian CP, Connelly JP, Pruett-Miller SM. A Survey of Validation Strategies for CRISPR-Cas9 Editing. Sci Rep. 2018;8:888.

133. Vouillot L, Thélie A, Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. G3 (Bethesda). 2015;5:407-415.

134. Lill R. Function and biogenesis of iron-sulphur proteins. Nature. 2009;460:831-838.

135. Anderson CP, Shen M, Eisenstein RS, Leibold EA. Mammalian iron metabolism and its control by iron regulatory proteins. Biochim Biophys Acta. 2012;1823:1468-1483.

136. Walden WE et al. Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. Science. 2006;314:1903-1908.

137. Volz K. The functional duality of iron regulatory protein 1. Curr Opin Struct Biol. 2008;18:106-111.

138. Yoshiyama-Yanagawa T et al. The conserved Rieske oxygenase DAF-36/Neverland is a novel cholesterol-metabolizing enzyme. J Biol Chem. 2011;286:25756-25762.

139. Wang L et al. Structure and Evolution of Glycogen Branching Enzyme N-Termini From Bacteria. Front Microbiol. 2018;9:3354.

140. Harlaar L et al. Large variation in effects during 10 years of enzyme therapy in adults with Pompe disease. Neurology. 2019;93:e1756-e1767.

141. Wei A, Ma H, Li Z, Zhang L, Zhang R, Wang T. Type IV Glycogen Storage Disease Associated With Hemophagocytic Lymphohistiocytosis: A Case Report. J Pediatr Hematol Oncol. 2019

142. Bilal H, Cheema HA, Fayyaz Z, Saeed A, Batool Hamdani SS. Hepatic Glycogenosis In Children: Spectrum Of Presentation And Diagnostic Modalities. J Ayub Med Coll Abbottabad. 2019;31:368-371.

143. Butler DC, Glen WB, Schandl C, Phillips A. Glycogen Storage Disease Type IV Diagnosed at Fetal Autopsy. Pediatr Dev Pathol. 20191093526619890224.

144. Cenacchi G et al. Update on polyglucosan storage diseases. Virchows Arch. 2019;475:671-686.

145. Buart S et al. Transcriptional response to hypoxic stress in melanoma and prognostic potential of GBE1 and BNIP3. Oncotarget. 2017;8:108786-108801.

146. Ebersole JL et al. Hypoxia-inducible transcription factors, HIF1A and HIF2A, increase in aging mucosal tissues. Immunology. 2018;154:452-464.

147. Semenza GL. HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations. J Clin Invest. 2013;123:3664-3671.

148. Wu D, Yotnda P. Induction and testing of hypoxia in cell culture. J Vis Exp. 2011

149. Zhao J, Chen H, Davidson T, Kluz T, Zhang Q, Costa M. Nickel-induced 1,4-alpha-glucan branching enzyme 1 up-regulation via the hypoxic signaling pathway. Toxicol Appl Pharmacol. 2004;196:404-409.

150. Simpson RJ. Effect of hypoxic exposure on iron absorption in heterozygous hypotransferrinaemic mice. Ann Hematol. 1992;65:260-264.

151. Zhang AS, Enns CA. Molecular mechanisms of normal iron homeostasis. Hematology Am Soc Hematol Educ Program. 2009207-214.

152. Christova T, Templeton DM. Effect of hypoxia on the binding and subcellular distribution of iron regulatory proteins. Mol Cell Biochem. 2007;301:21-32.

153. Sanchez M, Galy B, Muckenthaler MU, Hentze MW. Iron-regulatory proteins limit hypoxia-inducible factor-2alpha expression in iron deficiency. Nat Struct Mol Biol. 2007;14:420-426.

154. Kerins MJ, Ooi A. The Roles of NRF2 in Modulating Cellular Iron Homeostasis. Antioxid Redox Signal. 2018;29:1756-1773.

155. Uruno A et al. Nrf2-Mediated Regulation of Skeletal Muscle Glycogen Metabolism. Mol Cell Biol. 2016;36:1655-1672.

156. Heinemann IU, Jahn M, Jahn D. The biochemistry of heme biosynthesis. Arch Biochem Biophys. 2008;474:238-251.

157. Nagababu E, Mohanty JG, Bhamidipaty S, Ostera GR, Rifkind JM. Role of the membrane in the formation of heme degradation products in red blood cells. Life Sci. 2010;86:133-138.

158. Cáceres L, Necakov AS, Schwartz C, Kimber S, Roberts IJ, Krause HM. Nitric oxide coordinates metabolism, growth, and development via the nuclear receptor E75. Genes Dev. 2011;25:1476-1485.

159. Cappellini MD, Brancaleoni V, Graziadei G, Tavazzi D, Di Pierro E. Porphyrias at a glance: diagnosis and treatment. Intern Emerg Med. 2010;5 Suppl 1:S73-80.

160. Orhan Akman H et al. A novel mouse model that recapitulates adult-onset glycogenosis type 4. Hum Mol Genet. 2015;24:6801-6810.

161. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006;34:D535-9.

162. Stark C et al. The BioGRID Interaction Database: 2011 update. Nucleic Acids Res. 2011;39:D698-704.

163. Wan C et al. Panorama of ancient metazoan macromolecular complexes. Nature. 2015;525:339-344.

164. Philpott CC, Klausner RD, Rouault TA. The bifunctional iron-responsive element binding protein/cytosolic aconitase: the role of active-site residues in ligand binding and regulation. Proc Natl Acad Sci U S A. 1994;91:7321-7325.

165. Regev-Rudzki N, Karniely S, Ben-Haim NN, Pines O. Yeast aconitase in two locations and two metabolic pathways: seeing small amounts is believing. Mol Biol Cell. 2005;16:4163-4171.

166. Huynh N, Ou Q, Cox P, Lill R, King-Jones K. Glycogen branching enzyme controls cellular iron homeostasis via Iron Regulatory Protein 1 and mitoNEET. Nat Commun. 2019;10:5463.

167. Muckenthaler M, Gray NK, Hentze MW. IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. Mol Cell. 1998;2:383-388.

168. Paraskeva E, Gray NK, Schläger B, Wehr K, Hentze MW. Ribosomal pausing and scanning arrest as mechanisms of translational regulation from cap-distal iron-responsive elements. Mol Cell Biol. 1999;19:807-816.

169. Mittler R et al. NEET Proteins: A New Link Between Iron Metabolism, Reactive Oxygen Species, and Cancer. Antioxid Redox Signal. 2019;30:1083-1095.

170. Lipper CH, Paddock ML, Onuchic JN, Mittler R, Nechushtai R, Jennings PA. Cancer-Related NEET Proteins Transfer 2Fe-2S Clusters to Anamorsin, a Protein Required for Cytosolic Iron-Sulfur Cluster Biogenesis. PLoS One. 2015;10:e0139699.

171. Tan G et al. His-87 ligand in mitoNEET is crucial for the transfer of iron sulfur clusters from mitochondria to cytosolic aconitase. Biochem Biophys Res Commun. 2016;470:226-232.

172. Zuris JA et al. Facile transfer of [2Fe-2S] clusters from the diabetes drug target mitoNEET to an apo-acceptor protein. Proc Natl Acad Sci U S A. 2011;108:13047-13052.

173. Inupakutika MA et al. Phylogenetic analysis of eukaryotic NEET proteins uncovers a link between a key gene duplication event and the evolution of vertebrates. Sci Rep. 2017;7:42571.

174. Ferecatu I et al. The diabetes drug target MitoNEET governs a novel trafficking pathway to rebuild an Fe-S cluster into cytosolic aconitase/iron regulatory protein 1. J Biol Chem. 2014;289:28070-28086.

175. Roy A, Solodovnikova N, Nicholson T, Antholine W, Walden WE. A novel eukaryotic factor for cytosolic Fe-S cluster assembly. EMBO J. 2003;22:4826-4835.

176. Brown NM et al. Novel role of phosphorylation in Fe-S cluster stability revealed by phosphomimetic mutations at Ser-138 of iron regulatory protein 1. Proc Natl Acad Sci U S A. 1998;95:15235-15240.

177. Brown NM, Kennedy MC, Antholine WE, Eisenstein RS, Walden WE. Detection of a [3Fe-4S] cluster intermediate of cytosolic aconitase in yeast expressing iron regulatory protein 1. Insights into the mechanism of Fe-S cluster cycling. J Biol Chem. 2002;277:7246-7254.

178. Chintapalli VR, Wang J, Dow JA. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat Genet. 2007;39:715-720.

179. Nilsson J et al. Polyglucosan body myopathy caused by defective ubiquitin ligase RBCK1. Ann Neurol. 2013;74:914-919.

180. Yamanaka K et al. Identification of the ubiquitin-protein ligase that recognizes oxidized IRP2. Nat Cell Biol. 2003;5:336-340.

181. Netz DJ, Mascarenhas J, Stehling O, Pierik AJ, Lill R. Maturation of cytosolic and nuclear iron-sulfur proteins. Trends Cell Biol. 2014;24:303-312.

182. Wellen KE, Hatzivassiliou G, Sachdeva UM, Bui TV, Cross JR, Thompson CB. ATP-citrate lyase links cellular metabolism to histone acetylation. Science. 2009;324:1076-1080.

183. Moussaieff A et al. Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. Cell Metab. 2015;21:392-402.

184. Sutendra G et al. A nuclear pyruvate dehydrogenase complex is important for the generation of acetyl-CoA and histone acetylation. Cell. 2014;158:84-97.

185. Li SC et al. Glycogen storage disease type IV: novel mutations and molecular characterization of a heterogeneous disorder. J Inherit Metab Dis. 2010;33 Suppl 3:S83-90.

186. Pronicka E et al. New perspective in diagnostics of mitochondrial disorders: two years' experience with whole-exome sequencing at a national paediatric centre. J Transl Med. 2016;14:174.

187. Nolte KW, Janecke AR, Vorgerd M, Weis J, Schröder JM. Congenital type IV glycogenosis: the spectrum of pleomorphic polyglucosan bodies in muscle, nerve, and spinal cord with two novel mutations in the GBE1 gene. Acta Neuropathol. 2008;116:491-506.

188. Pescador N et al. Hypoxia promotes glycogen accumulation through hypoxia inducible factor (HIF)-mediated induction of glycogen synthase 1. PLoS One. 2010;5:e9644.

189. Hu Y et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011;12:357.

190. Anderson SA et al. The IRP1-HIF-2α axis coordinates iron and oxygen sensing with erythropoiesis and iron absorption. Cell Metab. 2013;17:282-290.

191. Galy B, Ferring D, Benesova M, Benes V, Hentze MW. Targeted mutagenesis of the murine IRP1 and IRP2 genes reveals context-dependent RNA processing differences in vivo. RNA. 2004;10:1019-1025.

192. Ghosh MC et al. Deletion of iron regulatory protein 1 causes polycythemia and pulmonary hypertension in mice through translational derepression of HIF2α. Cell Metab. 2013;17:271-281.

193. LaVaute T et al. Targeted deletion of the gene encoding iron regulatory protein-2 causes misregulation of iron metabolism and neurodegenerative disease in mice. Nat Genet. 2001;27:209-214.

194. Wilkinson N, Pantopoulos K. IRP1 regulates erythropoiesis and systemic iron homeostasis by controlling HIF2α mRNA translation. Blood. 2013;122:1658-1668.

195. Kristensen AR, Gsponer J, Foster LJ. A high-throughput approach for measuring temporal changes in the interactome. Nat Methods. 2012;9:907-909.

196. Patton SM, Piñero DJ, Surguladze N, Beard J, Connor JR. Subcellular localization of iron regulatory proteins to Golgi and ER membranes. J Cell Sci. 2005;118:4365-4373.

197. Kohler SA, Henderson BR, Kühn LC. Succinate dehydrogenase b mRNA of Drosophila melanogaster has a functional iron-responsive element in its 5'-untranslated region. J Biol Chem. 1995;270:30781-30786.

198. Surdej P, Richman L, Kühn LC. Differential translational regulation of IRE-containing mRNAs in Drosophila melanogaster by endogenous IRP and a constitutive human IRP1 mutant. Insect Biochem Mol Biol. 2008;38:891-894.

199. Huynh N, Zeng J, Liu W, King-Jones K. A Drosophila CRISPR/Cas9 Toolkit for Conditionally Manipulating Gene Expression in the Prothoracic Gland as a Test Case for Polytene Tissues. G3 (Bethesda). 2018;8:3593-3605.

200. Rouault TA. The role of iron regulatory proteins in mammalian iron homeostasis and disease. Nat Chem Biol. 2006;2:406-414.

201. Wang J, Pantopoulos K. Regulation of cellular iron metabolism. Biochem J. 2011;434:365-381.

202. Lill R. From the discovery to molecular understanding of cellular iron-sulfur protein biogenesis. Biol Chem. 2020

203. Tong WH, Rouault TA. Metabolic regulation of citrate and iron by aconitases: role of iron-sulfur cluster biogenesis. Biometals. 2007;20:549-564.

204. Lawlis VB, Roche TE. Effect of micromolar Ca^{2+} on NADH inhibition of bovine kidney alpha-ketoglutarate dehydrogenase complex and possible role of Ca^{2+} in signal amplification. Mol Cell Biochem. 1980;32:147-152.

205. Bisaccia F, De Palma A, Palmieri F. Identification and purification of the tricarboxylate carrier from rat liver mitochondria. Biochim Biophys Acta. 1989;977:171-176.

206. Koh HJ et al. Cytosolic NADP+-dependent isocitrate dehydrogenase plays a key role in lipid metabolism. J Biol Chem. 2004;279:39968-39974.

207. Minard KI, McAlister-Henn L. Sources of NADPH in yeast vary with carbon source. J Biol Chem. 2005;280:39890-39896.

208. Denton RM, Randle PJ. Citrate and the regulation of adipose-tissue phosphofructokinase. Biochem J. 1966;100:420-423.

209. Randle PJ. Regulatory interactions between lipids and carbohydrates: the glucose fatty acid cycle after 35 years. Diabetes Metab Rev. 1998;14:263-283.

210. Grootveld M, Bell JD, Halliwell B, Aruoma OI, Bomford A, Sadler PJ. Non-transferrin-bound iron in plasma or serum from patients with idiopathic hemochromatosis. Characterization by high performance liquid chromatography and nuclear magnetic resonance spectroscopy. J Biol Chem. 1989;264:4417-4422.

211. Chua AC, Olynyk JK, Leedman PJ, Trinder D. Nontransferrin-bound iron uptake by hepatocytes is increased in the Hfe knockout mouse model of hereditary hemochromatosis. Blood. 2004;104:1519-1525.

212. Melefors O, Hentze MW. Translational regulation by mRNA/protein interactions in eukaryotic cells: ferritin and beyond. Bioessays. 1993;15:85-90.

213. Martin RB. Citrate binding of Al3+ and Fe3+. J Inorg Biochem. 1986;28:181-187.

214. Pierre JL, Gautier-Luneau I. Iron and citric acid: a fuzzy chemistry of ubiquitous biological relevance. Biometals. 2000;13:91-96.

215. Chen OS, Hemenway S, Kaplan J. Genetic analysis of iron citrate toxicity in yeast: implications for mammalian iron homeostasis. Proc Natl Acad Sci U S A. 2002;99:16922-16927.

216. Hubert N et al. RNAs mediating cotranslational insertion of selenocysteine in eukaryotic selenoproteins. Biochimie. 1996;78:590-596.

217. Walczak R, Hubert N, Carbon P, Krol A. Solution structure of SECIS, the mRNA element required for eukaryotic selenocysteine insertion--interaction studies with the SECIS-binding protein SBP. Biomed Environ Sci. 1997;10:177-181.

218. Nudler E, Mironov AS. The riboswitch control of bacterial metabolism. Trends Biochem Sci. 2004;29:11-17.

219. Winkler WC. Riboswitches and the role of noncoding RNAs in bacterial metabolic control. Curr Opin Chem Biol. 2005;9:594-602.

220. Casey JL, Koeller DM, Ramin VC, Klausner RD, Harford JB. Iron regulation of transferrin receptor mRNA levels requires iron-responsive elements and a rapid turnover determinant in the 3' untranslated region of the mRNA. EMBO J. 1989;8:3693-3699.

221. Hentze MW et al. Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. Science. 1987;238:1570-1573.

222. Müllner EW, Kühn LC. A stem-loop in the 3' untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. Cell. 1988;53:815-825.

223. Briat JF, Lobréaux S. Iron storage and ferritin in plants. Met Ions Biol Syst. 1998;35:563-584.

224. Briat JF, Duc C, Ravet K, Gaymard F. Ferritins and iron storage in plants. Biochim Biophys Acta. 2010;1800:806-814.

225. Rothenberger S, Müllner EW, Kühn LC. The mRNA-binding protein which controls ferritin and transferrin receptor expression is conserved during evolution. Nucleic Acids Res. 1990;18:1175-1179.

226. Wei J, Theil EC. Identification and characterization of the iron regulatory element in the ferritin gene of a plant (soybean). J Biol Chem. 2000;275:17488-17493.

227. Pantopoulos K. Iron metabolism and the IRE/IRP regulatory system: an update. Ann N Y Acad Sci. 2004;1012:1-13.

228. Piccinelli P, Samuelsson T. Evolution of the iron-responsive element. RNA. 2007;13:952-966.

229. Ma J et al. Fe2+ binds iron responsive element-RNA, selectively changing protein-binding affinities and regulating mRNA repression and activation. Proc Natl Acad Sci U S A. 2012;109:8417-8422.

230. Cairo G, Pietrangelo A. Iron regulatory proteins in pathobiology. Biochem J. 2000;352 Pt 2:241-250.

231. Hanson ES, Leibold EA. Regulation of the iron regulatory proteins by reactive nitrogen and oxygen species. Gene Expr. 1999;7:367-376.

232. Hentze MW, Kühn LC. Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress. Proc Natl Acad Sci U S A. 1996;93:8175-8182.

233. Hentze MW, Muckenthaler MU, Andrews NC. Balancing acts: molecular control of mammalian iron metabolism. Cell. 2004;117:285-297.

234. Theil EC, Eisenstein RS. Combinatorial mRNA regulation: iron regulatory proteins and iso-iron-responsive elements (Iso-IREs). J Biol Chem. 2000;275:40659-40662.

235. Wallander ML, Leibold EA, Eisenstein RS. Molecular control of vertebrate iron homeostasis by iron regulatory proteins. Biochim Biophys Acta. 2006;1763:668-689.

236. Gunshin H et al. Iron-dependent regulation of the divalent metal ion transporter. FEBS Lett. 2001;509:309-316.

237. Cmejla R, Petrak J, Cmejlova J. A novel iron responsive element in the 3'UTR of human MRCKalpha. Biochem Biophys Res Commun. 2006;341:158-166.

238. Ke Y, Sierzputowska-Gracz H, Gdaniec Z, Theil EC. Internal loop/bulge and hairpin loop of the iron-responsive element of ferritin mRNA contribute to maximal iron regulatory protein 2 binding and translational regulation in the iso-iron-responsive element/iso-iron regulatory protein family. Biochemistry. 2000;39:6235-6242.

239. Bettany AJ, Eisenstein RS, Munro HN. Mutagenesis of the iron-regulatory element further defines a role for RNA secondary structure in the regulation of ferritin and transferrin receptor expression. J Biol Chem. 1992;267:16531-16537.

240. Butt J et al. Differences in the RNA binding sites of iron regulatory proteins and potential target diversity. Proc Natl Acad Sci U S A. 1996;93:4345-4349.

241. Henderson BR, Menotti E, Bonnard C, Kühn LC. Optimal sequence and structure of iron-responsive elements. Selection of RNA stem-loops with high affinity for iron regulatory factor. J Biol Chem. 1994;269:17481-17489.

242. Ke Y, Theil EC. An mRNA loop/bulge in the ferritin iron-responsive element forms in vivo and Was detected by radical probing with Cu-1,10-phenantholine and iron regulatory protein footprinting. J Biol Chem. 2002;277:2373-2376.

243. Sierzputowska-Gracz H, McKenzie RA, Theil EC. The importance of a single G in the hairpin loop of the iron responsive element (IRE) in ferritin mRNA for structure: an NMR spectroscopy study. Nucleic Acids Res. 1995;23:146-153.

244. Wang YH, Sczekan SR, Theil EC. Structure of the 5' untranslated regulatory region of ferritin mRNA studied in solution. Nucleic Acids Res. 1990;18:4463-4468.

245. Addess KJ, Basilion JP, Klausner RD, Rouault TA, Pardi A. Structure and dynamics of the iron responsive element RNA: implications for binding of the RNA by iron regulatory binding proteins. J Mol Biol. 1997;274:72-83.

246. Erlitzki R, Long JC, Theil EC. Multiple, conserved iron-responsive elements in the 3'-untranslated region of transferrin receptor mRNA enhance binding of iron regulatory protein 2. J Biol Chem. 2002;277:42579-42587.

247. Surdej P, Richman L, Kühn LC. Differential translational regulation of IRE-containing mRNAs in Drosophila melanogaster by endogenous IRP and a constitutive human IRP1 mutant. Insect Biochem Mol Biol. 2008;38:891-894.

248. Lind MI et al. Of two cytosolic aconitases expressed in Drosophila, only one functions as an iron-regulatory protein. J Biol Chem. 2006;281:18707-18714.

249. Dashti ZJ, Gamieldien J, Christoffels A. Computational characterization of Iron metabolism in the Tsetse disease vector, Glossina morsitans: IRE stem-loops. BMC Genomics. 2016;17:561.

250. Mellacheruvu D et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods. 2013;10:730-736.

251. Morris JH et al. Affinity purification-mass spectrometry and network analysis to understand protein-protein interactions. Nat Protoc. 2014;9:2539-2554.

252. Chung HR et al. The effect of micrococcal nuclease digestion on nucleosome positioning data. PLoS One. 2010;5:e15754.

253. Vernì F, Cenci G. The Drosophila histone variant H2A.V works in concert with HP1 to promote kinetochore-driven microtubule formation. Cell Cycle. 2015;14:577-588.

254. Kolkhof P et al. A Luciferase-fragment Complementation Assay to Detect Lipid Droplet-associated Protein-Protein Interactions. Mol Cell Proteomics. 2017;16:329-345.

255. Hanai K, Furuhashi H, Yamamoto T, Akasaka K, Hirose S. RSF governs silent chromatin formation via histone H2Av replacement. PLoS Genet. 2008;4:e1000011.

256. Thomas C et al. Hit and run versus long-term activation of PARP-1 by its different domains fine-tunes nuclear processes. Proc Natl Acad Sci U S A. 2019;116:9941-9946.

257. Thomas CJ et al. Kinase-mediated changes in nucleosome conformation trigger chromatin decondensation via poly(ADP-ribosyl)ation. Mol Cell. 2014;53:831-842.

258. Charlesworth A et al. Isolation and properties of Drosophila melanogaster ferritin--molecular cloning of a cDNA that encodes one subunit, and localization of the gene on the third chromosome. Eur J Biochem. 1997;247:470-475.

259. Georgieva T, Dunkov BC, Harizanova N, Ralchev K, Law JH. Iron availability dramatically alters the distribution of ferritin subunit messages in Drosophila melanogaster. Proc Natl Acad Sci U S A. 1999;96:2716-2721.

260. Lind MI, Ekengren S, Melefors O, Söderhäll K. Drosophila ferritin mRNA: alternative RNA splicing regulates the presence of the iron-responsive element. FEBS Lett. 1998;436:476-482.

261. Sanchez M et al. Iron regulatory protein-1 and -2: transcriptome-wide definition of binding mRNAs and shaping of the cellular proteome by iron regulatory proteins. Blood. 2011;118:e168-79.

262. Henderson BR, Menotti E, Kühn LC. Iron regulatory proteins 1 and 2 bind distinct sets of RNA target sequences. J Biol Chem. 1996;271:4900-4908.

263. Cairo G, Ronchi R, Recalcati S, Campanella A, Minotti G. Nitric oxide and peroxynitrite activate the iron regulatory protein-1 of J774A.1 macrophages by direct disassembly of the Fe-S cluster of cytoplasmic aconitase. Biochemistry. 2002;41:7435-7442.

264. Marcos R, Lloberas J, Creus A, Xamena N, Cabré O. Effect of cycloheximide on different stages of Drosophila melanogaster. Toxicol Lett. 1982;13:105-112.

265. Chafe SC, Pierce JB, Mangroo D. Nuclear-cytoplasmic trafficking of NTF2, the nuclear import receptor for the RanGTPase, is subjected to regulation. PLoS One. 2012;7:e42501.

266. Murphy GA et al. A T42A Ran mutation: differential interactions with effectors and regulators, and defect in nuclear protein import. Mol Biol Cell. 1997;8:2591-2604.

267. Dai F, Lin X, Chang C, Feng XH. Nuclear export of Smad2 and Smad3 by RanBP3 facilitates termination of TGF-beta signaling. Dev Cell. 2009;16:345-357.

268. Stüven T, Hartmann E, Görlich D. Exportin 6: a novel nuclear export receptor that is specific for profilin.actin complexes. EMBO J. 2003;22:5928-5940.

269. Ruth JL, Wassner SJ. Body composition: salt and water. Pediatr Rev. 2006;27:181-7; quiz 188.

270. Bourque CW. Central mechanisms of osmosensation and systemic osmoregulation. Nat Rev Neurosci. 2008;9:519-531.

271. Garty H, Benos DJ. Characteristics and regulatory mechanisms of the amiloride-blockable Na+ channel. Physiol Rev. 1988;68:309-373.

272. Garty H, Palmer LG. Epithelial sodium channels: function, structure, and regulation. Physiol Rev. 1997;77:359-396.

273. Palmer LG. Epithelial Na channels: function and diversity. Annu Rev Physiol. 1992;54:51-66.

274. Kashlan OB et al. Constraint-based, homology model of the extracellular domain of the epithelial Na+ channel α subunit reveals a mechanism of channel activation by proteases. J Biol Chem. 2011;286:649-660.

275. Kellenberger S, Schild L. International Union of Basic and Clinical Pharmacology. XCI. structure, function, and pharmacology of acid-sensing ion channels and the epithelial Na+ channel. Pharmacol Rev. 2015;67:1-35.

276. Weisz OA, Johnson JP. Noncoordinate regulation of ENaC: paradigm lost. Am J Physiol Renal Physiol. 2003;285:F833-42.

277. Büsst CJ. Blood pressure regulation via the epithelial sodium channel: from gene to kidney and beyond. Clin Exp Pharmacol Physiol. 2013;40:495-503.

278. Rossier BC, Baker ME, Studer RA. Epithelial sodium transport and its control by aldosterone: the story of our internal environment revisited. Physiol Rev. 2015;95:297-340.

279. Ghosh A, Boucher RC, Tarran R. Airway hydration and COPD. Cell Mol Life Sci. 2015;72:3637-3652.

280. Haq IJ, Gray MA, Garnett JP, Ward C, Brodlie M. Airway surface liquid homeostasis in cystic fibrosis: pathophysiology and therapeutic targets. Thorax. 2016;71:284-287.

281. Barrett KE. New ways of thinking about (and teaching about) intestinal epithelial function. Adv Physiol Educ. 2008;32:25-34.

282. Gaillard EA, Kota P, Gentzsch M, Dokholyan NV, Stutts MJ, Tarran R. Regulation of the epithelial Na+ channel and airway surface liquid volume by serine proteases. Pflugers Arch. 2010;460:1-17.

283. Asher C, Wald H, Rossier BC, Garty H. Aldosterone-induced increase in the abundance of Na+ channel subunits. Am J Physiol. 1996;271:C605-11.

284. Bhalla V, Hallows KR. Mechanisms of ENaC regulation and clinical implications. J Am Soc Nephrol. 2008;19:1845-1854.

285. Hanukoglu I, Hanukoglu A. Epithelial sodium channel (ENaC) family: Phylogeny, structure-function, tissue distribution, and associated inherited diseases. Gene. 2016;579:95-132.

286. Deval E, Lingueglia E. Acid-Sensing Ion Channels and nociception in the peripheral and central nervous systems. Neuropharmacology. 2015;94:49-57.

287. Lin SH, Sun WH, Chen CC. Genetic exploration of the role of acid-sensing ion channels. Neuropharmacology. 2015;94:99-118.

288. Omerbašić D, Schuhmacher LN, Bernal Sierra YA, Smith ES, Lewin GR. ASICs and mammalian mechanoreceptor function. Neuropharmacology. 2015;94:80-86.

289. Waldmann R, Lazdunski M. H(+)-gated cation channels: neuronal acid sensors in the NaC/DEG family of ion channels. Curr Opin Neurobiol. 1998;8:418-424.

290. Zelle KM, Lu B, Pyfrom SC, Ben-Shahar Y. The genetic architecture of degenerin/epithelial sodium channels in Drosophila. G3 (Bethesda). 2013;3:441-450.

291. Eastwood AL, Goodman MB. Insight into DEG/ENaC channel gating from genetics and structure. Physiology (Bethesda). 2012;27:282-290.

292. Gründer S, Assmann M. Peptide-gated ion channels and the simple nervous system of Hydra. J Exp Biol. 2015;218:551-561.

293. Eastwood AL, Goodman MB. Insight into DEG/ENaC channel gating from genetics and structure. Physiology (Bethesda). 2012;27:282-290.

294. Hughey RP et al. Maturation of the epithelial Na+ channel involves proteolytic processing of the alpha- and gamma-subunits. J Biol Chem. 2003;278:37073-37082.

295. Ji HL, Zhao R, Komissarov AA, Chang Y, Liu Y, Matthay MA. Proteolytic regulation of epithelial sodium channels by urokinase plasminogen activator: cutting edge and cleavage sites. J Biol Chem. 2015;290:5241-5255.

296. Kleyman TR, Carattino MD, Hughey RP. ENaC at the cutting edge: regulation of epithelial sodium channels by proteases. J Biol Chem. 2009;284:20447-20451.

297. Jasti J, Furukawa H, Gonzales EB, Gouaux E. Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. Nature. 2007;449:316-323.

298. Soundararajan R, Lu M, Pearce D. Organization of the ENaC-regulatory machinery. Crit Rev Biochem Mol Biol. 2012;47:349-359.

299. Heidrich E, Carattino MD, Hughey RP, Pilewski JM, Kleyman TR, Myerburg MM. Intracellular Na+ regulates epithelial Na+ channel maturation. J Biol Chem. 2015;290:11569-11577.

300. Rotin D, Kanelis V, Schild L. Trafficking and cell surface stability of ENaC. Am J Physiol Renal Physiol. 2001;281:F391-9.

301. Staub O et al. Regulation of stability and function of the epithelial Na+ channel (ENaC) by ubiquitination. EMBO J. 1997;16:6325-6336.

302. Staub O et al. Regulation of the epithelial Na+ channel by Nedd4 and ubiquitination. Kidney Int. 2000;57:809-815.

303. Butterworth MB, Edinger RS, Frizzell RA, Johnson JP. Regulation of the epithelial sodium channel by membrane trafficking. Am J Physiol Renal Physiol. 2009;296:F10-24.

304. Butterworth MB. Regulation of the epithelial sodium channel (ENaC) by membrane trafficking. Biochim Biophys Acta. 2010;1802:1166-1177.

305. Farinha CM, Matos P. Rab GTPases regulate the trafficking of channels and transporters - a focus on cystic fibrosis. Small GTPases. 2018;9:136-144.

306. Saxena S, Singh M, Engisch K, Fukuda M, Kaur S. Rab proteins regulate epithelial sodium channel activity in colonic epithelial HT-29 cells. Biochem Biophys Res Commun. 2005;337:1219-1223.

307. Loffing J et al. Differential subcellular localization of ENaC subunits in mouse kidney in response to high- and low-Na diets. Am J Physiol Renal Physiol. 2000;279:F252-8.

308. Masilamani S, Kim GH, Mitchell C, Wade JB, Knepper MA. Aldosterone-mediated regulation of ENaC alpha, beta, and gamma subunit proteins in rat kidney. J Clin Invest. 1999;104:R19-23.

309. Kleyman TR et al. Cell surface expression and turnover of the alpha-subunit of the epithelial sodium channel. Am J Physiol Renal Physiol. 2001;281:F213-21.

310. Prince LS, Welsh MJ. Effect of subunit composition and Liddle's syndrome mutations on biosynthesis of ENaC. Am J Physiol. 1999;276:C1346-51.

311. Weisz OA, Wang JM, Edinger RS, Johnson JP. Non-coordinate regulation of endogenous epithelial sodium channel (ENaC) subunit expression at the apical membrane of A6 cells in response to various transporting conditions. J Biol Chem. 2000;275:39886-39893.

312. Bazopoulou D, Tavernarakis N. Mechanosensitive Ion Channels in Caenorhabditis elegans. Curr Top Membr. 2007;59:49-79.

313. Ben-Shahar Y. Sensory functions for degenerin/epithelial sodium channels (DEG/ENaC). Adv Genet. 2011;76:1-26.

314. Liu L et al. Contribution of Drosophila DEG/ENaC genes to salt taste. Neuron. 2003;39:133-146.

315. Liu L, Johnson WA, Welsh MJ. Drosophila DEG/ENaC pickpocket genes are expressed in the tracheal system, where they may be involved in liquid clearance. Proc Natl Acad Sci U S A. 2003;100:2128-2133.

316. Studer RA, Person E, Robinson-Rechavi M, Rossier BC. Evolution of the epithelial sodium channel and the sodium pump as limiting factors of aldosterone action on sodium transport. Physiol Genomics. 2011;43:844-854.

317. Knight KK, Olson DR, Zhou R, Snyder PM. Liddle's syndrome mutations increase Na+ transport through dual effects on epithelial Na+ channel surface expression and proteolytic cleavage. Proc Natl Acad Sci U S A. 2006;103:2805-2808.

318. Loffing J et al. Aldosterone induces rapid apical translocation of ENaC in early portion of renal collecting system: possible role of SGK. Am J Physiol Renal Physiol. 2001;280:F675-82.

319. Salker MS et al. LEFTYA Activates the Epithelial Na+ Channel (ENaC) in Endometrial Cells via Serum and Glucocorticoid Inducible Kinase SGK1. Cell Physiol Biochem. 2016;39:1295-1306.

320. Sayegh R et al. Glucocorticoid induction of epithelial sodium channel expression in lung and renal epithelia occurs via trans-activation of a hormone response element in the 5'-flanking region of the human epithelial sodium channel alpha subunit gene. J Biol Chem. 1999;274:12431-12437.

321. Shigaev A, Asher C, Latter H, Garty H, Reuveny E. Regulation of sgk by aldosterone and its effects on the epithelial Na(+) channel. Am J Physiol Renal Physiol. 2000;278:F613-9.

322. Pearce D, Kleyman TR. Salt, sodium channels, and SGK1. J Clin Invest. 2007;117:592-595.

323. Bhalla V, Soundararajan R, Pao AC, Li H, Pearce D. Disinhibitory pathways for control of sodium transport: regulation of ENaC by SGK1 and GILZ. Am J Physiol Renal Physiol. 2006;291:F714-21.

324. King-Jones K, Thummel CS. Nuclear receptors--a perspective from Drosophila. Nat Rev Genet. 2005;6:311-323.

325. Verrey F. Early aldosterone action: toward filling the gap between transcription and transport. Am J Physiol. 1999;277:F319-27.

326. Grossmann C, Freudinger R, Mildenberger S, Krug AW, Gekle M. Evidence for epidermal growth factor receptor as negative-feedback control in aldosterone-induced Na+ reabsorption. Am J Physiol Renal Physiol. 2004;286:F1226-31.

327. Doll CJ, Hochachka PW, Reiner PB. Effects of anoxia and metabolic arrest on turtle and rat cortical neurons. Am J Physiol. 1991;260:R747-55.

328. Jiang C, Xia Y, Haddad GG. Role of ATP-sensitive K+ channels during anoxia: major differences between rat (newborn and adult) and turtle neurons. J Physiol. 1992;448:599-612.

329. Johnston IA. Anaerobic metabolism in the carp (Carassius carassius L.). Comp Biochem Physiol B. 1975;51:235-241.

330. Simon LM, Robin ED. Changes in heart and skeletal muscle cytochrome oxidase activity during anaerobiosis in the freshwater turtle Pseudemys scripta elegans. Comp Biochem Physiol. 1970;37:437-443.

331. HADDAD GG, WYMAN RJ, MOHSENIN A, SUN Y, KRISHNAN SN. Behavioral and Electrophysiologic Responses of Drosophila melanogaster to Prolonged Periods of Anoxia. J Insect Physiol. 1997;43:203-210.

332. Haddad GG, Sun Y, Wyman RJ, Xu T. Genetic basis of tolerance to O2 deprivation in Drosophila melanogaster. Proc Natl Acad Sci U S A. 1997;94:10809-10812.

333. Zhou D, Visk DW, Haddad GG. Drosophila, a golden bug, for the dissection of the genetic basis of tolerance and susceptibility to hypoxia. Pediatr Res. 2009;66:239-247.

334. Haddad GG, Gandhi MR, Mellins RB. Maturation of ventilatory response to hypoxia in puppies during sleep. J Appl Physiol Respir Environ Exerc Physiol. 1982;52:309-314.

335. Schwartz AJ et al. Hepatic hepcidin/intestinal HIF-2α axis maintains iron absorption during iron deficiency and overload. J Clin Invest. 2019;129:336-348.

336. Mastrogiannaki M, Matak P, Delga S, Deschemin JC, Vaulont S, Peyssonnaux C. Deletion of HIF-2α in the enterocytes decreases the severity of tissue iron loading in hepcidin knockout mice. Blood. 2012;119:587-590.

337. Wallace EM et al. A Small-Molecule Antagonist of HIF2α Is Efficacious in Preclinical Models of Renal Cell Carcinoma. Cancer Res. 2016;76:5491-5500.

338. Turi JL, Piantadosi CA, Stonehuerner JD, Ghio AJ. Iron accumulation in bronchial epithelial cells is dependent on concurrent sodium transport. Biometals. 2008;21:571-580.

339. Bischof J, Sheils EM, Björklund M, Basler K. Generation of a transgenic ORFeome library in Drosophila. Nat Protoc. 2014;9:1607-1620.

340. Ge DT, Tipping C, Brodsky MH, Zamore PD. Rapid Screening for CRISPR-Directed Editing of the Drosophila Genome Using white Coconversion. G3 (Bethesda). 2016;6:3197-3206.

341. Yuan XM, Li W, Baird SK, Carlsson M, Melefors O. Secretion of ferritin by iron-laden macrophages and influence of lipoproteins. Free Radic Res. 2004;38:1133-1142.

342. Hernández-Gallardo AK, Missirlis F. Loss of ferritin in developing wing cells: Apoptosis and ferroptosis coincide. PLoS Genet. 2020;16:e1008503.

343. Edinger RS, Bertrand CA, Rondandino C, Apodaca GA, Johnson JP, Butterworth MB. The epithelial sodium channel (ENaC) establishes a trafficking vesicle pool responsible for its regulation. PLoS One. 2012;7:e46593.

344. Ware AW, Cheung TT, Rasulov S, Burstein E, McDonald FJ. Epithelial Na$^+$ Channel: Reciprocal Control by COMMD10 and Nedd4-2. Front Physiol. 2018;9:793.

345. Akmammedov A, Geigges M, Paro R. Single vector non-leaky gene expression system for Drosophila melanogaster. Sci Rep. 2017;7:6899.

346. Liu Y, Lehmann M. A genomic response to the yeast transcription factor GAL4 in Drosophila. Fly (Austin). 2008;2:92-98.

347. Ohhara Y, Kobayashi S, Yamanaka N. Nutrient-Dependent Endocycling in Steroidogenic Tissue Dictates Timing of Metamorphosis in Drosophila melanogaster. PLoS Genet. 2017;13:e1006583.

348. Andrew DJ, Henderson KD, Seshaiah P. Salivary gland development in Drosophila melanogaster. Mech Dev. 2000;92:5-17.

349. Hochstrasser M. Chromosome structure in four wild-type polytene tissues of Drosophila melanogaster. The 87A and 87C heat shock loci are induced unequally in the midgut in a manner dependent on growth temperature. Chromosoma. 1987;95:197-208.

350. Webb CH, Lupták A. HDV-like self-cleaving ribozymes. RNA Biol. 2011;8:719-727.

351. Rewitz KF, Yamanaka N, Gilbert LI, O'Connor MB. The insect neuropeptide PTTH activates receptor tyrosine kinase torso to initiate metamorphosis. Science. 2009;326:1403-1405.

352. Komura-Kawa T et al. The Drosophila Zinc Finger Transcription Factor Ouija Board Controls Ecdysteroid Biosynthesis through Specific Regulation of spookier. PLoS Genet. 2015;11:e1005712.

353. Xue Z et al. CRISPR/Cas9 mediates efficient conditional mutagenesis in Drosophila. G3 (Bethesda). 2014;4:2167-2173.

354. Nicholson L, Singh GK, Osterwalder T, Roman GW, Davis RL, Keshishian H. Spatial and temporal control of gene expression in Drosophila using the inducible GeneSwitch GAL4 system. I. Screen for larval nervous system drivers. Genetics. 2008;178:215-234.

355. Ashburner M, Richards G. Sequential gene activation by ecdysone in polytene chromosomes of Drosophila melanogaster. III. Consequences of ecdysone withdrawal. Dev Biol. 1976;54:241-255.

356. Ono H et al. Spook and Spookier code for stage-specific components of the ecdysone biosynthetic pathway in Diptera. Dev Biol. 2006;298:555-570.

357. Sztal T, Chung H, Gramzow L, Daborn PJ, Batterham P, Robin C. Two independent duplications forming the Cyp307a genes in Drosophila. Insect Biochem Mol Biol. 2007;37:1044-1053.

358. Niwa R et al. CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila. J Biol Chem. 2004;279:35942-35949.

359. Warren JT et al. Phantom encodes the 25-hydroxylase of Drosophila melanogaster and Bombyx mori: a P450 enzyme critical in ecdysone biosynthesis. Insect Biochem Mol Biol. 2004;34:991-1010.

360. Niwa YS, Niwa R. Transcriptional regulation of insect steroid hormone biosynthesis and its role in controlling timing of molting and metamorphosis. Dev Growth Differ. 2016;58:94-105.

361. Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci U S A. 2015;112:3570-3575.

362. Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK. Structure of the multimodular endonuclease FokI bound to DNA. Nature. 1997;388:97-100.

363. Busson D, Pret AM. GAL4/UAS targeted gene expression for studying Drosophila Hedgehog signaling. Methods Mol Biol. 2007;397:161-201.

364. Pirrotta V. Vectors for P-mediated transformation in Drosophila. Biotechnology. 1988;10:437-456.

365. Stapleton M et al. The Drosophila gene collection: identification of putative full-length cDNAs for 70% of D. melanogaster genes. Genome Res. 2002;12:1294-1300.

366. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol. 2016;17:5-15.

367. Manivannan SN, Lai LB, Gopalan V, Simcox A. Transcriptional control of an essential ribozyme in Drosophila reveals an ancient evolutionary divide in animals. PLoS Genet. 2015;11:e1004893.

368. Arimbasseri AG, Rijal K, Maraia RJ. Transcription termination by the eukaryotic RNA polymerase III. Biochim Biophys Acta. 2013;1829:318-330.

369. Darnell JE. Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. RNA. 2013;19:443-460.

370. Lee RT, Ng AS, Ingham PW. Ribozyme Mediated gRNA Generation for In Vitro and In Vivo CRISPR/Cas9 Mutagenesis. PLoS One. 2016;11:e0166020.

371. Zhang T, Gao Y, Wang R, Zhao Y. Production of Guide RNAs in vitro and in vivo for CRISPR Using Ribozymes and RNA Polymerase II Promoters. Bio Protoc. 2017;7

372. Chavez A et al. Comparison of Cas9 activators in multiple species. Nat Methods. 2016;13:563-567.

373. Jia C et al. New applications of CRISPR/Cas9 system on mutant DNA detection. Gene. 2018;641:55-62.

374. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol. 2014;32:577-582.

375. Liu Z et al. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep. 2017;7:2193.

376. Banaszynski LA, Liu CW, Wandless TJ. Characterization of the FKBP.rapamycin.FRB ternary complex. J Am Chem Soc. 2005;127:4715-4721.

377. Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol. 2015;33:139-142.

378. Haghighat-Khah RE et al. Site-specific cassette exchange systems in the Aedes aegypti mosquito and the Plutella xylostella moth. PLoS One. 2015;10:e0121097.

379. Long D et al. In vivo site-specific integration of transgene in silkworm via PhiC31 integrase-mediated cassette exchange. Insect Biochem Mol Biol. 2013;43:997-1008.

380. Nimmo DD, Alphey L, Meredith JM, Eggleston P. High efficiency site-specific genetic engineering of the mosquito genome. Insect Mol Biol. 2006;15:129-136.

381. Poe AR et al. Robust CRISPR/Cas9-Mediated Tissue-Specific Mutagenesis Reveals Gene Redundancy and Perdurance in Drosophila. Genetics. 2019;211:459-472.

382. Jiang F, Doudna JA. CRISPR-Cas9 Structures and Mechanisms. Annu Rev Biophys. 2017;46:505-529.

383. Nishimasu H et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014;156:935-949.

384. Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor nuclease. Biotechniques. 2004;36:702-707.

385. Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev. 2013;65:1357-1369.

386. Chávez VM et al. The Drosophila disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. Development. 2000;127:4115-4126.

387. Siegmund T, Korge G. Innervation of the ring gland of Drosophila melanogaster. J Comp Neurol. 2001;431:481-491.

388. Ford D et al. Alteration of Drosophila life span using conditional, tissue-specific expression of transgenes triggered by doxycyline or RU486/Mifepristone. Exp Gerontol. 2007;42:483-497.

389. Osterwalder T, Yoon KS, White BH, Keshishian H. A conditional tissue-specific transgene expression system using inducible GAL4. Proc Natl Acad Sci U S A. 2001;98:12596-12601.

390. Chen J, Zheng XF, Brown EJ, Schreiber SL. Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. Proc Natl Acad Sci U S A. 1995;92:4947-4951.

391. Harding MW, Galat A, Uehling DE, Schreiber SL. A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase. Nature. 1989;341:758-760.

392. Wang Y, Wang F, Wang R, Zhao P, Xia Q. 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. Sci Rep. 2015;5:16273.

393. Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol Direct. 2011;6:38.

394. Makarova KS, Wolf YI, Koonin EV. Classification and Nomenclature of CRISPR-Cas Systems: Where from Here. CRISPR J. 2018;1:325-336.

395. Makarova KS et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol. 2019

396. Makarova KS et al. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol. 2015;13:722-736.

397. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci. 2009;34:401-407.

398. Meltzer H et al. Tissue-specific (ts)CRISPR as an efficient strategy for in vivo screening in Drosophila. Nat Commun. 2019;10:2113.

399. Xing HL et al. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol. 2014;14:327.

400. Port F et al. A large-scale resource for tissue-specific CRISPR mutagenesis in <em>Drosophila</em&gt. bioRxiv. 2019636076.

401. Epinat JC et al. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res. 2003;31:2952-2962.

402. Silva GH, Belfort M, Wende W, Pingoud A. From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes. J Mol Biol. 2006;361:744-754.

403. Mandell JG, Barbas CF. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. Nucleic Acids Res. 2006;34:W516-23.

404. Nakatsukasa T, Shiraishi Y, Negi S, Imanishi M, Futaki S, Sugiura Y. Site-specific DNA cleavage by artificial zinc finger-type nuclease with cerium-binding peptide. Biochem Biophys Res Commun. 2005;330:247-252.

405. Urnov FD et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005;435:646-651.

406. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol. 2013;14:49-55.

407. Miller JC et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29:143-148.

408. Abudayyeh OO et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016;353:aaf5573.

409. O'Connell MR. Molecular Mechanisms of RNA Targeting by Cas13-containing Type VI CRISPR-Cas Systems. J Mol Biol. 2019;431:66-87.

410. Bassett AR, Tibbit C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. 2013;4:220-228.

411. Bassett AR, Liu JL. CRISPR/Cas9 and genome editing in Drosophila. J Genet Genomics. 2014;41:7-19.

412. Brown K, Samarsky D. RNAi off-targeting: Light at the end of the tunnel. J RNAi Gene Silencing. 2006;2:175-177.

413. Perrimon N, Mathey-Prevot B. Matter arising: off-targets and genome-scale RNAi screens in Drosophila. Fly (Austin). 2007;1:1-5.

414. Bellen HJ et al. The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics. 2004;167:761-781.

415. Bischof J, Björklund M, Furger E, Schertel C, Taipale J, Basler K. A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. Development. 2013;140:2434-2442.

416. Kennerdell JR, Carthew RW. Heritable gene silencing in Drosophila using double-stranded RNA. Nat Biotechnol. 2000;18:896-898.

417. Yamamoto S et al. A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. Cell. 2014;159:200-214.

418. Gratz SJ, Wildonger J, Harrison MM, O'Connor-Giles KM. CRISPR/Cas9-mediated genome engineering and the promise of designer flies on demand. Fly (Austin). 2013;7:249-255.

419. Gratz SJ, Rubinstein CD, Harrison MM, Wildonger J, O'Connor-Giles KM. CRISPR-Cas9 Genome Editing in Drosophila. Curr Protoc Mol Biol. 2015;111:31.2.1-31.2.20.

420. Ren X et al. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. Proc Natl Acad Sci U S A. 2013;110:19012-19017.

421. Xu J et al. A Toolkit of CRISPR-Based Genome Editing Systems in Drosophila. J Genet Genomics. 2015;42:141-149.

422. Port F et al. A large-scale resource for tissue-specific CRISPR mutagenesis in Drosophila. Elife. 2020;9

423. Zirin J et al. Large-scale transgenic Drosophila resource collections for loss- and gain-of-function studies. bioRxiv. 2019852376.

424. Abudayyeh OO et al. A cytosine deaminase for programmable single-base RNA editing. Science. 2019;365:382-386.

425. Abudayyeh OO, Gootenberg JS, Kellner MJ, Zhang F. Nucleic Acid Detection of Plant Genes Using CRISPR-Cas13. CRISPR J. 2019;2:165-171.

426. Freije CA et al. Programmable Inhibition and Detection of RNA Viruses Using Cas13. Mol Cell. 2019;76:826-837.e11.

427. Cherbas L et al. Tools for Targeted Genome Engineering of Established Drosophila Cell Lines. Genetics. 2015;201:1307-1318.

428. Cherbas L et al. The transcriptional diversity of 25 Drosophila cell lines. Genome Res. 2011;21:301-314.

429. Laban A, Tobin JF, Curotto de Lafaille MA, Wirth DF. Stable expression of the bacterial neor gene in Leishmania enriettii. Nature. 1990;343:572-574.

430. Lacalle RA, Pulido D, Vara J, Zalacaín M, Jiménez A. Molecular analysis of the pac gene encoding a puromycin N-acetyl transferase from Streptomyces alboniger. Gene. 1989;79:375-380.

431. Lacalle RA, Tercero JA, Jiménez A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. EMBO J. 1992;11:785-792.

432. Huynh N, Wang S, King-Jones K. Spatial and temporal control of gene manipulation in Drosophila via drug-activated Cas9 nucleases. Insect Biochem Mol Biol. 2020103336.

433. Port F, Muschalik N, Bullock SL. Systematic Evaluation of Drosophila CRISPR Tools Reveals Safe and Robust Alternatives to Autonomous Gene Drives in Basic Research. G3 (Bethesda). 2015;5:1493-1502.

434. Wang Q et al. The CRISPR-Cas13a Gene-Editing System Induces Collateral Cleavage of RNA in Glioma Cells. Adv Sci (Weinh). 2019;6:1901299.

435. Caplen NJ, Parrish S, Imani F, Fire A, Morgan RA. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. Proc Natl Acad Sci U S A. 2001;98:9742-9747.

436. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001;411:494-498.

437. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391:806-811.

438. Jackson AL et al. Expression profiling reveals off-target gene regulation by RNAi. Nat Biotechnol. 2003;21:635-637.

439. Semizarov D, Frost L, Sarthy A, Kroeger P, Halbert DN, Fesik SW. Specificity of short interfering RNA determined through gene expression signatures. Proc Natl Acad Sci U S A. 2003;100:6347-6352.

440. Xu D et al. A CRISPR/Cas13-based approach demonstrates biological relevance of vlinc class of long non-coding RNAs in anticancer drug response. Sci Rep. 2020;10:1794.

441. Yan F, Wang W, Zhang J. CRISPR-Cas12 and Cas13: the lesser known siblings of CRISPR-Cas9. Cell Biol Toxicol. 2019;35:489-492.

442. Yang LZ et al. Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems. Mol Cell. 2019;76:981-997.e7.

443. Gray NK, Pantopoulos K, Dandekar T, Ackrell BA, Hentze MW. Translational regulation of mammalian and Drosophila citric acid cycle enzymes via iron-responsive elements. Proc Natl Acad Sci U S A. 1996;93:4925-4930.

444. Melefors O. Translational regulation in vivo of the Drosophila melanogaster mRNA encoding succinate dehydrogenase iron protein via iron responsive elements. Biochem Biophys Res Commun. 1996;221:437-441.

445. Chiaratti MR et al. Maternal transmission of mitochondrial diseases. Genet Mol Biol. 2020;43:e20190095.

446. Russell O, Turnbull D. Mitochondrial DNA disease-molecular insights and potential routes to a cure. Exp Cell Res. 2014;325:38-43.

447. Viscomi C, Zeviani M. Strategies for fighting mitochondrial diseases. J Intern Med. 2020

448. Chen Z et al. Genetic mosaic analysis of a deleterious mitochondrial DNA mutation in Drosophila reveals novel aspects of mitochondrial regulation and function. Mol Biol Cell. 2015;26:674-684.

449. Hill JH, Chen Z, Xu H. Selective propagation of functional mitochondrial DNA during oogenesis restricts the transmission of a deleterious mitochondrial variant. Nat Genet. 2014;46:389-392.

450. Ma H, Xu H, O'Farrell PH. Transmission of mitochondrial mutations and action of purifying selection in Drosophila melanogaster. Nat Genet. 2014;46:393-397.

451. Jepson JE, Savva YA, Yokose C, Sugden AU, Sahin A, Reenan RA. Engineered alterations in RNA editing modulate complex behavior in Drosophila: regulatory diversity of adenosine deaminase acting on RNA (ADAR) targets. J Biol Chem. 2011;286:8325-8337.

452. Buchman AB, Brogan DJ, Sun R, Yang T, Hsu PD, Akbari OS. Programmable RNA Targeting Using CasRx in Flies. CRISPR J. 2020;3:164-176.

453. Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. Science. 2018;360:439-444.

454. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F. Author Correction: SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc. 2020;15:1311.

455. Metsky HC, Freije CA, Kosoko-Thoroddsen T-SF, Sabeti PC, Myhrvold C. CRISPR-based COVID-19 surveillance using a genomically-comprehensive machine learning approach. bioRxiv. 20202020.02.26.967026.

Appendices

A.1 Modified protocol for the current embryo injection system

A.1.1 Materials

A.1.1.1 Fly stocks

Depending on experiment purposes, different fly strains were be used for injection. List of fly strains being used during my program can be found in Appendix A.3 and modified methods in corresponding chapters.

List of flies and corresponding applications in embryo injection

application	genotype	characteristics	source
cDNA transgenic line	*y,w^{1118},P(y[+t7.7]=nos-phiC31\int.NLS)X; P(y[+t7.7]=carryP)attP40*	2nd chromosome attP docking site for phiC31 integrase-mediated transformation	BestGene Inc.
	y,w^{1118},P(y[+t7.7]=nos-phiC31\int.NLS)X; P(y[+t7.7]=carryP)attP2	3rd chromosome attP docking site for phiC31 integrase-mediated transformation	BestGene Inc.
	w^{1118}	To be used in G0 cross	Bloomington #3605
	w^{1118}; roi/CyO; Sb/TM6B Hu, Tb	Balancers for generating stable lines	Nhan Huynh
	w^{1118}; roi/Sp; Sb/TM6B Hu, TB	Balancers for generating stable lines	Kirst King-Jones lab
	w^{1118}; roi/Sp; Sb/TM3 Ser^{1}	Balancers for generating stable lines	Kirst King-Jones lab
	w^{1118}; roi/Sp; Sb/TM3 Ser^{1}.GFP	Balancers for generating stable lines	Kirst King-Jones lab

gRNA transgenic line	$y^1,v^1,P(y[+t7.7]=nos-phiC31\backslash int.NLS)X;$ $P(y[+t7.7]=carryP)attP40$	2nd chromosome attP docking site for phiC31 integrase-mediated transformation	Bloomington stock center #25709
	$y^1,v^1,P(y[+t7.7]=nos-phiC31\backslash int.NLS)X;$ $P(y[+t7.7]=carryP)attP2$	3rd chromosome attP docking site for phiC31 integrase-mediated transformation	Bloomington stock center #25710.
	y^2,cho^2,v^1	To be used in G0 cross	NIG TBX-0004
	$y^2,cho^2,v^1; Sco/CyO$	Balancers for generating stable lines	NIG TBX-0007
	$y^2,cho^2,v^1; Sb/TM6B\ Hu\ Tb$	Balancers for generating stable lines	Nhan Huynh
CRISPR line	$y^1,M(vas-Cas9)ZH-2A$ $w^{1118}/FM7c$	Expresses Cas9 protein in the ovary or germ cells under control of vas regulatory system	Bloomington #51323
	$y^1,w^{1118};$ $M(nos-Cas9)/CyO$	Expresses Cas9 protein in the ovary or germ cells under control of nos regulatory system	GenetiVision

A.1.1.2   Reagents

. Fly cages (Diamed GEN59-100).

. Grape juice plates for embryo collection.

. Grape juice plates for embryo alignment.

. Embryo collection cages.

. Large weigh boats.

. Paint brushes.

. Fresh 50% bleach.

. Slides with coverslips attached.

. Injection needles (Borosilicate glass B120-90-10)

. Dissection forceps

. Extracted glue

. Slide holder cage

. Halocarbon oil S200, S700 and S400 (a combination of S200 and S700 at the ratio of 3:2 v:v)

. Vaseline

. Slides holder cages

A.1.2 Procedure

Plasmids to be injected were cultured in DH5α *E.coli* and isolated using plasmid midi kit (QIAGEN #12143) and eluted in nuclease-free water at the final concentration of 500 ng/µl transgenic constructs including cDNA, tissue-specific Cas9, conditional CRISPR gRNAs, crRNAs. For classic CRISPR approach, donor template plasmid was mixed together with gRNA plasmids at the final concentration of 500 ng/µl for donor template and 100 ng/µl for individual gRNA plasmids. Plasmids were store at -20^{0}C prior until needed.

In order to stabilize embryos on injection slides, glue was prepared by incubating double-sided tape (Scotch #665) with heptane (Sigma H9629) in a tiny hybridization glass bottle overnight until viscous enough. For proper embryo injection, I used the borosilicate glass tubes from Sutter instrument company with outer diameter of 1.2 mm, inner diameter of 0.90 mm and the length of 10cm (Sutter B120-90-10). Needles were prepared using the Flaming/Brown micropipetter puller P-87 from Sutter instrument company at different settings based on injection approach. For dechorionated embryos, the following settings were used: heat-590, pull-250, vel-250, time-170, ram-639, pressure 550, delay-1, looping-none. For non-dechorionated embryos, the settings were

changed to: heat-505, pull-20, vel-60, time-250, pressure-300, delay-n/a, looping-none. Needles were then broken opened by moving against the sharp broken slide edge. Using the gel-loading tips (Sigma CLS4853), needles were filled with 2-3 µl of plasmid and submerged in halo carbon oil until needed. During injection, when the needle is no longer good, the new needle is prepared following the same procedure.

4-8 bottles of newly eclosed flies for injection (1-2 days old) were combined into the fly cage and fed with yeast enriched fly food for 2 days. On the day of injection, petri discs containing fly food were replaced with grape juice plates and changed every 30-35 minutes for about 3 hours prior to injection to reduce embryos retention. On the day of injection, dechodrionation and desiccation time were evaluated prior to injection since these values vary based on fly genotypes and room humidity. Embryos were collected every 35 minutes on the grape juice agar plates and gently washed under tap water. Embryos were then dechodrionated using fresh 50% bleach for the right amount of time as described earlier. Bleached embryos were then washed twice in distilled water for 30 seconds and briefly dried using kimwipe. On the other hand, the non-dechodrionation approach will omit the bleaching steps described above. Approximately 50-70 embyos were then aligned on a cambered grape juice plates before being transferred to a glued injection slides and allowed to dry in the desiccation chamber (Sigma #456071) for the right amount of time, this period often lasts from 3-7 minutes depending on room humidity. Desiccated embryos were quickly covered with a drop of halocarbon oil 700S (Sigma #H8898), just enough to cover embryos. Each injection was done within 5-7 minutes, old embryos were screened and killed immediately. Injected embryos were fenced with Vaseline and covered with a thin layer of 200S halocarbon oil (Cedarlane #25073). Sometimes, halocarbon oil 400S can be replaced for both 700S and 200S. Injected embryos were allowed to recover in 18^0C incubator with high enough humidity.

About 24-48 hours after injection, embryos were checked, and any hatched larvae were transferred to vials containing fly food and raised in 25^0C with 60-70% humidity. Survived adults were used for later crosses to establish the stable lines carry successful transformants.

A.2   Generation of Cas9-compatible gRNA

1. Designing the construct:

1a. Finding target site(s) for gRNA:

→ Finding target site(s):

- o Option 1: Use FlyBase database: copy your sequence of interest and look for optimal target in http://tools.flycrispr.molbio.wisc.edu/targetFinder/index.php.

- o Option 2: Copy your gene location and look for optimal target in http://www.flyrnai.org/crispr/.

- o There might also be other options as CRISPR tools are being developed.

1b. Sequencing to verify if target site(s) exists:

→ Extract DNA from fly stocks to be used for embryo injection, G0 cross, and balancing to make sure target site(s) all exists in those lines. Those include:

    . Lines to be injected (your choice):   BL25709 (attP site on 2nd chromosome)

                                            BL25710 (attP site on 3rd chromosome)

    . Line for G0 cross: TBX-0004

    . Lines for balancing: TBX-0007, TBX-0008 or TBX-0009 for 2nd chromosome

                                    TBX-0010 for 3rd chromosome

→ Design PCR primers that cover region of interest, make sure target sites are inside that region.

→ Perform PCR + sequencing to verify the accuracy of target site(s).

2. Choose the appropriate gRNA:

I recommend choosing either pCFD5 or PG.gRNA, especially pCFD5 for most application.

Refer to the summary table for current available gRNA plasmids.

3. Prepare primers following this table:

plasmid	No. of gRNA	primer name	sequence
pCFD3	1	Top oligos	5'- GTC**G**-(N)$_{19/20}$ - 3'
		Bottom oligos	5'- AAAC-(N)$_{19/20}$ reverse complement - 3'
pCFD4	1	Top oligos	5'- agatatccgggtgaacttcg -(N)$_{19/20}$ - 3'
		Bottom oligos	5'- tgctatttctagctctaaaac-(N)$_{19/20}$ reverse complement - 3'
	2	Forward primer	5'- agatatccgggtgaacttcg -(N)$_{19/20}$ **for target No.1**- gttttagagctagaaatag - 3'
		Reverse primer	5'- tgctatttctagctctaaaac -(N)$_{19/20}$ reverse **complement for target No.2** - gacgttaaattgaaaatag - 3'
pCFD5 or PG.gRNA	1	Top oligos	5' – TGCA – (N)$_{20}$ – 3'
		Bottom oligos	5' – AAAC – (N)$_{20}$ reverse complement – 3'
	2	5' Forward primer	5' – TTCGATTCCCGGCCGATGC – A(N)$_{20}$ for **target No.1**– GTTTTAGAGCTAGAAATAGC – 3'
		3' Reverse primer	5' – CTATTTCTAGCTCTAAAAC – (N)$_{20}$ reverse complement **for target No.2** – TGCACCAGCCGGGAATCGAAC – 3'
	>2	5' Forward primer	5' – TTCGATTCCCGGCCGATGC – A(N)$_{20}$ **for target No.1**– GTTTTAGAGCTAGAAATAGC – 3'
		PCR1 reverse primer	5' – (N)$_{20}$ reverse complement **for target No.2** – TGCACCAGCCGGGAATCGAACCC 3'
		PCR2 forward primer	5' – (N)$_{20}$ **for target No.2** – GTTTTAGAGCTAGAAATAGCAAG – 3'
		PCR2 reverse primer	5' – (N)$_{20}$ reverse complement **for target No.3** - TGCACCAGCCGGGAATCGAACCC 3'

		Keep going	
		PCR n^{th} forward primer	5' – (N)$_{20}$ **for target No.nth** – GTTTTAGAGCTAGAAATAGCAAG – 3'
		3' Reverse primer	5' – CTATTTCTAGCTCTAAAAC – (N)$_{20}$ reverse complement **for target No.(n+1)th** – TGCACCAGCCGGGAATCGAAC – 3'

→ At the same time, design the map of your desire plasmid using <u>SnapGene Viewer</u>.

4. Prepare backbone (for pCFD3, pCFD4, pCFD5 and PG.gRNA):

→ No matter what gRNA plasmid was chosen, protocol for preparing backbone is the same.

→ Digest gRNA plasmid with BbsI restriction enzyme:

reagents	μL
10X CutSmart Buffer	2.0
gRNA plasmid	X (5μg)
BbsI HF (NEB R3539S)	1.0
Nuclease-free water	Y
Total	20.0

→ Incubate reaction overnight at 37^0C.

→ Optional: before proceeding to the next step, add 1μL of CIP (NEB M0290S) and incubate at 37^0C for 1 hour. This step can reduce self-ligation of the two phosphorylated ends of the backbone after digestion.

→ Purify digested backbone using either ethanol precipitation, gel extraction or magnetic beads. Either way worked in my hand. For tips on increasing gel extraction efficiency, talk to Nhan.

→ Elute backbone and bring to final concentration of 50-100ng/uL in water. Final amount will be good enough for several cloning reactions later.

5. Prepare gRNA fragment(s)

5a. Single gRNA cloning ONLY (apply for single gRNA into pCFD3, pCFD4, pCFD5 and PG.gRNA):

→ Re-suspend oligos in nuclease-free water to a concentration of 100µM.

→ Phosphorylation and annealing reaction:

reagents	µL
Top oligo (100µM)	1.0
Bottom oligo (100µM)	1.0
10X T4 ligation buffer (NEB)	1.0
Nuclease-free water	6.0
T4 Polynucleotide Kinase (NEB M0201S)	1.0
Total	10.0

→ Incubate reaction in a thermocycler:

T^{0}C	Time
37	30 min
95	5 min
Ramp down to 25^{0}C at the rate 5^{0}C/min	
4	Forever

→ Ligation reaction: at room temperature for 1 hour

reagents	µL
BbsI digested backbone (50ng/µL-100ng/µL)	1.0
Annealed oligos from previous step	1.0
10X T4 ligation buffer (NEB)	1.0
Nuclease-free water	6.0
T4 DNA ligase (NEB M0202S)	1.0
Total	10.0

→ Proceed to transformation. Note: all gRNA plasmids mentioned in this protocol are Ampicillin resistant.

5b. Double or multiple gRNAs cloning (apply for pCFD4, pCFD5, PG.gRNA)

→ Set up PCR reaction using High Fidelity Polymerase.

Example protocol

reagents	μL
nuclease-free water	10.75
5X Q5 reaction buffer	5.0
5X Q5 high GC enhancer buffer	5.0
10mM dNTPs	0.5
10uM forward primer	1.25
10uM reverse primer	1.25
gRNA plasmid (your choice) 10ng/μL	1.0
Q5 high fidelity DNA polymerase (NEB M0491S)	0.25
Total	25.0

→ Run PCR in a thermocycler:

step	T^{0}C	time
1	98	0:30
2	98	0:30
3	Annealing temperature	0:30
4	72	30 sec/kb
Go to step 2 25-35 cycles		
5	72	2:00
6	4	Forever

→ Check PCR product on gel

→ Perform DpnI digestion for PCR product overnight at 37^{0}C to eliminate original template.

→ Purify PCR fragment by ethanol precipitation

6. Generating final gRNA construct:

→ For multiple gRNA cloning (more than 4 fragments), it is recommended to do gene soeing to combine PCR fragments with overlap region and reduce number of fragments for Gibson reaction later.

→ Set up Gibson assembly reaction:

. Calculate amount of each fragment to be used: https://nebiocalculator.neb.com/#!/ligation

. Recommended ratio for PCR fragment: backbone = 3:1 to 5:1 (Molar ratio)

. Prepare Gibson reaction: This is an example of Gibson assembly for 4 PCR fragments into backbone

reagents	μL
PCR fragment 1	X
PCR fragment 2	Y
PCR fragment 3	Z
PCR fragment 4	W
Backbone (50ng/uL-100ng/uL)	1.0
1.33X Gibson master mix	3*(X+Y+Z+W+1)

. Incubate reaction in a thermocycler at 50^0C for 1-4 hours.

. Proceed to transformation. Note: all gRNA plasmids mentioned in this protocol are Ampicillin resistant.

7. Screening for positive cloning: Apply for all cloning, no matter it is single for multiple gRNAs.

→ Pick individual colonies and grow in individual liquid culture for at least 8 hours.

→ Purify plasmid and send samples for sequencing:

primer name	sequence (5' – 3')	plasmids

pCFDseq RP	GCA CAA TTG TCT AGA ATG CAT AC	pCFD3, pCFD4, pCFD5, PG.gRNA
U63seq FP	ACG TTT TAT AAC TTA TGC CCC TAA G	pCFD3, pCFD5
U61seq FP	GAC ACA GCG CGT ACG TCC TTC G	pCFD4
tRNA FP	ATA GTA CCC TGC CAC GGT AC	pCFD5, PG.gRNA

- Note: For multiple gRNA cloning case: pCFDseqRP and tRNA FP can cover a region up to 1.6kb, relevant for 6 gRNAs. For more than 6 gRNA cloning, these primers might not be able to cover the whole gRNA region, it is recommended to design your own primer to have a full sequence.

→ It is recommended to sequence at least 4 colonies. For multiple gRNA cloning, the more the better.

→ For MBSU sequencing, follow this link for more information:

https://www.ualberta.ca/biological-sciences/services/mbsu/what-we-do

→ If you want to do your own sequencing for the first time, contact Nhan for proper training.

8+9+10. Embryo injection and selection of positive transformants:

→ Follow modified protocol in Appendix A.1.

→ It is recommended to use injected male only and screen young adult (newly eclosed flies)

11. Balancing positive transformant(s)

→ Cross positive transformant(s) with balancer line(s): TBX-0007, TBX-0008, TBX-0009 or TBX-0010

→ Generate stable line via genetic cross series.

12. Verify transgenic line(s)

→ To verify insertion, perform PCR using the following primers:

 . attB_gypsy FP: 5' – GGC ATG TCG ACA AGC CGA ATT G – 3'

 . gypsy_attB RP: 5' – GAT CGG CTA AAT GGT ATG GC – 3'

→ To verify existence of gRNA, perform PCR, following by sequencing using primers in step 7.

A.3   Generation of Cas13-compatible crRNA

1. Designing the construct:

1a. Choosing target site(s) for crRNA:

→ Evaluate target transcript secondary structure using either or both of the following online tools:

RNAfold: http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi

RNAstructure: https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html

→ Optional: one can also use siRNA design tool RNAxs (http://rna.tbi.univie.ac.at/cgi-bin/RNAxs/RNAxs.cgi) to find the regions of transcripts with good accessibility.

1b. Sequencing to verify if your target site(s) exists:

→ Extract DNA from fly stocks used to perform embryo injection, G0 cross, and balancing to ensure target site(s) all exists in those lines.

For our case, we used the following:

    . Lines to be injected: # 25709 (Bloomington) (attP site on 2nd chromosome)

                                   #25710 (Bloomington) (attP site on 3rd chromosome)

    . Line for G0 cross: TBX-0004 (Japan National Institute of Genetics_NIG)

    . Lines for balancing: TBX-0007 (NIG)

                                   TBX-0008 (NIG)

                                   TBX-0009 (NIG)

                                   TBX-0010 (NIG)

→ Design PCR primers that cover region of interest, ensuring that target sites are inside that region.

→ Perform PCR + sequencing to verify the accuracy of your target site(s).

2. Design crRNA primers:

<u>2a. Single crRNA:</u>

→ Order standard desalted oligos:

  For pC13B:

    Forward primer: 5' - CACC (antisense sequence)$_{30}$ - 3'

    Reverse primer: 5' - CAAC (sense sequence)$_{30}$ - 3'

  For pC13X:

    Forward primer: 5' - AAAC (antisense sequence)$_{30}$ - 3'

    Reverse primer: 5' - AAAA (sense sequence)$_{30}$ - 3'

<u>2b. 2 crRNAs:</u>

→ Order oligos:

  <u>For pC13B</u>

- Forward primer: 5' - GTTCGATTCCCGGCCGATGCcacc (antisense sequence 1)$_{30}$ GTTGGGACTGCTCTCACTTTG - 3'

- Reverse primer: 5' - CAAAGTGAGAGCAGTCCCAAC (sense sequence 2)$_{30}$ ggtgTGCACCAGCCGGGAATC - 3'

  <u>For pC13X</u>

- Forward primer: 5' - AACCCCTACCAACTGGTCGGGGTTTGAAAC (antisense sequence 1)$_{30}$ taacaaagcaccagtggtctag - 3'

- Reverse primer: 5' - aggctccaggtaggcaaaaaa (sense sequence 2)$_{30}$ GTTTCAAACCCCGACCAGTTG - 3'

3c. ≥ 3 crRNAs

→ Order oligos:

  For pC13B

- Distal forward primer: 5' - GTTCGATTCCCGGCCGATGCcacc (antisense sequence 1)30 GTTGGGACTGCTCTCACTTTG - 3'

- F1 reverse primer: 5' - (sense sequence 2)30 ggtgTGCACCAGCCGGGAATC - 3'

- F2 forward primer: 5' - (antisense sequence 2)30 GTTGGGACTGCTCTCACTTTGAAGGGTATTCACAACtaacaaagcaccagtggtc - 3'

- F2 reverse primer: 5' - (sense sequence 2)30 ggtgTGCACCAGCCGGGAATC- 3'

- Fn forward primer: 5' - (antisense sequence (n-1))30 GTTGGGACTGCTCTCACTTTGAAGGGTATTCACAACtaacaaagcaccagtggtc - 3'

- Distal reverse primer: 5' - CAAAGTGAGAGCAGTCCCAAC (sense sequence n)30 ggtgTGCACCAGCCGGGAATC - 3'

  For pC13X

- Distal forward primer: 5' - AACCCCTACCAACTGGTCGGGGTTTGAAAC (antisense sequence 1)$_{30}$ taacaaagcaccagtggtctag - 3'

- F1 reverse primer: 5' - (sense sequence 2)30 GTTTCAAACCCCGACCAGTTG- 3'

- F2 forward primer: 5' - (antisense sequence 2)30 taacaaagcaccagtggtctagtg - 3'

- F2 reverse primer: 5' - (sense sequence 2)30 GTTTCAAACCCCGACCAGTTG- 3'

- Fn forward primer: 5'-(antisense sequence (n-1))30 taacaaagcaccagtggtctagtg-3'

- Distal reverse primer: 5' - aggctccaggtaggcaaaaaa (sense sequence n)$_{30}$ GTTTCAAACCCCGACCAGTTG - 3'

406

3. Prepare backbone:

→ Digest pC13B or pC13X plasmid with BbsI restriction enzyme:

reagents	amount added
10x CutSmart buffer	2.0 μl
plasmid	X (5 μg)
BbsI HF (NEB R3539S)	1.0 μl
nuclease-free water	Y μl (to total volume of 20 μl)
total	20.0 μl

→ Incubate reaction overnight at 37°C.

→ Purify digested backbone using either gel extraction or magnetic beads method.

→ Elute backbone to the final concentration of 50-100 ng/μl in nuclease-free water.

4. Prepare crRNA fragment(s):

4a. Single crRNA cloning

→ Resuspend oligos in nuclease-free water to a concentration of 100 μM

→ Phosphorylation and annealing reaction

reagents	amount added
forward primer (100 μM)	1.0 μl
reverse primer (100 μM)	1.0 μl
10x T4 ligation buffer (NEB)	1.0 μl
nuclease-free water	6.0 μl
T4 Polynucleotide Kinase (NEB M0201S)	1.0 μl
total	10.0 μl

- Incubate reaction in a thermocycler:

temperature (^{0}C)	time
37	30 minutes
95	5 minutes
ramp down to 25^{0}C at the rate 5^{0}C/minute	
4	forever

- Ligation reaction is done at room temperature for at least 1 hour:

reagents	amount added
BbsI-digested plasmid backbone	1.0 µl (50-100 ng)
annealed oligos	1.0 µl
10X T4 ligation buffer (NEB)	1.0 µl
nuclease-free water	6.0 µl
T4 DNA ligase (NEB M0202S)	1.0 µl
total	10.0 µl

→ Proceed to transformation.

4b. Double or multiple crRNAs

→ Set up PCR reaction using high fidelity polymerase

reagents	amount added (µl)
nuclease-free water	10.75
5x Q5 reaction buffer	5.0
5x Q5 high GC enhancer buffer	5.0
10mM dNTPs	0.5
10µM forward primer	1.25
10µM reverse primer	1.25
crRNA plasmid backbone (10ng/µl)	1.0
Q5 high fidelity DNA polymerase	0.25
Total	25.0

→ Run PCR in a thermocycler

step	temperature (^{0}C)	time
1	98	0:30
2	98	0:15
3	annealing temperature	0:15
4	72	30 sec / kb
go to step 2 for 25-35 cycles		
5	72	2:00
6	4	forever

→ Check PCR on electrophoresis gel

→ Perform DpnI digestion overnight at 37^{0}C or gel extraction to eliminate original template.

5. Generation of final crRNA construct (2 or more crRNAs)

→ Set up Gibson assembly reaction:

    . Calculate amount of each fragment: : https://nebiocalculator.neb.com/#!/ligation

    . Recommended ratio for PCR fragment : digested backbone = 3:1 to 5:1 (molar ratio)

    . Prepare Gibson reaction: This is an example of Gibson assembly for 4 PCR fragments

into the same backbone

reagents	amount added (µl)
PCR fragment 1	X
PCR fragment 2	Y
PCR fragment 3	Z
PCR fragment 4	W
digested backbone	1.0 (50-100ng)

409

| 1.33x Gibson assembly | $3 * (X + Y + Z + W + 1)$ |

. Incubate reaction in a thermocycler at 50^0C for 1-4 hours.

$\rightarrow$ Proceed to transformation.

$\rightarrow$ To screen for positive cloning, use the following primer:

5' - ACGTTTTATAACTTATGCCCCTAAG - 3'

A.4 Whole body mass spectrometry via Flag-tagged protein immunoprecipitation

A.4.1 Reagents

Prepare the following reagents:

1. 2x Buffer G:

component	stock concentration	final concentration	amount to be added
Na-HEPES, pH 7.5	1.0M	50mM	5.0mL
NaCl	5.0M	150mM	3.0mL
EDTA	0.5M	1mM	0.2mL
MiliQ water			To 50mL

→ Filter-sterilize

→ Store at 4^0C

2. 1x Lysis Buffer:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	10.0mL
Triton X-100		0.1%	0.2mL
Glycerol	50%	10%	4.0mL
Protease inhibitor cocktail	25X	1X	0.8mL
MiliQ water			To 20mL

3. Wash Buffer 1:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	5.0mL
Triton X-100		0.1%	0.1mL
Glycerol	50%	5%	1.0mL
MiliQ water			To 10mL

4. Wash Buffer 2:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	5.0mL
Glycerol	50%	5%	1.0mL
MiliQ water			To 10mL

A.4.2   During experiment:

→ Wear Latex gloves and change frequently

→ Work in laminar flow chamber whenever possible

→ Wear lab coat during the experiment.

A.4.3   Sample collection

→ At least 2 groups of samples to be collected: w^{1118} (control) and experimental line carries tagged version of target protein.

→ Collect 40hr L3 larva in 1X PBS, leave animal in the buffer for 2-5 minutes. Animals will vomit food from gut.

→ Transfer animals to new 1X PBS two more times to ensure the gut is clean.

→ Transfer animals to 1.5mL tube contain ice-cold 1x PBS pH 7.4 with 0.1% Triton X-100 using either clean forceps or cut P-1000 pipette tips (I prefer the first way) .

→ Prepare fresh 1X PBS with 0.2% Formaldehyde (fixative).

→ Remove as much PBST from sample as possible

→ Add fresh fixative to dissected BRGC and incubate at room temperature,

→ Mix gently on a rotary shaker for 10 min.

→ Let sample sink to the bottom of the tube.

→ Remove fixative solution as soon as possible

→ Add 0.25M Glycine in PBST to quench crosslinking reaction.

→ Incubate on shaker for 5 min.

→ Remove quench solution

→ Wash sample for PBST three times by gently inverting the tube 5 times.

→ Before storage, remove as much PBST as possible

→ Flash freeze samples using liquid Nitrogen and store in -80^{0}C until getting enough samples.

→ Aim: collect at least 150 larvae for each sample.

A.4.4   Protein preparation using anti-Flag M2 affinity gel

→ Remove samples from -80^{0}C freezer and thaw on ice.

→ At the same time, pre-rinse Dounce homogenizer with lysis buffer

→ Combine samples if needed and homogenize using a Dounce homogenizer in 1mL lysis buffer

on ice for about 10 strokes.

→ Repeat homogenization every 15 minutes for about 1 hour.

→ Transfer lysate to 1.5mL centrifuge tubes

→ Incubate for 15 min to ensure thorough lysis

→ Centrifuge lysate at 16,000 x g for 30 min at 4^{0}C

→ While waiting for centrifuge, prepare anti-Flag M2 affinity gel.

→ Resuspend gel well

→ Pipet 40uL of anti-Flag M2 gel into spin column using cut pipette tip

→ Wash gel by add 300uL of lysis buffer

→ Spin down quickly and discard supernatant

→ Repeat washing for another 4 times,

→ For the 6th time, plug the spin column bottom part and add 300uL of lysis buffer, store on ice until needed.

→ Right before use of gel in spin column, remove plug and spin column and discard supernatant, place the plug back.

→ After lysate centrifuge, transfer supernatant into the spin column, try not to interfere the fat layer on top. Leave some lysate to measure protein concentration

→ Incubate on a rotating platform for 2 hours at 4^0C

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash column using 300uL wash buffer 1

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash for total 3 times

→ Wash column using 300uL wash buffer 2

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash for total 3 times

→ Add 30-40uL elution buffer

→ Boil at 95-100^{0}C for 5 minutes, note: slightly release column cap otherwise elution buffer will boil and leak thru bottom plug.

→ Quick centrifuge at 12,000 x g and collect the flow through, this is your final sample.

→ Prepare fresh gel prior to run

→ Load sample and run gel for 1cm on separating gel (concentration of your choice, I used 12.5% in my experiment)

→ Proceed to commassie blue staining: https://www.cytographica.com/lab/protocols/gel_destain.html

→ When samples are ready to be submitted for analyzing:

     ✎ Fill out this form: http://apm.biochem.ualberta.ca/forms/In%20Gel%20Protein%20ID.pdf

     ✎ Email form to this email: jmoore@ualberta.ca (Jack Moore)

     ✎ Submit sample at 4096 Katz Group Centre For Pharmacy and Health Research Building (KGR).

A.4.5   Protein preparation using IgG magnetic beads

→ Remove samples from -80^{0}C freezer and thaw on ice.

→ At the same time, pre-rinse Dounce homogenizer with lysis buffer

→ Combine samples if needed and homogenize using a Dounce homogenizer in 1mL lysis buffer on ice for about 10 strokes.

→ Repeat homogenization every 15 minutes for about 1 hour.

→ Transfer lystate to 1.5mL centrifuge tubes

→ Incubate for 15 min to ensure thorough lysis

→ Centrifuge lysate at 16,000 x g for 30 min at 4^{0}C

→ While waiting for centrifuge, prepare IgG magnetic beads: https://www.thermofisher.com/order/catalog/product/10004D?SID=srch-srp-10004D

→ Prepare fresh gel prior to run

→ Load sample and run gel for 1cm on separating gel (concentration of your choice, I used 12.5% in my experiment)

→ Proceed to commassie blue staining: https://www.cytographica.com/lab/protocols/gel_destain.html

→ When samples are ready to be submitted for analyzing:

  ✍ Fill out this form: http://apm.biochem.ualberta.ca/forms/In%20Gel%20Protein%20ID.pdf

  ✍ Email form to this email: jmoore@ualberta.ca (Jack Moore)

  ✍ Submit sample at 4096 Katz Group Centre For Pharmacy and Health Research Building (KGR).

A.5 Prothoracic gland-specific mass spectrometry using Flag-tagged protein immunoprecipitation

A.5.1 Reagents

Prepare the following reagents:

1. 2x Buffer G:

component	stock concentration	final concentration	amount to be added
Na-HEPES, pH 7.5	1.0M	50mM	5.0mL
NaCl	5.0M	150mM	3.0mL
EDTA	0.5M	1mM	0.2mL
MiliQ water			To 50mL

→ Filter-sterilize

→ Store at 4^0C

2. 1x Lysis Buffer:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	10.0mL
Triton X-100		0.1%	0.2mL
Glycerol	50%	10%	4.0mL
Protease Inhibitor cocktail	25X	1X	0.8mL
MiliQ water			To 20mL

3. Wash Buffer 1:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	5.0mL
Triton X-100		0.1%	0.1mL
Glycerol	50%	5%	1.0mL

component	stock concentration	final concentration	amount to be added
MiliQ water			To 10mL

4. Wash Buffer 2:

component	stock concentration	final concentration	amount to be added
2x Buffer G	2x	1x	5.0mL
Glycerol	50%	5%	1.0mL
MiliQ water			To 10mL

5. *Ex vivo* media: Make fresh before use

Schneider medium containing 10% Heat inactivated FBS, 1% Streptomycine-Penicilin, insulin (10ug/mL) and ecdysone (2ug/mL).

6. 1x PBS with 0.1% Triton (PBST)

7. Fixative (make fresh before use): 0.02% Formaldehyde in 1xPBST

8. Quenching solution

→ 20mL of 1xPBST

→ 0.75g Glycine

→ 1xPBST to 40mL final volume to make quenching solution (0.25M Glycine in 40mL 1xPBST)

→ Store at room temperature

9. PBST + 25% Glycerol

10. Calcium-Magnesium free (CMF) buffer:

component	stock concentration	final concentration	amount to be added
Na-HEPES, pH 7.5	1.0M	15mM	15mL
NaH_2PO_4	Powder	400mg/L	0.040g
NaCl	Powder	800mg/L	0.080g
KCl	Powder	1200mg/L	0.120g
$NaHCO_3$	Powder	800mg/L	0.080g

Glucose	Powder	240mg/L	0.024g
BSA	Powder	1%	1.000g
MiliQ water			To 100mL

→ Filter-sterilize + aliquot

→ Store at 4^0C

11. Cells dissociation buffer:

→ 2mL CMF buffer

→ 5mg collagenase

→ 5mg papain

→ Mix well

→ CMF buffer to 5mL

A.5.2   Safety

→ Wear Latex gloves and change frequently

→ Work in laminar flow chamber whenever possible

→ Wear lab coat during the experiment.

A.5.3   Sample preparation:

→ Fly line carry tagged version of your protein.

→ w^{1118}.

→ If planning for Prothoracic Gland (PG) cells MS: combine phm22, UAS-mCD8.GFP (or UAS-mCD8.RFP) with the line carries your target proteins.

→ Grow samples following the desire of your experiment (for example: L2 or L3 larvae? 12hr or 36hr after L2/L3 molting?)

A.5.4   **Sample collection:**

→ Remove staged larvae from food and **wash** in *ex vivo* media containing 10% Heat inactivated FBS, 1% Streptomycine-Penicilin, insulin (10ug/mL) and ecdysone (2ug/mL) for three times (3 minutes each) with gentle shaking.

→ Dissect brain-ring gland complex (BRC) in *ex vivo* media, quickly transfer dissected samples into 1.5mL collection tubes contains 500uL *ex vivo* media with 1x protease inhibitor at room temperature. Finish **dissection within 1 hour** to avoid any physiological change. (It's already been proved this approach will not affect anything for about 48 hours, so don't be too stressful about time).

→ Remove as much media as possible without losing too many samples (you know, losing a lot means you have to collect a lot more).

→ **Wash** dissected samples in *ex vivo* media with 1x protease inhibitor for 3 minutes with gentle shaking.

→ Remove as much media as possible without losing too many samples.

→ **Wash** dissected samples in PBST for 5 minutes, let samples sink to bottom of tube.

→ Remove as much solution as possible without losing too many samples (you know, losing a lot means you have to collect a lot more).

→ **Wash** dissected samples in PBST for 3 minutes, let samples sink to bottom of tube.

→ Remove as much solution as possible without losing too many samples (you know, losing a lot means you have to collect a lot more).

→ Prepare fresh fixative

→ **Add fixative** to your samples, incubate at room temperature for exactly **10 minutes** (incubation + let samples sit) by gently mixing. Fixing for more than 10 minutes may increase non-specific binding, make you work more difficult later.

→ Remove as much solution as possible.

→ **Add quenching solution** to your samples, incubate at room temperature for exactly **5 minutes** (incubation + let samples sit) by gently mixing.

→ Remove as much solution as possible.

→ **Wash** dissected samples in PBST three times, let samples sink to bottom of tube.

→ Remove as much solution as possible.

→ Add **PBST + 25% Glycero**l to sample.

→ **Flash freeze** samples using liquid Nitrogen and store in -80^{0}C until getting enough samples.

→ Aim: collect at least 500-750 uL of each sample (not including PBST 25% Glycerol).

   . If plan to use BRC for MS: at least 300uL is recommended (roughly 400-500 BRC), proceed to protein extraction

   . If plan to purify prothoracic gland cells only, the more, the better. Proceed to step E.

A.5.5   Cells dissociation:

→ Remove preserved samples form -80^{0}C

→ Thaw samples through a graded thawing process

   ⇨ -20^{0}C for 15 minutes

⇨ 4^{0}C until completely thaw (10-15 minutes, depending on how much PBST 25% Glycerol has been added)

→ Combine samples into a same collection tube (5mL tube is recommended) using cut pipette tips (wait, I mean samples of the same kind, don't mix your experimental samples with control samples)

→ Remove as much solution as possible.

→ Add cells dissociation buffer enough to cover your samples.

→ Incubate at 30^{0}C for 30 min with gentle shaking.

→ To stop your experiment, add 4x volume of CMF as your cells dissociation buffer.

→ Let samples sit at room temperature for 5 minutes.

→ Centrifuge at the speed of 1000 x g for 1 minutes (room temperature)

→ Remove as much solution as possible without affecting cells

→ Wash cells in PBST 3 times

→ Proceed to step F.

A.5.6   PG cells isolation:

. Note: PG cells contain mCD8-eGFP (or RFP, depending on your choice), mCD8 is a mouse protein, this will make the purification step easier.

→ Prepare IgG beads following manufacture protocol

→ Incubate secondary anti-mouse antibody with beads

→ Crosslink using BS3

→ Add to cells population

→ Incubate with gentle shaking for 30 minutes

→ Purify using magnetic stand.

→ Wash

→ Purify using elution buffer

→ Centrifuge, collect PG cells

→ Proceed to protein preparation (step G).

A.5.7   Protein preparation using anti-Flag M2 affinity gel

✍ Pre-rinse Dounce homogenizer with lysis buffer

→ Combine samples if needed and homogenize using a Dounce homogenizer in 1mL lysis buffer on ice for about 10 strokes.

→ Repeat homogenization every 15 minutes for about 1 hour.

→ Transfer lysate to 1.5mL centrifuge tubes

→ Incubate for 15 min to ensure thorough lysis

→ Centrifuge lysate at 16,000 x g for 30 min at 4^0C

→ While waiting for centrifuge, prepare anti-Flag M2 affinity gel.

→ Resuspend gel well

→ Pipet 40uL of anti-Flag M2 gel into spin column using cut pipette tip

→ Wash gel by add 300uL of lysis buffer

→ Spin down quickly and discard supernatant

→ Repeat washing for another 4 times,

→ For the 6th time, plug the spin column bottom part and add 300uL of lysis buffer, store on ice until needed.

→ Right before use of gel in spin column, remove plug and spin column and discard supernatant, place the plug back.

→ After lysate centrifuge, transfer supernatant into the spin column, try not to interfere the fat layer on top. Leave some lysate to measure protein concentration

→ Incubate on a rotating platform for 2 hours at 4^0C

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash column using 300uL wash buffer 1

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash for total 3 times

→ Wash column using 300uL wash buffer 2

→ Centrifuge at 12,000 x g for 10 sec

→ Discard the flow thru

→ Wash for total 3 times

→ Add 30-40uL elution buffer

→ Boil at 95-100^0C for 5 minutes, note: slightly release column cap otherwise elution buffer will boil and leak thru bottom plug.

→ Quick centrifuge at 12,000 x g and collect the flow thru, this is your final sample.

→ Prepare fresh gel prior to run

→ Load sample and run gel for 1cm on separating gel (concentration of your choice, I used 12.5% in my experiment)

$\rightarrow$ Proceed to commassie blue staining: https://www.cytographica.com/lab/protocols/gel_destain.html

$\rightarrow$ When samples are ready to be submitted for analyzing:

$\Rightarrow$ Fill out this form: http://apm.biochem.ualberta.ca/forms/In%20Gel%20Protein%20ID.pdf

$\Rightarrow$ Email form to this email: jmoore@ualberta.ca (Jack Moore)

$\rightarrow$ Submit sample at 4096 Katz Group Centre For Pharmacy and Health Research Building (KGR).

A.6 RNA-immunoprecipitation (RIP) protocol

. Antibody being used: anti-Flag antibody

. Protein being IP: flag-tagged protein

. Tissue type: whole L3 larval body (40 hours after L2/L3 molt)

For other proteins, other antibodies or other tissue types, modification might be needed.

A.6.1 Reagents

1. Lysis buffer:

component	stock concentration	final concentration
KCl		150mM
Tris-HCl pH 7.4		25mM
EDTA		5mM
NP40		0.5% v/v
DTT		5mM
PMSF		1mM
Proteinase inhibitor (add fresh)	25X	1X
RNAse inhibitor (add fresh)		100U/mL

2. Wash buffer

component	stock concentration	final concentration
KCl		150mM
Tris-HCl pH 7.4		25mM
EDTA		5mM
Proteinase inhibitor (add fresh)	25X	1X
RNAse inhibitor (add fresh)		100U/mL

A.6.2   Samples preparation

→ At least 2 groups of samples to be collected: w^{1118} (control) and experimental line carries tagged version of target protein.

→ Collect 500 40hr L3 larvae per replicate with 3 replicates per condition. Number can vary based on experiment purposes.

→ Leave animals in 1X PBS for 2-5 minutes. Animals will vomit food from gut.

→ Transfer animals to new 1X PBS two more times to ensure the gut is clean.

→ Transfer animals to 1.5mL tube contain ice-cold 1x PBS pH 7.4 with 0.1% Triton X-100 using either clean forceps or cut P-1000 pipette tips (I prefer the first way) .

→ Prepare fresh 1X PBS with 0.2% Formaldehyde (fixative).

→ Remove as much PBST from sample as possible

→ Add fresh fixative to dissected BRGC and incubate at room temperature,

  ⇨ Mix gently on a rotary shaker for 10 min.

→ Let sample sink to the bottom of the tube.

→ Remove fixative solution as soon as possible

→ Add 0.25M Glycine in PBST to quench crosslinking reaction.

→ Incubate on shaker for 5 min.

→ Remove quench solution

→ Wash sample for PBST three times by gently inverting the tube 5 times.

→ Before storage, remove as much PBST as possible

→ Flash freeze samples using liquid Nitrogen and store in -80^{0}C until getting enough samples.

A.6.3 Protein extraction and immunoprecipitation

→ Remove samples from -80^{0}C freezer and thaw on ice.

→ At the same time, pre-rinse Dounce homogenizer with ice-cold lysis buffer

→ Combine samples if needed and homogenize using a Dounce homogenizer in 1mL lysis buffer on ice for about 10 strokes.

→ Repeat homogenization every 15 minutes for about 1 hour. Samples need to be kept on ice all the time

→ Transfer lysate to pre-chilled 1.5mL centrifuge tubes

→ Incubate for another 30 min on shaker to ensure thorough lysis

→ Centrifuge for 30 min at 12,000 x g at 4^{0}C.

→ Transfer supernatants to new pre-chilled 1.5mL centrifuge tubes without disturbing the pellets.

→ Filter supernatants through a 0.45μm syringe filter and collect through new pre-chilled 1.5mL centrifuge tubes.

→ Prepare anti-Flag M2 affinity gel by well resuspension in ice-cold lysis buffer.

→ Get the right amount needed and equilibrate gel by washing with lysis buffer with 10x volume for 3 times.

→ Remove the old lysis buffer and add the same amount of buffer as the original amount of gel being used.

→ Mix the lysate with 300 μL of equilibrated anti-Flag gel. This results in a total reaction volume of approximately 1.2-1.3mL.

→ Add 13 μL of DNAse H.

→ Incubate with gentle shaking at 4^{0}C overnight (8-10 hours).

→ Wash with 10x volume of lysis buffer for 2x 5 minutes.

→ Divide each sample to two parts: (i) with 5% total volume (for SDS Page analysis) and (ii) with 95% total volume (for qPCR)

→ Flash frozen and store the gel at -80^{0}C if needed

→ For part (i): add elution buffer and run on SDS gel to check for protein signal

→ For part (ii): add trizol and proceed to RNA extraction (chapter 2).

→ Extracted RNA needs to be analyzed for quality control using Bioanalyzer.

→ Library preparation is done using RNA-seq-compatible kit following manufacturer's instruction.

A.7　　　Cross-linking RNA-immunoprecipitation (CLIP) protocol

. Antibody being used: anti-Flag antibody

. Protein being IP: flag-tagged protein

. Tissue type: whole L3 larval body (40 hours after L2/L3 molt)

For other proteins, other antibodies or other tissue types, modification might be needed.

A.7.1　　　Reagents

1. Lysis buffer:

component	stock concentration	final concentration
KCl		150mM
Tris-HCl pH 7.4		25mM
EDTA		5mM
NP40		0.5% v/v
DTT		5mM
PMSF		1mM
Proteinase inhibitor (add fresh)	25X	1X
RNAse inhibitor (add fresh)		100U/mL

2. Wash buffer:

component	stock concentration	final concentration
KCl		150mM
Tris-HCl pH 7.4		25mM
EDTA		5mM
Proteinase inhibitor (add fresh)	25X	1X
RNAse inhibitor (add fresh)		100U/mL

3. Ligation mix:

component	amount added
10x ligation buffer	2 μl

430

DMSO	2 μl
RNAsin	1 μl
truncated T4 Rnl2	2 μl
preadenylated linker	1 μl
nuclease-free water	12 μl

A.7.2   Samples preparation

→ At least 2 groups of sample to be collected: w^{1118} (control) and experimental line carries tagged version of target protein.

→ Collect 500 40hr L3 larvae per replicate with 3 replicates per condition. Number can vary based on experiment purposes.

→ Leave animals in 1X PBS for 2-5 minutes. Animals will vomit food from gut.

→ Transfer animals to new 1X PBS two more times to ensure the gut is clean.

→ Transfer animals to 1.5mL tube contain ice-cold 1x PBS pH 7.4 with 0.1% Triton X-100 using either clean forceps or cut P-1000 pipette tips (I prefer the first way) .

→ Prepare fresh 1X PBS with 0.2% Formaldehyde (fixative).

→ Remove as much PBST from sample as possible

→ Add fresh fixative to dissected BRGC and incubate at room temperature,

⇨ Mix gently on a rotary shaker for 10 min.

→ Let sample sink to the bottom of the tube.

→ Remove fixative solution as soon as possible

→ Add 0.25M Glycine in PBST to quench crosslinking reaction.

→ Incubate on shaker for 5 min.

→ Remove quench solution

→ Wash sample for PBST three times by gently inverting the tube 5 times.

→ Remove as much PBST as possible

→ Transfer samples to a 10mL UV-compatible tube filled with ice-cold 1x PBST, 0.1% Tween 20.

→ UV-crosslink on ice with 6x 5-minute pulses at 254nm in a UV-crosslinker. Gently shake the tube for 5-10 seconds between each pulse. No recovery pause is needed.

→ Wash samples in PBST for three times by gently inverting the tube 5 times per wash

→ Freeze samples using liquid Nitrogen and store in -80^{0}C until getting enough samples.

A.7.3 Protein extraction and immunoprecipitation

→ Remove samples from -80^{0}C freezer and thaw on ice.

→ At the same time, pre-rinse Dounce homogenizer with ice-cold lysis buffer

→ Combine samples if needed to the Dounce homogenizer and add 1mL lysis buffer

→ Add RNAse I for partial RNAse digestion at the ratio of 1:3000

→ Homogenize using a Dounce homogenizer in 1mL lysis buffer on ice for about 10 strokes.

→ Repeat homogenization every 15 minutes for about 1 hour. Samples need to be kept on ice all the time.

→ Transfer lysate to pre-chilled 1.5mL centrifuge tubes

→ Incubate for another 30 min on shaker to ensure thorough lysis

→ Centrifuge for 30 min at 12,000 x g at 4^{0}C.

→ Transfer supernatants to new pre-chilled 1.5mL centrifuge tubes without disturbing the pellets.

→ Filter supernatants through a 0.45μm syringe filter and collect through new pre-chilled 1.5mL centrifuge tubes.

→ Prepare anti-Flag M2 affinity gel by well resuspension in ice-cold lysis buffer.

→ Get the right amount needed and equilibrate gel by washing with lysis buffer with 10x volume for 3 times.

→ Remove the old lysis buffer and add the same amount of buffer as the original amount of gel being used.

→ Mix the lysate with 300 μL of equilibrated anti-Flag gel. This results in a total reaction volume of approximately 1.2-1.3mL.

→ Add 13 μL of DNAse H.

→ Incubate with gentle shaking at 4^0C overnight (8-10 hours).

→ Move samples to room temperature and centrifuge at 12,000 x g for 8 second, discard flow through

→ Incubate agarose beads with 20 μl of ligation mix for 3 hours with gentle shaking

→ Wash with 10x volume of wash buffer for 2x 5 minutes.

→ Divide each sample to two parts: (i) with 5% total volume (for SDS Page analysis) and (ii) with 95% total volume (for qPCR)

→ Flash frozen and store the gel at -80^0C if needed

→ For part (i): add elution buffer and run on SDS gel to check for protein signal

→ For part (ii): add trizol and proceed to RNA extraction (chapter 2).

→ Extracted RNA needs to be analyzed for quality control using Bioanalyzer.

→ Library preparation is done using RNA-seq-compatible kit following manufacturer's instruction.

A.8 List of transgenic and CRISPR resources generated by me during my program

genes/alleles	description	genotype
IRP1A$^{WT.3F}$	Endogenous IRP1A with C-terminal 3xFlag tag (CRISPR)	w^{1118}; IRP1A$^{WT.3F}$ (III)
IRP1A^{FCF}	Mutant FRT flanked IRP1A with C-terminal 3xFlag tag (CRISPR)	w^{1118}; IRP1A^{FCF}/TM6B, Hu, Tb (III)
IRP1A^{KO}	IRP1A (CRISPR) deletion	IRP1A^{KO}/TM6B, Hu, Tb (III)
IRP1A-Gal4	IRP1A Gal4 (PhiC31 P-element)	IRP1A-Gal4 (III)
UAS-IRP1A^{3M}	C-terminal 3xMyc tagged wildtype IRP1A cDNA (PhiC31 P-element)	w^{1118};UAS-IRP1A$^{WT.3M}$ (II)
IRP1A^{gR}	Somatic double gRNA targeting IRP1A CDS region (PhiC31 P-element).	y^1,v^1;IRP1A^{gR} (II)
IRP1A^{13B}	CasFB-compatible crRNA targeting IRP1A mRNA (PhiC31 P-element)	y^1,v^1;IRP1A^{13B} (II)
IRP1A^{13X}	CasFX-compatible crRNA targeting IRP1A mRNA (PhiC31 P-element)	y^1,v^1;IRP1A^{13X} (II)
IRP1A$^{C450S.3F}$	Mutant IRP1A with C450S mutation, result in apo form IRP1A only, has C-terminal 3xFlag tag (CRISPR)	w^{1118};IRP1A$^{C450S.3F}$/TM6B, Hu, Tb (III)
UAS-IRP1A$^{3F.C450S.NLS}$	N-terminal 3xFlag tagged IRP1A^{C450S} cDNA fused with Nuclear Localization Signal (NLS) (PhiC31 P-element). IRP1A^{C450S} is predicted to be apo-form only.	w^{1118};UAS-IRP1A$^{3F.C450S.NLS}$ (II)
UAS-IRP1A$^{3F.C516S}$	N-terminal 3xFlag tagged IRP1A^{C516S} cDNA (PhiC31 P-element). IRP1A C516S is predicted to be apo-form only.	w^{1118};UAS-IRP1A$^{3F.C516S}$ (II)
UAS-IRP1A$^{C516S.3M}$	C-terminal 3xMyc tagged IRP1A^{C516S} cDNA (PhiC31 P-element). IRP1A C516S is predicted to be apo-form only.	w^{1118};UAS-IRP1A$^{C516S.3M}$ (III)
UAS-IRP1A$^{3M.C516S}$	N-terminal 3xMyc tagged IRP1A^{C516S} cDNA (PhiC31 P-element). IRP1A C516S is predicted to be apo-form only.	w^{1118};UAS-IRP1A$^{3M.C516S}$ (III)
UAS-IRP1A$^{3F.3R3Q}$	N-terminal 3xFlag tagged IRP1A^{3R3Q} cDNA (PhiC31 P-element). IRP1A 3R3Q is predicted to be holo-form only.	w^{1118};UAS-IRP1A$^{3F.3R3Q}$ (II)

genes/alleles	description	genotype
UAS-IRP1A$^{3R3Q.3M}$	N-terminal 3xMyc tagged IRP1A^{3R3Q} cDNA (PhiC31 P-element). IRP1A 3R3Q is predicted to be holo-form only.	w^{1118};UAS-IRP1A$^{3R3Q.3M}$ (II)
IRP1A$^{C516S.3F}$	Mutant IRP1A with C516S mutation, result in apo form IRP1A only, has C-terminal 3xFlag tag (CRISPR)	w^{1118};IRP1A$^{C516S.3F}$/TM6B, Hu, Tb (III)
IRP1A$^{C519S.3F}$	Mutant IRP1A with C519S mutation, result in apo form IRP1A only, has C-terminal 3xFlag tag (CRISPR)	w^{1118};IRP1A$^{C519S.3F}$/TM6B, Hu, Tb (III)
IRP1A$^{R793Q.3F}$	C-terminal 3xFlag tagged IRP1A^{R793Q} cDNA (PhiC31 P-element). IRP1A^{R793Q} is predicted to almost abolish IRE binding activity.	w^{1118};UAS-IRP1A$^{R793Q.3F}$ (II)
IRP1A$^{R549Q.3F}$	C-terminal 3xFlag tagged IRP1A^{R549Q} cDNA (PhiC31 P-element). IRP1A^{R549Q} is predicted to have reduced IRE binding activity.	w^{1118};UAS-IRP1A$^{R549Q.3F}$ (II)
IRP1A$^{R554Q.3F}$	C-terminal 3xFlag tagged IRP1A^{R554Q} cDNA (PhiC31 P-element). IRP1A^{R554Q} is predicted to have reduced IRE binding activity.	w^{1118};UAS-IRP1A$^{R554Q.3F}$ (II)
IRP1A$^{2R2Q1.3F}$	C-terminal 3xFlag tagged IRP1A^{2R2Q1} cDNA (PhiC31 P-element). IRP1A^{2R2Q1} has two mutated residues (R549Q, R554Q) and is predicted to have reduced IRE binding activity.	w^{1118};UAS-IRP1A$^{2R2Q1.3F}$ (II)
IRP1A$^{2R2Q2.3F}$	C-terminal 3xFlag tagged IRP1A^{2R2Q2} cDNA (PhiC31 P-element). IRP1A^{2R2Q2} has two mutated residues (R549Q, R793Q) and is predicted to have reduced IRE binding activity.	w^{1118};UAS-IRP1A$^{2R2Q2.3F}$ (II)
IRP1A$^{2R2Q3.3F}$	C-terminal 3xFlag tagged IRP1A^{2R2Q3} cDNA (PhiC31 P-element). IRP1A^{2R2Q3} has two mutated residues (R554Q, R793Q) and is predicted to almost abolish IRE binding activity.	w^{1118};UAS-IRP1A$^{2R2Q3.3F}$ (II)

genes/alleles	description	genotype
IRP1A$^{3R3Q.3F}$	C-terminal 3xFlag tagged IRP1A^{3R3Q} cDNA (PhiC31 P-element). IRP1A^{3R3Q} has three mutated residues (R549Q, R554Q, R793Q) and is predicted to almost abolish IRE binding activity.	w^{1118};UAS-IRP1A$^{3R3Q.3F}$ (II)
UAS-YAco1$^{WT.3F}$	C-terminal 3xFlag tagged wildtype yeast Aco1 cDNA (PhiC31 P-element), predicted to localize both in mitochondria and cytosol.	w^{1118};UAS-YAco1$^{WT.3F}$ (II)
UAS-YAco1^{WT}	Wildtype yeast Aco1 cDNA (PhiC31 P-element), predicted to localize both in mitochondria and cytosol.	w^{1118};UAS-YAco1^{WT} (II)
UAS-YAco1$^{ΔSp.3F}$	C-terminal 3xFlag tagged mutant yeast Aco1^{ΔSp} cDNA (PhiC31 P-element), predicted to localize only in cytosol.	w^{1118};UAS-YAco1$^{ΔSp.3F}$ (II)
UAS-YAco1^{ΔSp}	Mutant yeast Aco1^{ΔSp} cDNA (PhiC31 P-element), predicted to localize only in cytosol.	w^{1118};UAS-YAco1^{ΔSp} (II)
UAS-hIRP1	Wildtype human IRP1 cDNA (PhiC31 P-element)	w^{1118};UAS-hIRP1 (II)
UAS-hIRP1$^{.3F}$	N-terminal 3xFlag tagged human IRP1 cDNA (PhiC31 P-element)	w^{1118};UAS-hIRP1$^{.3F}$ (II)
UAS-hIRP2	Wildtype human IRP2 cDNA (PhiC31 P-element)	w^{1118};UAS-hIRP2 (II)
UAS-hIRP2$^{.3F}$	N-terminal 3xFlag tagged human IRP2 cDNA (PhiC31 P-element)	w^{1118};UAS-hIRP2$^{.3F}$ (II)
IRP1B$^{WT.3F}$	Endogenous IRP1B with C-terminal 3xFlag tag (CRISPR)	w^{1118}; IRP1B$^{WT.3F}$ (III)
IRP1B^{KO}	IRP1B (CRISPR) deletion	w^{1118}; IRP1B^{KO} (III)
IRP1B^{gR}	Somatic double gRNA targeting IRP1B CDS (PhiC31 P-element).	w^{1118};IRP1B^{gR} (II)
IRP1A^{gR},IRP1B^{gR}	Somatic double gRNA targeting IRP1A CDS and somatic double gRNA targeting IRP1B CDS (PhiC31 P-element).	y^{1},v^{1};IRP1A^{gR},IRP1B^{gR} (II)
UAS-IRP1B$^{C447S.3F}$	N-terminal 3xFlag tagged IRP1B^{C447S} cDNA (PhiC31 P-element), predicted to lose aconitase activity, the ability to bind mRNA remains unknown.	w^{1118}; UAS-IRP1B$^{C447S.3F}$ (II)

genes/alleles	description	genotype
IRP1B$^{C447S.3F}$	Mutant IRP1B^{C447S} with C-terminal 3xFlag tag (CRISPR), predicted to lose aconitase activity, the ability to bind mRNA remains unknown.	w^{1118}; IRP1B$^{C447S.3F}$ (III)
UAS-IRP1B$^{3R3Q.3F}$	N-terminal 3xFlag tagged IRP1B^{3R3Q} cDNA (PhiC31 P-element), predicted to further disable the ability to bind IRE mRNAs.	w^{1118}; UAS-IRP1B$^{3R3Q.3F}$ (II)
IRP1B$^{3R3Q.3F}$	Mutant IRP1B^{3R3Q} with C-terminal 3xFlag tag (CRISPR), predicted to further disable the ability to bind IRE mRNAs.	w^{1118}; IRP1B$^{3R3Q.3F}$ (III)
Fer1HCH$^{sfGFP(C)}$	Mutant Fer1HCH with C-terminal sfGFP tag (CRISPR), predicted to generate sfGFP tag version of all Fer1HCH isoforms, except PE isoform.	w^{1118}; Fer1HCH$^{sfGFP(C)}$/TM6B,Hu,Tb (III)
AGBE^{FCM}	FRT-flanked AGBE allele with C-terminal 3xMyc tag (CRISPR)	w^{1118}; AGBE^{FCM} (II)
AGBE^{FCF}	FRT-flanked AGBE allele with C-terminal 3xFlag tag (CRISPR)	w^{1118}; AGBE^{FCF} (II)
UAS-AGBE^{3M}	N-terminal 3xMyc tagged AGBE cDNA (PhiC31 P-element)	w^{1118}; UAS-AGBE^{3M} (III)
UAS-AGBE^{3F}	N-terminal 3xFlag tagged AGBE cDNA (PhiC31 P-element)	w^{1118}; UAS-AGBE^{3F} (III)
UAS-AGBE	Wildtype AGBE cDNA (PhiC31 P-element)	w^{1118}; UAS-AGBE (II)
AGBE^{gR}	Somatic double gRNA targeting AGBE CDS (PhiC31 P-element)	y^{1}, v^{1}; AGBE^{gR} (II)
AGBE^{13X}	CasFX-compatible crRNA targeting AGBE mRNA (PhiC31 P-element)	y^{1}, v^{1}; AGBE^{13X} (II)
ppk20^{FCH}	FRT-flanked ppk20 allele with C-terminal 3xHA tag (CRISPR)	w^{1118}; ppk20^{FCH}/TM6B,Hu,Tb
ppk20^{KO}	ppk20 deletion (CRISPR)	w^{1118}; ppk20^{KO}/TM6B, Hu, Tb
ppk20^{gR}	Somatic double gRNA targeting ppk20 CDS (PhiC31 P-element)	y^{1},v^{1}; ppk20^{gR}/CyO (II)
UAS-ppk20$^{PA.mVenus}$	C-terminal mVenus tagged ppk20 PA isoform cDNA (PhiC31 P-element)	w^{1118}; UAS-ppk20$^{PA.mVenus}$ (III)
UAS-ppk20$^{3H.PA}$	N-terminal 3xHA tagged ppk20 PA isoform cDNA (PhiC31 P-element)	w^{1118}; UAS-ppk20$^{3H.PA}$ (III)
UAS-ppk20$^{PC.mVenus}$	C-terminal mVenus tagged ppk20 PC isoform cDNA (PhiC31 P-element)	w^{1118}; UAS-ppk20$^{PC.mVenus}$ (III)
UAS-ppk20$^{3H.PC}$	N-terminal 3xHA tagged ppk20 PC isoform cDNA (PhiC31 P-element)	w^{1118}; UAS-ppk20$^{3H.PC}$ (III)
spz5^{KO}	spz5 deletion (CRISPR)	w^{1118}; spz5^{KO} (III)

genes/alleles	description	genotype
spz5^{FCG}	FRT-flanked spz5 with C-terminal sfGFP tag (CRISPR)	w^{1118}; spz5^{FCG} (III)
spz5^{gR}	Somatic double gRNA targeting spz5 CDS (PhiC31 P-element)	y^1,v^1; spz5^{gR} (II)
spok-DmC	*spookier*-driven fly optimized Cas9 (PhiC31 P-element) for gene disruption	w^{1118}; spok-DmC/TM3 Ser.GFP (III)
spok-HsC	*spookier*-driven human optimized Cas9 (PhiC31 P-element) for gene disruption	w^{1118}; spok-HsC/CyO.GFP (II)
spok-dFC	*spookier*-driven nuclease dead Cas9 (dCas9) fused with FokI nuclease domain (PhiC31 P-element) for gene disruption	w^{1118}; spok-dFC/CyO.GFP (II)
spok-GSD	*spookier*-driven dCas9 fused with human Progesterone Receptor ligand-binding domain (hPR LBD) and FokI nuclease domain (PhiC31 P-element) for RU-486 geneswitch gene disruption (GSD)	w^{1118}; spok-GSD/TM6B, Hu, Tb (III)
act-GSD	*actin5C* (*ac5*)-driven dCas9 fused with hPR LBD and FokI nuclease domain (PhiC31 P-element) for RU-486 geneswitch gene disruption (GSD)	w^{1118}, act-GSD (X)
UAS-GSD	Gal4/UAS-driven dCas9 fused with hPR LBD and FokI nuclease domain (PhiC31 P-element) for RU-486 geneswitch gene disruption (GSD)	w^{1118}; UAS-GSD (II)
spok-RDX	*spookier*-driven split Cas9 for rapamycin dependent gene disruption (RDX) (PhiC31 P-element)	w^{1118}; spok-RDX (III)
act-RDX	*act*-driven split Cas9 for rapamycin dependent gene disruption (RDX) (PhiC31 P-element)	w^{1118}, act-RDX (II)
UAS-RDX	Gal4/UAS-driven split Cas9 for rapamycin dependent gene disruption (RDX) (PhiC31 P-element)	w^{1118}; UAS-RDX (II)
spok-dI	*spookier*-driven nuclease dead Cas9 (dCas9) (PhiC31 P-element)	w^{1118}; spok-dI/TM3 Ser.GFP (III)
spok-64bO	*spookier*-driven dCas9 fused with VP64 activation domain for overexpression purpose (PhiC31 P-element)	w^{1118}; spok-64bO/TM6,Hu,Tb (III)

438

genes/alleles	description	genotype
spok-VPR	*spookier*-driven dCas9 fused with VP64, p65 and Rta activation domains for overexpression purpose (PhiC31 P-element)	w^{1118}; *spok-VPR/TM6,Hu,Tb (III)*
spok-GSO	*spookier*-driven dCas9 fused with human Progesterone Receptor ligand-binding domain for temporal overexpression purpose (PhiC31 P-element)	w^{1118}; *spok-GSO (III)*
spok-RDO	*spookier*-driven split dCas9 for rapamycin dependent gene overexpression (PhiC31 P-element)	w^{1118}; *spok-RDO (III)*
act-RDO	*act*-driven split dCas9 for rapamycin dependent gene overexpression (PhiC31 P-element)	w^{1118}; *act-RDO (II)*
UAS-RDO	Gal4/UAS-driven split dCas9 for rapamycin dependent gene overexpression (PhiC31 P-element)	w^{1118}; *UAS-RDO (III)*
phm^{gR1}	Somatic double gRNA targeting *phm* CDS (PhiC31 P-element)	y^1, v^1; *phm^{gR1} (II)*
phm^{gR2}	Somatic quadruple gRNA targeting *phm* CDS (PhiC31 P-element)	y^1, v^1; *phm^{gR2} (II)*
phm$^{PG.gR1}$	PG-specific somatic double gRNA targeting *phm* CDS (PhiC31 P-element)	y^1, v^1; *phm$^{PG.gR1}$ (II)*
phm$^{PG.gR2}$	PG-specific somatic quadruple gRNA targeting *phm* CDS (PhiC31 P-element)	y^1, v^1; *phm$^{PG.gR2}$ (II)*
phm TSS^{-174}	Somatic sgRNA targeting 174 bp upstream *phm* transcription start site (PhiC31 P-element)	y^1, v^1; *phm^{-174} (II)*
phm TSS^{-423}	Somatic sgRNA targeting 423 bp upstream *phm* transcription start site (PhiC31 P-element)	y^1, v^1; *phm^{-174} (II)*
phm TSS$^{PG.-174}$	PG-specific somatic sgRNA targeting 174 bp upstream *phm* transcription start site (PhiC31 P-element)	y^1, v^1; *phm$^{PG.-174}$ (II)*
phm TSS$^{PG.-423}$	PG-specific somatic sgRNA targeting 423 bp upstream *phm* transcription start site (PhiC31 P-element)	y^1, v^1; *phm$^{PG.-174}$ (II)*
phm^{13B}	crRNA targeting *phm* transcripts for CasFB (PhiC31 P-element)	y^1, v^1; *phm^{13B} (III)*

genes/alleles	description	genotype
phm^{13X}	crRNA targeting *phm* transcripts for CasFX (PhiC31 P-element) with the same targeting site as phm^{13B}.	y^1, v^1; phm^{13X} (II)
dib^{gR1}	Somatic double gRNA targeting *dib* CDS (PhiC31 P-element)	y^1, v^1; dib^{gR1} (II)
dib^{gR2}	Somatic quadruple gRNA targeting *dib* CDS (PhiC31 P-element)	y^1, v^1; dib^{gR2} (II)
dib$^{PG.gR1}$	PG-specific somatic double gRNA targeting *dib* CDS (PhiC31 P-element)	y^1, v^1; $dib^{PG.gR1}$ (II)
dib $^{PG.gR2}$	PG-specific somatic quadruple gRNA targeting *dib* CDS (PhiC31 P-element)	y^1, v^1; $dib^{PG.gR2}$ (II)
dib TSS^{-110}	Somatic sgRNA targeting 110 bp upstream *dib* transcription start site (PhiC31 P-element)	y^1, v^1; dib^{-110} (II)
dib TSS^{-482}	Somatic sgRNA targeting 482 bp upstream *dib* transcription start site (PhiC31 P-element)	y^1, v^1; dib^{-482} (II)
dib TSS $^{PG.-110}$	PG-specific somatic sgRNA targeting 110 bp upstream *dib* transcription start site (PhiC31 P-element)	y^1, v^1; $dib^{PG.-110}$ (II)
dib TSS $^{PG.-482}$	PG-specific somatic sgRNA targeting 482 bp upstream *dib* transcription start site (PhiC31 P-element)	y^1, v^1; $dib^{PG.-182}$ (II)
dib^{13B}	crRNA targeting *dib* transcript for CasFB (PhiC31 P-element)	y^1, v^1; dib^{13B} (III)
dib^{13X}	crRNA targeting *dib* transcript for CasFX (PhiC31 P-element) with the same targeting site as dib^{13B}.	y^1, v^1; dib^{13X} (II)
dI^{13X}	double crRNA targeting *dib* and *IRP1A* transcripts for CasFX (PhiC31 P-element) with the same targeting sites as dib^{13X} and IRP1A^{13X}	y^1,v^1; dI^{13X} (II)
pCFD3^{gR}	Blank pCFD3 gRNA for control purpose (PhiC31 P-element)	y^1,v^1; $pCFD3^{gR}$ (II)
pCFD4^{gR}	Blank pCFD4 gRNA for control purpose (PhiC31 P-element)	y^1,v^1; $pCFD4^{gR}$ (II)
pCFD5^{gR}	Blank pCFD5 gRNA for control purpose (PhiC31 P-element)	y^1,v^1; $pCFD5^{gR}$ (II)
pPG1^{gR}	Blank PG.gRNA 1 for control purpose (PhiC31 P-element)	y^1,v^1; $pPG1^{gR}$ (II)

genes/alleles	description	genotype
pPG2^{gR}	Blank PG.gRNA 2 for control purpose (PhiC31 P-element)	$y^1,v^1; pPG2^{gR}$ (II)
pPG3^{gR}	Blank PG.gRNA 3 for control purpose (PhiC31 P-element)	$y^1,v^1; pPG3^{gR}$ (II)
Nos^{gR}	Somatic double gRNA targeting *Nos* CDS (PhiC31 P-element)	$y^1, v^1; Nos^{gR}$ (II)
HO^{gR}	Somatic double gRNA targeting *HO* CDS (PhiC31 P-element)	$y^1, v^1; HO^{gR}$ (II)
CG7650^{gR}	Somatic double gRNA targeting *CG7650* CDS (PhiC31 P-element)	$y^1, v^1; CG7650^{gR}$ (II)
CG7955^{gR}	Somatic double gRNA targeting *CG7955* CDS (PhiC31 P-element)	$y^1, v^1; CG7955^{gR}$ (II)
CG6222^{gR}	Somatic double gRNA targeting *CG6222* CDS (PhiC31 P-element)	$y^1, v^1; CG6222^{gR}$ (II)
CG12373^{gR}	Somatic double gRNA targeting *CG12373* CDS (PhiC31 P-element)	$y^1, v^1; CG12373^{gR}$ (II)
CG11771^{gR}	Somatic double gRNA targeting *CG11771* CDS (PhiC31 P-element)	$y^1, v^1; CG11771^{gR}$ (II)
spz5,tig^{gR}	Somatic quadruple gRNA, two gRNA targeting *spz5* CDS and the other two targeting *tig* CDS (PhiC31 P-element)	$y^1, v^1; spz5,tig^{gR}$ (II)
tig^{gR}	Somatic double gRNA targeting *tig* CDS (PhiC31 P-element)	$y^1, v^1; tig^{gR}$ (II)
CG17985^{gR}	Somatic double gRNA targeting *CG17985* CDS (PhiC31 P-element)	$y^1, v^1; CG17985^{gR}$ (II)
CG10531^{gR}	Somatic double gRNA targeting *CG10531* CDS (PhiC31 P-element)	$y^1, v^1; CG10531^{gR}$ (II)
CG14750^{gR}	Somatic double gRNA targeting *CG14750* CDS (PhiC31 P-element)	$y^1, v^1; CG14750^{gR}$ (II)
CG2887^{gR}	Somatic double gRNA targeting *CG2887* CDS (PhiC31 P-element)	$y^1, v^1; CG2887^{gR}$ (II)
CG13465^{gR}	Somatic double gRNA targeting *CG13465* CDS (PhiC31 P-element)	$y^1, v^1; CG13465^{gR}$ (II)
CG13837^{gR}	Somatic double gRNA targeting *CG13837* CDS (PhiC31 P-element)	$y^1, v^1; CG13837^{gR}$ (II)
CG10372^{gR}	Somatic double gRNA targeting *CG10372* CDS (PhiC31 P-element)	$y^1, v^1; CG10372^{gR}$ (II)
CG30410^{gR}	Somatic double gRNA targeting *CG30410* CDS (PhiC31 P-element)	$y^1, v^1; CG30410^{gR}$ (II)
VG^{gR}	Somatic single gRNA targeting upstream of TSS of pBID-UASc-VG construct (PhiC31 P-element)	$y^1, v^1; VG^{gR}$ (III)

genes/alleles	description	genotype
FG^{gR}	Somatic single gRNA targeting upstream of TSS of pBID-UASc-FG construct (PhiC31 P-element)	$y^1, v^1; FG^{gR}$ (III)
MRG^{gR}	Somatic single gRNA targeting upstream of TSS of pBID-UASc-MRG construct (PhiC31 P-element)	$y^1, v^1; VG^{gR}$ (III)
act-CasFB	Ubiquitous (*act*)-driven *Drosophila* codon-optimized CasFB for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*act-CasFB* (III)
UAS-CasFB	Gal4/UAS-driven *Drosophila* codon-optimized CasFB for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*UAS-CasFB* (II)
act-CasFX	Ubiquitous (*act*)-driven *Drosophila* codon-optimized CasFX for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*act-CasFX* (III)
UAS-CasFX	Gal4/UAS-driven *Drosophila* codon-optimized CasFX for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*UAS-CasFX* (II)
act-CasFX^{MT}	Ubiquitous (*act*)-driven mitochondrial localized CasFX^{MT} for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*act-CasFX^{MT}* (III)
UAS-CasFX^{MT}	Gal4/UAS-driven mitochondrial localized CasFX^{MT} for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*UAS-CasFX^{MT}* (II)
act-dCasFX	Ubiquitous (*act*)-driven nuclease-dead dCasFX for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*act-dCasFX* (III)
UAS-CasFX	Gal4/UAS-driven nuclease-dead dCasFX for RNA cleavage purpose (PhiC31 P-element)	w^{1118};*UAS-dCasFX* (II)
UAS-FREPAIRv2	Gal4/UAS-driven REPAIRv2 system in *Drosophila* (PhiC31 P-element)	w^{1118};*UAS-dCasFX* (II)
pC13B^{Bl}	Blank pC13B with no target crRNA for control purpose (PhiC31 P-element)	y^1v^1;*pC13B* (II)
pC13X^{Bl}	Blank pC13X with no target crRNA for control purpose (PhiC31 P-element)	y^1v^1;*pC13X* (II)

genes/alleles	description	genotype
IRP1A$^{C450S.Rv2\ cr}$	crRNA with single mismatch on IRP1A^{C450S} for REPAIR purpose (PhiC31 P-element)	y^1v^1; IRP1A$^{C450S.13Xrv2}$ (II)
Nos^{13X}	CasFX-compatible crRNA targeting *Nos* mRNA (PhiC31 P-element)	y^1, v^1; Nos^{13X} (II)
spz5^{13X}	CasFX-compatible crRNA targeting *spz5* mRNA (PhiC31 P-element)	y^1, v^1; spz5^{13X} (II)
HO^{13X}	CasFX-compatible crRNA targeting *HO* mRNA (PhiC31 P-element)	y^1, v^1; HO^{13X} (II)
tig^{13X}	CasFX-compatible crRNA targeting *tig* mRNA (PhiC31 P-element)	y^1, v^1; tig^{13X} (II)
spz5^{13X},tig^{13X}	CasFX-compatible double crRNA each either targets *spz5* or *tig* mRNAs (PhiC31 P-element)	y^1v^1;spz5^{13X},tig^{13X} (III)
spz5^{13X}, HO^{13X}	CasFX-compatible double crRNA each either targets *spz5* or *HO* mRNAs (PhiC31 P-element)	y^1v^1;spz5^{13X},HO^{13X} (II)
spz5^{13X},HO^{13X},tig^{13X}	CasFX-compatible triple crRNA each either target *spz5*, *HO* or *tig* mRNAs (PhiC31 P-element)	y^1v^1;spz5^{13X},HO^{13X},tig^{13X} (III)
CG7650^{13X}	CasFX-compatible crRNA targeting *CG7650* mRNA (PhiC31 P-element)	y^1, v^1; CG7650^{13X} (II)
CG7955^{13X}	CasFX-compatible crRNA targeting *CG7955* mRNA (PhiC31 P-element)	y^1, v^1; CG7955^{13X} (II)
CG6222^{13X}	CasFX-compatible crRNA targeting *CG6222* mRNA (PhiC31 P-element)	y^1, v^1; CG6222^{13X} (II)
CG12373^{13X}	CasFX-compatible crRNA targeting *CG12373* mRNA (PhiC31 P-element)	y^1, v^1; CG12373^{13X} (II)
CG11771^{13X}	CasFX-compatible crRNA targeting *CG11771* mRNA (PhiC31 P-element)	y^1, v^1; CG11771^{13X} (II)
CG17985^{13X}	CasFX-compatible crRNA targeting *CG17985* mRNA (PhiC31 P-element)	y^1, v^1; CG17985^{13X} (II)
CG10531^{13X}	CasFX-compatible crRNA targeting *CG10531* mRNA (PhiC31 P-element)	y^1, v^1; CG10531^{13X} (II)
CG14750^{13X}	CasFX-compatible crRNA targeting *CG14750* mRNA (PhiC31 P-element)	y^1, v^1; CG14750^{13X} (II)
CG2887^{13X}	CasFX-compatible crRNA targeting *CG2887* mRNA (PhiC31 P-element)	y^1, v^1; CG2887^{13X} (II)
CG13465^{13X}	CasFX-compatible crRNA targeting *CG13465* mRNA (PhiC31 P-element)	y^1, v^1; CG13465^{13X} (II)
CG13837^{13X}	CasFX-compatible crRNA targeting *CG13837* mRNA (PhiC31 P-element)	y^1, v^1; CG13837^{13X} (II)

genes/alleles	description	genotype
CG10372^{13X}	CasFX-compatible crRNA targeting *CG10372* mRNA (PhiC31 P-element)	*y^1, v^1; CG10372^{13X} (II)*
CG30410^{13X}	CasFX-compatible crRNA targeting *CG30410* mRNA (PhiC31 P-element)	*y^1, v^1; CG30410^{13X} (II)*

A.9 List of plasmid resources generated during my program

plasmid name	description	applications
pNH01	modified pAc-Stable2 plasmid that contains IRP1A^{WT} cDNA and a gateway cassette as well as a multiple cloning site (MCS), each region is separated by a T2A viral sequence. Can be used for coIP to test interaction with 1-2 candidates	AGBE-Cisd2-IRP1A interactions validation (chapter 3)
pNH02	modified pAc-Stable2 plasmid that contains IRP1A^{3R3Q} cDNA and a gateway cassette as well as a multiple cloning site (MCS), each region is separated by a T2A viral sequence. Can be used for coIP to test interaction with 1 candidate of interest	AGBE-IRP1A interaction validation (chapter 3)
pNH03	modified pAc-Stable2 plasmid that contains IRP1A^{C450S} cDNA and a gateway cassette as well as a multiple cloning site (MCS), each region is separated by a T2A viral sequence. Can be used for coIP to test interaction with 1 candidate of interest	AGBE-IRP1A interaction validation (chapter 3)
pNH04	modified pAc-Stable2 plasmid that contains eGFP cDNA and a gateway cassette as well as a multiple cloning site (MCS), each region is separated by a T2A viral sequence. Can be used for coIP to test interaction with 1 candidate of interest	coIP control (chapter 3)
pNH05	modified pHD-DsRed plasmid with mVenus replaces DsRed in classic CRISPR	chapter 3, chapter 5
pPG-gRNA 01	modified pCFD5 plasmid with *spok* regulatory region replaces the dU6:3 to drive the expression of gRNA	chapter 6
pPG-gRNA 02	modified pCFD5 plasmid with *spok* regulatory region replaces the dU6:3 to drive the expression of gRNA. Plasmid also contains HDV ribozyme sequence	chapter 6
pPG-gRNA 03	modified pCFD5 plasmid with *spok* regulatory region replaces the dU6:3 to drive the expression of gRNA. Plasmid also contains HH and HDV ribozyme sequences	chapter 6
pPG-gRMS2	modified pCFD5 plasmid with *spok* regulatory region replaces the dU6:3 to drive the expression of gRNA while the gRNA scaffold also carries a MS2 sequence to recruit transcription machinery for overexpression purposes	chapter 6

plasmid name	description	applications
gG-DmC	Plasmid carries *Drosophila*-optimized Cas9 as well as gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-HsC	Plasmid carries human-optimized Cas9 as well as gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-dFC	Plasmid carries nuclease-dead Cas9 fused with nuclease domain of FokI restriction enzyme as well as gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
en-GSD	Plasmid carries nuclease-dead Cas9 fused with human progesterone receptor ligand-binding domain (hPR LBD) and nuclease domain of FokI. Construct also carries the gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
en-RDX	Plasmid carries split Cas9 as wel as the gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-dI	Plasmid carries nuclease-dead Cas9 as well as gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-64bO	Plasmid carries dCas9 fused with VP64 for transcription machinery recruitment. This plasmid also carries a gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-VPRO	Plasmid carries dCas9 fused with VPR complex for transcription machinery recruitment. This plasmid also carries a gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-GSO	Plasmid carries dCas9 fused with hPR LBD for transcription machinery recruitment. This plasmid also carries a gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
gG-RDO	Plasmid carries split dCas9 fused with hPR LBD, and the gateway cassette to insert promoter or enhancer sequences of interest	chapter 6
pC13cr01	modified pAc-Stable2 plasmid that contains gateway cassette for Cas13 variant and dU6:3-driven crRNA cassette for cell culture Cas13 targeting	chapter 7
pC13B	dU6:3-driven CasFB-compatible crRNA plasmid	chapter 7

plasmid name	description	applications
pC13X	dU6:3-driven CasFX-compatible crRNA plasmid	chapter 7
pC13mtcr01	modified pAc-Stable 2 plasmid that carries *ac5*-driven CasFX^{mt} with mitochondrial localization signal as well as crRNA backbone.	chapter 7

These plasmids were listed since they can be easily modified for further applications. Some other plasmids were not listed since they were constructed for a specific purpose during my study.

A.10   Phenotype validation of 34 candidates with unknown function in iron or heme metabolism

A.10.1   Introduction

In an effort to identify genes with undiscovered roles in heme or iron homeostasis, the KKJ lab performed a genome-wide RNAi screening in collaboration with two other labs, namely Michael O'Connor's lab at University of Minnesota, US and Kim Furbo Rewitz lab at University of Copenhagen, Denmark. In this approach, RNAi against each gene was crossed with *phm22*, a PG-specific Gal4, and observed any developmental defects. This approach summed up a list of 800 genes which RNAi gave developmental defect at either larval arrested, pupariation delay. Later, a secondary RNAi screening focus on thess 800 genes were conducted in KKJ lab where people looked for abnormal tissue morphology, either enlarged or enlarged and also red ring gland. This secondary screening gave us a list of 34 hits for further investigation. Among these 34 hits are some genes already known to be involved in heme biosynthesis like *ALAS, UPDO, PPOX, FeCH* or iron-sulfur cluster assembly like *ABCB7*. The overall list seemed to be quite diverse with respect to cellular functions, including transcription factors, enzymes, signaling factors and receptors. In an effort to understand their functions further, I fed these RNAi animals with different supplements along the Ecdysone synthesis pathway, or heme/iron homeostasis. On the other hand, the phenotypes from these 34 hits were observed when using a single RNAi line. Further evident is required to make sure the observed phenotype is real. In KKJ lab, people are trying different approaches to validate the observed phenotypes from the secondary screening. Couple genes have been chosen by different lab members leaving about 13 genes untouched. I generated the conditional CRISPR gRNA for these genes as well as some other genes from previous lab members

for phenotype validation, namely *spz5*, *HO*, *Tig* and *Nos*. This project only provides fundamental insights for future studies. As a result, I decided to leave it in Appendix sections.

A.10.2   Results

A.10.2.1   Each RNAi behaves differently from each other when being raised on different supplements

				Nutri-Fly		sterol rescue			Iron treatment	
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
zfrp8	zinc-finger protein	V 11521		Major L3 arrest (~5% adult)	Red Giant	17% adult	32% adult	5% adult	2% adult	0% adult
		NIG 3260R-2		NOP	NOP					
		NIG 3260R-6		NOP	NOP					
CG8145	zinc-finger protein	CRISPR mutant		L1, L2 lethal (0% adult)	Giant					
		V 100854	**YES**	Major L3 arrest (~7% adult)	Moderate Red Giant	26% adult	72% adult	68% adult	38% adult	34% adult
		V 35840		Major L3 arrest	Moderate Red Giant					
		T 43551		Partial L3 arrest (~50% adult)	Moderate Red Giant					
PPOX	heme biosynthetic enzyme	V 100577/CyO, GFP	**YES**	L3 arrest	Red Giant	17% adult	4% adult	13% adult	0% adult	0% adult
		V 40607		L3 arrest	Red Giant					

				Nutri-Fly		sterol rescue			Iron treatment	
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
spz5	growth factor	V 102389		Major L3 arrest (0~2% adult)	Red Giant	2% adult	8% adult	70% adult	2% adult	2% adult
		V 41295		Major L3 arrest (~30% adult)	Red Giant					
CG30493	ubiquinone biosynthesis	**V 105722**	**YES?**	Major L3 arrest (0~10% adult)	Red Giant	76% adult	46%	44% adult	28% adult	Partial 42% adult
		V 43131		Major L3 arrest (~10% adult)	Red Giant					
		T 53336		Partial L3 arrest (10~20% **arrested**)	maybe Red? Giant					
		CRISPR mutant (small del)		L1 lethal (0% adult)	N/A					

				Nutri-Fly		sterol rescue			Iron treatment	
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
AGBE	glycogen biosynthetic process	**V 108087**	**YES**	Major L3 arrest (0~21% adult)	Red Giant	15% adult	60% adult (normal ring gland)	33.33% adult	between 78% to 82% adult starting from 200uM to 1mM FAC, normal ring gland when dissected at 200uM FAC	6% adult at 100uM
		T 40860		Major L3 arrest (0~27% adult)	Giant	25% adult	75% adult	40% adult	1mM FAC, 85% adult	25% adult in 100uM
		T 42753		Development al delay (see Note)	Red Giant	No rescue on developme	Rescue on DD (Day 7 PPF)	No rescue on	1mM FAC: Day 7 PPF	100uM: Day 9 PPF

452

gene name	function	RNAi lines tested	validated?	Nutri-Fly		sterol rescue			Iron treatment	
				phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
						ntal delay (Day 9 PPF)		developmental delay (Day 9 PPF)		
		NIG 4023R-4		NOP	NOP					
Dh 31R 1	GPCR	V 101995		Major L3 arrest	Red Giant					
		T 25925/Sb		NOP	NOP					
CG11771	proteolysis	**V 108188**		Major L3 arrest	Red Giant	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
		V18946		NOP	NOP				No info	No info
		NIG 11771R-1		NOP	NOP				No info	No info
		NIG 11771R-2		NOP	NOP				No info	No info

gene name	function	RNAi lines tested	validated?	Nutri-Fly		sterol rescue			Iron treatment	
				phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
CG5910	protein kinase-like domain	V101674		Partial L3 arrest, developmental delay	Red Giant	No rescue	No rescue	No rescue	Delay, no acceleration	Delay, no acceleration
		BL mutant		L1 lethal	N/A				No info	No info
		NIG 5910R-2/CyO		NOP	NOP				No info	No info
		NIG 5910R-3								
CG34404	Usher's syndrome protein 1	**V 103335**		Major L3 arrest (~11% adult)	Red Giant	34% adult	74% adult	0% adult	79% adult	9% adult
		V 22258		NOP	NOP				No info	No info
		V 108594		NOP	NOP	Lose phenotype	Lose phenotype	Lose phenotype	No phenotype	No phenotype
		NIG 6156R-1								
		NIG 6156R-4		NOP	NOP					
Updo	heme biosynthetic enzyme	V 105642	**YES**	L2 arrest (0% adult)	Red Giant	0% adult	0% adult	0% adult	0% adult	0% adult

454

				Nutri-Fly		sterol rescue			Iron treatment	
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
pickpocket20	sodium ion transport	V 36659	**YES**	L2 arrest (~1% adult)	Red Giant	0% adult	8% adult	13% adult	0% adult	0% adult
		T 25897		L3 arrest	Red Giant					
CG17985	Peptidoglycan-binding Lysin subgroup	**V 44908**		L3 arrest (~7% adult)	Red Giant	61% adult (25% larvae normal ring gland 50% larvae partial red ring gland 25% larvae red RG)	61% adult (normal ring gland)	14% adult	0% adult	6% adult
		T 36907/Sb		Partial developmental delay	Giant					
		NIG 17985R-1								
		NIG 17985R-4		NOP	NOP					
Chitase 9	glycoside hydrolase	V 104258		Major L3 arrest	Red Giant	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype

gene name	function	RNAi lines tested	validated?	Nutri-Fly		sterol rescue			Iron treatment	
				phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
		V16072		NOP	NOP					
		NIG 10531R-1		NOP	NOP					
		NIG 10531R-4/X		NOP	NOP					
RanBP3	Ran binding protein 3, negative regulator of nucleocytoplasmic transport of STAT92E	V 104432	YES?	L3 arrest (0% adult)	Red Giant	1% adult	2% adult	1% adult	3% adult	1% adult
				N/A	N/A					
		T 40948		NOP	NOP					
		NIG 10225R-4		Partial developmental delay	Normal					
CG13837	Chitin metabolic pathway	V 109774		Major L3 arrest (~2.5% adult)	Red Giant	30% adult	86% adult	52% adult	85% adult	76% adult
				N/A	N/A					
		T 55980		NOP	NOP					
Evi5	RabGTPase GAP activity	V 105146	YES	Major L3 arrest	Red Giant					

| | | | | Nutri-Fly | | sterol rescue | | | Iron treatment | |
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
		T 38350		Partial L3 arrest	Red Giant					
CG2887	DnaJ domain, heat response	V 106083		Majore L3 arrest	Red Giant	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
				N/A	N/A					
toe (twin of eyeless)	TF, negative regulator of RNA PolII promoter	**V 107893**		Major L3 arrest (~10% adult)	Red Giant	3% adult	29% adult	5% adult	2% adult	1% adult
		V 46515		NOP	NOP					
		T 29345		NOP	NOP					
		T 50660		NOP	NOP					
CG13465	unknown function	V 49141		Major L3 arrest (~1% adult)	Red Giant	10% adult	63% adult	2% adult	8% adult	0% adult
split-end	nucleic acid binding	**V 108828**	**YES**	Major L3 arrest	Malformed					
		V 49542		Parial L3 arrest	Malformed					
		T 50529		Partial L3 arrest	Malformed					

gene name	function	RNAi lines tested	validated?	Nutri-Fly phenotype	tissue phenotype	sterol rescue cholesterol	7dC	20E	Iron treatment FAC	BPS
		T 33398		Partial L3 arrest	Malformed					
PBGS	heme enzyme	V 107988	**YES**	Partial L3 arrest (see Note)	Giant	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
ABCB7	iron-sulfur cluster transporter?	V 106039		Major L3 arrest	Giant					
		V 40839		NOP	NOP					
mRpL18	mitochondial structure protein	V 105995		Major L3 arrest	Giant	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
		T 14789		NOP	NOP					
		NIG 12373R-1		Partial L3 arrest	Giant					
		NIG 12373R-3		Partial L3 arrest	Giant					
Suppressor of variegation 2-10	DNA-binding	**V 100813**	**YES**	Major L3 arrest (~3.5% adult)	Giant	73% pupae lethal, 2% adult	58.67% adult	4% adult	1.3% adult	2.5% adult

gene name	function	RNAi lines tested	validated?	Nutri-Fly		sterol rescue			Iron treatment	
				phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
		T 32915		Major L3 arrest (~15% adult)	Giant	19% adult	45% adult	39% adult	23% adult	26% adult
		T 29448		Major L3 arrest	Giant	No	Yes	No	No	No
		T 58067		Major L3 arrest	Giant					
Vacuolar protein sorting 25	Notch signaling pathway	V 108105	YES	Major L3 arrest (~7% adult)	Giant	77% adult	83% adult	45% adult	0% adult	1% adult
				N/A	N/A					
		T 26286		Partial L3 arrest	Giant					
		T 54831/CyO		Partial L3 arrest	Giant					
				N/A	N/A					
				N/A	N/A					
CG7650	phototransduction	V 41714		Major L3 arrest (~3% adult)	Giant	41% adult	67% adult	2% adult	3% adult	1% adult
				N/A	N/A					

				Nutri-Fly		sterol rescue			Iron treatment	
gene name	function	RNAi lines tested	validated?	phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
Fas-associated factor	JNK signaling?	V 107414		Major L3 arrest (~1% adult)	Giant	83% adult	75% adult	0% adult	53% adult	2% adult
		T 43224		NOP	NOP					
				N/A	N/A					
SRY interacting protein 1		V 109289		Major L3 arrest (~22% adult)	Giant	74% adult	64% adult	29% adult	36% adult	15% adult
		T 12034		NOP	NOP					
		NIG 10939R-1		NOP	NOP					
		NIG 10939R-3/CyO		NOP	NOP					
Ribose-5-phosphate isomerase	pentose-phosphate shunt, non-oxidative branch	**V100275**		Major L3 arrest (~1% adult)	Giant	77% adult	73% adult	1% adult	0% adult	3% adult
				N/A	N/A					
		T 62196/CyO		NOP	NOP					
ALAS	heme biosynthesis		**YES**	Major L3 arrest (0% adult)	Giant					
Integrator 2		T 60396		NOP	NOP					

gene name	function	RNAi lines tested	validated?	Nutri-Fly		sterol rescue			Iron treatment	
				phenotype	tissue phenotype	cholesterol	7dC	20E	FAC	BPS
	snRNA 3'-end processing; neurogenesis	V 105887		Partial L3 arrest	Necrosis	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
CG6222	suppressor of sable	**V 10853/X**		Major L3 arrest	Malformed	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype	Lose phenotype
				N/A	N/A					
		T 33982		NOP?						

461

A.10.2.2   Phenotypes validation of 13 candidates

gene name	function	original RNAi phenotype	Cas9/gRNA phenotype	Cas13/crRNA phenotype
AGBE	glycogen biosynthetic process	L3 arrest, red RG	L3 arrest, red RG	L3 arrest, red RG
IRP1A	iron metabolism	normal on regular fly food, porphyria phenotype with L3 arrest when growth on iron-deplete condition for 3 generation	L3 arrest, red RG on regular fly food	L3 arrest, red RG on regular fly food
Cisd2	repair damaged iron-sulfur cluster on IRP1A	normal on regular fly food, red RG with L3 arrest on iron-deplete condition	normal on regular fly food, red RG with L3 arrest on iron-deplete condition	normal on regular fly food, red RG with L3 arrest on iron-deplete condition
Nos	Nitric oxide synthase	L3 arrest, red RG	no phenotype	no phenotype
HO	Heme oxygenase	no phenotype	no phenotype	no phenotype
spz5	growth factor	L3 arrest, red RG	no phenotype	no phenotype
Tig	heme oxygenase-like	no phenotype	no phenotype	no phenotype
spz5 + HO + Tig		L3 arrest, red RG	L3 arrest, red RG	L3 arrest, red RG
CG7650	phototransduction	L3 arrest, giant RG	no phenotype	no phenotype
ABCB7	iron-sulfur cluster transporter	L3 arrest, giant RG	no phenotype	no phenotype
CG6222	suppressor of sable	L3 arrest, malformed RG	no phenotype	no phenotype
mRpL18	mitochondrial ribosomal protein L18	L3 arrest, giant RG	no phenotype	no phenotype
CG11771	proteolysis	L3 arrest, red RG	L3 arrest, red RG	L3 arrest, red RG
CG17985	peptidoglycan-binding lysin group	L3 arrest, red RG	no phenotype	no phenotype

Chitase9	glycoside hydroxilase	L3 arrest, red RG	no phenotype	no phenotype
Vps25	vacuolar sorting	L3 arrest, giant RG	L3 arrest, giant RG	L3 arrest, giant RG
CG2887	DnaJ domain, heat response	L3 arrest, red RG	no phenotype	no phenotype
CG13465	unknown	L3 arrest, red RG	L3 arrest, red RG	L3 arrest, red RG
CG13837	Chitin metabolic pathway	L3 arrest, red RG	L3 arrest, red RG	L3 arrest, red RG
Faf2	Fas-associated factor 2	pupariation delay, giant RG	no phenotype	no phenotype
Rpi	Ribose-5-phosphate isomerase	L3 arrest, giant RG	no phenotype	no phenotype