

Regulation of the *Drosophila* innate immune response by SUMO conjugation of amino-acyl tRNA Synthetases

**A Thesis
Submitted in partial fulfilment of the requirements
of the degree of
Doctor of Philosophy
By**

**Prajna Nayak
20143322**



**Indian Institute of Science Education and Research, Pune
2023**

Certificate

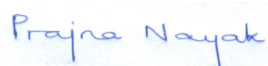
Certified that the work incorporated in the thesis entitled “**Regulation of the *Drosophila* innate immune response by SUMO conjugation of amino-acyl tRNA Synthetases**”, submitted by Prajna Nayak was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.



Dr. Girish Ratnaparkhi

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgment of collaborative research and discussions. The work was done under the guidance of Dr. Girish Ratnaparkhi, at the Indian Institute of Science Education and Research, Pune.



Prajna Nayak

20143322

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.



Dr. Girish Ratnaparkhi

*To Mama, Baba, and my little sister, Lipi.
For their advice, their patience, and their faith,
because they always understood.*

Acknowledgement

Completing this Ph.D. thesis has been an incredibly challenging and rewarding journey, and I could not have done it without the support and encouragement of so many people.

First and foremost, I want to express my heartfelt gratitude to my advisor, *Dr. Girish Ratnaparkhi*, whose unwavering support, guidance, and patience have been nothing short of remarkable. He has been a true mentor, friend, and advocate, and their commitment to my success has been inspiring.

I would also like to thank my Research Advisory Committee members (*Dr. Krishanpal Karmodia*, *Dr. Saikrishnan Kayarat*, and *Dr. Vasudevan Seshadri*) and Thesis Review Committee members for their valuable feedback, constructive criticism, and insightful suggestions. Their expertise and knowledge have helped me to navigate the complexities of my research and to broaden my understanding of the field.

I am deeply grateful to my colleagues and collaborators for their support, encouragement, and camaraderie. Their friendship and collaboration have been a source of motivation and inspiration, and I am grateful for the many hours of discussion, debate, and brainstorming that have enriched my work.

I would also like to extend my appreciation to the staff and faculty of IISER, Pune, whose dedication to academic excellence and support for graduate students have been instrumental in making this journey possible. The resources, opportunities, and community they have provided have enabled me to pursue my research and achieve my goals.

Finally, I want to express my deepest gratitude to my family and friends. Their love, encouragement, and support have sustained me through the long and challenging process of completing this thesis. Their unwavering belief in me has been a source of strength and inspiration, and I am humbled and grateful for their presence in my life.

I offer my heartfelt thanks to everyone who has played a role in my journey. Your support, encouragement, and friendship have made this experience truly unforgettable.

I would like to thank the following people without whom I wouldn't have sustained so far (In order of their appearance during the journey).

(Special thanks to people in Italics for their unconditional support)

Ravi

Diana

Rakhee

Gajanan

Madhumita

Krutula

Mithila

Vallari

Senthil

Bhagyashree

Kriti

Swati

Amar

Shweta

Sushmitha

Aparna

Aarti

Aishwarya

Lovleen

Kundan

Bhavin

Snehal

Geetanjali

Sanhita

Subhradip

Namrata

Somya

Amrita

Alex

Mayank

Akanksha

Table of contents

Chapter I

The Multi-Aminoacyl tRNA Synthetase (MARS) complex in Innate Immunity

Summary	1
1. The <i>Drosophila</i> innate immune response	1
2. The physiological role of AARSs: Protein synthesis	2
2.1 Types based on the architecture of active sites.....	2
2.2 Species-specific structural diversity (Domain analysis)	4
2.3 The evolution of AARSs.....	4
2.4 Alternative splicing.....	6
2.5 Post-translational modification (PTM).....	7
2.6 Proteolytic cleavage.....	8
2.7 The connection between catalytic and non-catalytic functions.....	8
2.8 AARSs in disease.....	9
3. MARS complex	12
4. MARS as a regulator of immune response	17
4.1 EPRS as a regulator of immune response.....	17
4.1.1 GAIT dependent immune function.....	17
4.1.2 GAIT independent immune function.....	20
4.2 KRS as a regulator of immune response.....	21
References	25-34

Chapter II

The MARS and GAIT complex respond to infection: A Mass Spectrometric study

Summary	35
1. Introduction	35
2. Aim	35
3. Materials and Methods	36
3.1 <i>Drosophila</i> husbandry.....	36
3.2 Infection Assay.....	36
3.3 Immuno-precipitation.....	37
3.4 In-gel trypsin digestion and LC-MS/MS analysis.....	37
4. Results	38
4.1 EPRS immune-precipitates pull down the multi-aminoacyl tRNA synthetase complex in <i>Drosophila</i>	38

4.2 Sexual dimorphism at the molecular level, EPRS interaction is different between males and females.....	40
4.3 Kinetics of EPRS and the MARS complex in <i>Drosophila</i> post infection.....	40
4.4 A GAIT-like complex is formed in response to infection.....	42
4.5 Association/Dissociation of non-GAIT, non-MARS elements with EPRS.....	43
5. Contributions & Acknowledgements.....	44
References.....	45-46

Chapter III

SUMOylation of Arginyl tRNA Synthetase Modulates the *Drosophila* Innate Immune Response

Summary.....	47
1. Introduction.....	47
2. Materials and Methods.....	49
2.1 SUMO conjugation assay.....	49
2.2 SUMO-binding-motif and SIM-motif prediction.....	49
2.3 Identification of evolutionarily conserved SUMO target lysine residues in-silico.....	49
2.4 Homology model for <i>Drosophila</i> RRS.....	49
2.5 Generation of Δ RRS using CRISPR Cas9 technology.....	49
2.6 pUASp AttB fly lines/strains.....	50
2.7 Fly infections.....	50
2.8 Total RNA extraction cDNA library construction and sequencing.....	51
2.9 Demultiplexing, adapter trimming, read mapping, counts generation and differential expression analysis.....	51
2.10 Survival Analysis.....	51
3. Results.....	52
3.1 The MARS Complex is a target for SUMO machinery.....	52
3.2 Generation of a Δ RRS line using CRISPR Cas9 genome editing.....	57
3.3 Generation of a transgenic RRS ^{SCR} line.....	61
3.4 Transcriptomics of immune challenged, RRS ^{WT} and RRS ^{SCR} transgenic animals.....	63
3.5 Modulation of the immune transcriptome in RRS ^{SCR} transgenics.....	70
4. Discussion.....	72
5. Contributions.....	73
6. Acknowledgement.....	73
References.....	74-76

Chapter IV

Does SUMOylation of EPRS regulate innate immune response?

Summary	77
1. Introduction	77
2. Materials and Methods	81
2.1 SUMO conjugation assay.....	81
2.2 SUMO-binding-motif and SIM-motif prediction.....	81
2.3 Identification of evolutionarily conserved SUMO target lysine residues in-silico.....	81
2.4 Generation of EPRS ^{SCR} using CRISPR Cas9 technology.....	81
2.5 Generation of Δ EPRS using CRISPR Cas9 technology.....	82
3. Results	82
3.1 EPRS is a target for SUMO machinery.....	82
3.2 Generation of EPRS ^{SCR} using CRISPR Cas9 genome editing.....	84
3.3 Generation of a Δ EPRS line using CRISPR Cas9 genome editing.....	89
4. Discussion	91
5. Acknowledgements	91
References	92
Appendix I	95-128
Appendix II	129-134

List of Figures

Chapter I

Figure 1.1: A. Structure of the MARS complex	16
B. Structure of the components of the MARS complex and their corresponding domains	16
Figure 1.2: Schematic diagram of translational silencing of a subset of functionally related genes by Interferon γ activated inhibitor of translation (GAIT) complex in human myeloid cells	20
Figure 1.3: GAIT complex independent function.	21
Figure 1.4: IgE-Ag binds to surface FcϵRI and activates the mast cells. Activation of mast cells induces phosphorylation of KRS at Ser 207 y MAPK cascade and its subsequent release from the MARS complex. Free KRS partially translocate into the nucleus and interacts with proteins; MITF and HINT to form a multi-protein complex. Phosphorylated KRS also synthesizes Ap4A in the vicinity of the multi-	

protein complex. Ap4A sequesters HINT from the complex; unmasking the DNA binding sites on MITF. Along with phosphorylated KRS, MITF engages in MITF dependent gene expression.....23

Figure 1.5 : Schematic diagram of the utilization of a dynamic MARS complex by HIV1 in human host cells to enhance its replication. During HIV1 assembly, tRNA^{Lys,3} acts as a primer for reverse transcription. tRNA^{Lys,3} is packaged into virion by its interaction with hKRS and viral proteins; Gag polyprotein and Gag pol precursor. HIV1 infection induces phosphorylation of hKRS at Ser 207 and its subsequent dissociation from the MARS complex. Free KRS partially translocate into the nucleus and the remaining takes part in HIV1 virion assembly in the cytosol.....24

Chapter II

Figure 2.1: A. Dynamics of the MARS and GAIT complex, in mammals, in response to infection.....35

B. Formation of the GAIT complex in *Drosophila*.....35

Figure 2.2: A. Infection by Gram positive (*M. luteus*) and gram negative (*Ecc15*).....37

B. Cellular and humoral response to infection.....37

Figure 2.3: Uncovering interactors of EPRS using immunoprecipitation.....39

A. IP and in-gel digestion.....39

B. MARS complex components affinity purified along with EPRS.....39

Figure 2.4: Sexual Dimorphism in the formation of the MARS complex.40

A. Components of the MARS complex.....40

B. Heat map of peptide count for different members of the MARS complex.....40

Figure 2.5: Kinetics of the EPRS dimer.....41

Figure 2.6: Kinetics of the MARS complex in response to infection.....41

Figure 2.7: No change in transcript levels after infection for genes coding for proteins in the MARS complex.....42

Figure 2.8: A GAIT-like complex in *Drosophila*.....42

Figure 2.9: Non-GAIT, non-MARS players that interact with EPRS.....43

Figure 2.10: Dissociating partners in response to infection.....44

Chapter III

Figure 3.1: *Drosophila* MARS Complex is a target of SUMO conjugation machinery.....52

Figure 3.2: A. RRS is SUMO conjugated.....55

B. RRS is SUMO conjugated at K147 and K383.....55

Figure 3.3: A. Schematic of the RRS and the MARS complex.....56

Figure 3.4: A. Schematic of the structure for the human QRS:RRS:AIMP1 complex (PDB-ID 4R3Z)57

B. Homology model of *Drosophila* RRS.....57

C. SUMO conjugation site is conserved from flies to mammals.....	57
Figure 3.5: An RRS-null (ΔRRS) line generated using CRISPR Cas9 genome editing.....	58
A. Design of the dual guide-RNA for excision of the RRS locus.....	58
B. Excision of the RRS locus to generate a Δ RRS line.....	58
Figure 3.6: A. ΔRRS lines generated by disruption of the 5'UTR.....	59
B. Schematic of the mutations in the Δ RRS ^{6B1} line.....	59
Figure 3.7: Structural consequences of CRISPR/Cas9 mediated insertions in ΔRRS line 6B1.....	60
Figure 3.8: ΔRRS flies are haplo-sufficient.....	61
Figure 3.9: A. ΔRRS^{6B1} line shows lower transcript levels of RRS as compared to wildtype.....	62
B. Rescue of Δ RRS ^{6B1} by ectopic expression of RRS using UAS-Gal4 system.....	62
Figure 3.10: Survival plots for RRS^{WT} and RRS^{SCR} upon <i>M. luteus</i>, <i>Ecc15</i> and <i>S.saprophyticus</i>	62
infection.....	62
Figure 3.11: RRS^{WT} and RRS^{SCR} show a robust immune response to bacterial infection.....	64
Figure 3.12: A. Total Number of transcripts upregulated and downregulated in response to infection.	
B. Differential expression of genes.....	65
Figure 3.13: Transcriptome changes for MARS Complex genes.....	69
Figure 3.14: Expression of RRS, Toll pathway target gene <i>Drosomycin</i> and Immune regulated catalase	
(<i>Irc</i>) in RRS^{WT} and RRS^{SCR} upon <i>M. luteus</i> infection across 0-48 Hr.....	69
Figure 3.15: Expression of RRS, Imd pathway target genes <i>Diptericin B</i> and <i>Attacin D</i> in RRS^{WT} and	
RRS^{SCR} upon <i>Ecc15</i> infection across 0-24 Hr.....	70
Figure 3.16: A. SUMO Conjugation of RRS.....	71
B. Model for immune regulation by RRS.....	71

Chapter IV

Figure 4.1: Schematic of EPRS.....	78
Figure 4.2: A. Schematic of EPRS.....	83
B. Schematic of the MARS complex.....	83
C. The structure for the human DRS:EPRS:AIMP2 complex (PDB-ID 6IY6).....	83
Figure 4.3: A. EPRS is SUMO conjugated at K957, K1063, K1083, K1106 and K1198.....	84
B. EPRS Is SUMO conjugated in-vivo.....	84
Figure 4.4: A. Design of single gRNA for generation of EPRS^{SCR}.....	86
B. Strategy for generation of EPRS ^{SCR}	86
Figure 4.5: An EPRS-null (ΔEPRS) line generated using CRISPR Cas9 genome editing.....	90
A. Design of the dual guide-RNA for excision of the EPRS locus.....	90
B. Excision of the EPRS locus to generate a Δ EPRS line.....	90

Appendix I

Fig A.3.1 Multiple Sequence Alignment of <i>Drosophila</i> RRS and Δ RRS lines.	95
A. Forward and reverse sequencing of line 6B1.....	95
B. Forward and reverse sequencing of line 18B1.....	96
Figure A.3.2: Gene Ontology Enrichment Analysis (GOEA) of the significantly differentially expressed genes in RRS ^{WT} and RRS ^{SCR} post infection with <i>M. luteus</i> and <i>Ecc15</i>	119
Figure A.3.3: GOEA is done for 4 different categories (B1) Biological Process, (B2) Molecular Function, (B3) Cellular Component and (B4) Protein Class for RRS ^{WT} and RRS ^{SCR} post infection with <i>Ecc15</i>	120
Figure A.3.4: Gene Ontology (GO) Analysis for significantly differentially expressed genes for RRS ^{SCR/WT} post infection with <i>M. luteus</i> (A) and <i>Ecc15</i> (B).....	121
Figure A.3.5: Gene Ontology Enrichment Analysis (GOEA) for significantly differentially expressed genes for RRS ^{SCR/WT} post infection with <i>M. luteus</i> and <i>Ecc15</i>	122
Figure A.3.6: GOEA is done for 2 different categories (B1) Biological process and (B2) Cellular Component and for RRS ^{SCR/WT} post infection with <i>Ecc15</i>	123

Appendix II

Figure B.1: A <i>brwl</i> loss of function line was generated using CRISPR Cas9 genome editing	131
Figure B.2: PCR based screening to detect a successful <i>brwl</i> gene knockout after a CRISPR genome editing experiment.....	133

List of Tables

Chapter I

Table 1.1: <i>Immuno-regulatory roles of MARS Complex components</i>	18
---	-----------

Chapter III

Table 3.1: <i>SUMO conjugated proteins based on Proteomic studies</i>	53
--	-----------

Table 3.2: <i>RRS^{WT} and RRS^{SCR} show differential expression of immune target genes.A-B. Tabulation of differentially expressed genes</i>	66
---	-----------

Chapter IV

Table 4.1: <i>Multiple Sequence Alignment of Drosophila EPRS and EPRS^{SCR} lines</i>	87
--	-----------

Table 4.2: <i>Tabulation of mutations confirmed upon sequencing, listed against the corresponding lines positive for presence of DsRED in their eyes</i>	88
---	-----------

Appendix I

Table A.3.1: <i>Gene expression levels for RRS WT and RRS SCR upon M.luteus infection were measured using the counts generated by HTSeq-count v 0.6.0</i>	99
--	-----------

A. <i>List of genes significantly differentially expressed in RRS SCR as compared to RRS WT 0 hours post M.luteus infection</i>	99
---	-----------

B. <i>List of genes significantly differentially expressed in RRS SCR as compared to RRS WT 0 hours post Ecc15 infection</i>	99
--	-----------

Table A.3.2: <i>Gene expression levels for RRS WT and RRS SCR upon Ecc15 infection were measured using the counts generated by HTSeq-count v 0.6.0</i>	100
---	------------

A. <i>List of genes uniquely expressed for RRS WT at 12 hours post Ecc15 infection compared to its baseline</i>	100
---	------------

B. <i>List of genes uniquely expressed for RRS SCR at 12 hours post Ecc15 infection compared to its baseline</i>	101
--	------------

C. <i>List of genes uniquely expressed for RRS WT at 12 hours post M.luteus infection compared to its baseline</i>	102
--	------------

D. <i>List of genes uniquely expressed for RRS SCR at 12 hours post M.luteus infection compared to its baseline</i>	112
---	------------

Resource Table	125-128
-----------------------------	----------------

Abbreviations

MARS complex - multi-aminoacyl tRNA synthetase complex	Ap4A - diadenosine tetra-phosphate
NF-κB – Nuclear Factor kappa B	IFN β - Interferon β
AMP - Anti-microbial peptide	PCBP2 - poly(rC) binding protein 2
IMD pathway - Immune Deficiency pathway	MAVS - Mitochondrial anti-viral signalling protein
tRNA – Transfer RNA	GAIT - interferon (IFN)- γ -activated inhibitor of translation
AMP – Adenosine Monophosphate	IFN γ - Interferon γ
ApnA - diadenosine oligophosphate	Cdk5 - Cyclin-dependent kinase 5
ATP – Adenosine Triphosphate	ERK2 - Extracellular signal-regulated kinase
RF - Rossmann ATP binding Fold	mTORC1 - mammalian target of rapamycin complex 1
HIGH - (His-Ile-Gly-His)	S6K1 - mTOR Substrate S6 Kinase 1
KMSKS - (Lys-Met-Ser-Lys-Ser)	SYNCRIP - Synaptotagmin binding, cytoplasmic RNA interacting protein
CP1 - connecting peptide 1	L13a - 60S ribosomal protein L13a
ABDs - Anticodon binding domains	DAPK - Death associated protein kinase
LZ - Leucine zipper	ZIPK - Zipper associated protein kinase
EMAP II - Endothelial monocyte activating polypeptide II	GAPDH - Glyceraldehyde 3-phosphate dehydrogenase
GST - Glutathione-S-transferase	3' UTR - 3' Untranslated region
AIMPs - aminoacyl tRNA synthetase interacting proteins	eIF4G - eukaryotic initiation factor 4G
FL - full length	eIF3 – eukaryotic initiation factor 3
PTM - post-translational modification	hnRNP - Heterogeneous nuclear ribonucleoproteins
VEGFA – Vascular Endothelial Growth Factor A	RBPs - RNA-binding proteins
SMURF2 – SMAD Specific E3 Ubiquitin Protein Ligase 2	IRES - internal ribosome entry site
TGFβ - transforming growth factor β	IFN β - Interferon β
FATP1 – Fatty acid transport protein 1	PCBP2 - poly(rC) binding protein 2
LCFA - long-chain fatty acids	MAVS - Mitochondrial anti-viral signalling protein
UV - Ultraviolet	SUMO – Small Ubiquitin-like Modifier
ROS - reactive oxygen species	CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats
CXCR1 - CXC-chemoreceptor 1	UAS – Upstream Activating Sequence
OSCC - oral squamous cell carcinoma	EPRS - Glutamyl-Prolyl tRNA Synthetase
HLD - hypomyelinating leukodystrophy	RRS - Arginyl tRNA Synthetase
NFL - neurofilament light protein	SCR - SUMO conjugation resistant
CMT - Charcot-Marie-Tooth	DGRC - <i>Drosophila</i> Genome Resource Centre
NRP1 - neuropilin 1	COTF - CRISPR Optimal Target Finder
HDAC6 - Histone deacetylase 6	LB - Luria-Bertani
ALS – Amyotrophic lateral sclerosis	PBS – Phosphate Buffer Saline
SOD1 - superoxide dismutase 1	LZD - Leucine zipper domain
MSC – Multi-aminoacyl tRNA Synthetase complex	PCR – Polymerase Chain Reaction
IgE - immunoglobulin E	GFP – Green Fluorescent Protein
IgE-FcϵRI - Immunoglobulin E high-affinity receptors	WT - Wildtype
MEK - MAPK kinase kinase	ASSD - anti-synthetase syndrome
MITF - microphthalmia-associated transcription factor	NLS – Nuclear localisation Signal
HINT - Histidine triad nucleotide binding protein	

Synopsis

Regulation of the *Drosophila* innate immune response by SUMO conjugation of amino-acyl tRNA Synthetases

For my Ph.D. project, I worked on a large macromolecular complex that exists in all eukaryotic cells, called the multi-amino acyl tRNA synthetase complex (MARS) which is an association of at least eight aminoacyl tRNA synthetases (AARS), with three non-AARS proteins. The MARS complex exists from *Drosophila* to mammals but the requirement for its formation in the cell remains a mystery. The Ratnaparkhi laboratory works on uncovering the function of SUMO-conjugated proteins and what caught my interest was that many tRNA synthetases, including those in the MARS complex, were potential targets of SUMO conjugation. Of these, the Glutamyl prolyl tRNA Synthetase (EPRS) had been shown to be SUMO conjugated, but the biological roles for SUMO conjugation are uncertain.

I, therefore, attempted to first experimentally define the presence of the MARS complex in *Drosophila* (Chapter II) and then studied the response of the complex to infection (Chapter II), all using mass spectrometry as a tool. Once I established that the MARS complex existed, I screened, using *in-bacto* SUMO conjugation assay, the number, and extent of SUMO conjugation in the proteins that make up the MARS complex (Chapter III). Of these, I first focused on RRS, a SUMO target, and uncovered possible roles for RRS in the immune response, as well the subtle effects of SUMO conjugation (Chapter III). In Chapter IV, I moved on to my second target, EPRS, but my attempts to generate an *EPRS*^{SCR} animal were unsuccessful. Below, I detail the contents of each Chapter in my thesis.

In **Chapter I**, I introduce the aminoacyl tRNA Synthetases (AARSs) and their canonical function; aminoacylation/tRNA charging in protein synthesis along with the description of the assembly of certain AARSs into a cytoplasmic, stable, supramolecular complex called the multi-amino-acyl tRNA Synthetase (MARS) complex; comprising of eight tRNA synthetases and three auxiliary proteins called the Amino acyl tRNA synthetase interacting proteins (AIMPs). I also provide an insight into the evolution of these AARSs over the course of time in terms of acquiring additional appended domains required for carrying out non-canonical/moonlighting functions. I also list out all the non-canonical functions of AARSs in the context of immunity.

In **Chapter II**, our goal was to isolate/purify the MARS complex and study its role in response to infection. In previous studies, various tools have been used to monitor the MARS complex like gel-filtration, immunoprecipitation of the components of the MARS complex by using an antibody raised against one of the members of the MARS complex or by using an antibody against tagged versions of the same. In our study, we have used an antibody raised against Glutamyl prolyl tRNA Synthetase (EPRS) for immuno-precipitating the MARS complex. EPRS exists as a homodimer as a part of the Sub-complex I and its closest interacting member is Iso-leucyl tRNA Synthetase. We could capture all the members of the MARS complex except for Lysyl tRNA Synthetase (KRS) suggesting that KRS is farthest away from EPRS in the MARS complex. Moreover, we could demonstrate that the MARS complex is sexually dimorphic; nearly the entire pool of EPRS associates with the MARS complex in males whereas in females only one-third associate with the MARS complex, and the rest of the two-thirds either exist as free-standing protein or in association with other interacting partners. To elucidate its role in immunity we infected the wildtype flies with Gram-positive bacteria, *Micrococcus luteus*, and Gram-negative bacteria, *Erwinia carotovora carotovora* (*Ecc15*). In response to infection, the MARS complex seems to be stabilized; with more units being assembled upon infection, indicating a plausible role for the complex to fight infection. We could also demonstrate the existence of a mammalian Gamma Inhibitor of Translation (GAIT) like complex in *Drosophila*. Similar to that in mammals, four hours post infection EPRS interacts with Syncrip (NSAP1) to form a pre-GAIT complex and later with Gapdh to form the GAIT complex. EPRS interactome is modulated upon infection; some prominent interactions are lost while some others are gained.

In **Chapter III**, the goal of our study was to identify a role for SUMOylation of Arginyl tRNA Synthetase (RRS) in the innate immunity of the fly. We found that RRS, a member of subcomplex II of the MARS complex, is a SUMO target. We demonstrated that the sites for SUMO conjugation are Lys 147 and 383. Furthermore, the replacement of these residues by Arg, RRS^{K147R, K383R} created a SUMO conjugation-resistant variant (RRS^{SCR}). We generated the transgenic *Drosophila* lines for RRS^{WT} and RRS^{SCR} by expressing these variants in RRS loss of function (RRS^{lof}) animals, using a UAS-Gal4 system. We created the RRS^{lof} line by dual gRNA strategy-based CRISPR Cas9 genome editing technology. We found that both RRS^{WT} and RRS^{SCR} rescue the RRS^{lof} lethality. Adult animals expressing RRS^{WT} and RRS^{SCR} were compared and contrasted for their response to bacterial infection by Gram-positive *M. luteus* and Gram-negative *Ecc15*. We found that RRS^{SCR} when compared to RRS^{WT} showed

modulation of transcriptional response, as measured by quantitative 3' mRNA sequencing. Our study thus uncovers a possible non-canonical role SUMOylation of RRS, an amino-acyl tRNA Synthetase.

Chapter IV deals with our attempt to elucidate the role of SUMOylation of EPRS and its role in response to infection. Past studies and our work validate that EPRS, a member of the MARS complex, is a SUMO target. We confirmed that EPRS gets SUMOylated at Quintuple lysines namely K957, K1063, K1083, K1106, and K1198. We generated the transgenic *Drosophila* line for EPRS^{SCR} by using SCARLESS-based CRISPR Cas9 genomic editing technology. All the SUMO conjugation-resistant EPRS^{SCR} lines were unstable and reverted to wild type in a few months post-generation of the line.

Publications

1. Nayak P, Kejriwal A, Ratnaparkhi GS. SUMOylation of Arginyl tRNA Synthetase Modulates the *Drosophila* Innate Immune Response. *Front Cell Dev Biol.* 2021 Sep 30;9:695630. doi: 10.3389/fcell.2021.695630. PMID: 34660574; PMCID: PMC8514731.
2. Gene knockouts in *Drosophila* using CRISPR-Cas9 based genome editing. Hegde S, Nayak P, Trivedi D, Ratnaparkhi GS (2021). In 'Experiments with *Drosophila* for Biology Courses'. Editors, SC Lakhotia and HA Ranganath. Indian Academy of Sciences. eBook.
3. Shukla, Vallari, Neena K. Dhiman, Prajna Nayak, Neelesh Dahanukar, Girish Deshpande and Girish S. Ratnaparkhi. "Stonewall and Brickwall: Two Partially Redundant Determinants Required for the Maintenance of Female Germline in *Drosophila*." *G3: Genes/Genomes/Genetics* 8 (2018): 2027 - 2041.

Abstract

Post-translational modification of a substrate protein by SUMO (Small Ubiquitin-related modifier) can modify its activity, localization, interaction or function. A large number of SUMO targets in cells have been identified by proteomic studies, but the biological roles for SUMO conjugation for most targets remain elusive.

Multi-aminoacyl tRNA Synthetase complex (MARS) is a large cytoplasmic 1.2 MDa signalling hub that acts as a sensor and regulator of the immune response. MARS consists of eight Amino-acyl tRNA Synthetases (AARS) and three non-synthetase adaptors (AIMP1-3). Using quantitative proteomics, we have determined that the members of the MARS complex showed enhanced SUMO conjugation in response to an immune challenge (Handu *et. al.*, 2015). Subsequently, I could demonstrate that eight of its eleven members were SUMO conjugated using *in-vitro* SUMOylation assays. Immunoprecipitation of the MARS complex, followed by mass spectrometry, suggests that the complex was stabilized in response to both gram-positive and gram-negative infection in adult flies, underscoring a role for MARS in the *Drosophila* immune response.

Glutamyl-Prolyl tRNA Synthetase (EPRS), a member of MARS sub-complex I is SUMO conjugated at its WHEP domain, which is involved in non-canonical roles. In mammals, EPRS dissociates from the MARS complex in response to infection, to form a secondary 'GAIT' complex, that regulates translation. In order to study roles for SUMO conjugation of EPRS, I have used CRISPR Cas9 genome editing technology to generate a SUMO conjugation resistant (SCR) variant (EPRS^{SCR}; EPRS^{K957R, K1063R, K1083R, K1106R, K1198R}). The transgenic lines generated were unstable, precluding the exploration of immune regulation in EPRS^{SCR} flies. Arginyl tRNA Synthetase (RRS), a member of sub-complex II of MARS is also SUMO conjugated. A SCR variant (RRS^{SCR}; RRS^{K147R, K383R}) was uncovered by a combination of *in-bacto* SUMOylation assay with Lys mutagenesis. Transgenic *Drosophila* lines of RRS^{WT} and RRS^{SCR} were made by expressing these variants in an RRS null (*DRRS*) animal, using the UAS-Gal4 system. The *DRRS* line was itself generated using CRISPR Cas9 genome editing, using a dual guide-RNA system. Both RRS^{WT} and RRS^{SCR} rescue the *DRRS* lethality. Adult animals expressing RRS^{WT} and RRS^{SCR} were compared and contrasted for their response to bacterial infection. Interestingly, RRS^{SCR} animals show an upregulated immune response upon infection, in comparison to RRS^{WT}, as measured by the activation of defence genes using

quantitative RNA sequencing. This suggests that SUMOylation of RRS is necessary to restrain aberrant NFkB signalling upon an immune challenge.

My research highlights the significance of SUMO conjugation of tRNA synthetases in host defence and uncovers a non-canonical role for SUMOylation of RRS, a member of the MARS complex, in the *Drosophila* immune response.

The Multi-Aminoacyl tRNA Synthetase (MARS) complex in Innate Immunity

Summary

In this introductory chapter, I review the current literature about the multi-aminoacyl tRNA synthetase (MARS) complex in terms of its role in innate immunity. I describe the 1.4 Mega Da MARS complex, its role, and of its components, highlighting roles in the regulation of the animal innate immune response.

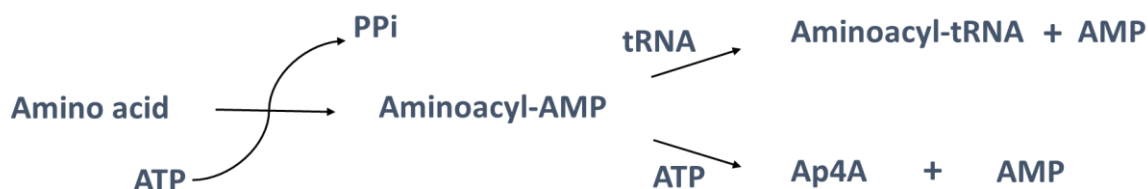
1. The *Drosophila* innate immune response

Drosophila, like other insects, co-exists with microorganisms in the native environment. In the wild, larvae develop in decaying organic matter, and adults sometimes serve as vectors for microorganisms causing diseases in plants and animals. Hence *Drosophila*, like other holometabolous insects, has developed sensitive mechanisms for the recognition of pathogens and strategies to fight off attacks against bacteria, fungi, parasites, and viruses. It combats infection via multiple innate immune defence reactions, which are somewhat similar to those of higher organisms. *Drosophila*, like its other insect counterparts, is incapable of mounting an adaptive immune response and relies solely on its innate immunity to combat infection.

Innate immune response in *Drosophila* can be classified into three categories a) 1st line of defence comprising of physical barriers in the form of chitin exoskeleton, tracheal and intestinal epithelia referred to as epithelial immunity to protect the fly from the invading pathogens, b) 2nd line of defence -Activation of proteolytic cascades leading to coagulation and melanization also referred to as the cellular response. Three types of hemocytes participate in the cellular response i) plasmatocytes which act as professional phagocytes. ii) crystal cells that are non-phagocytic and participate in melanization reactions. iii) lamellocytes, formed in response to wasp infection or wounding and are found only in larval stages. Both crystal cells and lamellocytes participate in the melanization reactions, but the role of crystal cells in encapsulation is not yet deciphered. c) 3rd line of defence – Activation of immune signalling pathways and systemic production of immune effectors, also referred to as the humoral response. Systemic production of a battery of Anti-Microbial Peptides (AMPs) by the fat body into the surrounding haemolymph is considered the hallmark of the humoral immune response. The production of AMPs is controlled by Toll and Imd signalling pathways. These pathways are activated by nuclear translocation of NF- κ B transcription factors Dif/Dorsal or Relish

followed by induction of a differential set of AMPs and other immune-responsive genes (Vanha-Aho, Valanne, and Rämetsä 2016; Buchon, Silverman, and Cherry 2014).

2. The physiological role of AARSs: Protein synthesis



Amino-acyl tRNA synthetases (AARS) catalysed aminoacylation reaction proceeds in two steps. In the first step, AARS catalyses the condensation of substrates cognate amino acid and ATP to form an enzyme-bound reaction intermediate, aminoacyl adenylate. In the second step, the cognate tRNA is docked into the catalytic site of AARS. The activated amino-acyl moiety is transferred to the 3' terminal adenosine acceptor end of the cognate tRNA accompanied by the release of AMP. AARSs also take part in additional secondary chemical reactions, i.e., the synthesis of diadenosine oligo phosphate (ApnA). Enzyme-bound amino-acyl adenylate is attacked by pyrophosphate moiety of an additional ATP instead of a tRNA resulting in the formation of ApnA. ApnA acts as a second messenger for cell regulation.

2.1 Types Based on the architecture of active sites

The AARSs can be categorized into two major classes based on the architecture of their active sites (Cusack et al. 1990; Eriani et al. 1995; Burbaum and Schimmel 1991; Cusack 1997; P. Schimmel and Ribas de Pouplana 2001; O'Donoghue and Luthey-Schulten 2003). In Class I synthetases, the catalytic domain bears a dinucleotide or Rossmann ATP binding Fold (RF) featuring a five-stranded parallel β sheet connected by α helices and is usually located at or near the N-terminus of the protein. The Rossmann Fold comprises two highly conserved motifs, HIGH (His-Ile-Gly-His) and KMSKS(Lys-Met-Ser-Lys-Ser), that mediate interaction with ATP (J. G. Arnez et al. 1995; Brick, Bhat, and Blow 1989; Schmidt and Schimmel 1994), separated by a connecting domain called Connecting Peptide 1(CP1) (Starzyk, Webster, and Schimmel 1987). In Class II synthetases, the catalytic domain bears seven-stranded anti-parallel β sheets flanked by α helices and feature three motifs which show a lesser degree of conservation as compared to those in Class I (Cusack et al. 1990; Z. M. Arnez et al. 1995; Eriani et al. 1995). Both classes also exhibit pronounced differences in their modes of substrate

binding. Class I AARSs bind to the minor groove of the tRNA acceptor stem (with the exception of TrpRS and TyrRS) and aminoacylate 2'-OH group of the ribose of A76. On the other hand, Class II AARSs approach tRNA from the major groove and aminoacylate 3'-OH group (with the exception of PheRS) (Sprinzl and Cramer 1975; Ruff and Weissman 1991; M. Ibba, Stathopoulos, and Söll 2001). Due to the difference in active site structure, the two classes of AARSs also differ in their mode of ATP binding. Class I AARSs bind ATP in an extended conformation (Brick and Blow 1987; Brick, Bhat, and Blow 1989; Rould et al. 1989), whereas Class II bind in a bent configuration with the γ phosphate folding back over adenine ring (John J. Perona and Hadd 2012). The two classes also vary concerning the kinetics of aminoacylation reaction. For Class I AARSs, aminoacyl-tRNA release is the rate-limiting step (with the exception of IleRS and some GluRS), whereas, for Class II, it is the amino-acid activation rate (Fersht 1977; J. J. Perona et al. 1991; Kaminska, Shalak, and Mirande 2001). In contrast to the Class I enzymes, which tend to be monomeric, Class II AARSs are almost exclusively oligomeric and are usually in the dimeric or tetrameric form.

Class I and Class II are further subdivided into different sub-groups based on phylogenetic analysis, mechanistic properties, anticodon binding domain characteristics, and organization of conserved structural motifs. Class II synthetases are subdivided into three subgroups (a, b, and c); the classification of class I is more complex, classified into three subgroups by some authors (P. Schimmel and Ribas de Pouplana 2001), while others propound up to five subclasses (John J. Perona and Hadd 2012; Valencia-Sánchez et al. 2016). AARSs are grouped into sub-classes based on their amino-acid substrates. Subclass Ia recognizes aliphatic amino acids such as Ile, Leu, and Val and thiolated amino acids such as Met and Cys, while class Ic AARSs activate the aromatic amino acids Trp and Tyr. Class Ib enzymes activate charged amino acids such as Lys, Glu, and Gln, while their class IIb counterparts activate polar amino acids Lys, Asp, and Asn.

The AARSs are structurally diverse as they recognize structurally and chemically different cognate substrates to avoid mischarging similar or non-cognate substrates. For this purpose, some AARSs have acquired insertions or appendages to the canonical catalytic cores. In many instances, the inserted domains enhance enzyme specificity and fidelity of the aminoacylation reaction. To prevent mischarging tRNAs in protein synthesis, some synthetases have evolved editing activities that specifically target and hydrolyse mis-activated amino acids and/or mis-acylated tRNAs. The editing activity may be carried out by separate editing domains present

near the catalytic site or by separate freestanding proteins. In Class I synthetases, the editing domains are usually located as the connecting peptide CP1 within the active site, while in Class II editing activity can be confined to different domains (Schmidt and Schimmel 1994; Schmidt and Schimmel 1995; Lin and Schimmel 1996; Nureki et al. 1998; Dock-Bregeon et al. 2000; Giegé, Sissler, and Florentz 1998).

2.2 Species-specific structural diversity (Domain analysis)

The catalytic domains of AARSs show a high degree of conservation across species. The structural and functional differences are more pronounced between classes than between species (S. Kim 2014; Guo, Yang, and Schimmel 2010). The protein sequences in Class I AARSs are phylogenetically well conserved within the same subclass (Ia, Ib, and Ic), pointing towards a common ancestor (Fournier et al. 2011). The editing domains of Class I AARSs show homology in structure and sequence. The CP1 editing domains of LeuRS, ValRS, and IleRS are highly conserved from *Escherichia coli* to humans (S. Kim 2014; Beuning and Musier-Forsyth 2001). The INS domain exists in prokaryotic ProRS but is absent in the eukaryotic one. The N2 editing domain of ThrRS is well conserved amid bacteria and eukaryotes but non-existent in archaea (Beuning and Musier-Forsyth 2001; Beebe, Ribas De Pouplana, and Schimmel 2003). On the other hand, the editing domain of the bacterial PheRS does not share any apparent sequence similarity with its archaeal and eukaryotic counterparts (Sasaki et al. 2006). Even though the catalytic core has remained relatively well conserved, AARSs have acquired new sequence motifs and appended domains throughout evolution, ensuing the expansion of their functional capacities to adapt to the increasing complexity (Guo, Yang, and Schimmel 2010; Guo and Yang 2014).

2.3 The evolution of AARSs

Even though the catalytic function of AARSs has been extensively studied and understood, the regulatory functions beyond translation have recently surfaced (S. Kim, You, and Hwang 2011; Yao and Fox 2013). The evolution of AARSs was most likely first driven by the need for enhancement of catalytic efficiency and fidelity and then by the accretion of novel functions to meet the growing complexity with the evolution of eukaryotes.

Catalytic evolution: The functional domains of AARSs probably evolved in a piecemeal fashion, beginning with an ancient core enzyme that activated the amino acid and mediated the docking of RNA oligonucleotides for aminoacylation (Paul Schimmel 2018; P. Schimmel and

Ribas De Pouplana 2000). RNA oligonucleotides, in all likelihood, originated in a primitive acceptor stem form that evolved to become the current tRNA structure via the acquisition of anticodon stem and loop parts. Through coevolution with tRNAs, AARSs acquired Anticodon binding domains (ABDs) to facilitate additional tRNA interactions and integrated editing domains. Mis-aminoacylation can be detrimental to the organism. As a result of which, editing domains may have emerged in the last common primogenitor before the divergence of the three kingdoms (Guo, Yang, and Schimmel 2010).

Non-catalytic evolution: AARSs are involved in protein-protein interactions via the appended domains acquired during evolution. Appended domains exist in the form of N-terminal helix appendix, Leucine zipper (LZ), Endothelial monocyte activating polypeptide II (EMAP II), Glutathione-S-transferase (GST), and WHEP domains (Guo, Yang, and Schimmel 2010; Guo and Yang 2014). AARSs are also associated with transacting factors like aminoacyl tRNA synthetase interacting proteins (AIMPs). The emergence of auxiliary domains coincides with the expansion of AARSs function beyond its catalytic role in protein synthesis.

N-terminal helix: N-helix is anchored to an anticodon binding domain (ABD) of eukaryotic AspRS (DRS) and LysRS (KRS). It interacts with the elbow region of the tRNA and augments binding specificity. Apart from its role in catalysis, the N-helix of KRS is also involved in the packaging of HIV by delivering tRNA^{Lys3} (an iso-acceptor of tRNA^{Lys}) into a virion. It enables KRS to translocate to the plasma membrane and aids in laminin-mediated cell migration by interacting with and stabilizing the Lamin receptor (also known as 40S ribosomal protein SA) (D. G. Kim et al. 2012, 2014).

Leucine Zipper: The leucine zipper domain is present at the N-terminus of ArgRS (RRS), AIMP1, and AIMP2 and is conserved from insects to mammals. It enables these proteins to assemble into the MARS complex (Fu et al. 2014).

Endothelial monocyte activating protein II (EMAP II): EMAP II domain is present in the C-terminus of AIMP1 and TyrRS (YRS) and is conserved from insects to mammals. It is also found in MetRS (MRS) in *Caenorhabditis elegans* and a yeast orthologue of AIMP1 called ARS cofactor 1 (Arcp1). It facilitates interaction with tRNA and is also known to be secreted in response to immune challenges and angiogenesis (Guo and Yang 2014; Wakasugi and Schimmel 1999; D. Kim, Kwon, and Kim 2014).

Glutathione-S-transferase (GST): GST is appended in GluProRS (EPRS), MetRS (MRS), ValRS (VRS), CysRS, arcp1, AIMP2, and AIMP3. GST domains of EPRS, MRS, AIMP1, and AIMP2 come together to form a heterotetramer aiding in the assembly of the MARS complex (Guo, Yang, and Schimmel 2010; Cho et al. 2015).

WHEP domain: The WHEP domain is composed of a helix turn helix structure. It has diverse sequences in different AARSs and hence is capable of interacting specifically with multiple partners. WHEP domain is present in TrpRS (WRS), HisRS (HRS), GluProRS (EPRS), MetRS (MRS), and GlyRS (GRS). WHEP domain was first discovered in the first four proteins and hence the name. In EPRS, it is involved in the assembly of INF γ activated inhibition of translation (GAIT) complex, which inhibits translation of specific pro-inflammatory mRNA transcripts like those of Vascular endothelial growth factor A (VEGFA) (Arif et al. 2009; Jia et al. 2008). WHEP domain of WRS interacts with the catalytic unit of DNA-dependent protein kinase (DNA-PKc) and poly (ADP-ribose) polymerase I (PARPI), which in turn activates p53 in the nucleus (Sajish et al. 2012). Initial 154 amino acids at the N-terminus of the WHEP domain respond to infection and activate macrophages. It interacts with a heterodimer formed by Toll-like receptor 4 (TLR4) and myeloid differentiation factor 2 (MD2) (Ahn et al. 2016).

Uniquely attached sequence motif (UNE): UNE in PheRS (FRS), AsnRS (NRS), and GlnRS (QRS) is involved in tRNA binding and other protein-protein interactions. UNE-I₂ of IleRS (IRS) assists in MARS complex assembly and other functions (Guo and Yang 2014). In LeuRS (LRS), a variable C-terminal domain (VC) and UNE-L domain interact with RAS-related GTP binding protein D (RAGD) to regulate mTORC1 (Han et al. 2012). In yeast, the CP1 domain acts as a GTPase (Bonfils et al. 2012). In Zebrafish, UNE-S of SerRS (SRS) encodes a nuclear localization signal (NLS), translocates into the nucleus, and attenuates VEGFA expression (X. Xu et al. 2012).

2.4 Alternative splicing

HRS Δ CD, a splice variant of HisRS, lacks the catalytic domain and constitutes a structure formed by the fusion of the N-terminal WHEP domain and C-terminal anticodon binding domain (Z. Xu et al. 2012). HRS Δ CD attains a different conformation in comparison to the full-length (FL) variant by which it exists as a monomer (The native FL variant exists as a homodimer). This conformation loosens the connection between the WHEP domain and the anticodon binding domain. A splice variant of WRS exists in the form of a truncated N-terminal

WHEP domain and functions as an angiostatic factor upon secretion (Tolstrup et al. 1995). In another instance, a splice variant of YRS, YRSSV-N13, stimulates megakaryopoiesis (Kanaji et al. 2018). CRS also has a splice variant wherein an additional peptide is inserted into the native protein, which mediates its interaction with Elongation factor 1 γ (EF1 γ) (J. E. Kim et al. 2000). Besides the AARSs, the AIMP2s also exhibit the presence of splice variants involved in an alternative function. AIMP2-DX2, a splice variant of AIMP2, lacks the leucine zipper domain. It acts in an antagonistic fashion to the FL variant by interfering with the anti-proliferative signaling of AIMP2 (Yao et al. 2012; D. G. Kim et al. 2016). EPRS^{N1}, a splice variant of EPRS (N-terminally truncated EPRS fragment), is generated by post-translational modification (PTM) wherein an alternative polyadenylation event within EPRS mRNA recodes Tyr codon to a stop codon; leading to deletion of entire PRS and partial WHEP domains. EPRS^{N1} is shown to associate with VEGFA transcripts but not with the constituents of the GAIT complex, inhibiting repression of VEGFA translation (Yao et al. 2012).

2.5 Post-translational modification (PTM)

AIMP2 gets phosphorylated on Ser¹⁵⁶, inducing its translocation into the nucleus wherein it interacts with SMURF2, leading to an increase in transforming growth factor β (TGF β) signaling (D. G. Kim et al. 2016). LysRS (KRS) undergoes a conformational change and dissociates from the MARS complex upon being phosphorylated on either Thr⁵² or Ser²⁰⁷. Thr⁵² phosphorylation enables KRS to translocate to the plasma membrane and associate with a 67kDa laminin receptor (67LR), thereby promoting cell migration. Ser²⁰⁷ phosphorylation permits KRS to enter into the nucleus where it is involved in the localized production of second messenger Ap₄A and activates transcription factor MITF, furthermore expression of MITF-induced genes (D. G. Kim et al. 2014; Ofir-Birin et al. 2013).

In human monocyte cells of the myeloid lineage, sequential phosphorylation of Ser⁸⁸⁶ and Ser⁹⁹⁹ residues on EPRS embarks on its release from the MARS complex and formation of the GAIT complex. Insulin stimulation on adipocytes results in phosphorylation of EPRS on Ser⁹⁹⁰ by S6K1, inducing its interaction with FATP1, consequently transporting it to the plasma membrane where it enhances uptake of long-chain fatty acids (LCFA) (Arif et al. 2017).

UV irradiation ablates the tRNA binding capability of MetRS (MRS) by phosphorylating it on Ser⁶⁶², therefore, preventing protein synthesis (Kwon et al. 2011). This phosphorylation is mediated by GCN2 (E2AK4). MRS undergoes dual phosphorylation on Ser²⁰⁹ and Ser⁸²⁵ when

the cell encounters oxidative stress. It decreases MRS specificity for tRNA^{Met} and enhances the misincorporation of methionine into nascent proteins(J. Y. Lee et al. 2014). Methionine residues on the surface of the proteins scavenge reactive oxygen species (ROS) hence Met mistranslation acts as a protective mechanism for the cell(J. Y. Lee et al. 2014; Luo and Levine 2009).

2.6 Proteolytic cleavage

Secreted forms of YRS and WRS are cleaved by elastase or plasmin(Wakasugi and Schimmel 1999; Otani et al. 2002; Tzima et al. 2005). The ELR cytokine motif of YRS is exposed upon removal of the EMAP II domain. It aids in its interaction with CXC-chemoreceptor 1 (CXCR1). Mini-YRS promotes cell migration and angiogenesis(D. Kim, Kwon, and Kim 2014; Vo, Yang, and Schimmel 2011). WRS undergoes N-terminal cleavage to form two different variants of WRS, T1-WRS (70 amino acids deletion at the N-terminus) and T2-WRS (93 amino acids deletion at the N-terminus). T1-WRS has an activity antagonistic to T2-WRS; it interacts with vascular endothelial cadherin of endothelial cells(Tzima et al. 2005; S. B. Kim et al. 2017).

hKRS undergoes caspase 8 mediated cleavage resulting in the deletion of 12 amino acids of 1st subunit from the N-terminus, enabling interaction of the N-terminal region of one KRS with the C-terminal region of another KRS and vice versa, forming a homodimer. The formation of the homodimer facilitates its interaction with syntenin assisting in its incorporation into secondary exosomes(S. B. Kim et al. 2017).

2.7 The connection between catalytic and non-catalytic functions

A mutant of AlaRS, Ala734Glu (charges Serine to tRNA^{Ala} instead of alanine), interacts with ANKRD16 (a vertebrate-specific protein containing ankyrin repeats) via catalytic domain and anticodon binding domain(Vo et al. 2018). It helps ANKRD16 in pre-transfer editing resulting in the sequestration of mis-activated serine and preventing protein aggregation and cell death.

Fragmented forms of WRS (T2-WRS) and YRS have amino acid binding pockets enabling them to develop multifunctionality. T2-WRS interacts with Trp residues of VE-cadherin, ensuing angiostasis(Q. Zhou et al. 2010). Truncated YRS interacts with the Tyr analogue of resveratrol, redirecting YRS to bind to PARP1, thus activating DNA damage response(Sajish and Schimmel 2015).

2.8 AARSs in disease

Pathological expression: Expression levels of AARSs can be used as a prognostic tool for the repertoire of cancers. It correlates with overall patient survival for individual cancer types. In most of the scenarios, high expression levels of AARSs correlates with lower patient survival with few exceptions like reduction in the levels of WRS and AIMP3 in case of ovarian cancer, AIMP2 in renal, AIMP3 and FRS α in cervical, and FRS α in stomach cancer. These observations can be attributed to the role of AARSs in the suppression of proliferation(Sajish et al. 2012; Choi et al. 2011; B.-J. Park et al. 2005, 2006; Choi, Um, et al. 2009). Enhanced expression of GRS is observed in renal, urothelial, liver, breast, and endometrial cancers(Thul and Lindskog 2018). Expression levels of MRS are augmented in the neoplastic region of the lung tissue. This increase in the levels of MRS correlates with advanced-stage cancer and a shortening of lifespan(E. Y. Kim et al. 2017). High expression of MRS has also been reported in malignant fibrous histiocytomas, lipoma, osteosarcomas, malignant gliomas, and glioblastomas(S. Kim, You, and Hwang 2011; Forus et al. 1994; Nilbert et al. 1995; Palmer et al. 1997; Reifenberger et al. 1996). A wide range of AARSs, namely FRS α , GRS, NRS, TRS, HRS, and WRS, are upregulated in prostate cancer(Vellaichamy et al., 2009). This occurrence can be ascribed to their involvement in androgen response(Vellaichamy et al. 2009). TRS, VEGFA, and mucin 1 have been reported to co-express and co-localize in advanced-stage epithelial ovarian cancer and pancreatic cancer, respectively(Wellman et al. 2014; Jeong et al. 2018). Elevated levels of WRS have been documented in oral squamous cell carcinoma (OSCC)(C.-W. Lee et al. 2015; Chi et al. 2009). In OSCC, the IFN γ pathway is significantly altered(Chi et al. 2009). IFN γ is known to regulate the splicing, expression, localization, and secretion of WRS. Expression levels of WRS correlate with OSCC tumor stages and degree of invasion and cell migration(Chi et al. 2009; Tolstrup et al. 1995; Liu et al. 2004; Turpaev et al. 1996). WRS is avowed to have two tandem promoters and five splice variants. The full-length form and the truncated variant mini-WRS (which has a deletion of 47 amino acids from the N-terminal end(Chi et al. 2009; Tolstrup et al. 1995; Liu et al. 2004; Turpaev et al. 1996) have an antagonistic effect concerning controlling angiogenesis(Otani et al. 2002) and immune stimulation(Ahn et al. 2016).

Disease-specific mutations: Patients suffering from autosomal recessive non-syndromic sensorineural deafness harbours mutations in KARS(Cusack et al. 1990; van Meel et al. 2013).

Mis-sense mutations in KARS effectuate a decrease in its aminoacylation in cells of the inner ear. Mutations in a multitude of AARSs cause hypomyelinating leukodystrophy (HLD)(Garbern 2007). To date, seventeen forms of HLD have been catalogued, linked to various heterogeneous mutations. In all cases, the catalytic activity of AARSs is affected, subsequently eliciting a diminution of cellular translation, which might be the underlying cause of the disease(Nafisinia et al. 2017; Wolf et al. 2014). HLD 3,9,15,17 are caused by mutations in AIMP1, RARS, EPRS, and AIMP2, respectively(Nafisinia et al. 2017; Wolf et al. 2014; Mendes et al. 2018; Shukla et al. 2018; Iqbal et al. 2016). AIMP1 has been shown to maintain normal phosphorylation levels of neurofilament light protein (NFL). Phosphorylation level maintenance is required for the structural integrity of neurons(Zhu et al. 2009). In HLD3, owing to frameshift mutations, C-terminally truncated forms of AIMP1, lacking EMAP II domain, are transcribed. The truncated forms of AIMP1 retain the ability to interact with NFL, thereby promoting aggregation and punctate formation in cells(H. Xu et al. 2015). The mutant protein acts as a dominant-negative, and the competition between the wild type and the mutant protein is responsible for HLD3. These multifarious AARSs mutations lead to early infantile epileptic encephalopathy. Mutations in MARS and LARS bring about acute infantile liver failure and multiple organ dysfunction(Rosenthal 2018; van Meel et al. 2013). Heterozygous mutants of MRS show a reduction in aminoacylation activity. LARS mutations bear aberrations in CP1 domains which affects the catalytic function of LARS(Cusack et al. 1990; van Meel et al. 2013). Patients possessing these faulty forms of LARS show dysfunction in the mTORC1-related pathway, such as autophagy(Cusack et al. 1990; van Meel et al. 2013). Mutations in IARS are correlated to infantile hepatopathy(Kopajtich et al. 2016). Mutants of HRS show reduced catalytic activity and in patients induce hearing loss and visual impairment in those suffering from a rare genetic disorder called Usher syndrome(Puffenberger et al. 2012). Compound heterozygous mutations in QARS lead to defects in gene transcription and expression, ultimately leading to cerebral, cerebellar atrophy(Zhang et al. 2014).

Mutations in AARSs cause diseases characterized by defects in gene transcription and protein translation. Mutations pArg40Trp and pArg515Trp in QARS disrupt its catalytic activity. Interestingly mutation in FARS β affects the tRNA charging function out does not affect total protein synthesis(Cusack et al. 1990; Michael Ibba 2005). These mutations reduce the gene expression, hamper the stability of protein, and in retrospect, affect the levels of its binding partner FARS α . Thus, affects the rate of aminoacylation of Phe but does not affect total protein synthesis or cell proliferation (Cusack et al. 1990; Michael Ibba 2005).

Mutation-associated interactions: Mutations in mitochondrial AARSs are responsible for neurodegenerative disorders and also sporadic affect other organs such as skeletal muscle, kidney, lungs, and heart. Autosomal dominant mutations in GRS, YRS, AlaRS, KRS, HRS, and MRS are associated with Charcot-Marie-Tooth (CMT) disease. Close to 20 percent of CMT causes mutations to affect the catalytic activity of the corresponding AARSs(Datt and Sharma 2014). Patients with mutations in the GARS gene suffer from CMT type 2D (CMT2D) subtype distal muscular dystrophy (dSMA-V)(Martin, Mentis, and Tosolini 2021; Storkebaum 2016).

Mutation Pro234LysTyr in GARS in mice causes CMT2D. The mutation is located near the dimeric interface inducing a conformational change that unfolds a new protein interface. The mutant proteins interact more strongly with neuropilin 1 (NRP1) as compared to the wild type, thereby outcompeting VEGFA for NRP1 binding. This results in the induction of VEGFA signalling leading to caudal migration of facial motor neurons from the rhombomere during embryonic development(Schwarz et al. 2004). The disease phenotype manifests due to disruption of NRP1 signalling. Other mutations documented in GARS, i.e. Pro234LysTyr, Ser581Leu, and Gly598Ala. These mutant proteins bind to and enhance the activity of Histone deacetylase 6 (HDAC6) on α -tubulin(Mo et al. 2018). Deacetylation of α -tubulin impairs axonal transport in CMT2D mutants harbouring Pro234LysTyr in mice. Mice bearing Ser581Leu and Gly598Ala mutations endure severe distal weakness and wasting in the lower limbs. Mutations Cys201Arg and Cys157Arg in mice and humans, respectively, are gain-of-function mutants, which anomalously interact with Tyrosine kinase receptors leading to sensory deficits(Sleigh et al. 2017).

In ALS patients, mutants of superoxide dismutase 1 (SOD1) bind to mitochondrial KARS2 leading to misfolding and aggregation and eventually leading to degradation of KARS2 followed by an impaired mitochondrial translation(Kunst et al. 1997). MRS mutant attains the ability to interact with CDK4, accelerating the cell cycle. It also interacts with tumour suppressor p16^{INK4a}. p16^{INK4a} is inactivated by either a conformational change by gene deletion or by interaction with the mutant, rendering it inactive(Mehta and Siddik 2009).

Parkin, a known multifunctional E3 ligase polyubiquitinates substrates which, are then targeted for proteasome-mediated degradation(Vink and Nechifor 2011). Loss of function of parkin results in Parkinson's disease(Lahiri 2014). Parkin mutants fail to bind and ubiquitylate AIMP2. AIMP2 accumulates in Lewy body inclusions in the substantia nigra(Ko et al. 2005).

Enhanced AIMP2 induces age-dependent loss of dopaminergic neurons(Y. Lee et al. 2013; David et al. 2009).

Disease-specific variant production: AIMP2-DX2, a splice variant of AIMP2, lacks a leucine zipper (LZ) domain and hence cannot be incorporated into the MARS complex. It competes with AIMP2 for binding to the cellular target protein through the GST domain, thereby interfering with the functions of full-length AIMP2 protein, inhibiting its tumour suppressing activities. AIMP2-DX2 levels are elevated in cancers(M. C. Park et al. 2012; Choi et al. 2011). The expression of AIMP2-DX2 is induced by carcinogenic stresses. It procures the ability to interact with a tumour suppressor p14^{ARF}. p14^{ARF} vital for tumour growth in KRAS driven lung cancers(D. Kim, Kwon, and Kim 2014; Choi et al. 2011; Choi, Kim, et al. 2009; Choi et al. 2012; Oh et al. 2016).

Patients suffering from autoimmune diseases (such as interstitial lung disease, dermatomyositis, and arthritis) develop anti-AARS auto-antibodies contributing to antisynthetase syndrome(Lega et al. 2014; Cavagna et al. 2017; J. J. Zhou et al. 2014). They also show the presence of anti-JO1 auto-antibodies raised against the N-terminal WHEP domain of WRS. A variant of HRS, HRS Δ CD (constituting a WHEP domain), reacts with anti-JO1 antibodies in the serum of the patients(Z. Xu et al. 2012).

Disease-specific secretion: During infection, WRS is secreted from monocytes resulting in the priming of innate immunity(Ahn et al. 2016). The cleaved product is 154 amino acids long peptide and contains a WHEP domain that triggers immune stimulation. It acts as an immune stimulatory agent and can be used as a therapeutic biomarker for the early diagnosis of sepsis. WRS levels continually and significantly increase upon infection(Ahn et al. 2016).

GRS is secreted in tumour-associated macrophages inducing tumour cell death upon interaction with K cadherin (CDH6) leading to the deactivation of the ERK pathway(M. C. Park et al. 2012).

3. MARS complex

In mammalian cells, eight polypeptides corresponding to nine tRNA synthetase activities and three non-tRNA synthetase factors come together to form a 1.4/5 MDa complex (SEC-MALS

analysis) called the MARS complex(Kerjan, Triconnet, and Waller 1992; Kerjan et al. 1994). The MARS complex is first isolated from vertebrates from rat liver, rabbit liver, reticulocytes, sheep liver, human placenta, and mammalian cells in culture (Brevet et al. 1982; Kellermann, Heuser, and Mertens 1982; Cirakoglu and Waller 1985; Venema and Traugh 1991). The three non-synthetase factors (p43, p38, p18)(Quevillon and Mirande 1996; Quevillon et al. 1997, 1999) are involved in tRNA binding, complex assembly, and stability(J. Y. Kim et al. 2002; Saxholm and Pitot 1979). The tRNA synthetases within MARS belong to both class I (monomers of RRS, QRS, IRS, LRS, and MRS) and class II (dimers of DRS EPRS and KRS) synthetase types. Thus, the distinction between the complexed and non-complexed forms does not rely on the structural architecture of the active site and the amino acids that they activate(Quevillon and Mirande 1996; Quevillon et al. 1997, 1999). The MARS complex harbours tRNA synthetases involved in charging polar, hydrophobic, and non-aromatic amino acids. The synthetases charging the smallest and largest amino acids are absent in the complex. It is unknown why a distinct group of tRNA synthetases associate to form the complex while others exist as freestanding proteins. Several others, or likely all, associate transiently in the cell to form the complex, but the current model of MARS is probably a reflection of experimental limitations when co-purifying macromolecular complexes. A correlation between the size of the substrates and their connection with the citric acid cycle has been proposed(Eswarappa and Fox, 2013). Even though the composition of the complex has been deciphered four decades ago, the exact structural arrangement and assembly are uncertain. With the advent of biochemical, genetic, and cryo-electron microscopic analysis, a coherent picture of the MARS complex has been captured(Saxholm and Pitot 1979; Sihag and Deutscher 1983; M. T. Norcum 1991, 1989). Under electron and immune-electron microscopy, the MARS complex is visualized as a compact V-shaped structure(Mona T. Norcum and Boisset 2002; M. T. Norcum 1989). The structure can be divided into two arms joined together by a base. Arm I constitute DRS, MRS, and QRS; on the other hand, Arm II is made up of KRS and RRS. The two arms are affixed to the base comprising EPRS, IRS, and LRS. The assembly of the constituents is interdependent and is augmented by other binding events within the complex. The constituents undergo an ordered assembly. NaSCN and a high concentration of NaCl induce the release of KRS and DRS, respectively, from the complex. This study points toward the fact that KRS and DRS are located at the periphery of the complex. MRS, QRS, and RRS are relatively easy to remove and are also found as free forms(Mona Trempe Norcum, Anthony Warrington, and Wolfe 2006). Extensive pair-wise yeast two-hybrid screening of protein-protein interaction within the complex gives an account of sixty-four such interactions(Rho et

al. 1996; Quevillon et al. 1999). The assembly of MRS, LRS, and RRS is mediated via appended domains (Quevillon and Mirande 1996). Interestingly, KRS, DRS, and QRS associate through their catalytic domains. This statement is confirmed by a study involving the deletion of appended extensions resulting in the non-abolishment of interactions within MSC. The literature shows a striking example of inter-dependence of binding events where the association of QRS to p38 is enhanced by the presence of p43 and RRS resulting in the formation of a discreet quaternary complex before the formation of the MARS complex (S. G. Park et al. 1999) (Fig. 1.1).

The auxiliary protein provides stability to the complex and also aids in the binding of the tRNA. p38 is a scaffolding protein that interacts with most of the components of the complex; hence, it is an indispensable core protein required for assembly and stability (D. G. Kim et al. 2016). It exists as a dimer and does not have a homolog in yeast, bacteria, or archaea. It comprises of N-terminal Leucine zipper domain and a C-terminal GST domain (D. G. Kim et al. 2016). A splice variant of p38, lacking the LZ domain, does not associate with MARS. It promotes tumorigenesis via degradation of p53 (Choi et al. 2011). Mutations in the p38 gene in mice or shRNA-mediated knockdown in HeLa cells impair the assembly of MARS. p38 also has roles beyond the MARS complex. In the nucleus, it interacts with the FUSE binding protein (transcriptional activator of c-myc) (D. G. Kim et al. 2016). p38 associates with and acts as a substrate for Parkin-mediated ubiquitination and degradation of p38. Autosomal recessive loss-of-function mutations in the gene encoding Parkin result in Parkinson's disease. Parkinson's disease is characterized by the degeneration of dopaminergic neurons. Loss of Parkin leads to the accumulation of non-ubiquitinated p38 resulting in the formation of aggresome-like inclusions. The mechanism behind the accumulation of p38 and dopaminergic neurodegeneration is unknown (Y. Lee et al., 2013).

p43 interacts with p38 via the C-terminal EMAPII domain (Shalak et al. 2001). EMAPII has a tRNA binding property and is homologous to the yeast Arc1p which facilitates the binding of cognate tRNA substrates to MRS and QRS. EMAPII is cleaved off by apoptotic cleavage (caspase-7), and it acts as a pro-inflammatory cytokine downstream, resulting in enhanced expression of pro-inflammatory cytokine genes and an increase in chemotactic migration of polymorphonuclear leukocytes (Shalak et al. 2001). It has an oligonucleotide (OB) fold like that of bacterial Trbp III tRNA binding protein. It interacts with QRS and RRS via its N-terminal domain and transfers tRNA^{Arg} to RRS within the MARS complex (Shalak et al. 2001; Bottoni

et al. 2007). The multifunctionality of p38 and p43 provides an insight into the interconnection between translation and cellular processes outside protein synthesis.

p18 has a GST-like fold similar to that of yeast Arc1p. It promotes the association of p18 to MARS and the transfer of Met-tRNA_i^{Met} to eIF2 (Kang et al. 2012). The function of p18 is unknown. Structurally, it is homologous to human VRS and forms a complex with EF-1H (Negrutskii et al. 1999). EF-1A associates with KRS and utilizes it as a means to shuttle aminoacylated tRNA directly from the MARS complex to the ribosome during translation (Guzzo and Yang 2008).

Another study shows that the MARS complex is co-purified with polysomes. But it cannot be iterated for sure whether a complex is formed between them or they are transiently associated. tRNA synthetases within the complex show enhanced aminoacylation activities owing to an increase in efficiency of early steps of translation of aminoacyl tRNA synthetase and its subsequent delivery to the ribosome. In yeast, Arc1p interacts with ERS and MRS resulting in higher protein stability and restriction of associated tRNA synthetases in the cytoplasm. The MARS complex sequesters a definite pool of tRNA synthetases to be used in protein synthesis. In permeabilized CHO cells, even on the incorporation of radiolabelled aminoacyl tRNAs, it is not used as a substrate for protein synthesis. On the other hand, radiolabelled amino acids are efficiently utilized by the MARS complex components and are incorporated during translation. This study shows that endogenously aminoacylated tRNAs synthesized within MSC are required for protein synthesis. Pools of aminoacylated tRNAs do not mix freely. RRS has two isoforms; one complexed full-length form and another truncated at N-terminus. The two isoforms are translated from alternative start codons. The complexed form is involved in the synthesis of a pool of tRNAs which are preferentially utilized as substrates for protein synthesis *in vivo* whereas the free form synthesizes aa-tRNA, utilized for arginylation of proteins in the ubiquitin-dependent degradation pathway. EPRS is also known to have two isoforms; a full-length polypeptide and a truncated isoform encoding ERS followed by two WHEP domains. There is no evidence for truncated isoform to participate in protein synthesis. The MARS complex directly channels aminoacylated tRNA to MSC to elongating ribosomes. A larger complex of aminoacyl tRNA synthetases is observed in the nucleus with a role similar to the cytoplasmic one. It has been shown to associate with nuclear pore EF-1 α leading to tRNA export. Thus, MSC is involved in protein stability and restriction of aminoacyl tRNA synthetases to the cytoplasm, in addition to coordinating events during translation. Different

approaches have helped in the structural mapping of the MARS complex to determine topological relationships between different components like a) reverse chemical crosslinking, b) extensive two hybrid searches, and c) pull-down experiments of native proteins.

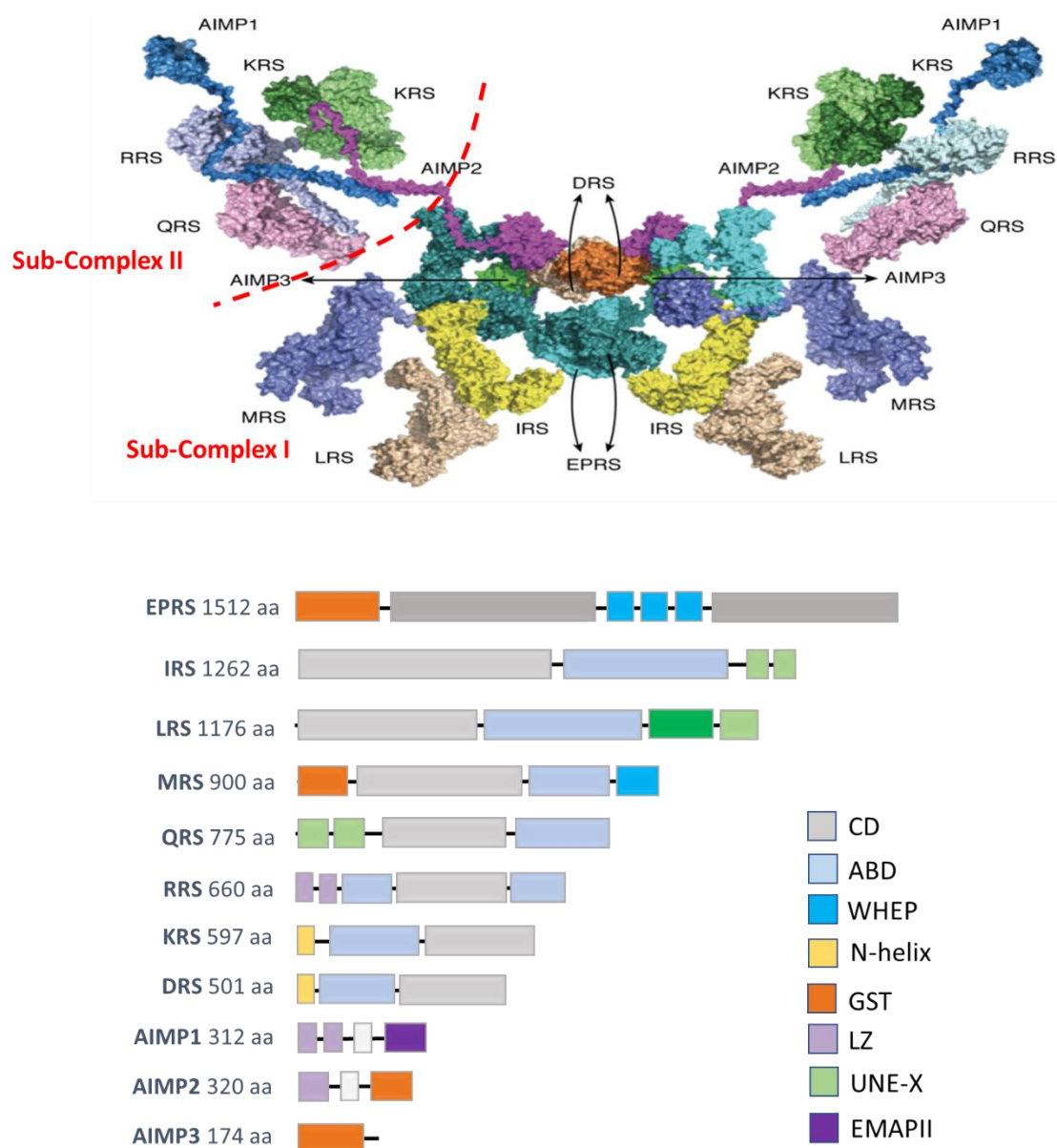


Figure 1.1: A. Structure of the MARS complex. The MARS complex consists of eight AARSs and three auxiliary AIMPs and is subdivided into Sub-complex I and Sub-complex II in the mammalian system. AARSs and AIMPs are represented by their substrate amino acids and numbers, respectively. AIMP1, RRS, and QRS are anchored N-terminal region of AIMP2, and the other components are linked to the C-terminal region of AIMP2. AIMP1 interacts with AIMP2 via a leucine zipper motif. In addition to these interactions, the complex is further stabilized by multidirectional interactions between the components. **B. Structure of the components of the MARS complex and their corresponding domains.** (CD: Catalytic domain; ABD: Anticodon Binding Domain; WHEP: WHEP-TRS domain; N-helix: N-terminal Helix; GST: Glutathione-S-transferase like domain; LZ: Leucine

Zipper domain; UNE-X: Unique followed by corresponding AARS specific domain; EMAPII: Endothelial monocyte activating polypeptide II). Adapted from Khan K, Baleanu-Gogonea C, Willard B, Gogonea V, Fox PL. The 3-Dimensional architecture of the human multi-tRNA synthetase complex. *Nucleic Acids Res.* 2020 Sep 4;48(15):8740-8754. doi: 10.1093/nar/gkaa569. PMID: 32644155; PMCID: PMC7470956.

4. MARS as a regulator of immune response

Multi-omics profiles affirm that AARSs are involved in the immune response.

4.1 EPRS as a Regulator of immune response

GAIT-dependent immune function: In higher eukaryotes (humans), EPRS is involved in a transcript selective translational inhibition of a group of functionally related genes involved in chronic inflammation (Jia et al. 2008; Arif and Fox 2017; Yao et al. 2012; Mukhopadhyay et al. 2009; Arif and Fox 2017). EPRS is the only known bifunctional tRNA synthetase comprising two synthetase domains at the termini corresponding to ERS and PRS, respectively (Ray et al. 2011). These catalytic domains are separated by a non-catalytic linker encompassing three consecutive WHEP domains via which EPRS interacts with AIMP3 and KRS in the MARS complex (Ray et al. 2011).

In cells of the myeloid lineage, EPRS is phosphorylated on Ser⁸⁸⁶ and Ser⁹⁹⁹ by two distinct kinase systems upon interferon γ (IFN γ) stimulation (Arif et al. 2009). Ser⁸⁸⁶ is phosphorylated by cyclin-dependent kinase 5 (Cdk5) in conjunction with Ras-dependent extracellular signal-regulated kinase (ERK2; a mitogen-activated protein kinase) and activator, p35, directly (Arif et al. 2011; Jia et al. 2008; Jia et al. 2008; Mukhopadhyay et al. 2009; Arif et al. 2018). On the other hand, Ser⁹⁹⁹ is phosphorylated by the mammalian target of rapamycin complex 1 (mTORC1) activated ribosomal protein, S6K1 (Arif et al. 2018; Jia et al. 2008; Mukhopadhyay et al. 2009; Arif et al. 2009). The phosphorylated EPRS undergoes a conformational change and dissociates from the MARS (Mukhopadhyay et al. 2009) complex. Upon release, it acquires new binding partners to form the interferon (IFN)- γ -activated inhibitor of translation (GAIT) complex (Arif et al. 2018; Jia et al. 2008; Mukhopadhyay et al. 2009).

Assembly of the GAIT complex occurs in two tightly regulated stages, with the recruitment of two independent and temporally distinct signalling pathways (i), Early induction of EPRS phosphorylation and (ii) delayed induction of L13a phosphorylation. 2-4 hours post-IFN γ stimulation EPRS is biphosphorylated and released from the MARS complex (Mukhopadhyay et al. 2009). Dissociated EPRS interacts with cytosolic synaptotagmin binding, cytoplasmic

RNA interacting protein (SYNCRIP/NSAP1) through the region containing phosphorylated Ser⁸⁸⁶ to form an inactive, binary pre-GAIT complex (Mukhopadhyay et al. 2009). 12-16 hours post-IFN γ stimulation, L13a, located at the surface of 60S ribosomal subunit, is phosphorylated by Death associated protein kinase (DAPK) activated Zipper-associated protein kinase ZIPK on Ser⁷⁷, leading to its release from the ribosome (Mukhopadhyay et al. 2009; Arif et al. 2009; Mukhopadhyay et al. 2008). Delayed release of phosphorylated L13a is the rate-determining step for GAIT complex formation (Mukhopadhyay et al. 2009; Arif et al. 2009). Phosphorylated L13a along with Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) interacts with the pre-GAIT complex to form an active hetero-tetrameric GAIT complex resulting in a conformational shift exposing the binding sites for GAIT element (Mukhopadhyay et al. 2009; Arif et al. 2009).

Table 1.1: Immuno-regulatory roles of MARS complex components.

AARS	Stimulus	PTM	Interactor	Function	cells
EPRS	-H1N1 Influenza virus -VSV -PolyI:C	Phosphorylation at Ser 990	PCBP2	Activation of MAVS mediated anti-viral signalling	U937 Raw264.7
EPRS	INF γ	Phosphorylation at Ser 886, Ser 999	NSAP1 GAPDH L13a	Inhibition of inflammatory gene translation by heterotetrameric GAIT complex formation	U937
EPRS	INF γ	Phosphorylation at Ser 999	GAPDH L13a	Inhibition of inflammatory gene translation by heterotrimeric GAIT complex formation	Raw264.7
KRS	IgE-Ag	Phosphorylation at Ser 207	MITF	Control of Ap4A levels in the nucleus Immune activation by MAPK signalling cascade	RBL Mast cell
AIMP1	LPS H3N2 Influenza virus	-	-	MAPK signalling activation T _H 1 polarization Adaptive immune activation	BMDC

Abbreviations: PTM, post-translational modification; VSV, vesicular stomatitis virus; GAIT, IFN- γ -activated inhibition of translation; PBMC, peripheral blood mononuclear cell; MAPK, mitogen-activated protein kinase; T_H1, T-helper type 1; BMDC, bone marrow-derived dendritic cell; IgE-Ag, immunoglobulin E-antigen; MITF, microphthalmia-associated transcription factor; Ap4A, diadenosine tetraphosphate.

The GAIT complex associates with GAIT element, a bipartite, 29-33 nucleotide stem-loop structure within the 3' untranslated region (3'UTR) of inflammation-responsive mRNAs, thereby inhibiting their translation. Recognition of GAIT element by the GAIT complex depends upon its tertiary structure and not on the primary sequence, with the exception of two relatively invariant nucleotides in the internal bulge (Vyas et al. 2009). Suppression of

translation of mRNA occurs via a process involving the circularisation of target mRNAs (Mazumder et al. 2001; Wells et al. 1998). EPRS is the lone component of the GAIT complex which interacts with GAIT elements of target mRNAs (Mukhopadhyay et al. 2009). Phosphorylated L13a associates with translation initiation factor eIF4G at or near the eIF3 binding site preventing its interaction with eIF3 containing 43S pre-initiation complex (Kapasi et al. 2007). The GAIT complex is also known to inhibit the translation of the kinases, namely *DAPK* and *ZIPK*, thus forming an auto-regulated feedback loop (Vyas et al. 2009; Mukhopadhyay et al. 2008; Singer 2008) (Fig. 1.2; Table 1.1).

Yao and co-workers have discovered a truncated isoform of EPRS christened EPRS^{N1} (generated by poly-adenylation pathway) in which UAU, coding for Tyr⁸⁶⁴ is converted to UAA, a stop codon (Yao et al. 2012). *EPRS^{N1}* encodes ERS domain at the N-terminus, followed by two WHEP domains (Yao et al. 2012). The region of full-length EPRS (EPRS-FL) that can be phosphorylated for the GAIT complex assembly is absent in EPRS^{N1} even though it can still bind to the GAIT element (Yao et al. 2012). EPRS^{N1} competes with its full-length isoform for the GAIT element binding, thereby maintaining basal expression of pro-inflammatory mRNAs (Yao et al. 2012).

Interestingly, total protein synthesis remains unaltered during this process (Arif et al., 2018). Inflammatory responses differ in humans and mice (Arif et al. 2012). The murine GAIT complex is hetero-trimeric and lacks SYNCRIP/NSAP1 due to the lack of conservation of Ser⁸⁸⁶ for binding (Arif et al. 2012).

The GAIT complex constitutes proteins with canonical functions distinct from their function in the GAIT system. EPRS and L13a are the ubiquitous components of the translational machinery, but in the GAIT system, these are part of a complex where these are involved in translational silencing by interacting with the GAIT element and eIF4G, respectively. GAPDH canonically takes part in glycolysis (conversion of glyceraldehyde-3-phosphate to bi-phosphoglycerate) (J.-W. Kim and Dang 2005; Seidler 2012) and energy production but in the GAIT system, it has a chaperone-like activity protecting L13a from proteasomal degradation. SYNCRIP/NSAP1, a member of the hnRNP family of RBPs, is known to be involved in the packaging of nuclear transcripts, RNA editing, splicing, stabilization, transport, mRNA decay, transcriptional control, and internal ribosome entry site (IRES) dependent translational control (Geuens, Bouhy, and Timmerman 2016; Bannai et al. 2004; Svitkin et al. 2013; S. M. Park et al. 2011; Blanc et al. 2001; Mourelatos et al. 2001) but in the GAIT system, it masks

the GAIT element binding sites of EPRS. Both MARS and ribosomal complexes act as a depot of stimulus-dependent releasable regulatory proteins, which perform auxiliary functions unrelated to their primary function within the parental complex (Fig. 1.2).

A 32-nucleotide RNA motif exhibiting structural similarity to the GAIT element of human transcripts is reported in the genome of the transmissible gastroenteritis coronavirus. The RNA motif selectively interacts with EPRS in a similar manner as seen in human cells. In-vitro studies show the interaction of EPRS with viral RNA leads to interference with the host defence system. It blocks the accessibility of EPRS to host GAIT elements, in turn affecting MDA5-mediated antiviral signalling.

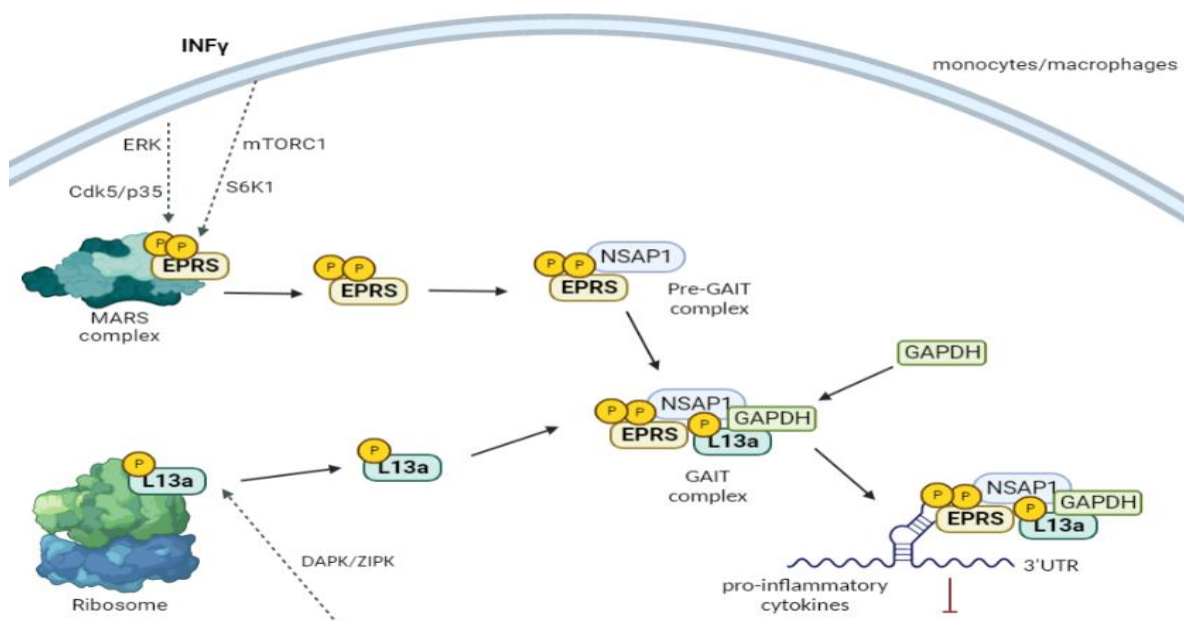


Figure 1.2: Schematic diagram of translational silencing of a subset of functionally related genes by Interferon γ activated inhibitor of translation (GAIT) complex in human myeloid cells. Assembly and activation of the GAIT complex occur in two tightly regulated stages. In the first stage, IFN γ induces two-step phosphorylation of EPRS at Ser⁸⁸⁶ and Ser⁹⁹⁹, in the linker region, by two kinase systems Cdk5/p35 and mTORC1/S6K1 respectively, 2-4 hours post IFN γ stimulation. Phosphorylated EPRS undergoes a conformational change and dissociates from the MARS complex. Unbound phosphorylated EPRS interacts with NSAP1 in the cytosol via phosphorylated Ser 886 and forms an inactive pre-GAIT complex. 12-16 hours later, L13a is phosphorylated by DAPK-activated ZIPK and is released from the 60S subunit of the ribosome. Free phosphorylated L13a interacts with GAPDH in the cytosol and subsequently associates with the pre-GAIT complex to form a functional GAIT complex. The complex binds to GAIT elements located in 3'UTR of transcripts of genes activated during inflammation via EPRS and inhibits their translation by a process involving the circularization of target mRNAs. L13a interferes with the binding of translation initiation factor eIF4G near the eIF3 binding site to block translation. The GAIT complex also represses the translation of DAPK and ZIPK, thereby comprising a feedback loop to control L13a phosphorylation and, in turn, the formation and assembly of the GAIT complex.

GAIT-independent immune function

The MARS complex is hypothesized to act as a surveillance system for infection. Upon RNA viral infection, the MARS complex senses invasion and takes part in antiviral signalling. Downregulation of EPRS in macrophages results in enhanced replication of RNA viruses like influenza A virus and stomatitis virus and reduced production of antiviral cytokines. Moreover, ectopic expression of EPRS attenuates viral replication and enhances the production of antiviral interferon β (IFN β). RNA viruses also show enhanced lethality in *EPRS*^{+/-} as compared to *EPRS*^{+/+} by virtue of the increase in rates of viral replication and weak antiviral cytokine response. Lee and co-workers have shown that upon viral infection, EPRS is phosphorylated on Ser⁹⁹⁰. Phosphorylation of EPRS leads to a conformational change and its subsequent dissociation from the MARS complex (E.-Y. Lee et al. 2016). Free EPRS interacts with poly(rC) binding protein 2 (PCBP2), a negative regulator of mitochondrial antiviral signalling protein (MAVS) in the cytosol. This interaction protects MAVS from Itchy E3 ubiquitin-protein ligase (ITCH) mediated ubiquitination and degradation, stabilizing MAVS to mount antiviral immunity (You et al. 2009). Interestingly, immunofluorescence data shows that the MARS complex components EPRS, MRS, and KRS and a few freestanding AARSs, SRS, and YRS are recruited to viral factories in the infected cell (Fig. 1.3; Table 1.1).

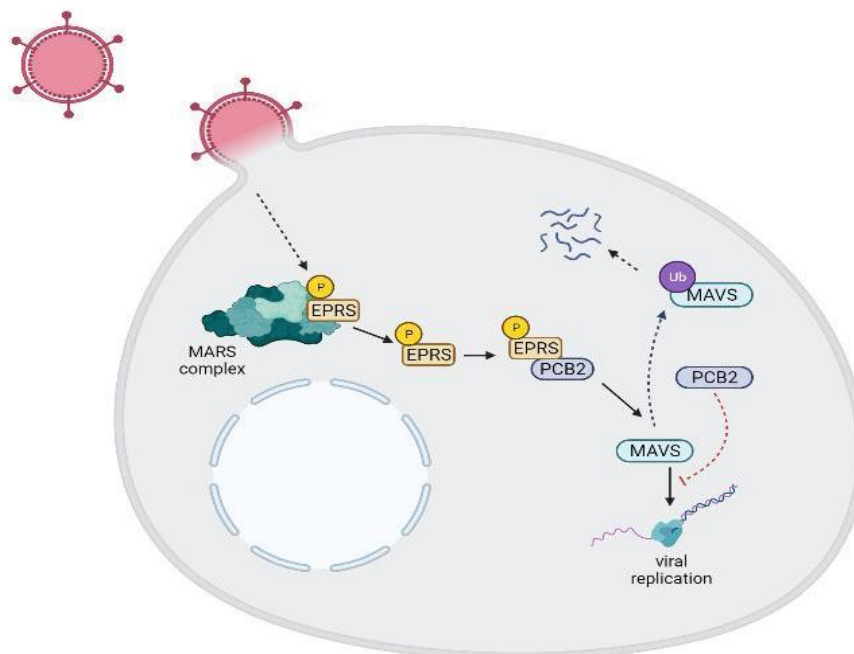


Figure 1.3: *GAIT* complex independent function. Viral infection induces phosphorylation of EPRS at Ser 990 in the linker region. Phosphorylated EPRS undergoes a conformational change and dissociates from the MARS complex. Free phosphorylated EPRS interacts with PCBP2 and sequesters it from the

cytosol and subsequently prevents ubiquitination-mediated degradation of MAVS. MAVS inhibits viral replication.

KRS as a regulator of immune response

Allergen; immunoglobulin E (IgE) acts as an antigen for mast cells. IgE recognition by Immunoglobulin E high-affinity receptors (IgE-FcεRI) leads to its aggregation at the cell surface resulting in the activation of mast cells. This congregation at the cell surface leads to phosphorylation of KRS at Ser²⁰⁷ (within the anticodon binding domain (ABD)) by MAPK-dependent kinases: MAPK kinase kinase (MEK)/ extracellular signal-regulated kinase (ERK). Phosphorylation of KRS results in change and dissociation from its binding partner AIMP2 within the MARS complex (Y.-N. Lee et al. 2004; Nechushtan and Razin 2002; Razin et al. 1999). Phosphorylation diminishes the aminoacylation activity of KRS. Free KRS in the cytosol translocates into the nucleus. Phosphorylated KRS interacts with microphthalmia-associated transcription factor (MITF) and Histidine triad nucleotide-binding protein (HINT), forming a multi-protein complex in the nucleus (Krause et al. 1996; Y.-N. Lee et al. 2004; Nechushtan and Razin 2002; Razin et al. 1999). Wherein phosphorylated KRS controls the levels of diadenosine tetra-phosphate (Ap4A) (Y.-N. Lee et al. 2004). Ap4A is a signalling molecule (popularly known as alarmone) that acts as a secondary messenger. Synthesis of Ap4A is associated with a subset of tRNA synthetases, but it is most robust with KRS. 70-80% cellular Ap4A is synthesized by phosphorylated KRS in the vicinity of the multi-protein complex inside the nucleus. Ap4A sequesters HINT from the multi-protein complex, thus unmasking the DNA binding sites on MITF (Razin et al. 1999; Razin et al. 1999; Y.-N. Lee et al. 2004). Phosphorylated KRS interacts with MITF via the C-terminal domain and guides it to carry out the transcription of a subset of genes (i.e., Tryptophan hydroxylase and Mast cell protease 5) (Carmi-Levy et al. 2008). In quiescent cells, KRS resides within the MARS complex and is involved in translation. Thus, a PTM-like phosphorylation enables host cell KRS to switch its function from translation to transcription. (Y.-N. Lee et al. 2004) (Fig. 1.4; Table 1.1).

Starvation or TNFα stimulation induces secretion of KRS from human cells. Secreted KRS activates macrophages, subsequently inducing TNFα secretion and cell migration. KRS has acquired appended domains at N and C-termini containing a caspase 8 specific cleavage site and a PDZ binding domain, respectively, during evolution. KRS is a class II tRNA synthetase, existing as a homodimer within the MARS complex, wherein the N-terminal extension of one masks the C-terminal PDZ binding domain of the other. Starvation activates the enzyme

caspase 8. Caspase 8 cleaves the N-terminal extension of KRS, rendering it a monomer. Monomeric KRS interacts with Syntenin through the PDZ binding domain. Syntenin guides KRS to the exosomes, thereby incorporating it into the lumen prior to secretion from the cell.

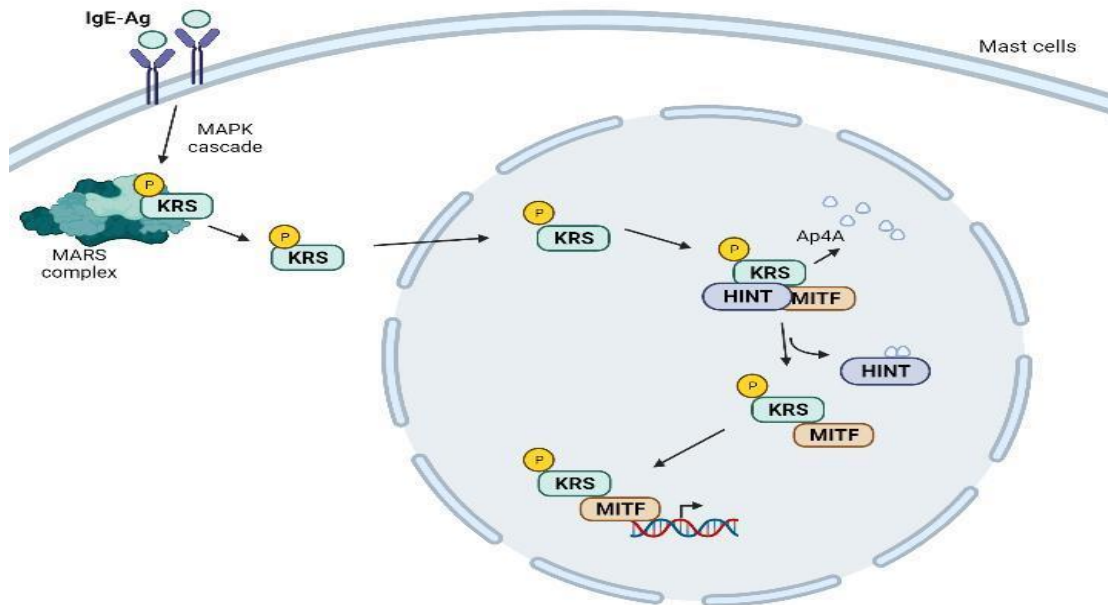


Figure 1.4: IgE-Ag binds to surface FcεRI and activates the mast cells. Activation of mast cells induces phosphorylation of KRS at Ser 207 by MAPK cascade and its subsequent release from the MARS complex. Free KRS partially translocate into the nucleus and interacts with proteins, MITF, and HINT to form a multi-protein complex. Phosphorylated KRS also synthesizes Ap4A in the vicinity of the multi-protein complex. Ap4A sequesters HINT from the complex, unmasking the DNA binding sites on MITF. Along with phosphorylated KRS, MITF engages in MITF-dependent gene expression.

Human Immunodeficiency Virus 1 (HIV1) exploits the dynamic nature of the MARS complex and redirects its constituents from translation to co-opting for viral replication. The key step for viral replication is the reverse transcription of genomic RNA. Viruses use host tRNA isoacceptors as primers for reverse transcription. In HIV1 infection, tRNA^{Lys,3} is used as a primer for catalysing reverse transcription for (-) strand DNA synthesis. Studies show that human KRS is selectively packaged into the HIV virion along with the Gag protein, coding for structural proteins of the virus. Gag protein encodes Gagpol polyprotein and Gagpol precursor, fundamental building blocks of retroviral particles. tRNA^{Lys}, along with truncated KRS, is also encapsidated into the virion. In a separate study, upon HIV infection, KRS is phosphorylated on Ser²⁰⁷, resulting in its eviction from the MARS complex. Phosphorylated KRS loses its ability to charge tRNA even though it retains its tRNA binding ability. Free phosphorylated KRS in the cytosol partially translocates into the nucleus upon HIV infection; nonetheless, its relevance in viral infectivity and its function inside the nucleus is not known. It can be hypothesized that

phosphorylated KRS is probably required for virion packaging and progeny virus infectivity. Additionally, Duchon and co-workers have detected low levels of MARS complex-dissociated EPRS upon HIV infection. WRS is encapsided and used for viral replication upon Rous sarcoma virus infection in host cells. In another instance, Moloney murine leukaemia virus uses PRS as a primer for reverse transcription, but the process does not require PRS to be encapsided (Fig. 1.5; Table 1.1).

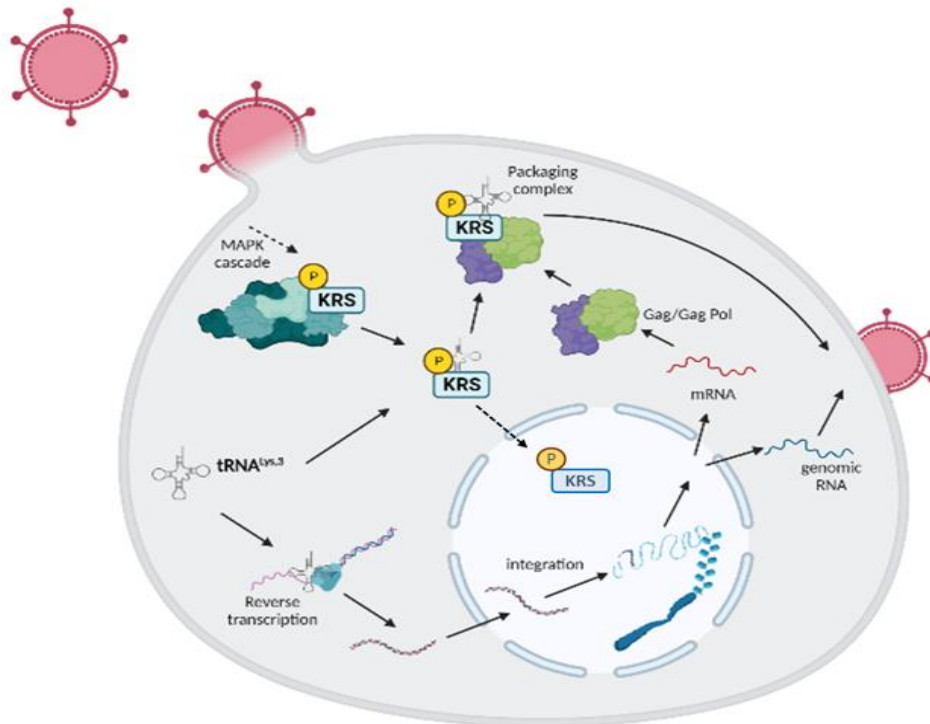


Figure 1.5: Schematic diagram of the utilization of a dynamic MARS complex by HIV1 in human host cells to enhance its replication. During HIV1 assembly, $tRNA^{Lys,3}$ acts as a primer for reverse transcription. $tRNA^{Lys,3}$ is packaged into virion by its interaction with hKRS and viral proteins, Gag polyprotein, and Gag pol precursor. HIV1 infection induces phosphorylation of hKRS at Ser 207 and its subsequent dissociation from the MARS complex. Free KRS partially translocate into the nucleus, and the remaining takes part in HIV1 virion assembly in the cytosol.

Laminin located in the extracellular matrix interacts with membranous integrins of surrounding organs. In cancer cells, the laminin-integrin complex prompts phosphorylation of KRS at Tyr⁵² by the p38-MAPK cascade, subsequently leading to its release from the MARS complex. Free cytosolic phosphorylated KRS translocates to the plasma membrane, where it associates with 67 Laminin receptor (67LR) via the N-terminal anticodon binding domain (ABD). The interaction inhibits Nedd4-mediated proteasomal degradation of 67LR. Hence the ligand-receptor interaction at the cell surface induces a cascade of reactions pertaining to a moonlighting function of KRS ensuing laminin-dependent cell migration and cell dissemination (hallmarks of metastasis).

References

- Ahn, Young Ha, Sunyoung Park, Jeong June Choi, Bo-Kyung Park, Kyung Hee Rhee, Eunjoo Kang, Soyeon Ahn, et al. 2016. "Secreted Tryptophanyl-tRNA Synthetase as a Primary Defence System against Infection." *Nature Microbiology* 2 (October): 16191.
- Arif, Abul, Piyali Chatterjee, Robyn A. Moodt, and Paul L. Fox. 2012. "Heterotrimeric GAIT Complex Drives Transcript-Selective Translation Inhibition in Murine Macrophages." *Molecular and Cellular Biology* 32 (24): 5046–55.
- Arif, Abul, and Paul L. Fox. 2017. "Unexpected Metabolic Function of a tRNA Synthetase." *Cell Cycle* .
- Arif, Abul, Jie Jia, Robyn A. Moodt, Paul E. DiCorleto, and Paul L. Fox. 2011. "Phosphorylation of Glutamyl-Prolyl tRNA Synthetase by Cyclin-Dependent Kinase 5 Dictates Transcript-Selective Translational Control." *Proceedings of the National Academy of Sciences of the United States of America* 108 (4): 1415–20.
- Arif, Abul, Jie Jia, Rupak Mukhopadhyay, Belinda Willard, Michael Kinter, and Paul L. Fox. 2009. "Two-Site Phosphorylation of EPRS Coordinates Multimodal Regulation of Noncanonical Translational Control Activity." *Molecular Cell* 35 (2): 164–80.
- Arif, Abul, Fulvia Terenzi, Alka A. Potdar, Jie Jia, Jessica Sacks, Arnab China, Dalia Halawani, et al. 2017. "EPRS Is a Critical mTORC1-S6K1 Effector That Influences Adiposity in Mice." *Nature* 542 (7641): 357–61.
- Arif, Abul, Peng Yao, Fulvia Terenzi, Jie Jia, Partho Sarothi Ray, and Paul L. Fox. 2018. "The GAIT Translational Control System." *Wiley Interdisciplinary Reviews. RNA* 9 (2). <https://doi.org/10.1002/wrna.1441>.
- Arnez, J. G., D. C. Harris, A. Mitschler, B. Rees, C. S. Francklyn, and D. Moras. 1995. "Crystal Structure of Histidyl-tRNA Synthetase from Escherichia Coli Complexed with Histidyl-Adenylate." *The EMBO Journal* 14 (17): 4143–55.
- Arnez, Z. M., L. Valdatta, M. P. Tyler, and F. Planinsek. 1995. "Anatomy of the Internal Mammary Veins and Their Use in Free TRAM Flap Breast Reconstruction." *British Journal of Plastic Surgery* 48 (8): 540–45.
- Bannai, Hiroko, Kazumi Fukatsu, Akihiro Mizutani, Tohru Natsume, Shun-Ichiro Iemura, Tohru Ikegami, Takafumi Inoue, and Katsuhiko Mikoshiba. 2004. "An RNA-Interacting Protein, SYNCRIP (heterogeneous Nuclear Ribonuclear Protein Q1/NSAP1) Is a Component of mRNA Granule Transported with Inositol 1,4,5-Trisphosphate Receptor Type 1 mRNA in Neuronal Dendrites." *The Journal of Biological Chemistry* 279 (51): 53427–34.
- Beebe, Kirk, Lluís Ribas De Pouplana, and Paul Schimmel. 2003. "Elucidation of tRNA-Dependent Editing by a Class II tRNA Synthetase and Significance for Cell Viability." *The EMBO Journal* 22 (3): 668–75.
- Beuning, P. J., and K. Musier-Forsyth. 2001. "Species-Specific Differences in Amino Acid Editing by Class II Prolyl-tRNA Synthetase." *The Journal of Biological Chemistry* 276 (33): 30779–85.
- Blanc, Valerie, Naveenan Navaratnam, Jeffrey O. Henderson, Shrikant Anant, Susan Kennedy, Adam Jarmuz, James Scott, and Nicholas O. Davidson. 2001. "Identification of GRY-RBP as an Apolipoprotein B RNA-Binding Protein That Interacts with Both Apobec-1 and Apobec-1 Complementation Factor to Modulate C to U Editing." *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.m006435200>.
- Bonfils, Grégory, Malika Jaquenoud, Séverine Bontron, Clemens Ostrowicz, Christian Ungermann, and Claudio De Virgilio. 2012. "Leucyl-tRNA Synthetase Controls TORC1 via the EGO Complex." *Molecular Cell* 46 (1): 105–10.
- Bottoni, Arianna, Cristina Vignali, Daniela Piccin, Federico Tagliati, Andrea Luchin, Maria Chiara Zatelli, and Ettore C. Degli Uberti. 2007. "Proteasomes and RARS Modulate AIMP1/EMAP II Secretion in Human Cancer Cell Lines." *Journal of Cellular Physiology* 212 (2): 293–97.
- Brevet, A., P. Plateau, B. Cirakoğlu, J. P. Pailliez, and S. Blanquet. 1982. "Zinc-Dependent Synthesis of 5',5'-Diadenosine Tetraphosphate by Sheep Liver Lysyl- and Phenylalanyl-tRNA Synthetases." *The Journal of Biological Chemistry* 257 (24): 14613–15.

- Brick, P., T. N. Bhat, and D. M. Blow. 1989. "Structure of Tyrosyl-tRNA Synthetase Refined at 2.3 Å Resolution. Interaction of the Enzyme with the Tyrosyl Adenylate Intermediate." *Journal of Molecular Biology* 208 (1): 83–98.
- Brick, P., and D. M. Blow. 1987. "Crystal Structure of a Deletion Mutant of a Tyrosyl-tRNA Synthetase Complexed with Tyrosine." *Journal of Molecular Biology* 194 (2): 287–97.
- Buchon, Nicolas, Neal Silverman, and Sara Cherry. 2014. "Immunity in *Drosophila Melanogaster*—from Microbial Recognition to Whole-Organism Physiology." *Nature Reviews. Immunology* 14 (12): 796–810.
- Burbaum, J. J., and P. Schimmel. 1991. "Structural Relationships and the Classification of Aminoacyl-tRNA Synthetases." *The Journal of Biological Chemistry* 266 (26): 16965–68.
- Carmi-Levy, Irit, Nurit Yannay-Cohen, Gillian Kay, Ehud Razin, and Hovav Nechushtan. 2008. "Diadenosine Tetraphosphate Hydrolase Is Part of the Transcriptional Regulation Network in Immunologically Activated Mast Cells." *Molecular and Cellular Biology*. <https://doi.org/10.1128/mcb.00106-08>.
- Cavagna, Lorenzo, on Behalf of AENEAS (American and European Network of Antisynthetase Syndrome) Collaborative Group, Laura Nuño, Carlo Alberto Scirè, Marcello Govoni, Francisco Javier Lopez Longo, Franco Franceschini, et al. 2017. "Serum Jo-1 Autoantibody and Isolated Arthritis in the Antisynthetase Syndrome: Review of the Literature and Report of the Experience of AENEAS Collaborative Group." *Clinical Reviews in Allergy & Immunology*. <https://doi.org/10.1007/s12016-016-8528-9>.
- Chi, Lang-Ming, Chien-Wei Lee, Kai-Ping Chang, Sheng-Po Hao, Hang-Mao Lee, Ying Liang, Chuen Hsueh, et al. 2009. "Enhanced Interferon Signaling Pathway in Oral Cancer Revealed by Quantitative Proteome Analysis of Microdissected Specimens Using 16O/18O Labeling and Integrated Two-Dimensional LC-ESI-MALDI Tandem MS." *Molecular & Cellular Proteomics: MCP* 8 (7): 1453–74.
- Cho, Ha Yeon, Seo Jin Maeng, Hyo Je Cho, Yoon Seo Choi, Jeong Min Chung, Sangmin Lee, Hoi Kyoung Kim, et al. 2015. "Assembly of Multi-tRNA Synthetase Complex via Heterotetrameric Glutathione Transferase-Homology Domains." *The Journal of Biological Chemistry* 290 (49): 29313–28.
- Choi, Jin Woo, Dae Gyu Kim, Al-Eum Lee, Hye Rim Kim, Jin Young Lee, Nam Hoon Kwon, Young Kee Shin, et al. 2011. "Cancer-Associated Splicing Variant of Tumor Suppressor AIMP2/p38: Pathological Implication in Tumorigenesis." *PLoS Genetics* 7 (3): e1001351.
- Choi, Jin Woo, Dae Gyu Kim, Min Chul Park, Jung Yeon Um, Jung Min Han, Sang Gyu Park, Eung-Chil Choi, and Sunghoon Kim. 2009. "AIMP2 Promotes TNFalpha-Dependent Apoptosis via Ubiquitin-Mediated Degradation of TRAF2." *Journal of Cell Science* 122 (Pt 15): 2710–15.
- Choi, Jin Woo, Jeong-Won Lee, Jun Ki Kim, Hye-Kyung Jeon, Jung-Joo Choi, Dae Gyu Kim, Byoung-Gie Kim, et al. 2012. "Splicing Variant of AIMP2 as an Effective Target against Chemoresistant Ovarian Cancer." *Journal of Molecular Cell Biology* 4 (3): 164–73.
- Choi, Jin Woo, Jung Yeon Um, Joydeb Kumar Kundu, Young-Joon Surh, and Sunghoon Kim. 2009. "Multidirectional Tumor-Suppressive Activity of AIMP2/p38 and the Enhanced Susceptibility of AIMP2 Heterozygous Mice to Carcinogenesis." *Carcinogenesis* 30 (9): 1638–44.
- Cirakoglu, B., and J. P. Waller. 1985. "Leucyl-tRNA and Lysyl-tRNA Synthetases, Derived from the High-Mr Complex of Sheep Liver, Are Hydrophobic Proteins." *European Journal of Biochemistry / FEBS* 151 (1): 101–10.
- Cusack, S. 1997. "Aminoacyl-tRNA Synthetases." *Current Opinion in Structural Biology* 7 (6): 881–89.
- Cusack, S., C. Berthet-Colominas, M. Härtlein, N. Nassar, and R. Leberman. 1990. "A Second Class of Synthetase Structure Revealed by X-Ray Analysis of *Escherichia Coli* Seryl-tRNA Synthetase at 2.5 Å." *Nature* 347 (6290): 249–55.
- Datt, Manish, and Amit Sharma. 2014. "Evolutionary and Structural Annotation of Disease-Associated Mutations in Human Aminoacyl-tRNA Synthetases." *BMC Genomics* 15 (1): 1063.
- David, Karen Kate, Shaida Ahmad Andrabi, Ted Murray Dawson, and Valina Lynn Dawson. 2009. "Parthanatos, a Messenger of Death." *Frontiers in Bioscience* 14 (3): 1116–28.
- Dock-Bregeon, A., R. Sankaranarayanan, P. Romby, J. Caillet, M. Springer, B. Rees, C. S. Francklyn, C. Ehresmann, and D. Moras. 2000. "Transfer RNA-Mediated Editing in Threonyl-tRNA Synthetase. The Class II Solution to the Double Discrimination Problem." *Cell* 103 (6): 877–84.

- Eriani, G., J. Cavarelli, F. Martin, L. Ador, B. Rees, J. C. Thierry, J. Gangloff, and D. Moras. 1995. "The Class II Aminoacyl-tRNA Synthetases and Their Active Site: Evolutionary Conservation of an ATP Binding Site." *Journal of Molecular Evolution* 40 (5): 499–508.
- Eswarappa, Sandeepa M., and Paul L. Fox. 2013. "Citric Acid Cycle and the Origin of MARS." *Trends in Biochemical Sciences* 38 (5): 222–28.
- Fersht, A. R. 1977. "Editing Mechanisms in Protein Synthesis. Rejection of Valine by the Isoleucyl-tRNA Synthetase." *Biochemistry* 16 (5): 1025–30.
- Forus, Anne, Vivi Ann Flørenes, Gunhild Mari Maeldandsmo, Øystein Fodstad, and Ola Myklebost. 1994. "The Protooncogene CHOP/GADD153, Involved in Growth Arrest and DNA Damage Response, Is Amplified in a Subset of Human Sarcomas." *Cancer Genetics and Cytogenetics*. [https://doi.org/10.1016/0165-4608\(94\)90085-x](https://doi.org/10.1016/0165-4608(94)90085-x).
- Fournier, Gregory P., Cheryl P. Andam, Eric J. Alm, and J. Peter Gogarten. 2011. "Molecular Evolution of Aminoacyl tRNA Synthetase Proteins in the Early History of Life." *Origins of Life and Evolution of the Biosphere: The Journal of the International Society for the Study of the Origin of Life* 41 (6): 621–32.
- Fu, Yaoyao, Youngran Kim, Kyeong Sik Jin, Hyun Sook Kim, Jong Hyun Kim, Dongming Wang, Minyoung Park, et al. 2014. "Structure of the ArgRS-GlnRS-AIMP1 Complex and Its Implications for Mammalian Translation." *Proceedings of the National Academy of Sciences of the United States of America* 111 (42): 15084–89.
- Garbern, J. Y. 2007. "Pelizaeus-Merzbacher Disease: Genetic and Cellular Pathogenesis." *Cellular and Molecular Life Sciences: CMLS* 64 (1): 50–65.
- Geuens, Thomas, Delphine Bouhy, and Vincent Timmerman. 2016. "The hnRNP Family: Insights into Their Role in Health and Disease." *Human Genetics* 135 (8): 851–67.
- Giegé, R., M. Sissler, and C. Florentz. 1998. "Universal Rules and Idiosyncratic Features in tRNA Identity." *Nucleic Acids Research* 26 (22): 5017–35.
- Guo, Min, and Xiang-Lei Yang. 2014. "Architecture and Metamorphosis." *Topics in Current Chemistry* 344: 89–118.
- Guo, Min, Xiang-Lei Yang, and Paul Schimmel. 2010. "New Functions of Aminoacyl-tRNA Synthetases beyond Translation." *Nature Reviews. Molecular Cell Biology* 11 (9): 668–74.
- Guzzo, Catherine M., and David C. H. Yang. 2008. "Lysyl-tRNA Synthetase Interacts with EF1alpha, Aspartyl-tRNA Synthetase and p38 in Vitro." *Biochemical and Biophysical Research Communications* 365 (4): 718–23.
- Han, Jung Min, Seung Jae Jeong, Min Chul Park, Gyuyoup Kim, Nam Hoon Kwon, Hoi Kyoung Kim, Sang Hoon Ha, Sung Ho Ryu, and Sunghoon Kim. 2012. "Leucyl-tRNA Synthetase Is an Intracellular Leucine Sensor for the mTORC1-Signaling Pathway." *Cell* 149 (2): 410–24.
- Ibba, Michael. 2005. *The Aminoacyl-tRNA Synthetases*. CRC Press.
- Ibba, M., C. Stathopoulos, and D. Söll. 2001. "Protein Synthesis: Twenty Three Amino Acids and Counting." *Current Biology: CB* 11 (14): R563–65.
- Iqbal, Zafar, Lucia Püttmann, Luciana Musante, Attia Razzaq, Muhammad Yasir Zahoor, Hao Hu, Thomas F. Wienker, et al. 2016. "Missense Variants in AIMP1 Gene Are Implicated in Autosomal Recessive Intellectual Disability without Neurodegeneration." *European Journal of Human Genetics: EJHG* 24 (3): 392–99.
- Jeong, Seung Jae, Jong Hyun Kim, Beom Jin Lim, Ina Yoon, Ji-Ae Song, Hee-Sun Moon, Doyeun Kim, Dong Ki Lee, and Sunghoon Kim. 2018. "Inhibition of MUC1 Biosynthesis via Threonyl-tRNA Synthetase Suppresses Pancreatic Cancer Cell Migration." *Experimental & Molecular Medicine* 50 (1): e424.
- Jia, Jie, Abul Arif, Partho S. Ray, and Paul L. Fox. 2008. "WHEP Domains Direct Noncanonical Function of Glutamyl-Prolyl tRNA Synthetase in Translational Control of Gene Expression." *Molecular Cell* 29 (6): 679–90.
- Kaminska, M., V. Shalak, and M. Mirande. 2001. "The Appended C-Domain of Human Methionyl-tRNA Synthetase Has a tRNA-Sequestering Function." *Biochemistry* 40 (47): 14309–16.
- Kanaji, Taisuke, My-Nuong Vo, Sachiko Kanaji, Alessandro Zarpellon, Ryan Shapiro, Yosuke Morodomi, Akinori Yuzuriha, et al. 2018. "Tyrosyl-tRNA Synthetase Stimulates Thrombopoietin-Independent Hematopoiesis Accelerating Recovery from Thrombocytopenia." *Proceedings of the National Academy of Sciences of the United States of America* 115 (35): E8228–35.

- Kang, Taehee, Nam Hoon Kwon, Jin Young Lee, Min Chul Park, Eunji Kang, Hyo Hyun Kim, Taek Jin Kang, and Sunghoon Kim. 2012. "AIMP3/p18 Controls Translational Initiation by Mediating the Delivery of Charged Initiator tRNA to Initiation Complex." *Journal of Molecular Biology* 423 (4): 475–81.
- Kapasi, Purvi, Sujan Chaudhuri, Keyur Vyas, Diane Baus, Anton A. Komar, Paul L. Fox, William C. Merrick, and Barsanjit Mazumder. 2007. "L13a Blocks 48S Assembly: Role of a General Initiation Factor in mRNA-Specific Translational Control." *Molecular Cell* 25 (1): 113–26.
- Kellermann, K., L. Heuser, and T. Mertens. 1982. "[Encephalitis of the basel ganglia associated with mumps]." *Monatsschrift Kinderheilkunde: Organ der Deutschen Gesellschaft für Kinderheilkunde* 130 (8): 624–27.
- Kerjan, P., C. Cerini, M. Sémériva, and M. Mirande. 1994. "The Multienzyme Complex Containing Nine Aminoacyl-tRNA Synthetases Is Ubiquitous from Drosophila to Mammals." *Biochimica et Biophysica Acta* 1199 (3): 293–97.
- Kerjan, P., M. Triconnet, and J. P. Waller. 1992. "Mammalian Prolyl-tRNA Synthetase Corresponds to the Approximately 150 kDa Subunit of the High-M(r) Aminoacyl-tRNA Synthetase Complex." *Biochimie* 74 (2): 195–205.
- Kim, Dae Gyu, Jin Woo Choi, Jin Young Lee, Hyerim Kim, Young Sun Oh, Jung Weon Lee, Yu Kyung Tak, et al. 2012. "Interaction of Two Translational Components, Lysyl-tRNA Synthetase and p40/37LRP, in Plasma Membrane Promotes Laminin-Dependent Cell Migration." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 26 (10): 4142–59.
- Kim, Dae Gyu, Jin Young Lee, Nam Hoon Kwon, Pengfei Fang, Qian Zhang, Jing Wang, Nicolas L. Young, et al. 2014. "Chemical Inhibition of Prometastatic Lysyl-tRNA Synthetase-Laminin Receptor Interaction." *Nature Chemical Biology* 10 (1): 29–34.
- Kim, Dae Gyu, Jin Young Lee, Ji-Hyun Lee, Ha Yeon Cho, Beom Sik Kang, Song-Yee Jang, Myung Hee Kim, et al. 2016. "Oncogenic Mutation of AIMP2/p38 Inhibits Its Tumor-Suppressive Interaction with Smurf2." *Cancer Research* 76 (11): 3422–36.
- Kim, Doyeun, Nam Hoon Kwon, and Sunghoon Kim. 2014. "Association of Aminoacyl-tRNA Synthetases with Cancer." *Topics in Current Chemistry* 344: 207–45.
- Kim, Eun Young, Ji Ye Jung, Arum Kim, Kwangsoo Kim, and Yoon Soo Chang. 2017. "Methionyl-tRNA Synthetase Overexpression Is Associated with Poor Clinical Outcomes in Non-Small Cell Lung Cancer." *BMC Cancer* 17 (1): 467.
- Kim, J. E., K. H. Kim, S. W. Lee, W. Seol, K. Shiba, and S. Kim. 2000. "An Elongation Factor-Associating Domain Is Inserted into Human CysteinyI-tRNA Synthetase by Alternative Splicing." *Nucleic Acids Research* 28 (15): 2866–72.
- Kim, Jin Young, Young-Sun Kang, Joong-Won Lee, Hyoung June Kim, Young Ha Ahn, Heonyong Park, Young-Gyu Ko, and Sunghoon Kim. 2002. "p38 Is Essential for the Assembly and Stability of Macromolecular tRNA Synthetase Complex: Implications for Its Physiological Significance." *Proceedings of the National Academy of Sciences of the United States of America* 99 (12): 7912–16.
- Kim, Jung-Whan, and Chi V. Dang. 2005. "Multifaceted Roles of Glycolytic Enzymes." *Trends in Biochemical Sciences* 30 (3): 142–50.
- Kim, Sang Bum, Hye Rim Kim, Min Chul Park, Seongmin Cho, Peter C. Goughnour, Daeyoung Han, Ina Yoon, et al. 2017. "Caspase-8 Controls the Secretion of Inflammatory Lysyl-tRNA Synthetase in Exosomes from Cancer Cells." *The Journal of Cell Biology* 216 (7): 2201–16.
- Kim, Sunghoon. 2014. *Aminoacyl-tRNA Synthetases in Biology and Medicine*. Springer.
- Kim, Sunghoon, Sungyong You, and Daehee Hwang. 2011. "Aminoacyl-tRNA Synthetases and Tumorigenesis: More than Housekeeping." *Nature Reviews. Cancer* 11 (10): 708–18.
- Ko, Han Seok, Rainer von Coelln, Sathya R. Sriram, Seong Who Kim, Kenny K. Chung, Olga Pletnikova, Juan Troncoso, et al. 2005. "Accumulation of the Authentic Parkin Substrate Aminoacyl-tRNA Synthetase Cofactor, p38/JTV-1, Leads to Catecholaminergic Cell Death." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25 (35): 7968–78.
- Kopajtich, Robert, Kei Murayama, Andreas R. Janecke, Tobias B. Haack, Maximilian Breuer, A. S. Knisely, Inga Harting, et al. 2016. "Biallelic IARS Mutations Cause Growth Retardation with Prenatal Onset, Intellectual Disability, Muscular Hypotonia, and Infantile Hepatopathy." *American Journal of Human Genetics* 99 (2): 414–22.

- Krause, Stefan W., Michael Rehli, Marina Kreutz, Lucia Schwarzfischer, Joseph D. Paulauskis, and Reinhard Andreesen. 1996. "Differential Screening Identifies Genetic Markers of Monocyte to Macrophage Maturation." *Journal of Leukocyte Biology*. <https://doi.org/10.1002/jlb.60.4.540>.
- Kunst, C. B., E. Mezey, M. J. Brownstein, and D. Patterson. 1997. "Mutations in SOD1 Associated with Amyotrophic Lateral Sclerosis Cause Novel Protein Interactions." *Nature Genetics* 15 (1): 91–94.
- Kwon, Nam Hoon, Taehee Kang, Jin Young Lee, Hyo Hyun Kim, Hye Rim Kim, Jeena Hong, Young Sun Oh, et al. 2011. "Dual Role of Methionyl-tRNA Synthetase in the Regulation of Translation and Tumor Suppressor Activity of Aminoacyl-tRNA Synthetase-Interacting Multifunctional Protein-3." *Proceedings of the National Academy of Sciences of the United States of America* 108 (49): 19635–40.
- Lahiri, Debomoy K. 2014. *Advances in Alzheimer's Research*. Bentham Science Publishers.
- Lee, Chien-Wei, Kai-Ping Chang, Yan-Yu Chen, Ying Liang, Chuen Hsueh, Jau-Song Yu, Yu-Sun Chang, and Chia-Jung Yu. 2015. "Overexpressed Tryptophanyl-tRNA Synthetase, an Angiostatic Protein, Enhances Oral Cancer Cell Invasiveness." *Oncotarget* 6 (26): 21979–92.
- Lee, Eun-Young, Hyun-Cheol Lee, Hyun-Kwan Kim, Song Yee Jang, Seong-Jun Park, Yong-Hoon Kim, Jong Hwan Kim, et al. 2016. "Infection-Specific Phosphorylation of Glutamyl-Prolyl tRNA Synthetase Induces Antiviral Immunity." *Nature Immunology*. <https://doi.org/10.1038/ni.3542>.
- Lee, Jin Young, Dae Gyu Kim, Byung-Gyu Kim, Won Suk Yang, Jeena Hong, Taehee Kang, Young Sun Oh, et al. 2014. "Promiscuous Methionyl-tRNA Synthetase Mediates Adaptive Mistranslation to Protect Cells against Oxidative Stress." *Journal of Cell Science* 127 (Pt 19): 4234–45.
- Lee, Yu-Nee, Hovav Nechushtan, Navah Figov, and Ehud Razin. 2004. "The Function of Lysyl-tRNA Synthetase and Ap4A as Signaling Regulators of MITF Activity in FcepsilonRI-Activated Mast Cells." *Immunity* 20 (2): 145–51.
- Lee, Yunjong, Senthilkumar S. Karuppagounder, Joo-Ho Shin, Yun-Il Lee, Han Seok Ko, Debbie Swing, Haisong Jiang, et al. 2013. "Parthanatos Mediates AIMP2-Activated Age-Dependent Dopaminergic Neuronal Loss." *Nature Neuroscience* 16 (10): 1392–1400.
- Lega, Jean-Christophe, Nicole Fabien, Quitterie Reynaud, Isabelle Durieu, Stéphane Durupt, Marine Dutertre, Jean-François Cordier, and Vincent Cottin. 2014. "The Clinical Phenotype Associated with Myositis-Specific and Associated Autoantibodies: A Meta-Analysis Revisiting the so-Called Antisynthetase Syndrome." *Autoimmunity Reviews* 13 (9): 883–91.
- Lin, Laura, and Paul Schimmel. 1996. "Mutational Analysis Suggests the Same Design for Editing Activities of Two tRNA Synthetases." *Biochemistry*. <https://doi.org/10.1021/bi960011y>.
- Liu, Jianming, Eveline Shue, Karla L. Ewalt, and Paul Schimmel. 2004. "A New Gamma-Interferon-Inducible Promoter and Splice Variants of an Anti-Angiogenic Human tRNA Synthetase." *Nucleic Acids Research* 32 (2): 719–27.
- Luo, Shen, and Rodney L. Levine. 2009. "Methionine in Proteins Defends against Oxidative Stress." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 23 (2): 464–72.
- Martin, John, George Mentis, and Andrew Paul Tosolini. 2021. *Dysfunction and Repair of Neural Circuits for Motor Control*. Frontiers Media SA.
- Mazumder, B., V. Seshadri, H. Imataka, N. Sonenberg, and P. L. Fox. 2001. "Translational Silencing of Ceruloplasmin Requires the Essential Elements of mRNA Circularization: poly(A) Tail, poly(A)-Binding Protein, and Eukaryotic Translation Initiation Factor 4G." *Molecular and Cellular Biology* 21 (19): 6440–49.
- Meel, Eline van, Daniel J. Wegner, Paul Cliften, Marcia C. Willing, Frances V. White, Stuart Kornfeld, and F. Sessions Cole. 2013. "Rare Recessive Loss-of-Function Methionyl-tRNA Synthetase Mutations Presenting as a Multi-Organ Phenotype." *BMC Medical Genetics* 14 (October): 106.
- Mehta, Kapil, and Zahid H. Siddik. 2009. *Drug Resistance in Cancer Cells*. Springer Science & Business Media.
- Mendes, Marisa I., Mariana Gutierrez Salazar, Kether Guerrero, Isabelle Thiffault, Gajja S. Salomons, Laurence Gauquelin, Luan T. Tran, et al. 2018. "Bi-Allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy." *The American Journal of Human Genetics*. <https://doi.org/10.1016/j.ajhg.2018.02.011>.
- Mourelatos, Z., L. Abel, J. Yong, N. Kataoka, and G. Dreyfuss. 2001. "SMN Interacts with a Novel Family of hnRNP and Spliceosomal Proteins." *The EMBO Journal* 20 (19): 5443–52.

- Mo, Zhongying, Xiaobei Zhao, Huaqing Liu, Qinghua Hu, Xu-Qiao Chen, Jessica Pham, Na Wei, et al. 2018. “Aberrant GlyRS-HDAC6 Interaction Linked to Axonal Transport Deficits in Charcot-Marie-Tooth Neuropathy.” *Nature Communications* 9 (1): 1007.
- Mukhopadhyay, Rupak, Jie Jia, Abul Arif, Partho Sarothi Ray, and Paul L. Fox. 2009. “The GAIT System: A Gatekeeper of Inflammatory Gene Expression.” *Trends in Biochemical Sciences* 34 (7): 324–31.
- Mukhopadhyay, Rupak, Partho Sarothi Ray, Abul Arif, Anna K. Brady, Michael Kinter, and Paul L. Fox. 2008. “DAPK-ZIPK-L13a Axis Constitutes a Negative-Feedback Module Regulating Inflammatory Gene Expression.” *Molecular Cell*. <https://doi.org/10.1016/j.molcel.2008.09.019>.
- Nafisinia, Michael, Nara Sobreira, Lisa Riley, Wendy Gold, Birgit Uhlenberg, Claudia Weiß, Corinne Boehm, Kristina Prelog, Robert Ouvrier, and John Christodoulou. 2017. “Mutations in RARS Cause a Hypomyelination Disorder Akin to Pelizaeus–Merzbacher Disease.” *European Journal of Human Genetics*. <https://doi.org/10.1038/ejhg.2017.119>.
- Nechushtan, Hovav, and Ehud Razin. 2002. “The Function of MITF and Associated Proteins in Mast Cells.” *Molecular Immunology*. [https://doi.org/10.1016/s0161-5890\(02\)00059-7](https://doi.org/10.1016/s0161-5890(02)00059-7).
- Negrutskii, B. S., V. F. Shalak, P. Kerjan, A. V. El’skaya, and M. Mirande. 1999. “Functional Interaction of Mammalian Valyl-tRNA Synthetase with Elongation Factor EF-1alpha in the Complex with EF-1H.” *The Journal of Biological Chemistry* 274 (8): 4545–50.
- Nilbert, M., A. Rydholm, F. Mitelman, P. S. Meltzer, and N. Mandahl. 1995. “Characterization of the 12q13-15 Amplicon in Soft Tissue Tumors.” *Cancer Genetics and Cytogenetics* 83 (1): 32–36.
- Norcum, Mona T., and Nicolas Boisset. 2002. “Three-Dimensional Architecture of the Eukaryotic Multisynthetase Complex Determined from Negatively Stained and Cryoelectron Micrographs.” *FEBS Letters* 512 (1-3): 298–302.
- Norcum, Mona Trempe, J. Anthony Warrington, and Cindy L. Wolfe. 2006. “Removal of arginyl-tRNA Synthetase Markedly Alters the Three-dimensional Structure of the Eukaryotic Multisynthetase Complex.” *The FASEB Journal*. <https://doi.org/10.1096/fasebj.20.4.a504-b>.
- Norcum, M. T. 1989. “Isolation and Electron Microscopic Characterization of the High Molecular Mass Aminoacyl-tRNA Synthetase Complex from Murine Erythroleukemia Cells.” *The Journal of Biological Chemistry* 264 (25): 15043–51.
- Nureki, O., D. G. Vassylyev, M. Tateno, A. Shimada, T. Nakama, S. Fukai, M. Konno, T. L. Hendrickson, P. Schimmel, and S. Yokoyama. 1998. “Enzyme Structure with Two Catalytic Sites for Double-Sieve Selection of Substrate.” *Science* 280 (5363): 578–82.
- O’Donoghue, Patrick, and Zaida Luthey-Schulten. 2003. “On the Evolution of Structure in Aminoacyl-tRNA Synthetases.” *Microbiology and Molecular Biology Reviews: MMBR* 67 (4): 550–73.
- Ofir-Birin, Yifat, Pengfei Fang, Steven P. Bennett, Hui-Min Zhang, Jing Wang, Inbal Rachmin, Ryan Shapiro, et al. 2013. “Structural Switch of Lysyl-tRNA Synthetase between Translation and Transcription.” *Molecular Cell* 49 (1): 30–42.
- Oh, Ah-Young, Youn Sang Jung, Jiseon Kim, Jee-Hyun Lee, Jung-Hyun Cho, Ho-Young Chun, Soyoung Park, et al. 2016. “Inhibiting DX2-p14/ARF Interaction Exerts Antitumor Effects in Lung Cancer and Delays Tumor Progression.” *Cancer Research* 76 (16): 4791–4804.
- Otani, Atsushi, Bonnie M. Slike, Michael I. Dorrell, John Hood, Karen Kinder, Karla L. Ewalt, David Cheresch, Paul Schimmel, and Martin Friedlander. 2002. “A Fragment of Human TrpRS as a Potent Antagonist of Ocular Angiogenesis.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (1): 178–83.
- Palmer, J. L., S. Masui, S. Pritchard, D. K. Kalousek, and P. H. Sorensen. 1997. “Cytogenetic and Molecular Genetic Analysis of a Pediatric Pleomorphic Sarcoma Reveals Similarities to Adult Malignant Fibrous Histiocytoma.” *Cancer Genetics and Cytogenetics* 95 (2): 141–47.
- Park, Bum-Joon, Jin Wook Kang, Sang Won Lee, So-Jung Choi, Young Kee Shin, Young Ha Ahn, Yun Hee Choi, Dongho Choi, Kwang Soo Lee, and Sunghoon Kim. 2005. “The Haploinsufficient Tumor Suppressor p18 Upregulates p53 via Interactions with ATM/ATR.” *Cell* 120 (2): 209–21.
- Park, Bum-Joon, Young Sun Oh, Seung Yong Park, So Jung Choi, Cornelia Rudolph, Brigitte Schlegelberger, and Sunghoon Kim. 2006. “AIMP3 Haploinsufficiency Disrupts Oncogene-Induced p53 Activation and Genomic Stability.” *Cancer Research* 66 (14): 6913–18.
- Park, Min Chul, Taehee Kang, Da Jin, Jung Min Han, Sang Bum Kim, Yun Jung Park, Kiwon Cho, et al. 2012. “Secreted Human Glycyl-tRNA Synthetase Implicated in Defense against ERK-Activated

- Tumorigenesis.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (11): E640–47.
- Park, S. G., K. H. Jung, J. S. Lee, Y. J. Jo, H. Motegi, S. Kim, and K. Shiba. 1999. “Precursor of pro-Apoptotic Cytokine Modulates Aminoacylation Activity of tRNA Synthetase.” *The Journal of Biological Chemistry* 274 (24): 16673–76.
- Park, Sung Mi, Ki Young Paek, Ka Young Hong, Christopher J. Jang, Sungchan Cho, Ji Hoon Park, Jong Heon Kim, Eric Jan, and Sung Key Jang. 2011. “Translation-Competent 48S Complex Formation on HCV IRES Requires the RNA-Binding Protein NSAP1.” *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkr509>.
- Perona, J. J., M. A. Rould, T. A. Steitz, J. L. Risler, C. Zelwer, and S. Brunie. 1991. “Structural Similarities in Glutamyl- and Methionyl-tRNA Synthetases Suggest a Common Overall Orientation of tRNA Binding.” *Proceedings of the National Academy of Sciences of the United States of America* 88 (7): 2903–7.
- Perona, John J., and Andrew Hadd. 2012. “Structural Diversity and Protein Engineering of the Aminoacyl-tRNA Synthetases.” *Biochemistry* 51 (44): 8705–29.
- Puffenberger, Erik G., Robert N. Jinks, Carrie Sougnez, Kristian Cibulskis, Rebecca A. Willert, Nathan P. Achilly, Ryan P. Cassidy, et al. 2012. “Genetic Mapping and Exome Sequencing Identify Variants Associated with Five Novel Diseases.” *PloS One* 7 (1): e28936.
- Quevillon, S., F. Agou, J. C. Robinson, and M. Mirande. 1997. “The p43 Component of the Mammalian Multi-Synthetase Complex Is Likely to Be the Precursor of the Endothelial Monocyte-Activating Polypeptide II Cytokine.” *The Journal of Biological Chemistry* 272 (51): 32573–79.
- Quevillon, S., and M. Mirande. 1996. “The p18 Component of the Multisynthetase Complex Shares a Protein Motif with the Beta and Gamma Subunits of Eukaryotic Elongation Factor 1.” *FEBS Letters* 395 (1): 63–67.
- Quevillon, S., J. C. Robinson, E. Berthonneau, M. Siatecka, and M. Mirande. 1999. “Macromolecular Assemblage of Aminoacyl-tRNA Synthetases: Identification of Protein-Protein Interactions and Characterization of a Core Protein.” *Journal of Molecular Biology* 285 (1): 183–95.
- Ray, Partho Sarothi, James C. Sullivan, Jie Jia, John Francis, John R. Finnerty, and Paul L. Fox. 2011. “Evolution of Function of a Fused Metazoan tRNA Synthetase.” *Molecular Biology and Evolution* 28 (1): 437–47.
- Razin, Ehud, Zhao Cheng Zhang, Hovav Nechushtan, Shahar Frenkel, Yu-Nee Lee, Ramachandran Arudchandran, and Juan Rivera. 1999. “Suppression of Microphthalmia Transcriptional Activity by Its Association with Protein Kinase C-Interacting Protein 1 in Mast Cells.” *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.274.48.34272>.
- Reifenberger, G., K. Ichimura, J. Reifenberger, A. G. Elkahloun, P. S. Meltzer, and V. P. Collins. 1996. “Refined Mapping of 12q13-q15 Amplicons in Human Malignant Gliomas Suggests CDK4/SAS and MDM2 as Independent Amplification Targets.” *Cancer Research* 56 (22): 5141–45.
- Rho, S. B., K. H. Lee, J. W. Kim, K. Shiba, Y. J. Jo, and S. Kim. 1996. “Interaction between Human tRNA Synthetases Involves Repeated Sequence Elements.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (19): 10128–33.
- Rosenthal, Philip. 2018. *Pediatric Hepatology, An Issue of Clinics in Liver Disease E-Book*. Elsevier Health Sciences.
- Rould, M. A., J. J. Perona, D. Söll, and T. A. Steitz. 1989. “Structure of E. Coli Glutamyl-tRNA Synthetase Complexed with tRNA(Gln) and ATP at 2.8 Å Resolution.” *Science* 246 (4934): 1135–42.
- Ruff, R. L., and J. Weissman. 1991. “Iodoacetate-Induced Contracture in Rat Skeletal Muscle: Possible Role of ADP.” *The American Journal of Physiology* 261 (5 Pt 1): C828–36.
- Sajish, Mathew, and Paul Schimmel. 2015. “A Human tRNA Synthetase Is a Potent PARP1-Activating Effector Target for Resveratrol.” *Nature* 519 (7543): 370–73.
- Sajish, Mathew, Quansheng Zhou, Shuji Kishi, Delgado M. Valdez Jr, Mili Kapoor, Min Guo, Sunhee Lee, Sunghoon Kim, Xiang-Lei Yang, and Paul Schimmel. 2012. “Trp-tRNA Synthetase Bridges DNA-PKcs to PARP-1 to Link IFN- γ and p53 Signaling.” *Nature Chemical Biology* 8 (6): 547–54.
- Sasaki, Hiroshi M., Shun-Ichi Sekine, Toru Sengoku, Ryuya Fukunaga, Motoyuki Hattori, Yukiko Utsunomiya, Chizu Kuroishi, Seiki Kuramitsu, Mikako Shirouzu, and Shigeyuki Yokoyama. 2006. “Structural and Mutational Studies of the Amino Acid-Editing Domain from Archaeal/eukaryal

- Phenylalanyl-tRNA Synthetase.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (40): 14744–49.
- Saxholm, H. J., and H. C. Pitot. 1979. “Characterization of a Proteolipid Complex of Aminoacyl-tRNA Synthetases and Transfer RNA from Rat Liver.” *Biochimica et Biophysica Acta* 562 (3): 386–99.
- Schimmel, Paul. 2018. “The Emerging Complexity of the tRNA World: Mammalian tRNAs beyond Protein Synthesis.” *Nature Reviews. Molecular Cell Biology* 19 (1): 45–58.
- Schimmel, P., and L. Ribas De Pouplana. 2000. “Footprints of Aminoacyl-tRNA Synthetases Are Everywhere.” *Trends in Biochemical Sciences* 25 (5): 207–9.
- Schimmel, P., and L. Ribas de Pouplana. 2001. “Formation of Two Classes of tRNA Synthetases in Relation to Editing Functions and Genetic Code.” *Cold Spring Harbor Symposia on Quantitative Biology* 66: 161–66.
- Schmidt, Eric, and Paul Schimmel. 1994. “Mutational Isolation of a Sieve for Editing in a Transfer RNA Synthetase.” *Science*. <https://doi.org/10.1126/science.8146659>.
- Schwarz, Quentin, Chenghua Gu, Hajime Fujisawa, Kimberly Sabelko, Marina Gertsenstein, Andras Nagy, Masahiko Taniguchi, et al. 2004. “Vascular Endothelial Growth Factor Controls Neuronal Migration and Cooperates with Sema3A to Pattern Distinct Compartments of the Facial Nerve.” *Genes & Development* 18 (22): 2822–34.
- Seidler, Norbert W. 2012. *GAPDH: Biological Properties and Diversity*. Springer Science & Business Media.
- Shalak, Vyacheslav, Monika Kaminska, Rita Mitnacht-Kraus, Peter Vandenabeele, Matthias Clauss, and Marc Mirande. 2001. “The EMAPII Cytokine Is Released from the Mammalian Multisynthetase Complex after Cleavage of Its p43/proEMAPII Component.” *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.m100489200>.
- Shukla, Anju, Aneek Das Bhowmik, Malavika Hebbar, Kadavigere V. Rajagopal, Katta M. Girisha, Neerja Gupta, and Ashwin Dalal. 2018. “Homozygosity for a Nonsense Variant in AIMP2 Is Associated with a Progressive Neurodevelopmental Disorder with Microcephaly, Seizures, and Spastic Quadriparesis.” *Journal of Human Genetics* 63 (1): 19–25.
- Sihag, R. K., and M. P. Deutscher. 1983. “Perturbation of the Aminoacyl-tRNA Synthetase Complex by Salts and Detergents. Importance of Hydrophobic Interactions and Possible Involvement of Lipids.” *The Journal of Biological Chemistry* 258 (19): 11846–50.
- Singer, Dinah. 2008. “Faculty Opinions Recommendation of DAPK-ZIPK-L13a Axis Constitutes a Negative-Feedback Module Regulating Inflammatory Gene Expression.” *Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature*. <https://doi.org/10.3410/f.1130849.587937>.
- Sleigh, James N., John M. Dawes, Steven J. West, Na Wei, Emily L. Spaulding, Adriana Gómez-Martín, Qian Zhang, et al. 2017. “Trk Receptor Signaling and Sensory Neuron Fate Are Perturbed in Human Neuropathy Caused by Mutations.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (16): E3324–33.
- Sprinzel, M., and F. Cramer. 1975. “Site of Aminoacylation of tRNAs from Escherichia Coli with Respect to the 2'- or 3'-Hydroxyl Group of the Terminal Adenosine.” *Proceedings of the National Academy of Sciences of the United States of America* 72 (8): 3049–53.
- Starzyk, R. M., T. A. Webster, and P. Schimmel. 1987. “Evidence for Dispensable Sequences Inserted into a Nucleotide Fold.” *Science* 237 (4822): 1614–18.
- Storkebaum, Erik. 2016. “Peripheral Neuropathy via Mutant tRNA Synthetases: Inhibition of Protein Translation Provides a Possible Explanation.” *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 38 (9): 818–29.
- Svitkin, Yuri V., Akiko Yanagiya, Alexey E. Karetnikov, Tommy Alain, Marc R. Fabian, Arkady Khoutorsky, Sandra Perreault, Ivan Topisirovic, and Nahum Sonenberg. 2013. “Control of Translation and miRNA-Dependent Repression by a Novel poly(A) Binding Protein, hnRNP-Q.” *PLoS Biology* 11 (5): e1001564.
- Thul, Peter J., and Cecilia Lindskog. 2018. “The Human Protein Atlas: A Spatial Map of the Human Proteome.” *Protein Science*. <https://doi.org/10.1002/pro.3307>.
- Tolstrup, A. B., A. Bejder, J. Fleckner, and J. Justesen. 1995. “Transcriptional Regulation of the Interferon-Gamma-Inducible Tryptophanyl-tRNA Synthetase Includes Alternative Splicing.” *The Journal of Biological Chemistry* 270 (1): 397–403.

- Turpaev, K. T., V. M. Zakhariev, I. V. Sokolova, A. N. Narovlyansky, A. M. Amchenkova, J. Justesen, and L. Y. Frolova. 1996. "Alternative Processing of the Tryptophanyl-tRNA Synthetase mRNA from Interferon-Treated Human Cells." *European Journal of Biochemistry / FEBS* 240 (3): 732–37.
- Tzima, Eleni, John S. Reader, Mohamad Irani-Tehrani, Karla L. Ewalt, Martin A. Schwartz, and Paul Schimmel. 2005. "VE-Cadherin Links tRNA Synthetase Cytokine to Anti-Angiogenic Function." *The Journal of Biological Chemistry* 280 (4): 2405–8.
- Valencia-Sánchez, Marco Igor, Annia Rodríguez-Hernández, Ruben Ferreira, Hugo Aníbal Santamaría-Suárez, Marcelino Arciniega, Anne-Catherine Dock-Bregeon, Dino Moras, et al. 2016. "Structural Insights into the Polyphyletic Origins of Glycyl tRNA Synthetases." *The Journal of Biological Chemistry* 291 (28): 14430–46.
- Vanha-Aho, Leena-Maija, Susanna Valanne, and Mika Rämet. 2016. "Cytokines in Drosophila Immunity." *Immunology Letters* 170 (February): 42–51.
- Vellaichamy, Adaikkalam, Arun Sreekumar, John R. Strahler, Theckelnaycke Rajendiran, Jindan Yu, Sooryanarayana Varambally, Yong Li, Gilbert S. Omenn, Arul M. Chinnaiyan, and Alexey I. Nesvizhskii. 2009. "Proteomic Interrogation of Androgen Action in Prostate Cancer Cells Reveals Roles of Aminoacyl tRNA Synthetases." *PLoS One* 4 (9): e7075.
- Venema, R. C., and J. A. Traugh. 1991. "Protein Kinase C Phosphorylates Glutamyl-tRNA Synthetase in Rabbit Reticulocytes Stimulated by Tumor Promoting Phorbol Esters." *The Journal of Biological Chemistry* 266 (8): 5298–5302.
- Vink, Robert, and Mihai Nechifor. 2011. *Magnesium in the Central Nervous System*. University of Adelaide Press.
- Vo, My-Nuong, Markus Terrey, Jeong Woong Lee, Bappaditya Roy, James J. Moresco, Litao Sun, Hongjun Fu, et al. 2018. "ANKRD16 Prevents Neuron Loss Caused by an Editing-Defective tRNA Synthetase." *Nature* 557 (7706): 510–15.
- Vo, My-Nuong, Xiang-Lei Yang, and Paul Schimmel. 2011. "Dissociating Quaternary Structure Regulates Cell-Signaling Functions of a Secreted Human tRNA Synthetase." *The Journal of Biological Chemistry* 286 (13): 11563–68.
- Vyas, Keyur, Sujan Chaudhuri, Douglas W. Leaman, Anton A. Komar, Alla Musiyenko, Sailen Barik, and Barsanjit Mazumder. 2009. "Genome-Wide Polysome Profiling Reveals an Inflammation-Responsive Posttranscriptional Operon in Gamma Interferon-Activated Monocytes." *Molecular and Cellular Biology* 29 (2): 458–70.
- Wakasugi, K., and P. Schimmel. 1999. "Two Distinct Cytokines Released from a Human Aminoacyl-tRNA Synthetase." *Science* 284 (5411): 147–51.
- Wellman, Theresa L., Midori Eckenstein, Cheung Wong, Mercedes Rincon, Takamaru Ashikaga, Sharon L. Mount, Christopher S. Francklyn, and Karen M. Lounsbury. 2014. "Threonyl-tRNA Synthetase Overexpression Correlates with Angiogenic Markers and Progression of Human Ovarian Cancer." *BMC Cancer* 14 (August): 620.
- Wells, S. E., P. E. Hillner, R. D. Vale, and A. B. Sachs. 1998. "Circularization of mRNA by Eukaryotic Translation Initiation Factors." *Molecular Cell* 2 (1): 135–40.
- Wolf, Nicole I., Gajja S. Salomons, Richard J. Rodenburg, Petra J. W. Pouwels, Jolanda H. Schieving, Terry G. J. Derks, Johanna M. Fock, et al. 2014. "Mutations in RARS Cause Hypomyelination." *Annals of Neurology* 76 (1): 134–39.
- Xu, Haiming, Nikolay L. Malinin, Niranjana Awasthi, Roderich E. Schwarz, and Margaret A. Schwarz. 2015. "The N Terminus of Pro-Endothelial Monocyte-Activating Polypeptide II (EMAP II) Regulates Its Binding with the C Terminus, Arginyl-tRNA Synthetase, and Neurofilament Light Protein." *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M114.630533>.
- Xu, Xiaoling, Yi Shi, Hui-Min Zhang, Eric C. Swindell, Alan G. Marshall, Min Guo, Shuji Kishi, and Xiang-Lei Yang. 2012. "Unique Domain Appended to Vertebrate tRNA Synthetase Is Essential for Vascular Development." *Nature Communications* 3 (February): 681.
- Xu, Zhiwen, Zhiyi Wei, Jie J. Zhou, Fei Ye, Wing-Sze Lo, Feng Wang, Ching-Fun Lau, et al. 2012. "Internally Deleted Human tRNA Synthetase Suggests Evolutionary Pressure for Repurposing." *Structure* 20 (9): 1470–77.
- Yao, Peng, and Paul L. Fox. 2013. "Aminoacyl-tRNA Synthetases in Medicine and Disease." *EMBO Molecular Medicine*. <https://doi.org/10.1002/emmm.201100626>.

- Yao, Peng, Alka A. Potdar, Abul Arif, Partho Sarothi Ray, Rupak Mukhopadhyay, Belinda Willard, Yichi Xu, Jun Yan, Gerald M. Saidel, and Paul L. Fox. 2012. "Coding Region Polyadenylation Generates a Truncated tRNA Synthetase That Counters Translation Repression." *Cell* 149 (1): 88–100.
- You, Fuping, Hui Sun, Xiang Zhou, Wenxiang Sun, Shimin Liang, Zhonghe Zhai, and Zhengfan Jiang. 2009. "PCBP2 Mediates Degradation of the Adaptor MAVS via the HECT Ubiquitin Ligase AIP4." *Nature Immunology*. <https://doi.org/10.1038/ni.1815>.
- Zhang, Xiaochang, Jiqiang Ling, Giulia Barcia, Lili Jing, Jiang Wu, Brenda J. Barry, Ganeshwaran H. Mochida, et al. 2014. "Mutations in QARS, Encoding Glutaminyl-tRNA Synthetase, Cause Progressive Microcephaly, Cerebral-Cerebellar Atrophy, and Intractable Seizures." *American Journal of Human Genetics* 94 (4): 547–58.
- Zhou, Jie J., Feng Wang, Zhiwen Xu, Wing-Sze Lo, Ching-Fun Lau, Kyle P. Chiang, Leslie A. Nangle, et al. 2014. "Secreted Histidyl-tRNA Synthetase Splice Variants Elaborate Major Epitopes for Autoantibodies in Inflammatory Myositis." *The Journal of Biological Chemistry* 289 (28): 19269–75.
- Zhou, Quansheng, Mili Kapoor, Min Guo, Rajesh Belani, Xiaoling Xu, William B. Kiosses, Melanie Hanan, et al. 2010. "Orthogonal Use of a Human tRNA Synthetase Active Site to Achieve Multifunctionality." *Nature Structural & Molecular Biology* 17 (1): 57–61.
- Zhu, Xiaodong, Yang Liu, Yanqing Yin, Aiyun Shao, Bo Zhang, Sunghoon Kim, and Jiawei Zhou. 2009. "MSC p43 Required for Axonal Development in Motor Neurons." *Proceedings of the National Academy of Sciences of the United States of America* 106 (37): 15944–49.

CHAPTER II

The MARS and GAIT complex respond to infection: A Mass Spectrometric study

Summary

In this section, I employ mass spectrometry, post affinity purification, utilising an anti-EPRS antibody for immunoprecipitation to track the kinetics of the MARS complex. The data sheds light on the dynamic changes to the MARS complex post-infection in adult animals. Interestingly, I find that the GAIT complex, described in the mammalian host defence response, is also stabilised as part of the host defence response of the adult fly.

1. Introduction

EPRS seems to serve as a key gatekeeper of inflammatory gene translation, evolving as an ‘off-switch’ to modulate protein production during the time the host/cell is dealing with an infection. EPRS appears to be part of two major complexes, the MARS Complex, and the GAIT Complex.

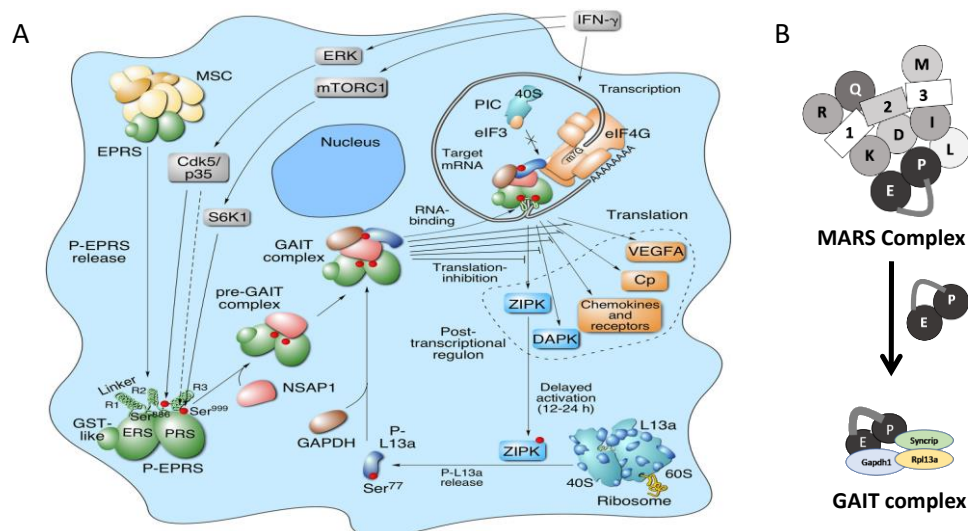


Figure 2.1: A. Dynamics of the MARS and GAIT complex, in mammals, in response to infection. EPRS, in response to infection, comes off the MARS complex and associates with NSAP1, GAPDH, and ribosomal protein L13a to form an interferon γ -activated inhibitor of translation (GAIT) complex. The phenomenon has been reviewed in detail (Mukhopadhyay et. al., 2009; Kim et. al, 2014; Arif et. al., 2017). **B. Formation of the GAIT complex in Drosophila.** Our data, described in this chapter, indicates that a GAIT-like complex is also formed in response to the infection of adult flies, with components orthologous to that of mammals. Figure Adapted from Arif et. al., 2017.

Experimental evidence for the MARS complex in flies is scarce, with limited information on the MARS complex (Shafer et. al., 1976; Cerini et. al., 1991; Lu et. al., 2015) and no experimental evidence on the existence of the GAIT complex. In mammals, the GAIT complex comprises EPRS, NS1-associated protein (NSAP1), ribosomal protein L13a, and GAPDH (Fig 2.1). The ribosomal protein L13a interacts with the Cp 3' UTR GAIT element such that L13a is required for translational silencing activity in IFN γ -treated cells. Here, EPRS binds the 3' UTR GAIT element in multiple proinflammatory transcripts (e.g., VEGF-A) and inhibits their translation in macrophages.

2. Aim

In this study, we attempt to experimentally verify the existence of the *Drosophila* MARS complex and study the change in the composition of the MARS complex in response to both gram-positive and gram-negative infection. Our validation is in the context of EPRS interacting with other members of the MARS complex. Also, interestingly we find that the EPRS also associates with a GAIT-like complex in response to infection.

3. Materials & Methods

3.1 *Drosophila* husbandry: All flies were raised, and crosses were conducted at 25°C in standard corn meal agar unless stated otherwise. *w¹¹¹⁸* fly lines were procured from Bloomington *Drosophila* Stock Centre (BDSC), Indiana.

3.2 Infection Assay: Overnight grown cultures of *Micrococcus luteus* and *Ecc15*, adjusted to OD₆₀₀ = 100, were used to infect 6-8 day old adult males and females separately (Fig 2.2). Flies were infected at the sterno-pleural plate with insect pins dipped in the bacterial culture. They were allowed to recover and were shifted to 29°C and sacrificed at different time points (0,4,8,16, 24,48 hours) post-infection.

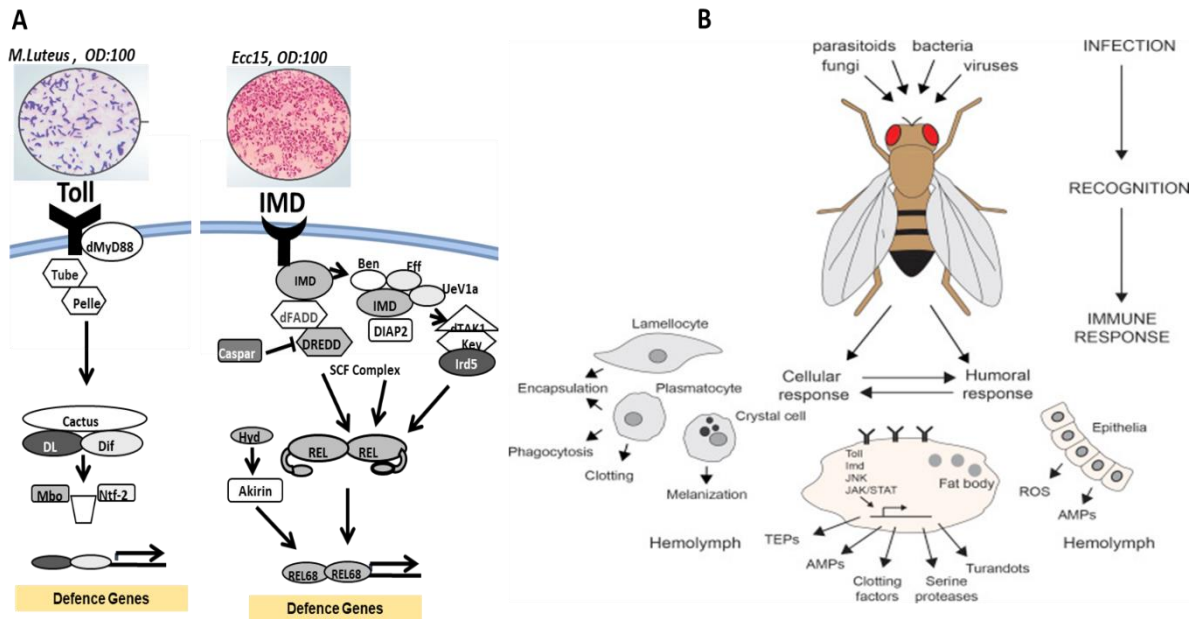


Figure 2.2: A. Infection by Gram-positive (*M. luteus*) and Gram-negative (*Ecc15*). Bacterial infection was used to activate the major host defence pathways in the adult animal, which include the Toll/NFκB and IMD/NFκB pathways. **B. Cellular and humoral response to infection.** The infection leads to the activation of both the cellular and humoral arms of the immune response.

3.3 Immuno-precipitation: In total, 6-10 day-old adult flies were lysed at different time points mentioned above in Co-IP Lysis Buffer (20 mM Tris pH 8.0, 137 mM NaCl, 1% IGEPAL, 2 mM EDTA, 1× PIC) using a Bead beater/Dounce homogeniser and centrifuged at 21 000g for 30 min. 3 mg of total fly lysate was incubated with 10 µg of primary antibody (Rb anti-EPRS) and 10 µg of Normal Rabbit IgG for 1-2 hours at room temperature. The Anti-EPRS antibody was generated in Courey Lab, UCLA. Antigen-antibody complexes were captured using 50 µl of BioRad SureBeads Protein A (1614013) at room temperature for 2 h. Beads were washed three times with Co-IP lysis buffer, and protein complexes were eluted by boiling in 1× Laemmli sample buffer. Eluted proteins were resolved on a 10% polyacrylamide gel followed by western blotting or in-gel trypsin digestion. Proteins separated by SDS-PAGE were stained with Coomassie stain.

3.4 In-gel trypsin digestion and LC-MS/MS analysis: Before in-gel trypsin digestion of the Co-IP eluate, the antibody was crosslinked to SureBeads using DMP (Sigma) according to the NEB crosslinking protocol to avoid elution of the antibody. After crosslinking 10 µg Caspar antibody, Co-IP was performed as described above. In-gel trypsin digestion was carried out as previously described. Briefly, Coomassie-stained bands on the gel were excised and cut into 1 mm cubes. Gel pieces were transferred to a clean microcentrifuge tube and destained with

buffer containing 50% acetonitrile in 50 mM Ammonium bicarbonate. Reduction and alkylation were carried out on the destained gel pieces by incubating with 10 mM dithiothreitol followed by incubating with 20 mM iodoacetamide. Gel pieces were saturated with sequencing-grade Trypsin (Promega) at a concentration of 10 ng/ μ l and incubated overnight at 37°C. Peptides were extracted by sequential addition of 100 μ l of 0.4% Trifluoroacetic acid (TFA) in 10% ACN, 100 μ l of 0.4% TFA in 60% ACN, and 100 μ l of ACN. The pooled extract was dried in a vacuum centrifuge and reconstituted with 50 μ l of 0.1% TFA. The peptides in TFA were purified using the StageTip protocol.

LC–MS/MS analysis was performed on the Sciex TripleTOF 6600 mass spectrometer interfaced with an Eksigent nano-LC 425. Tryptic peptides (1 μ g) were loaded onto an Eksigent C18 trap (5 μ g capacity) and subsequently eluted with a linear acetonitrile gradient on an Eksigent C18 analytical column (15 cm \times 75 μ m internal diameter). A typical LC run lasted 2 h post loading onto the trap at a constant flow rate of 300 nl/min with solvent A consisting of water + 0.1% formic acid and solvent B consisting of acetonitrile. The gradient schedule for the LC run was 5% (vol/vol) B for 10 min, a linear gradient of B from 0% to 80% (vol/vol) over 80 min, 80% (vol/vol) B for 15 min and equilibration with 5% (vol/vol) B for 15 min. Data were acquired in an information-dependent acquisition mode over a mass range of 300–2000 m/z . Each full MS survey scan was followed by MS/MS of the 15 most intense peptides. Dynamic exclusion was enabled for all experiments (repeat count 1; exclusion duration 6 s). Peptide identification and quantification were carried out with the SCIEX Protein Pilot software at a false discovery rate of 1%. A Ref Seq *Drosophila* protein database (release 6) was used for peptide identification. Proteins that were identified in two or more replicates and had two or more quantified peptides were tabulated.

4. Results

4.1 EPRS immune-precipitates pull down the multi-aminoacyl tRNA synthetase complex in *Drosophila*: To assess whether EPRS exists as a part of the multi-aminoacyl tRNA synthetase complex (MARS complex), we have used an EPRS antibody generated in Courey Lab to immuno-precipitate EPRS and its interactors. The anti-EPRS antibody recognises the two isoforms of EPRS running at 108 kDa and 180 kDa in adult whole fly lysates with specificity in a western blot, which is not seen in the immune precipitates of the pre-immuniser. EPRS immune-precipitates could enrich EPRS as visualised by western blotting. To identify the interactors of EPRS, we processed EPRS immune-precipitates through mass

spectrometry from adult male and female lysates individually and found that EPRS associates with cytoplasmic proteins involved in protein synthesis including the components of the MARS complex.

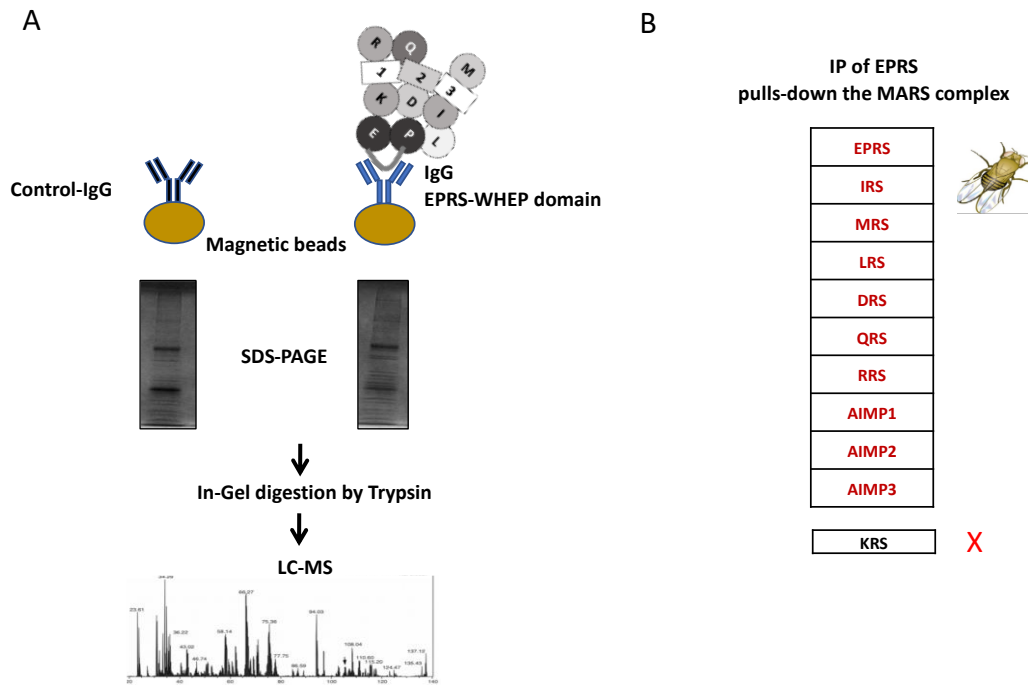


Figure 2.3: Uncovering interactors of EPRS using immunoprecipitation **A. IP and in-gel digestion.** The adult lysates were subject to IPs using both anti-EPRS and an IgG control. The affinity-purified proteins were separated on a 10% SDS-PAGE gel, and 10 slices were cut in each lane (control & experimental) at equivalent positions of m.w; the sliced gel fragments were subjected to trypsin cleavage as described in Materials & Methods with the eluted fractions subject to LC-MS. **B. MARS complex components affinity purified along with EPRS.** In general, in *Drosophila* adults, all orthologous MARS components could be identified in the LC-MS experiments except KRS.

These enzymes correspond to aminoacyl tRNA synthetases specific to amino acids Glu, Pro, Ile, Leu, Met, Gln, Lys, Arg, and Asp, and three non-synthetase proteins aminoacyl tRNA synthetase interacting proteins (AIMPs). The MARS complex in *Drosophila* is composed of eleven polypeptides ranging from 18 k Da to 180 k Da (Fig 2.3). Earlier, The MARS complex has been isolated from the Schneider cells; purified to homogeneity by Kerjan *et. al*, 1994 during their attempt to test the ubiquity of the MARS complex in all metazoan species, who searched for a complex like that obtained from mammalian cells along with a few other metazoans. They were the first group to show that the MARS complex is an idiosyncratic feature of all eukaryotic cells and that the composition of the MARS complex is conserved from *Drosophila* to humans.

4.2 Sexual dimorphism at the molecular level, EPRS interaction is different between males and females: IP experiments were conducted separately in both males and females. The experiments suggest that the composition of the MARS complex shows sexual dimorphism. In males, the data parallel data from mammals, with all members, except KRS, being significantly enriched along with EPRS. IRS, MRS, and LRS are enriched significantly, in agreement with these being part of sub-complex I in MARS. AIMP1 and AIMP3, and also members of sub-complex II (QRS, RRS, AIMP1), are enriched but with lower peptide counts (Fig 2.4), suggesting that subcomplex I is the most stable element of MARS in adult male flies.

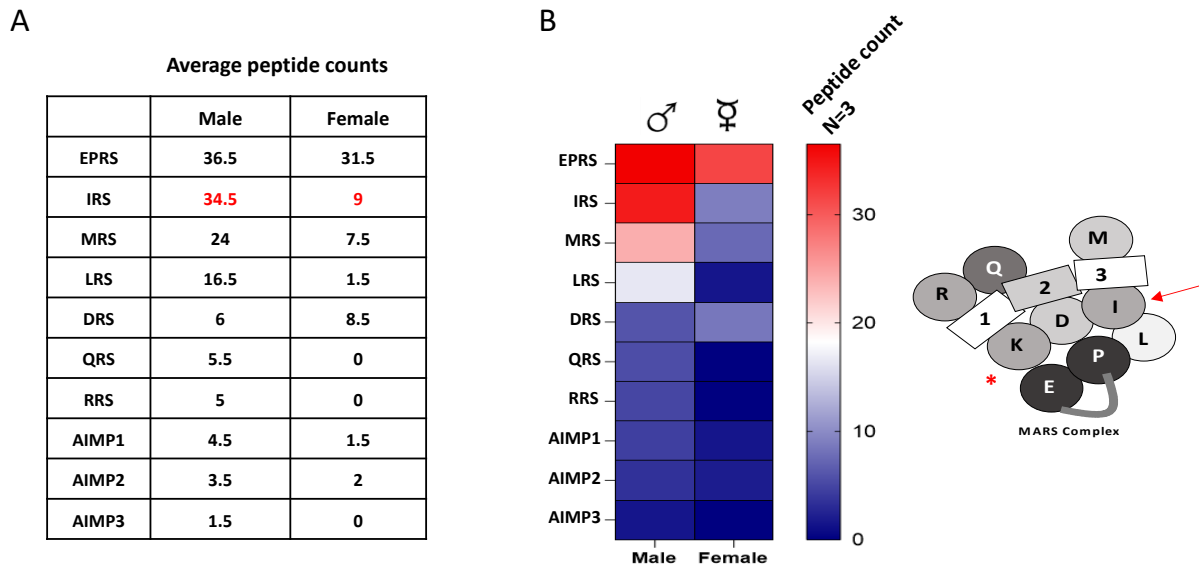


Figure 2.4: Sexual Dimorphism in the formation of the MARS complex. A. Components of the MARS complex. The IP data suggests that, unlike in males, in females, the MARS complex is relatively unstable. **B. Heat map of peptide counts for different members of the MARS complex.** Data suggests that the absence or lack of interaction of IRS (arrow) being a major cause for the reduced complex formation. KRS (*) is not enriched as a component of MARS in our hands.

In females, the data suggests that in comparison with males, subcomplex I is not as stable, and there being no association of EPRS with elements of subcomplex II (QRS, RRS, AIMP1, AIMP2), where counts are at or nearly zero (Fig 2.4).

4.3 Kinetics of EPRS and the MARS complex in *Drosophila* post-infection: Since EPRS exists as a dimer in the MARS complex, we measured the levels of EPRS enriched by IP over a 48-hour time period, post-infection, by both *M. luteus* and *Ecc15* (described in Materials & Methods) to the adult fly. IP with an IgG control was used as an affinity control. Our data suggest that EPRS is enriched post-infection. The kinetics appear to depend on infection with

a rapid increase in EPRS in the case of *M. luteus* infection and sustained retention of the stage even up to 48 hours (Fig 2.5).

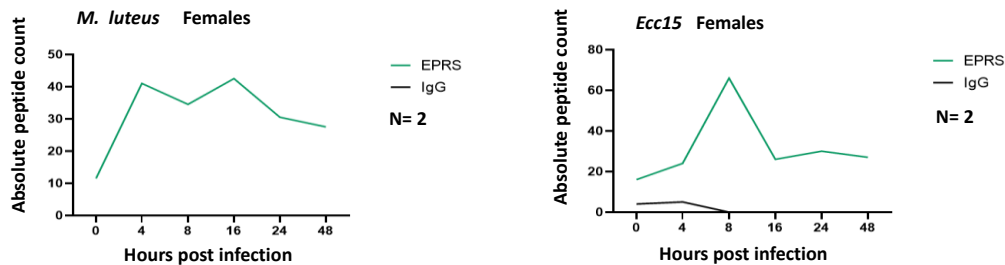


Fig 2.5: Kinetics of EPRS. Peptide counts of EPRS show an increase for both *M. luteus* and *Ecc15* infections. If the EPRS levels are assumed to be a proxy for the formation of the MARS complex, then the same is enriched in response to infection. The kinetics of complex formation differ for gram-positive vs gram-negative infections.

In the case of the weaker *Ecc15* infection, the upswing of infection is relatively slower, with a sharp peak at ~8 hours and a drop to baseline levels at 16 hours post-infection, suggesting that the pathogen is no longer a threat.

What about the interaction of other members of the MARS complex, with EPRS, in response to infection? The data (Fig 2.6) suggests that subcomplex I (IRS, DRS, QRS, MRS and LRS) are enriched in the case of *M. Luteus* infection, suggesting the stabilisation of the complex.

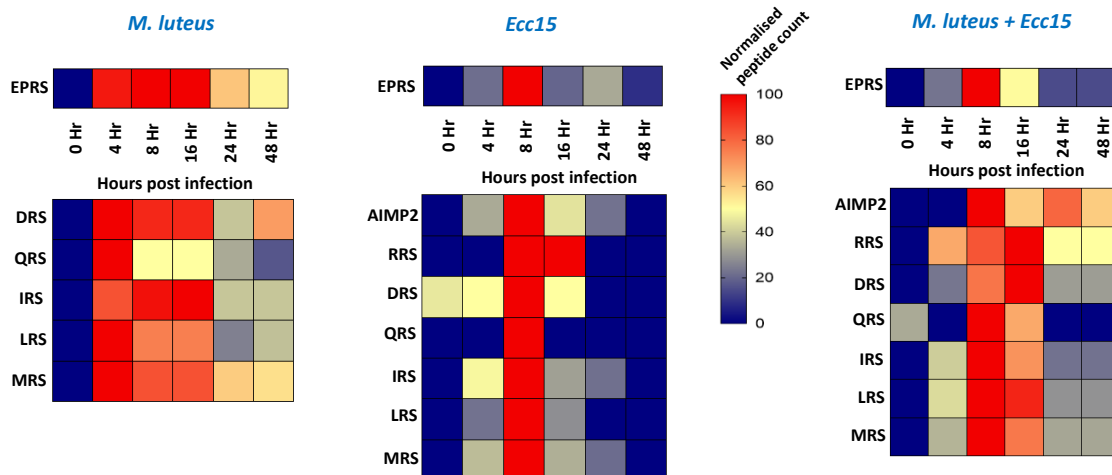


Figure 2.6: Kinetics of the MARS complex in response to infection. For *M. luteus*, the core members of the MARS complex are upregulated. The effects are weaker in the case of *Ecc15* infection. When both pathogens (*M. luteus*, *Ecc15*) are used for infection simultaneously, again, an upswing and downswing of complex formation is seen, using peptide counts as a readout.

For infection with *Ecc15*, post-infection, the data for complex formation is relatively weaker, with peptide counts falling by 50%. For a dual infection, the effects are intermediate.

Is the number of counts of peptide a reflection of increased transcription (and, therefore, translation)? This does not appear to be the case, as levels of mRNA for MARS complex genes remain stable (Fig 2.7) and are not upregulated at the peak of MARS complex formation.

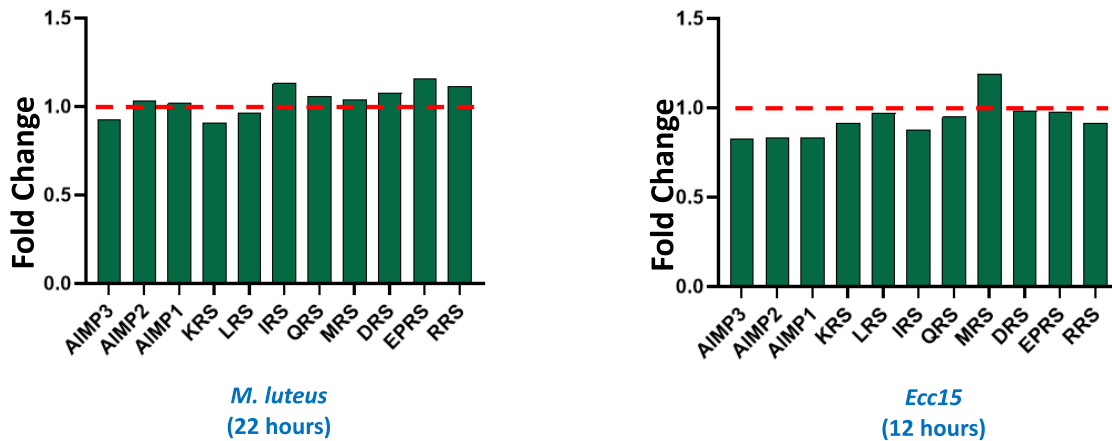


Figure 2.7: No change in transcript levels after infection for genes coding for proteins in the MARS complex. For both *M. luteus* & *Ecc15* infection, the levels of mRNA do not increase or decrease significantly, suggesting that the enrichment of the MARS complex is not a result of increased protein expression. The enrichment of the MARS complex is, in all probability, facilitated by post-translational modifications that mediate protein: protein interactions.

This suggests that the increase in enrichment of EPRS and its interactors is a result of infection-mediated signalling, with all probability that post-translational modification of EPRS or other members of the MARS complex is biasing the formation of the complex.

4.4 A GAIT-like complex is formed in response to infection: There has been no description of a GAIT complex in *Drosophila*. Here, we report that in agreement with the GAIT complex in humans, EPRS associates with GAPDH1, Rpl13a and Syncrip in response to infection (Fig 2.8).

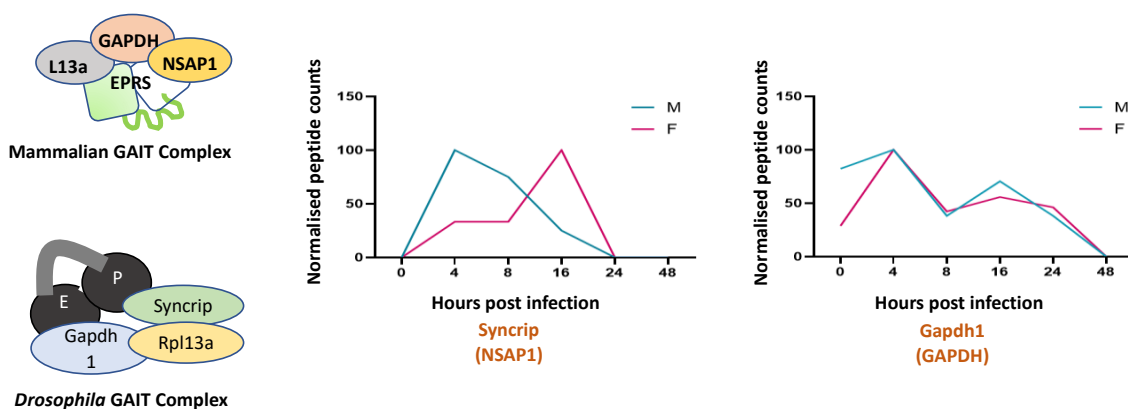


Figure 2.8: A GAIT-like complex in *Drosophila*. Post-infection, by monitoring the interaction of GAPDH and NSAP1/Syncrip, we monitored the possibility of the formation of a GAIT-like complex in response to infection. We find the association of both proteins increases by 4 hours post-infection

and reduces back to baseline by 48 hours post-infection. The kinetics appear to be different for males vs females.

Using Syncrip/NSAP1 as a reporter, we find that this protein is not associated with EPRS in the absence of infection, based on peptide counts in EPRS-enriched immune precipitates. Post-infection, as early as 4 hours, the Syncrip associates strongly with EPRS showing 40 (females) and 100 (males) peptide counts. A similar increase in GAPDH1 is seen post-infection. This suggests that a GAIT-like complex is formed in flies but only in response to infection.

4.5 Association/Dissociation of non-GAIT, non-MARS elements with EPRS: In addition, with the AARS and the AIMPs, a number of novel proteins are found to be associated with EPRS or its complexes in response to infection (Fig 2.9). Like other MARS interactors, the kinetics of the association with EPRS differs between males and females. These proteins with their functions are tabulated in panel B (Fig 2.9). None of these are listed as physical interactors under non-interactive conditions in Flybase, suggesting that the interactions are a specific response to immunity.

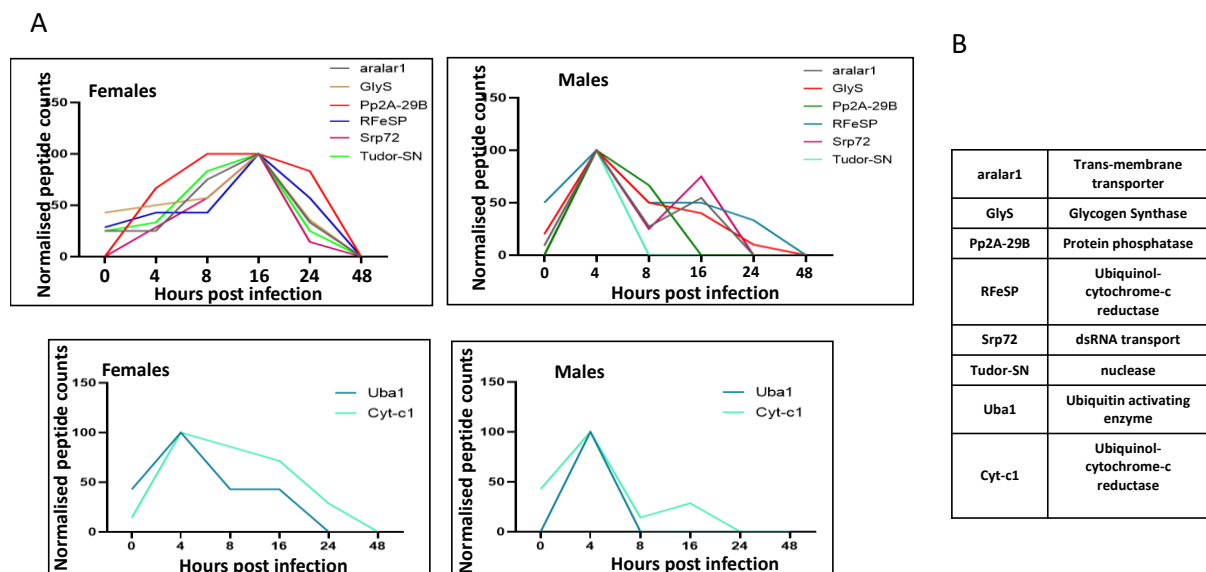


Figure 2.9: Non-GAIT, non-MARS players that interact with EPRS. A number of novel proteins that have earlier not been linked to EPRS (or the MARS complex) have been identified in our proteomics screen. These proteins engage with EPRS post-infection with differential kinetics (A), which vary between males and females. These proteins have varied functions in cell (B), suggesting connectivity between EPRS and the regulation of these functions.

Interestingly, Uba1 is one of the interactors enriched, indicating a possible role for SUMO conjugation of EPRS or other members of the MARS complex in response to infection.

Also interesting is a list of proteins that are interacting with EPRS (or one of its complexes) and which dissociate with infection. These include (Fig 2.10), Sod2, Ago1, Syt1, Pp1-87B and chic in females and Spn77Bc in males.

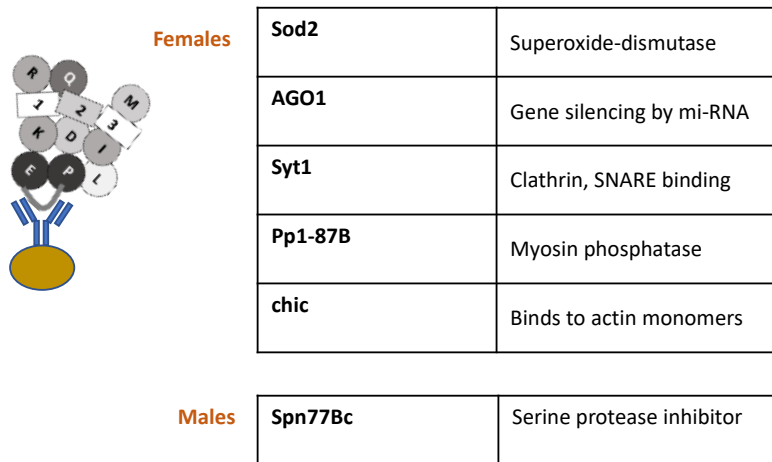


Figure 2.10: Dissociating partners in response to infection. A number of proteins cease their interaction with EPRS in response to infection. This suggests that under immune stress, EPRS steps away from a set of collaborative functions with these proteins that are resumed only after the host has secured itself against pathogenic attack.

The data of proteins that change their association status with EPRS in response to infection builds a picture of a dynamic EPRS protein:protein association landscape that is modified based on function. EPRS, either as part of the MARS/GAIT complex or on its own, appears to be collaborating in a number of cellular functions that are part of its moonlighting repertoire. Our data further highlight our limited understanding of the dynamic roles that EPRS plays within the cell.

Contributions & Acknowledgements

I thank Prof. Albert Courey, UCLA, for the gift of the *Drosophila* anti-EPRS antibody and the IISER Mass Spectrometry facility for access. Special thanks to Mr. Saddam at the LC-MS facility for data collection and also Kundan Kumar and Dr. Kamat for helping in the analysis of LC-MS data.

References

- Arif, Abul, Piyali Chatterjee, Robyn A. Moodt, and Paul L. Fox. 2012. "Heterotrimeric GAIT Complex Drives Transcript-Selective Translation Inhibition in Murine Macrophages." *Molecular and Cellular Biology* 32 (24): 5046–55.
- Arif, Abul, and Paul L. Fox. 2017. "Unexpected Metabolic Function of a tRNA Synthetase." *Cell Cycle* .
- Arif, Abul, Jie Jia, Robyn A. Moodt, Paul E. DiCorleto, and Paul L. Fox. 2011. "Phosphorylation of Glutamyl-Prolyl tRNA Synthetase by Cyclin-Dependent Kinase 5 Dictates Transcript-Selective Translational Control." *Proceedings of the National Academy of Sciences of the United States of America* 108 (4): 1415–20.
- Arif, Abul, Jie Jia, Rupak Mukhopadhyay, Belinda Willard, Michael Kinter, and Paul L. Fox. 2009. "Two-Site Phosphorylation of EPRS Coordinates Multimodal Regulation of Noncanonical Translational Control Activity." *Molecular Cell* 35 (2): 164–80.
- Arif, Abul, Peng Yao, Fulvia Terenzi, Jie Jia, Partho Sarothi Ray, and Paul L. Fox. 2018. "The GAIT Translational Control System." *Wiley Interdisciplinary Reviews. RNA* 9 (2). <https://doi.org/10.1002/wrna.1441>.
- Bannai, Hiroko, Kazumi Fukatsu, Akihiro Mizutani, Tohru Natsume, Shun-Ichiro Iemura, Tohru Ikegami, Takafumi Inoue, and Katsuhiko Mikoshiba. 2004. "An RNA-Interacting Protein, SYNCRIP (heterogeneous Nuclear Ribonuclear Protein Q1/NSAP1) Is a Component of mRNA Granule Transported with Inositol 1,4,5-Trisphosphate Receptor Type 1 mRNA in Neuronal Dendrites." *The Journal of Biological Chemistry* 279 (51): 53427–34.
- Blanc, Valerie, Naveenan Navaratnam, Jeffrey O. Henderson, Shrikant Anant, Susan Kennedy, Adam Jarmuz, James Scott, and Nicholas O. Davidson. 2001. "Identification of GRY-RBP as an Apolipoprotein B RNA-Binding Protein That Interacts with Both Apobec-1 and Apobec-1 Complementation Factor to Modulate C to U Editing." *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.m006435200>.
- Geuens, Thomas, Delphine Bouhy, and Vincent Timmerman. 2016. "The hnRNP Family: Insights into Their Role in Health and Disease." *Human Genetics* 135 (8): 851–67.
- Jia, Jie, Abul Arif, Partho S. Ray, and Paul L. Fox. 2008. "WHEP Domains Direct Noncanonical Function of Glutamyl-Prolyl tRNA Synthetase in Translational Control of Gene Expression." *Molecular Cell* 29 (6): 679–90.
- Kapasi, Purvi, Sujan Chaudhuri, Keyur Vyas, Diane Baus, Anton A. Komar, Paul L. Fox, William C. Merrick, and Barsanjit Mazumder. 2007. "L13a Blocks 48S Assembly: Role of a General Initiation Factor in mRNA-Specific Translational Control." *Molecular Cell* 25 (1): 113–26.
- Kim, Jung-Whan, and Chi V. Dang. 2005. "Multifaceted Roles of Glycolytic Enzymes." *Trends in Biochemical Sciences* 30 (3): 142–50.
- Lee, Eun-Young, Hyun-Cheol Lee, Hyun-Kwan Kim, Song Yee Jang, Seong-Jun Park, Yong-Hoon Kim, Jong Hwan Kim, et al. 2016. "Infection-Specific Phosphorylation of Glutamyl-Prolyl tRNA Synthetase Induces Antiviral Immunity." *Nature Immunology*. <https://doi.org/10.1038/ni.3542>.
- Mazumder, B., V. Seshadri, H. Imataka, N. Sonenberg, and P. L. Fox. 2001. "Translational Silencing of Ceruloplasmin Requires the Essential Elements of mRNA Circularisation: poly(A) Tail, poly(A)-Binding Protein, and Eukaryotic Translation Initiation Factor 4G." *Molecular and Cellular Biology* 21 (19): 6440–49.
- Mourelatos, Z., L. Abel, J. Yong, N. Kataoka, and G. Dreyfuss. 2001. "SMN Interacts with a Novel Family of hnRNP and Spliceosomal Proteins." *The EMBO Journal* 20 (19): 5443–52.
- Mukhopadhyay, Rupak, Jie Jia, Abul Arif, Partho Sarothi Ray, and Paul L. Fox. 2009. "The GAIT System: A Gatekeeper of Inflammatory Gene Expression." *Trends in Biochemical Sciences* 34 (7): 324–31.
- Mukhopadhyay, Rupak, Partho Sarothi Ray, Abul Arif, Anna K. Brady, Michael Kinter, and Paul L. Fox. 2008. "DAPK-ZIPK-L13a Axis Constitutes a Negative-Feedback Module Regulating Inflammatory Gene Expression." *Molecular Cell*. <https://doi.org/10.1016/j.molcel.2008.09.019>.
- Park, Sung Mi, Ki Young Paek, Ka Young Hong, Christopher J. Jang, Sungchan Cho, Ji Hoon Park, Jong Heon Kim, Eric Jan, and Sung Key Jang. 2011. "Translation-Competent 48S Complex Formation on HCV IRES Requires the RNA-Binding Protein NSAP1." *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkr509>.

- Ray, Partho Sarothi, James C. Sullivan, Jie Jia, John Francis, John R. Finnerty, and Paul L. Fox. 2011. "Evolution of Function of a Fused Metazoan tRNA Synthetase." *Molecular Biology and Evolution* 28 (1): 437–47.
- Seidler, Norbert W. 2012. *GAPDH: Biological Properties and Diversity*. Springer Science & Business Media.
- Singer, Dinah. 2008. "Faculty Opinions Recommendation of DAPK-ZIPK-L13a Axis Constitutes a Negative-Feedback Module Regulating Inflammatory Gene Expression." *Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature*. <https://doi.org/10.3410/f.1130849.587937>.
- Svitkin, Yuri V., Akiko Yanagiya, Alexey E. Karetnikov, Tommy Alain, Marc R. Fabian, Arkady Khoutorsky, Sandra Perreault, Ivan Topisirovic, and Nahum Sonenberg. 2013. "Control of Translation and miRNA-Dependent Repression by a Novel poly(A) Binding Protein, hnRNP-Q." *PLoS Biology* 11 (5): e1001564.
- Vyas, Keyur, Sujan Chaudhuri, Douglas W. Leaman, Anton A. Komar, Alla Musiyenko, Sailen Barik, and Barsanjit Mazumder. 2009. "Genome-Wide Polysome Profiling Reveals an Inflammation-Responsive Posttranscriptional Operon in Gamma Interferon-Activated Monocytes." *Molecular and Cellular Biology* 29 (2): 458–70.
- Wells, S. E., P. E. Hillner, R. D. Vale, and A. B. Sachs. 1998. "Circularisation of mRNA by Eukaryotic Translation Initiation Factors." *Molecular Cell* 2 (1): 135–40.
- Yao, Peng, Alka A. Potdar, Abul Arif, Partho Sarothi Ray, Rupak Mukhopadhyay, Belinda Willard, Yichi Xu, Jun Yan, Gerald M. Saidel, and Paul L. Fox. 2012. "Coding Region Polyadenylation Generates a Truncated tRNA Synthetase That Counters Translation Repression." *Cell* 149 (1): 88–100.
- You, Fuping, Hui Sun, Xiang Zhou, Wenxiang Sun, Shimin Liang, Zhonghe Zhai, and Zhengfan Jiang. 2009. "PCBP2 Mediates Degradation of the Adaptor MAVS via the HECT Ubiquitin Ligase AIP4." *Nature Immunology*. <https://doi.org/10.1038/ni.1815>.

CHAPTER III

SUMOylation of Arginyl tRNA Synthetase Modulates the *Drosophila* Innate Immune Response

Summary

SUMO conjugation of a substrate protein can modify its activity, localization, interaction, or function. A large number of SUMO targets in cells have been identified by Proteomics, but biological roles for the need for SUMO conjugation for most targets remain elusive.

In this study, I find that Arginyl tRNA Synthetase (RRS), an enzyme involved in charging Arginine to its cognate tRNA, is SUMO conjugated. The sites for SUMO conjugation are Lys 383 & 579. Replacement of these residues by Arg (RRS^{K383R, K579R}) creates a SUMO conjugation-resistant variant (RRS^{SCR}) of RRS. Further, to understand biological roles for SUMO conjugation of RRS, we have generated an RRS^{null} animal using CRISPR/Cas9 genome editing and have rescued the lethality of the RRS^{null} by expressing either RRS^{WT} or RRSSCR ubiquitously using the UAS-Gal4 system. Adult animals expressing RRS^{WT} and RRS^{SCR} are compared and contrasted for physiological changes.

RRS^{SCR} animals, when compared to RRS^{WT}, show increased inflammation on infection, as measured by the activation of anti-microbial genes using Quantitative RNA sequencing. Also, the assembly of the 1.2 MDa supramolecular cytoplasmic tRNA synthetase complex (MARS), of which RRS is a member, is enhanced on infection. Other members of the MARS complex, such as Glutamyl-Prolyl tRNA Synthetase also show SUMO conjugation on infection.

Our study thus uncovers a noncanonical role for RRS. Infection in *Drosophila* leads to SUMO conjugation of a subset of tRNA synthetases in the MARS complex, with SUMO conjugation of RRS leading to attenuation of transcription of anti-microbial defence genes.

1. Introduction

Aminoacyl-tRNA synthetases (ARSs) are ancient, evolutionarily conserved enzymes whose primary housekeeping function is to catalyse the aminoacylation of transfer RNAs (tRNAs) (Schimmel and Soll, 1979; Rubio Gomez and Ibba, 2020). In addition to their primary role of charging tRNA, ARSs also have noncanonical, ‘moonlighting’ functions (Guo and Schimmel, 2013; Yao et al., 2014). These secondary functions are driven by modifications to

the polypeptide chain by mutations, domain addition, or Post-Translational modifiers (PTMs)(Sampath et al., 2004). ARSs are a target of a variety of PTMs, with phosphorylation being studied extensively (Arif et al., 2017). The small ubiquitin-like modifier (SUMO; (Hay, 2005; Geiss-Friedlander and Melchior, 2007)) is one such PTM that targets ARSs. Proteomic studies on a wide range of eukaryotes have suggested (Panse et al., 2004; Golebiowski et al., 2009; Nie et al., 2009; Pirone et al., 2017; Hendriks and Vertegaal, 2016) that at least fourteen of the twenty ARSs are SUMO conjugated (SUMOylated) (**Table 3.1**).

In mammals, nine of the tRNA synthetases (Glu-Pro, Ile, Leu, Met, Gln, Lys, Arg, Asp) are part of a ~1.2 MDa Multi-acyl tRNA Synthetase (MARS) complex, along with three non-ARS components (AIMP1-3) (Khan et al., 2020). In addition to acting as a ‘depot’ or reservoir for tRNA synthetases and facilitating related translational functions, the release of individual components in response to stimulus, both internal and external, regulates the noncanonical functions of these proteins, inclusive of the AIMP. The released components can be secreted or relocated to a different cellular compartment (Ray and Fox, TIBS, 2007) (Park SG, Kim 2008, PNAS). The MARS complex is now perceived as a hub for many signalling networks within the cell (Park SG, Kim 2008, PNAS). The MARS complex is conserved from insects to mammals, with the *Drosophila* MARS complex (Kerjan et al., 1994; Havrylenko and Mirande, 2015) containing orthologs of the 11 components seen in mammals.

In an experiment to uncover proteins that are SUMO conjugated in response to infection, our laboratory identified 12 ARSs as potential targets using a quantitative proteomics screen (Handu et al., 2015). The study suggested that SUMOylation of ARSs was a response to immune signaling. Using an *in-bacto* SUMO conjugation assay (Nie et al., 2009), we validated a subset of *Drosophila* ARSs as being SUMOylated. Next, we focused our attention on one substrate, namely Arginyl tRNA synthetase (RRS). We determined that K147 and K383 in RRS were the targets of the SUMO machinery and generated transgenic wildtype and SUMO conjugation resistant (SCR) transgenic lines for RRS using a combination of CRISPR Cas9 genome editing and UAS-Gal4 system. A comparison of the transcriptome of *RRS^{WT}* versus *RRS^{SCR}* adult flies, in response to both gram-positive and gram-negative infection, led us to suggest that SUMOylation of RRS could modulate the host-defense response in *Drosophila*.

2. Materials & Methods

2.1 SUMO conjugation assay: SUMOylation of constituents of the MARS complex was tested by expressing the target/substrate protein simultaneously with the *Drosophila* SUMO cycle components based on a published protocol (Nie et al., 2019). Target proteins from the *Drosophila* Gold cDNA collection, procured from the *Drosophila* Genome Resource Centre (DGRC), Bloomington, Indiana, were sub-cloned into *pGEX-4T1* (Promega) and *pET-45b* and subsequently sequenced for validation. For visualization of SUMO conjugation, bacterial lysates were affinity purified using Glutathione-Agarose beads (Invitrogen) or Ni NTA-Agarose beads (Qiagen), run on an SDS-PAGE gel and monitored using mouse anti-GST antibody (sc53909, 1:5000; Santa-Cruz-Biotechnology), Rabbit anti-HA antibody (DW2, 1:3000; Millipore) and mouse anti-6X-His antibody (H1029, 1:1000; SIGMA) using Western blotting. The SUMO-conjugated forms appear as bands of a higher molecular weight.

2.2 SUMO-binding-motif and SIM-motif prediction: Putative SUMO acceptor lysines and SIM-motifs of all the MARS complex components of *Drosophila* were predicted *in-silico*, using Joined Advanced SUMOylation and Sim motif Analyzer (JASSA) tool with cut-off threshold criteria set at “high” (Beauclair et al., 2015).

2.3 Identification of evolutionarily conserved SUMO target lysine residues *in-silico*: FASTA sequences of RRS for model organisms belonging to different eukaryotic groups were procured from the UniProt protein database. Multiple sequence alignment (MSA) was done on the basis of homology extension using PSI-COFFEE (Chang et al., 2012). SUMO acceptor lysines were compared across different representative organisms post-alignment.

2.4 Homology model for *Drosophila* RRS: The automated SWISS-MODEL server (Waterhouse et al., 2018) was used to generate structural models (RRS^{WT}, Δ RRS) using default parameters. The human 4Q2T PDB structure (Kim et al., 2014a), solved at a resolution of 2.4 Å, containing a bound Arginine at the active site, was used as a template.

2.5 Generation of Δ RRS using CRISPR Cas9 technology: CRISPR Cas9 technology was employed to generate RRS null fly lines. Single guide (sg)-RNAs targeting the RRS coding region in the 5'UTR and Exon-5 were designed using CRISPR Optimal Target Finder (COTF;(Gratz et al., 2014)), a web tool for identifying CRISPR target sites and evaluating their specificity. The RRS gene region was sequenced prior to the experiment to design the gRNAs to account for SNPs at the sgRNA target sites. The sgRNAs were cloned into the *pU6-BbsI-*

chiRNA (Addgene # 45946) plasmid, which was then docked into $y^l v^l$; $P\{CaryP\}attP40$ *Drosophila* line (BDSC 36304), by transgenic injections, at the NCBS-CCAMP transgenic facility, Bangalore, India. The transgenic dual sgRNA line was crossed to the *nanos-Cas9* (BDSC 54591) line. The founder male progenies obtained were crossed to w^- ; *FM7a* balancer females wherein the Cas9-sgRNA complex is formed in the germline. In the next generation, three heterozygous female progenies from each cross (60 lines, each labelled A, B, and C) were maintained as a separate line over a *FM7a* balancer. Since the genomic *RRS* is located on the X chromosome, putative *RRS* null lines were screened for male lethality. Lines showing male lethality were chosen for PCR-based confirmation of the deletion. Single-fly genomic PCR for the extended gene region of *RRS* was performed on heterozygous females, and the mutations were confirmed through sequencing.

2.6 pUASp AttB fly lines/strains: *RRS-WT* and *RRS-SCR(K147,383→R)* were sub-cloned into *pUASp-attP2* using a homology-based recombination technique, a modification of the SLiCE protocol (Zhang et al., 2014). These were injected into *AttB* lines for the generation of transgenic fly lines. Fly lines were balanced with ubiquitously expressing Gal4s (*Actin-Gal4/Ubiquitin-Gal4*) of the following genotype *Actin-Gal4/+;UAS-RRS-WT/+*, *Actin-Gal4/+;UAS-RRS-SCR(2MT)*, *Ubiquitin-gal4/+;UAS-RRS-WT/+* and *Ubiquitin-Gal4/+;UAS-SCR(2MT)*. All experiments were carried out with the *Actin-Gal4* line.

Culturing and processing bacteria for infections: *M. luteus* and *Ecc15* were plated on Luria-Bertani (LB) agar plates and grown in LB broth under antibiotic selection. Bacteria were collected from the plate or pelleted and re-suspended in 1X PBS to make a concentrated solution.

2.7 Fly infections: 6–8-day-old males were collected and placed at 29 °C for 48 hours to acclimatize the flies to infection temperature. To cause septic injury, flies were pricked in the thorax with a needle dipped in the concentrated solution of bacteria. To activate the Toll pathway and Imd-pathway, flies were infected with *M. luteus* and *Ecc15*, respectively, at an Optical density of 100, measured at 600 nm. To measure gene expression levels, infected flies and non-infected controls were incubated at 29 °C for the time required, after which they were collected by snap-freezing them in liquid nitrogen and stored at –80°C until RNA extraction. Infectivity assays were done in three biological replicates, ten flies per replicate. For survival experiments, flies were pricked in the same way as for the gene expression measurements.

2.8 Total RNA extraction cDNA library construction and sequencing: Total RNA was extracted from adult flies with the following genotypes $\Delta RRS/Y$; *Actin Gal4/+*; *UAS-RRS* *WT/+* and $\Delta RRS/Y$; *Actin Gal4/+*; *UAS-RRS* *SCR/+*, 10 days post eclosion, in triplicates using RNeasy Plus Universal Kits (Qiagen; Part #74104) under control and infected conditions, according to manufacturer's instructions and RNA integrity was assessed. 3' mRNA-specific libraries were amplified using QuantSeq 3' mRNA-Seq Library Prep Kit FWD using the manufacturer's instructions. Quality assessment for the cDNA libraries was done using Bioanalyzer 2100 (Agilent Technologies). Single-end 75bp sequencing of the pooled libraries was performed on the Illumina NextSeq 500 platform.

2.9 Demultiplexing, adapter trimming, read mapping, counts generation, and differential expression analysis: On average, 4-5 million reads were generated per sample. The raw reads were demultiplexed using bcl2fastq, and the adapters were trimmed using bbduk v35.92. Sequencing quality was assessed using FastQC v0.11.5. Post quality control, the reads were mapped to the *Drosophila* genome (dm6) using STAR aligner v.2.5.2a (Dobin et al., 2013). Gene expression levels were measured using the counts generated by HTSeq-count v 0.6.0 (Anders et al., 2015). The gene expression counts were normalized for all samples together, and the biological conditions were compared pairwise using DESeq2 (Love et al., 2014). The Principle Component Analysis using the 'R' package of the regularized log counts was used to remove outliers from the final differential expression analysis. The regularized log-transformed counts of the transcripts from DESeq2 were used to determine upregulated and downregulated genes across biological samples. Genes with $\log_2(\text{FC})$ values ≥ 0.55 , ≤ -0.55 and $-\log_{10}(\text{FDR})$ values ≥ 2 were considered for further analysis. Gene Ontology analysis was done using a subroutine in Flybase.

Custom Venn diagrams were made using the Venneuler package in R to show the overlap and differences between the differentially expressed gene lists. Volcano plots were made using GraphPad Prism 8.0.2 for visual identification of genes with large fold changes that are also statistically significant.

2.10 Survival Analysis: Survival assays were carried out on *RRS*^{WT} and *RRS*^{SCR} flies. For each experiment, flies were infected 10 days post-eclosion. ~40 age-matched male flies of the desired genotype were collected, each vial containing 10 flies. Animals were flipped to a fresh vial every 5 days, with the number of flies recorded per vial daily. The survival data were plotted and analysed using the log-rank test in Prism 8.

Real Time-PCR: mRNA was extracted from 10-day-old adults post-infection using Qiagen RNeasy mini kit (74104). 500 ng of RNA was used for the cDNA synthesis using the High Capacity cDNA Reverse Transcriptase Kit (4368814) by Applied Biosystems. The qPCR reaction was carried out using KAPA SYBR FAST (KK4602) Sigma using Analytik Jena - qTOWER³ – Real-Time PCR Thermal Cycler. The experiments were carried out in triplicates with two technical replicates each. The relative fold change for each genotype was calculated by normalising it to housekeeping gene rp49. The data was analysed by Two-way ANOVA followed by Tukey's test for multiple comparisons. The primer pairs used are listed in the Resource Table (Resource Table).

3. Results

3.1 The MARS Complex is a target for SUMO machinery. Proteomics studies in a host of organisms suggest that members of the MARS Complex are SUMOylation targets (Table 3.1, Fig. 3.1) (Panse et al., 2004; Tatham et al., 2011; Handu et al., 2015), including studies in *Drosophila* (Handu et al., 2015; Pirone et al., 2017). Handu *et al.*, 2015 specifically enriched proteins that changed their SUMOylation status in response to a broad activation of immune pathways, with ARSs being significant targets. As a first step to validate the targets, we cloned members of the *Drosophila* MARS complex (Lu et al., 2015) into bacterial expression vectors and screened their ability to be SUMOylated in an *in-bacto* system (Nie et al., 2009), which uses *Drosophila* enzymes expressed in bacteria for SUMO conjugation. We find that five ARSs, EPRS, RRS, KRS, DRS, and one AIMP (AIMP1), were modified by SUMO (Fig. 3.1A-F).

Table 3.1: SUMO conjugated proteins based on Proteomic studies. Data in the literature suggests that elements of the MARS complex are SUMO Conjugated. The table compiles studies from yeast, flies, and mammals. Cells marked with grey indicate that the protein in the corresponding row is SUMO conjugated, based on Proteomics data.

		Yeast	Insect (<i>Drosophila</i>)			Mammals
			0-3hr embryos	S2 cells	S2 cells	Hela and U2OS cells
NAME		<i>Panse et al., 2004</i>	<i>Nie et al., 2009</i>	<i>Handu et al., 2015</i>	<i>Pirone et al., 2017</i>	<i>Hendricks et al., 2016</i>
MARS (tRNA Synthetases)	Lysyl-tRNA synthetase					
	Leucyl-tRNA synthetase					
	Arginyl-tRNA synthetase					
	Aspartyl-tRNA synthetase					
	Glutamyl-tRNA synthetase					
	Glutamyl-prolyl-tRNA synthetase					
	Isoleucyl-tRNA synthetase					
	Methionyl-tRNA synthetase					
MARS (AIMPs)	aaRS-interacting multifunctional protein 1					
	aaRS-interacting multifunctional protein 2					
	aaRS-interacting multifunctional protein 3					
Non-MARS (tRNA Synthetases)	Glycyl-tRNA synthetase					
	Histidyl-tRNA synthetase					
	Threonyl-tRNA synthetase					
	Alanyl-tRNA synthetase					
	Asparaginyl-tRNA synthetase					
	Cysteinyl-tRNA synthetase					
	Phenylalanyl-tRNA synthetase, alpha-subunit					
	Phenylalanyl-tRNA synthetase, beta-subunit					
	Seryl-tRNA synthetase					
	Tryptophanyl-tRNA synthetase					
	Tyrosyl-tRNA synthetase					
Valyl-tRNA synthetase						

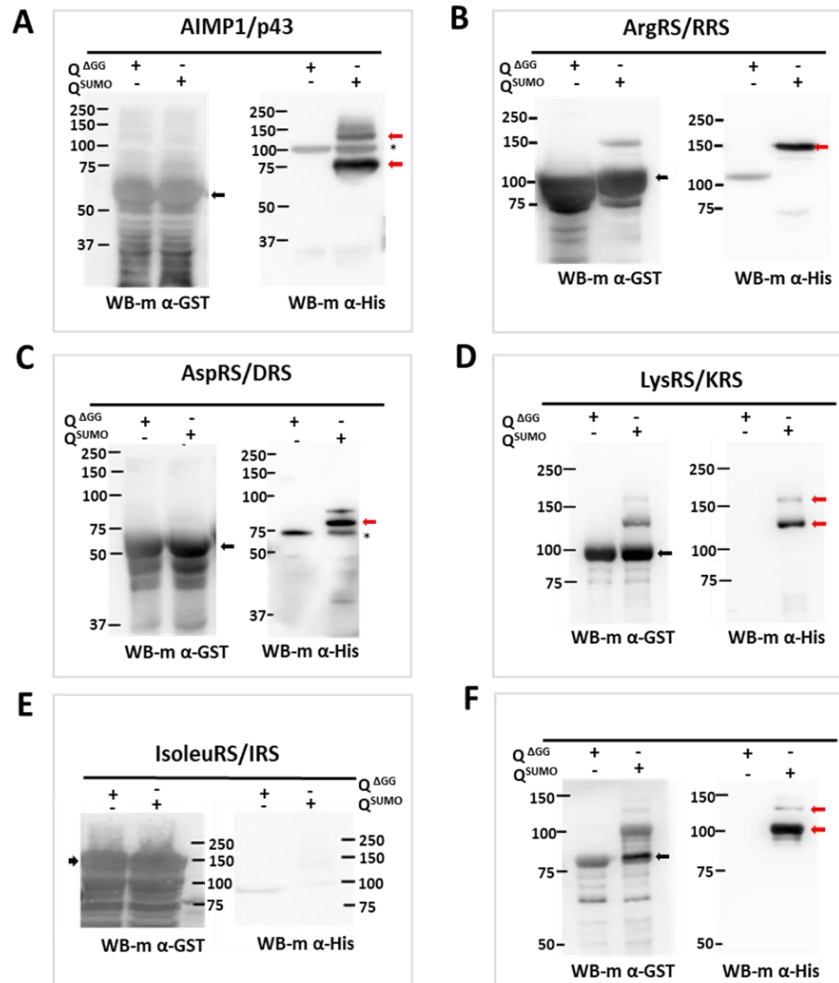


Figure 3.1: *Drosophila* MARS Complex is a target of SUMO conjugation machinery. A-F. Validation of SUMOylation by *in-bacto* SUMO conjugation. Genes coding for *Drosophila* MARS components were cloned (see Materials and Methods) and tested for SUMO conjugation using the *in-bacto* SUMOylation (Nie et. al., 2009). Here, the proteins to be tested are fused with GST and expressed in bacteria along with *Drosophila* SAE1/SAE2, Ubc9, and the matured form of SUMO, SUMO-GG (Nie et. al., 2009). AIMP1, RRS, DRS, and KRS were SUMO conjugated (panel A-D), while IRS did not show conjugation (panel E). The WHEP domain of EPRS was expressed to demonstrate the SUMOylation of EPRS (panel F). The black and red arrows denote the bands corresponding to non-SUMOylated (unmodified version) and SUMOylated species of the concerned protein. ‘*’ indicates a band corresponding to non-specific proteins.

MRS and LRS could not be expressed, while IRS was expressed and not SUMO conjugated. The SUMOylation status for QRS, AIMP2, and 3 was inconclusive due to low protein expression and high background in western blots. Of these, we choose RRS as a target to characterize, it being an understudied target showing robust SUMOylation.

RRS is a ubiquitous, cytoplasmic, Class I Aminoacyl tRNA Synthetase. It has three major domains an N-terminal coiled-coil leucine zipper domain, a central synthetase catalytic core domain, and a C-terminal all alpha-helical anticodon binding domain called DALR

domain. There is only one protein-coding transcript and only one polypeptide associated with the gene in flies, unlike in higher organisms like mammals, where the *RRS* gene codes for two isoforms via alternative initiation. The full-length complexed form provides arginyl tRNA for protein synthesis, and the free form, provides arginyl tRNA for N-terminal arginine modification (Arginylation) of proteins with the aid of protein Arginyl transferase. Prediction of SUMO conjugation sites (Beauclair et al., 2015) in the RRS sequence suggests that RRS has a strong consensus SUMO conjugation motif at K383. Our experimental data suggested that RRS can show up to two SUMO conjugates (Fig. 3.2A-B), and multiple rounds of mutagenesis followed by *in-bacto* SUMOylation led to the finding that a mutant RRS^{K147R, K383R} is SUMO conjugation resistant (RRS^{SCR}) (Fig. 3.2B).

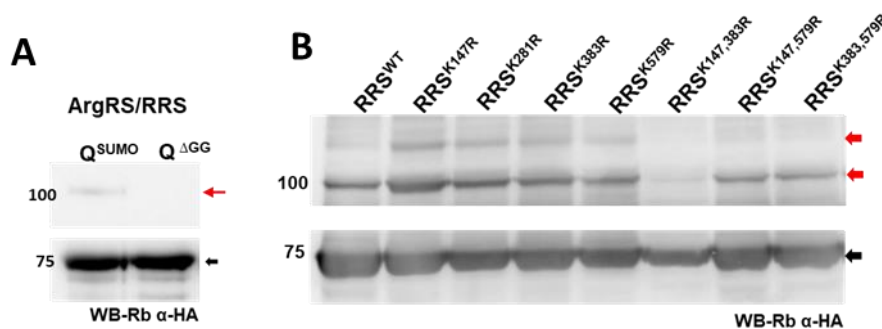


Figure 3.2: A. *RRS* is SUMO conjugated. When co-expressed with *Drosophila* E1, E2, and 6X-His:SUMO, RRS shows a single extra band running ~20 KDa higher than RRS itself in Western blots. The band also cross-reacts with an anti-His antibody (data not shown), confirming that it represents a SUMO-conjugated species. The band is not seen when a SUMO(Δ GG) variant, which is unable to conjugate to a substrate, is used. A second faint band seen in overexposed Western blots suggests that RRS may have a second SUMOylation site. **B. *RRS* is SUMO conjugated at K147 and K383.** Based on predictions of SUMO conjugation sites from JASSA (Beauclair et al., 2015), mutagenesis of four lysines was carried out one at a time. None of the single mutants showed loss of SUMOylation. Amongst double mutants, RRS^{K147R, K383R} double mutant was resistant to SUMO conjugation. The black and red arrows denote the bands corresponding to non-SUMOylated (unmodified version) and SUMOylated species of the concerned protein, respectively.

RRS is part of subcomplex-II (Fig. 3.3) in the MARS complex, associating intimately with QRS and AIMP1. Analysis of the crystal structure of sub-complex-II suggests that the equivalent amino acids in the human structure (Fig. 3.4A, 4R3Z, (Fu et al., 2014)) are not part of the protein: protein interface with either QRS or AIMP1.

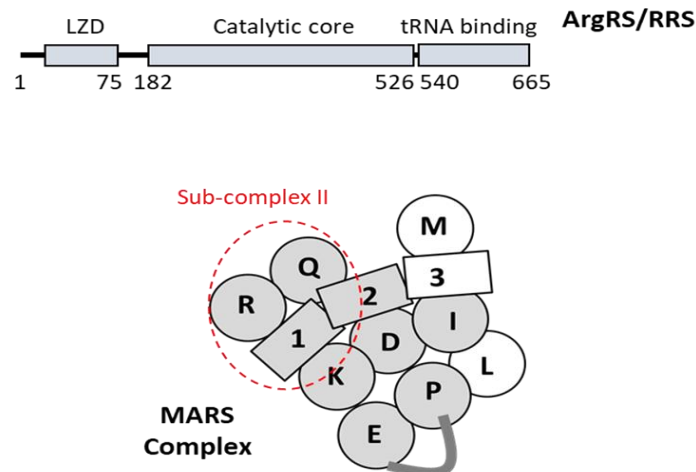


Figure 3.3: Schematic of the RRS and the MARS complex. *Drosophila* RRS consists of a Leucine zipper domain (LZD), a catalytic core, and C-terminal tRNA binding domain. In the MARS schematic, the ARSs are labelled with a single letter code, with the grey shading denoting mass spectrometric or *in-bacto* evidence for SUMO conjugation (Fig. 3.1). RRS is part of sub-complex II (marked with red dashed line), interacting with QRS and AIMPI.

We generated a structural model (Fig. 3.4B) of RRS using the automated SWISS-MODEL server (Waterhouse et al., 2018), using the 4Q2T PDB structure (Kim et al., 2014a) as a homology model and mapped the two conjugation sites onto the fly model (Fig. 3.4C). K147 is part of a low-scoring SUMO target motif (LKGH), at the end of a predicted helix, in a region that is not conserved (Fig. 3.4B-C). K383 is part of a high-scoring SUMO consensus motif (VKSD) in a conserved loop near the Arginine bound active site. The nearest residue which interacts with the bound Arg is F388.

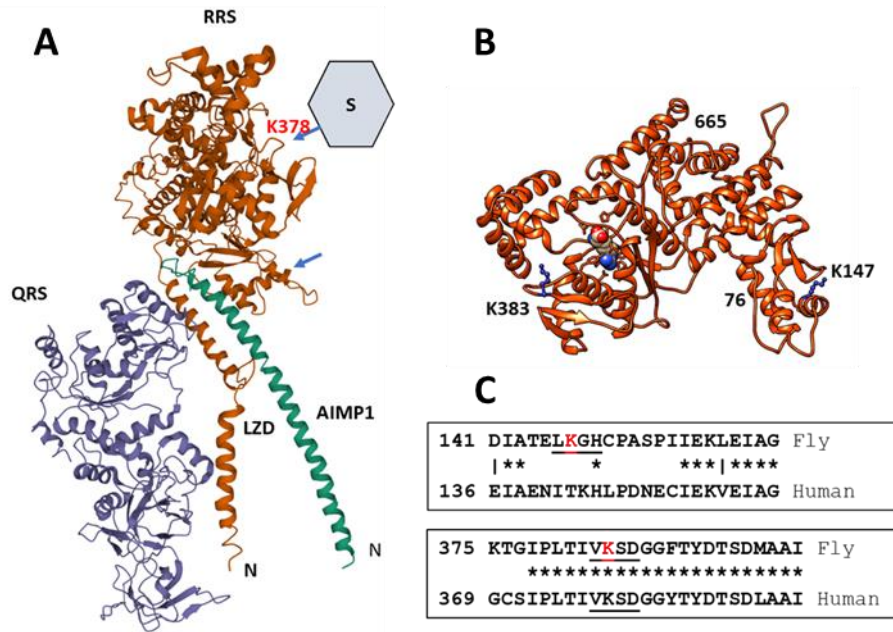


Figure 3.4: A. Schematic of the structure for the human QRS:RRS: AIMP1 complex (PDB-ID 4R3Z). The SUMOylation sites in the fly RRS were mapped to the human RRS structure after sequence alignment. The SUMOylation sites were distant from the binding regions of both QRS and AIMP1 and did not appear to interact with any component of MARS, based on current structural models (Khan et al., 2020). The fly K383 equivalent in humans, K378, is in a loop region. (arrow) and is a predicted SUMO conjugation site. **B. Homology model of *Drosophila* RRS.** A homology model of fly RRS, based on the Arg bound 4Q2T structure as a template. The structure includes residues 76-665 but not the N-terminal LZD (1-75). K383 is in a loop outside the Arg binding site, while K147 is at the end of a helix. **C. SUMO conjugation site is conserved from flies to mammals.** SUMO conjugation sites (K147, K383) for the fly RRS are underlined, with the target Lys marked in red. Based on the sequence alignment of fly and human RRS, the K383 site is in an evolutionarily conserved region, while the K147 is not.

3.2 Generation of a Δ RRS line using CRISPR Cas9 genome editing. The UAS-Gal4 system is an ideal system to express RRS^{WT} and RRS^{SCR} in an RRS-null (Δ RRS) background. Since the Δ RRS line is not available, as a first step, we used CRISPR Cas9 genome editing to generate the same. A transgenic dual-guide RNA line (*UAS-RRS^{dual-gRNA}*) was created (See Materials and Methods) to express *gRNA* that would recognize the 5' UTR and 3' end of the coding region of the RRS gene (inverted red triangles, Fig. 3.5A), near the translation start and stop sites. Our goal was to remove a major portion of the coding region to create a Δ RRS animal.

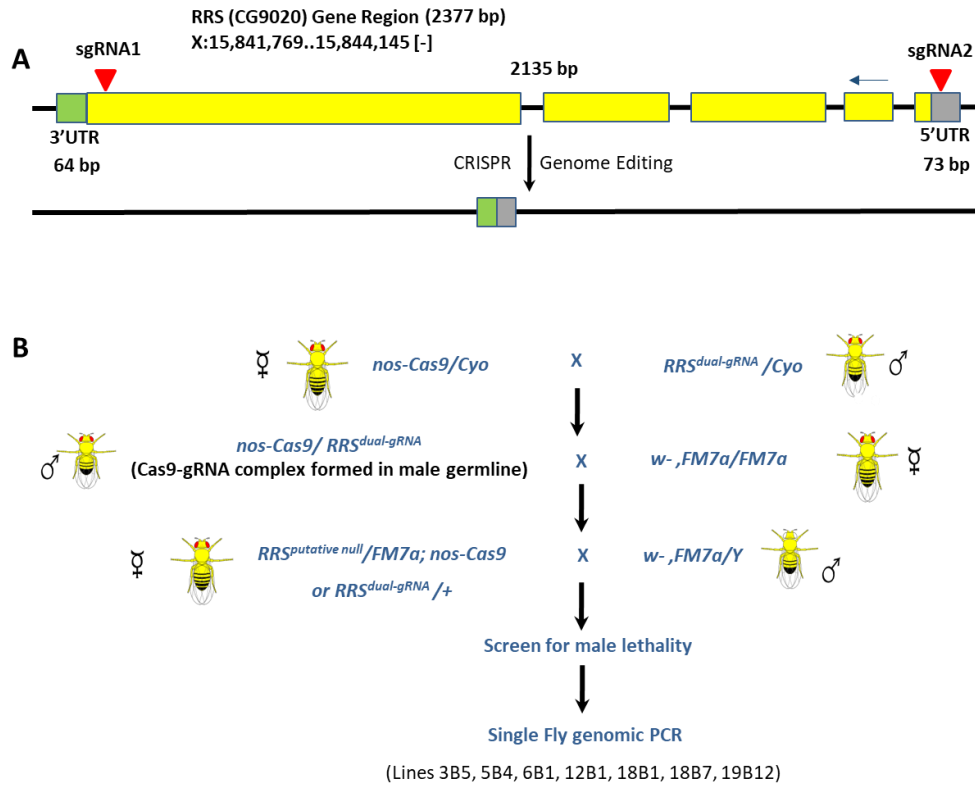


Figure 3.5: An RRS-null (ΔRRS) line generated using CRISPR Cas9 genome editing. A. Design of the dual guide-RNA for excision of the RRS locus. The RRS gene is located on the X chromosome. It has five exons, which code for a single annotated transcript that spans 2377 bp. Two gRNAs (inverted red triangles) were designed in the 5' UTR and 3' end of the coding region. Our goal was to excise most of the coding region and generate a RRS null line. B. Excision of the RRS locus to generate a ΔRRS line. A *UAS-RRS^{dual-gRNA}* line was generated (Materials & Methods) and crossed to *nos-Cas9* animals. Sixty lines were balanced over a first chromosome balancer and screened for male lethality. None of the seven male lethals had the expected deletion in the RRS locus, based on PCR.

The *UAS-RRS^{dual-gRNA}* line was crossed to a *nos-Cas9* animal (Fig. 3.5B), and sixty lines stabilized by balancing the putative nulls over an X chromosome balancer, FM7i where the balancer chromosome expresses GFP. Of these lines, seven were male lethal, which was indicative of a successful excision of the RRS locus since the absence of the RRS on the X chromosome would lead to lethality. Single-fly genomic PCRs were conducted on these lines, but the genomic PCR products did not show the expected 2.1 kb deletion that would be a consequence of the removal of the RRS genomic region. To probe the observed male lethality, we sequenced the genomic region of two lines, 6B1 and 18B1. To our surprise, we found that even though the coding region was not deleted, the gRNA activity caused changes to the sequence of the wildtype genome in the sites targeted by both gRNA (Fig. 3.6A-B), and these modifications presumably led to the generation of a functional RRS-null (ΔRRS). ΔRRS^{6B1} has

a 13 bp deletion in the 5'UTR region (Fig. 3.6A-B, Fig. A.3.1A), while in the case of ΔRRS^{18B1} , there appeared to be an 11 bp insertion in the same region (Fig. A.3.1B). In both cases, the 5'UTR is disrupted (Fig. 3.6A-B, Fig. A.3.1). The 5'UTR serves as the entry point for the ribosome during translation and can adopt elaborate RNA secondary and tertiary structures that may regulate translation initiation (Curran and Weiss, 2016; Leppik et al., 2018). To test the stability and/or expression of the transcripts, we measured mRNA levels using quantitative real-time PCR (qRT-PCR) in 1st Instar larvae. ΔRRS die during IInd instar larval stages, with embryonic survival till 1st Instar presumably driven by maternal RRS. ΔRRS homozygous larvae, identified by their lack of GFP fluorescence, show a 40% reduction in *RRS* transcripts as compared to *wt* (Fig. 3.9A-B). We believe that maternal RNA still perdures at this stage, and reduces as the animals transit to the 2nd Instar. The transcript levels measured are thus a sum of maternal and zygotic RNA.

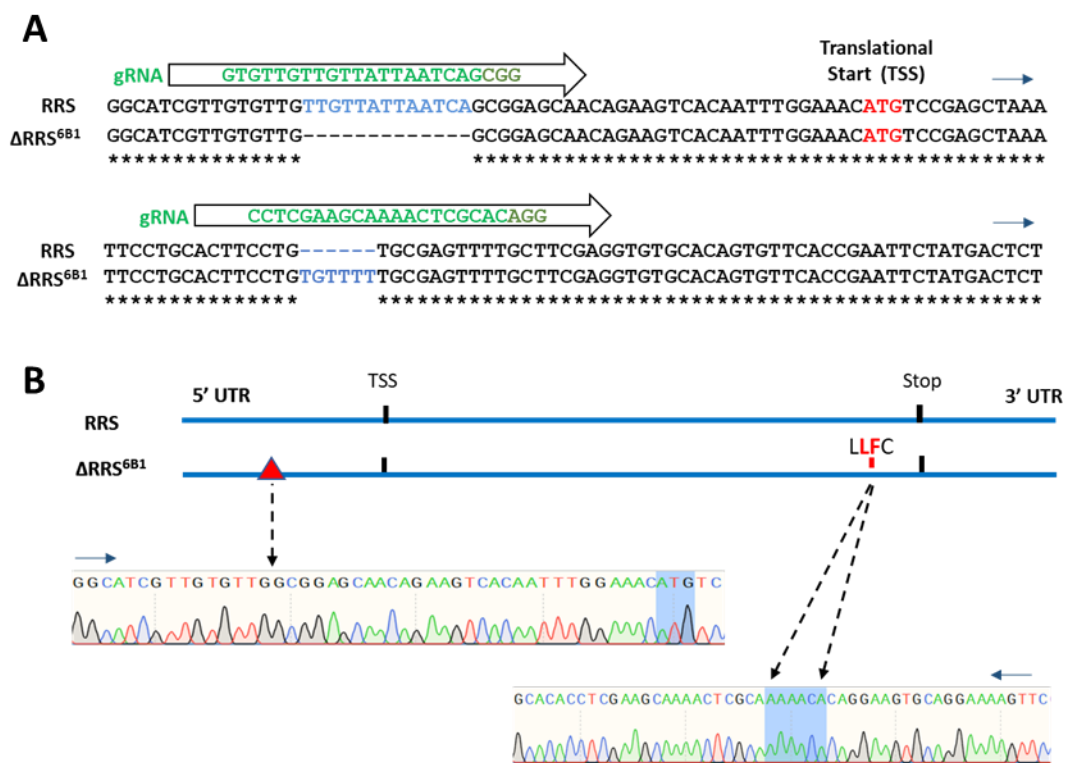


Figure 3.6: A. ΔRRS lines generated by disruption of the 5'UTR. Genomic DNA Sequencing of the 6B1 line in the *RRS* genomic region reveals deletions/insertions in the *RRS* locus at the gRNA binding site(s). A 13 bp deletion is seen in the 5'UTR and a 6 bp insertion near the 3' end of the coding region. **B. Schematic of the mutations in the ΔRRS^{6B1} line.** Schematic showing the deletions near the translation start site and insertion in the coding region near the translational stop site. DNA sequences for each perturbation are also shown.

Sequence changes in the coding region were also seen in both lines (Fig. A.3.1). For ΔRRS^{6B1} , a 6 bp insert would lead to the incorporation of a Leu and Phe (Fig. 3.6A-B, Fig. 3.7, Fig. A.3.1) in positions 604 and 605, within the RRS sequence; For ΔRRS^{18B1} , the sequence corresponding to the C-terminal domain could not be elucidated in spite of multiple sequencing attempts (Fig. A.3.1B). For ΔRRS^{6B1} , the insertion may perturb the structure of the C-terminal tRNA binding domain. One possible scenario is the disruption of the predicted (Craig and Dombkowski, 2013) C515:C604 disulfide bond (Fig. 3.7) in the *Drosophila* structural model, which could lead to significant destabilization of ΔRRS^{6B1} and make it a functional null. The RRS^{6B1} line (ΔRRS) with defined mutations in the 5' UTR and coding region and with homozygotes dying in the 1st to 2nd Instar transition was used for all further experiments. The $\Delta RRS/+$ lines are haplo-sufficient, showing normal lifespan at 25 and 29 °C (Fig. 3.8), and do not show any embryonic or larval lethality.

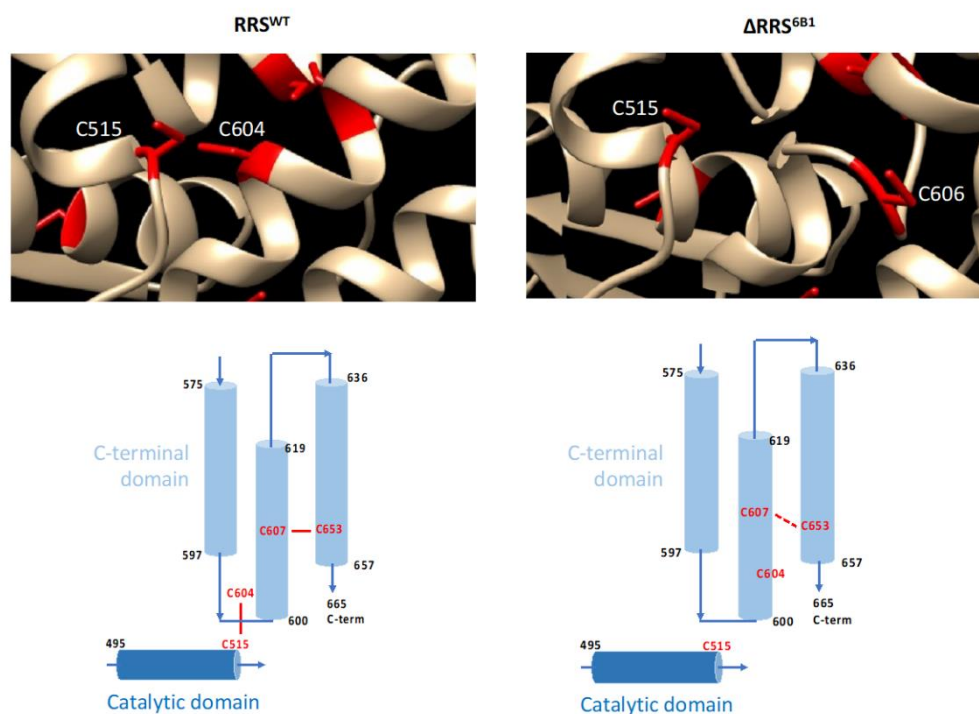


Figure 3.7: Structural consequences of CRISPR/Cas9 mediated insertions in ΔRRS line 6B1. 1-2. The *Drosophila* RRS structure, modelled by SWISS-MODEL, predicts a potential disulfide bond between C515:C604, two alpha helices that connect the catalytic and C-terminal tRNA binding domains. The prediction was based on analysis by the Disulfide by Designs server (<http://cptweb.cpt.wayne.edu/DbD2/>) (Craig et. al., 2013). This disulfide, if present, could be a significant factor in the folding and stability of RRS. A second potential disulfide is between C607:C653 (panel B2). 3-4. For line 6B1, An insertion of two amino acids, Leucine and Phenylalanine, in positions 604 and 605 would cause a shift in the Cystine (604 in wt) to position 606 on the opposite face of the helix, which would lead to a disruption of the disulfide bond. The insertion could also perturb the side-chain interactions and further add to the destabilization of RRS.

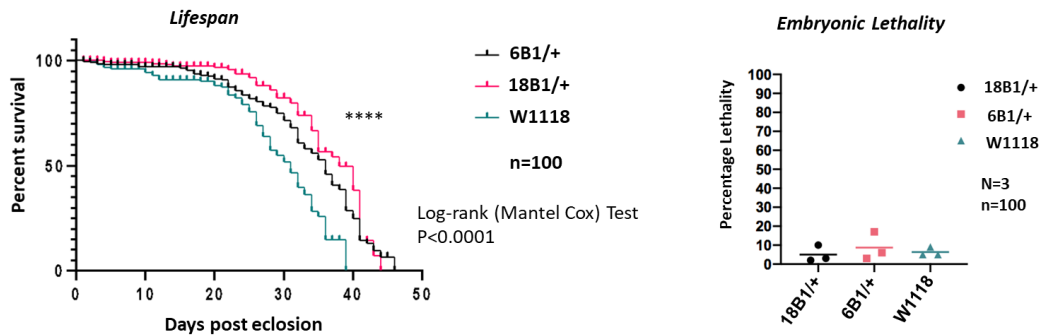


Figure 3.8: ΔRRS flies are haplo-sufficient. The ΔRRS animals, generated by CRISPR Cas9 genome editing, showed a lifespan (at 29 °C) similar to that of *w1118* flies (panel A) for genotypes $\Delta RRS/+$. Log-rank (Mantel-Cox) survival plot using GraphPad Prism 8.0.2 using Kaplan-Meier and Gehan-Wilcoxon tests suggests that $\Delta RRS/+$ flies show enhanced survival as compared to the control. (**** $P < 0.0001$). The $\Delta RRS/+$ flies also did not show any defects in embryonic development, based on the hatching of eggs, which was again at par with *w1118*. Homozygous ΔRRS flies are 2nd instar larval lethals.

3.3 Generation of a transgenic RRS^{SCR} line. The successful generation of the 6B1 RRS null (ΔRRS) line meant that the UAS-Gal4 system could be used to rescue the null. For this, *RRS-WT* and *RRS-SCR* sequences were cloned into a UAS vector (See materials and methods), and *UAS-RRS^{WT}* and *UAS-RRS^{SCR}* lines were created on the IIIrd chromosome. *Actin-Gal4; UAS-RRS^{WT}* and *Actin-Gal4; UAS-RRS^{SCR}* lines were balanced and crossed to $\Delta RRS/FM7i$ females. Both these lines could rescue the lethality of the ΔRRS male in the F1 generation, with the lines of the genotype, $\Delta RRS; Actin-Gal4; UAS-RRS^{WT}$ (referred to as RRS^{WT}) and $\Delta RRS; Actin-Gal4; UAS-RRS^{SCR}$ (referred to as RRS^{SCR}) being used for further experiments. A similar rescue was seen when *Ubiquitin-Gal4* was used instead of *Actin-Gal4*. Both the ‘rescued’ lines were homozygously viable and had a normal lifespan, suggesting that the SCR allele was functionally equivalent to the WT in terms of its canonical function. Western blots of adult males, rescued by expression of *UAS-RRS^{SCR}*, showed equal expression of RRS when compared to *UAS-RRS^{WT}* (Fig. 3.9B).

Drosophila reacts to immune challenges under laboratory conditions with a characteristic transcriptional upregulation and downregulation of defence genes. Infection with gram-positive *Micrococcus luteus* (*M. luteus*) and gram-negative *Erwinia carotovora carotovora* (*Ecc15*) were used to trigger the host-defence response. We measured the lifespan of RRS^{WT} and RRS^{SCR} animals post-infection. We find that there is no significant difference in lifespan

for *M. luteus* infections, while for *Ecc15*, there is an increase in lethality for *RRS^{SCR}* for younger animals (1-15 days), while not for older animals (20 days) (Fig. 3.10A-C).

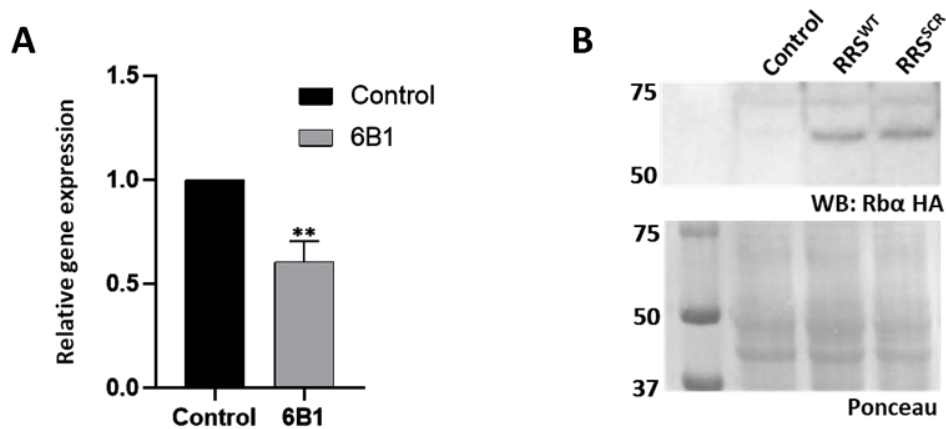


Figure 3.9: A. *ΔRRS^{6B1}* line shows lower transcript levels of RRS as compared to wildtype. *ΔRRS^{6B1}* shows a 40% reduction in transcript levels of RRS as compared to wildtype Control. Values on the Y-axis depict the fold change normalised to the housekeeping gene *rp49*. Values shown are Mean \pm SEM. N=3, n (larvae)=25. Statistical analysis by Unpaired t-test. * p<0.05, **p<0.01 ****p<0.001. **B. Rescue of *ΔRRS^{6B1}* by ectopic expression of RRS using UAS-Gal4 system.** Both *RRS^{WT}* and *RRS^{SCR}* lines show approximately equal expression of RRS when probed using an anti-HA antibody. Ponceau staining on the same blot is used to show equal loading.

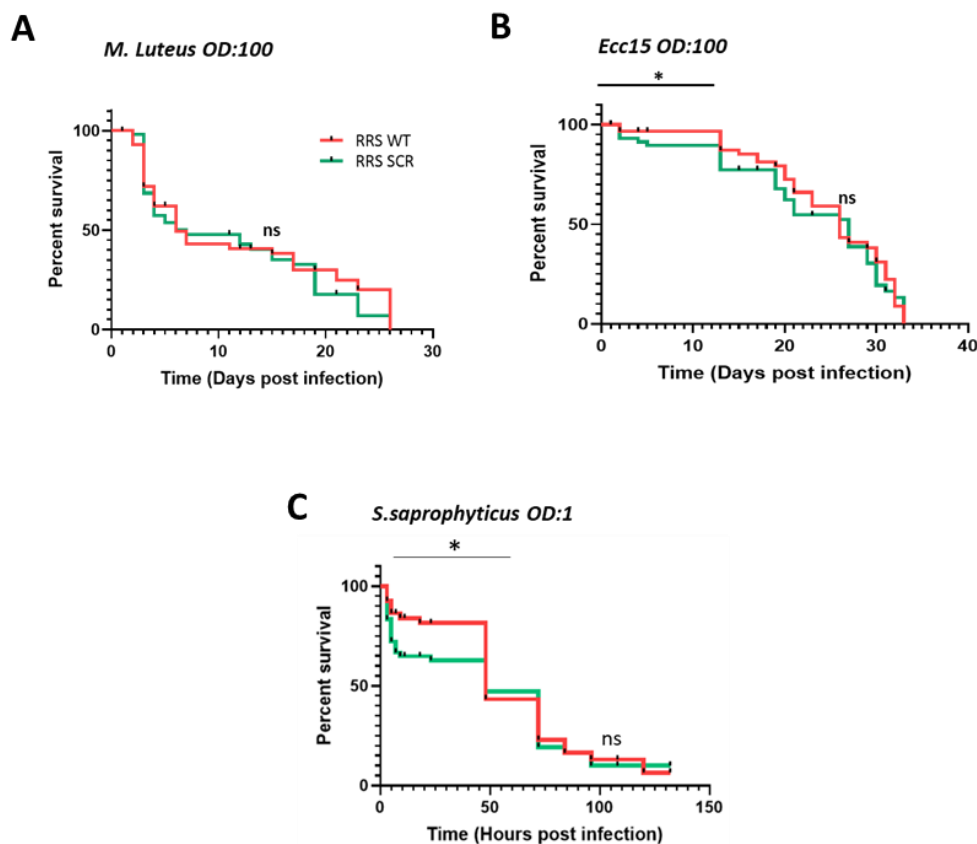


Figure 3.10: Survival plots for *RRS^{WT}* and *RRS^{SCR}* upon *M. luteus*, *Ecc15* and *S. saprophyticus* infection. Log-rank (Mantel cox) survival plot using Kalpan-Meier and Gehan-Wilcoxon tests suggests

that RRS^{WT} and RRS^{SCR} do not show a significant difference in lifespan overall post-infection with either *M. luteus* or *Ecc15*. However, post *Ecc15* and *S.saprophyticus* infection, RRS^{SCR} shows a significant (* $p < 0.05$) decrease in survival as compared to RRS^{WT} in the initial stages (0-15 days).

3.4 Transcriptomics of immune challenged, RRS^{WT} , and RRS^{SCR} transgenic animals. In order to uncover the role of SUMO conjugation in host-defense, we infected 7-8 day-old adult flies with *M. luteus* and *Ecc15* and measured transcript levels in both RRS^{WT} and RRS^{SCR} using quantitative 3' RNA sequencing (QuantSeq; Materials & Methods).

Infection with the bacteria gave a robust immune response (Fig. 3.11A-D). Gene Ontology analysis of the modulated genes revealed immune responsive genes associated with Gram Positive and Gram Negative infection showed both common and differentially expressed genes, as expected by Toll/NF κ B and Immune Deficient (Imd)/NF κ B pathway activation (De Gregorio et al. 2002)(Fig. A.3.2-3). For RRS^{WT} flies, infection by *M. luteus* led to an upregulation of 66 genes and a downregulation of 2 genes 22 hours post-infection. As expected, targets of the Toll pathway, such as *drosomycin* (*Drs*) and *metchnikowin* (*Mtk*), were upregulated (Fig. 3.12A, Table 3.2). For the RRS^{SCR} flies, 85 genes were upregulated, and 7 genes were downregulated. In a similar vein, infection by *Ecc15* led to 232 upregulated and 151 downregulated in RRS^{WT} and 209 upregulated and 80 downregulated in the RRS^{SCR} (Fig. 3.12B, Table 3.2). As expected, targets of the Imd pathway were strongly modulated. In order to examine the extent of overlap among upregulated and downregulated genes between different data sets, Venn diagrams were drawn (Fig. 3.12C-D). A majority of the genes were uniquely expressed among the data sets. Uniquely differentially expressed genes are listed in (Table A.3.1). Common genes between RRS^{WT} and RRS^{SCR} for each infection category were used for further analysis. At the basal level, before infection, RRS^{WT} and RRS^{SCR} did not show significant differences in their transcriptome (Table A.3.2).

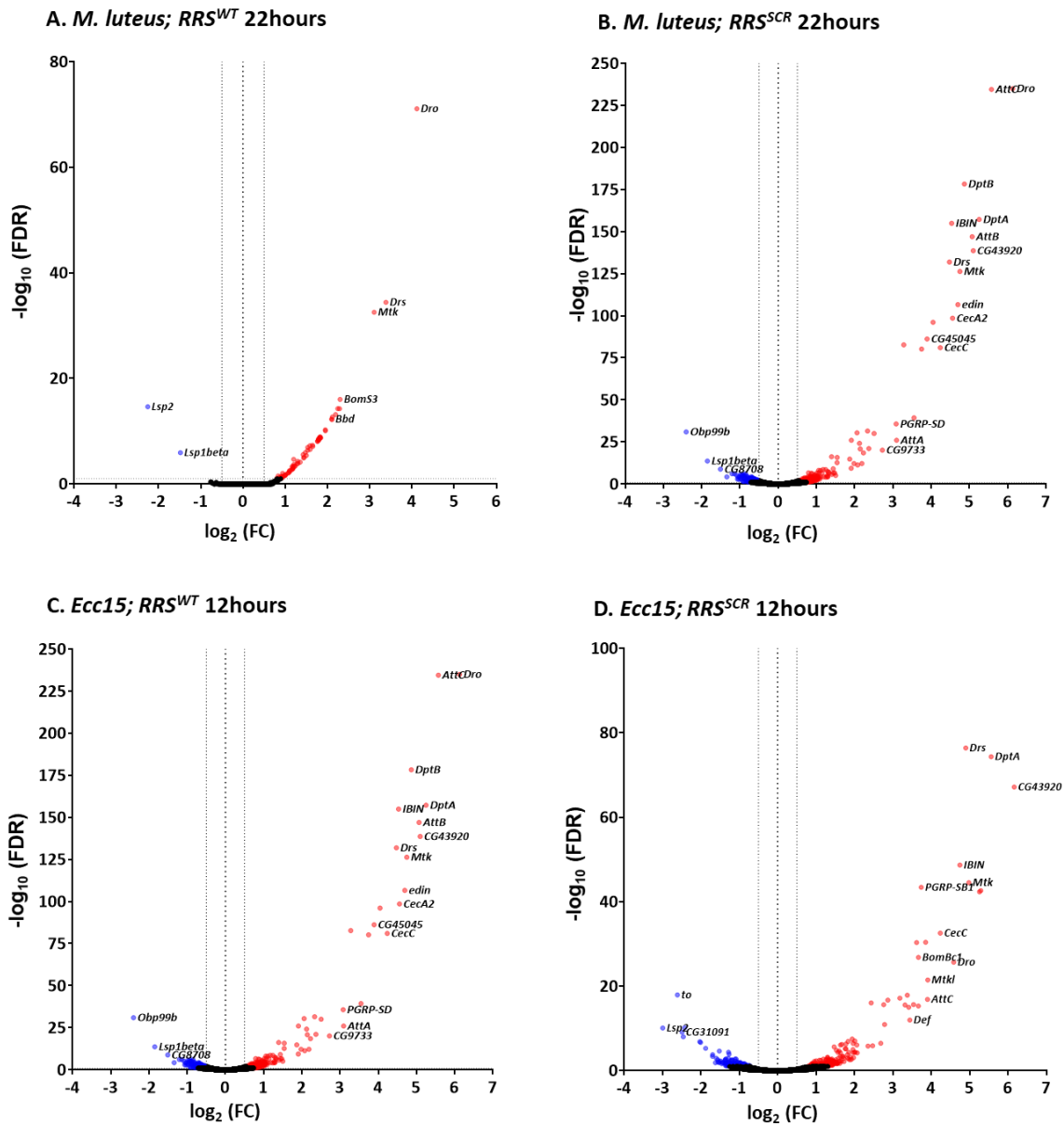


Figure 3.11: RRS^{WT} and RRS^{SCR} show a robust immune response to bacterial infection. A-D. Volcano plot(s) for the differentially expressed genes. $\log_2(\text{FC})$ for each gene is plotted against its $-\log_{10}(\text{FDR})$ value to display differentially expressed genes upon infection as compared to the baseline. Red and blue dots represent the genes that are significantly differentially expressed with $\log_2(\text{FC})$ of > 0.55 and < -0.55 , respectively, with $p\text{-value} < 0.05$ ($-\log_{10}(\text{FDR})$ of > 2) whereas black dots represent the genes that are uniformly expressed. Representative differentially expressed genes are mentioned on the right (upregulated) and left-hand (downregulated) corners of each plot. The time point for *M. luteus* infection is 22 hours, and for *Ecc15*, it is 12 hours.

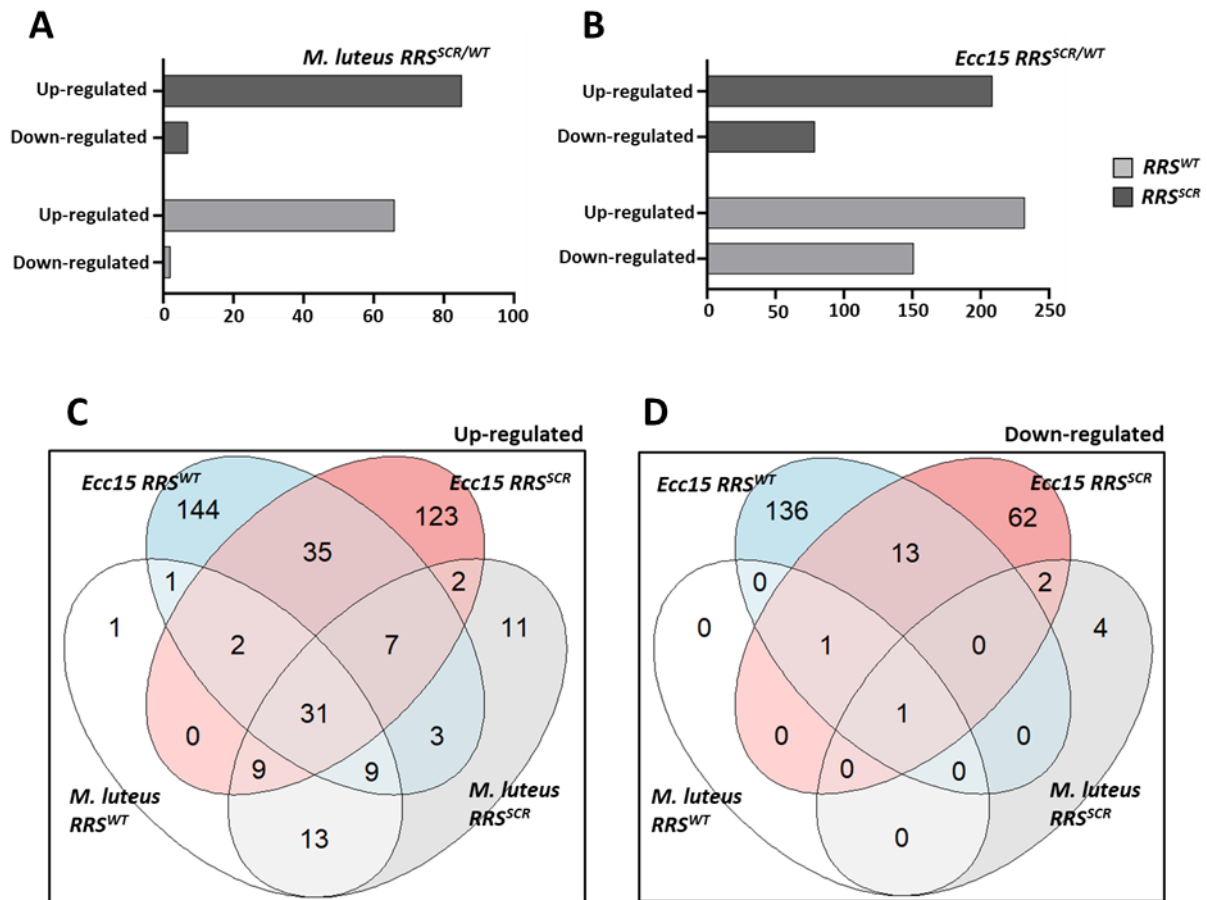


Figure 3.12: A. Total Number of transcripts upregulated and downregulated in response to infection. Genes modulated by infection by *M. luteus* (A) and *Ecc15* (B) for both RRS^{WT} and RRS^{SCR}. **B. Differential expression of genes.** Venn diagram showing sub-division of upregulated (C) and downregulated (D) genes for experiments conducted, as defined earlier.

Table 3.2: RRS^{WT} and RRS^{SCR} show differential expression of immune target genes.

A-B. Tabulation of differentially expressed genes. Representative genes with differential expression for RRS^{SCR} for *M. luteus* (A) and *Ecc15* (B) infections. For *M. luteus*, there is moderate upregulation for most genes, while for *Ecc15*, metabolic genes are both up and down-regulated.

 $RRS^{SCR/WT}$ *M. luteus* 22hours

TYPE	NAME	SYMBOL	RRS WT 22 hours	RRS SCR 22 hours	Difference
AMPs	<i>Drosomycin</i>	<i>Drs</i>	3.39	5.26	1.87
	<i>Attacin-C</i>	<i>AttC</i>	1.79	2.44	0.64
	<i>Cecropin A2</i>	<i>CecA2</i>	1.44	2.01	0.56
Bomanins	<i>Bomanin Bicipital 1</i>	<i>BomBc1</i>	2.29	4.43	2.14
	<i>Bomanin Short 1</i>	<i>BomS1</i>	1.95	4.01	2.06
	<i>Bomanin Tailed 1</i>	<i>BomT1</i>	1.59	3.54	1.95
	<i>Bomanin Short 2</i>	<i>BomS2</i>	1.83	3.06	1.23
	<i>Bomanin Bicipital 3</i>	<i>BomBc3</i>	1.77	2.60	0.83
Carbohydrate binding	<i>GGBP-like 3</i>	<i>GGBP-like3</i>	1.84	2.96	1.12
Catalase	<i>Immune-regulated catalase</i>	<i>Irc</i>	1.11	1.67	0.57
Cysteine-type endopeptidase	-	<i>CG11459</i>	1.80	2.65	0.85
Exonuclease	<i>Deoxyribonuclease II</i>	<i>DNaseII</i>	1.36	2.00	0.64
Serine Protease inhibitor	-	<i>CG42259</i>	1.79	2.94	1.15
Serine-type endopeptidase	<i>Serine protease homolog 93</i>	<i>SPH93</i>	2.19	4.01	1.82
Un-annotated	<i>Induced by Infection</i>	<i>IBIN</i>	2.11	3.15	1.04
	<i>Bombardier</i>	<i>Bbd</i>	2.10	2.79	0.70
	Daisho2	<i>Dso2</i>	2.17	2.70	0.58
	-	<i>CG16978</i>	1.06	1.63	0.57
Carboxyl esterase	-	<i>CG4757</i>	1.95	4.90	2.95
Dopachrome isomerase	<i>yellow-f</i>	<i>yellow-f</i>	1.76	2.33	0.57
Un-annotated	-	<i>CG16772</i>	1.50	2.09	0.59

RRS^{SCR/WT} Ecc15 12 hours

TYPE	NAME	SYMBOL	RRS WT 12 hours	RRS SCR 12 hours	Difference
AMPs	<i>Cecropin A1</i>	<i>CecA1</i>	1.50	3.32	1.82
	<i>Attacin-D</i>	<i>AttD</i>	1.98	3.54	1.56
	<i>Defensin</i>	<i>Def</i>	2.19	3.44	1.26
	<i>Cecropin A2</i>	<i>CecA2</i>	4.56	5.29	0.73
	<i>Drosocin</i>	<i>Dro</i>	6.13	4.59	-1.54
	<i>Attacin-C</i>	<i>AttC</i>	5.57	3.90	-1.67
	<i>Attacin-A</i>	<i>AttA</i>	3.09	1.40	-1.69
	<i>Diptericin B</i>	<i>DptB</i>	4.86	1.73	-3.14
Bomanins	<i>Bomanin Tailed 1</i>	<i>BomT1</i>	1.47	2.79	1.31
	<i>Bomanin Bicipital 1</i>	<i>BomBc1</i>	2.50	3.67	1.16
	<i>Bomanin Short 5</i>	<i>BomS5</i>	1.00	1.97	0.96
Carbohydrate binding	<i>GNBP-like 3</i>	<i>GNBP-like3</i>	2.06	2.87	0.81
Growth factor activity	<i>spatzle</i>	<i>spz</i>	0.95	1.60	0.65
Growth factor Receptor	<i>Nimrod B1</i>	<i>NimB1</i>	1.30	2.00	0.71
Peptidoglycan Recognition protein	<i>Peptidoglycan recognition protein LB</i>	<i>PGRP-LB</i>	0.62	1.22	0.59
	<i>Peptidoglycan recognition protein LF</i>	<i>PGRP-LF</i>	1.08	1.64	0.55
	<i>Peptidoglycan recognition protein SD</i>	<i>PGRP-SD</i>	3.08	3.86	0.77
Serine endopeptidase inhibitor	<i>Immune induced molecule 33</i>	<i>IM33</i>	0.66	1.30	0.64
Serine Protease	<i>Hayan</i>	<i>Hayan</i>	0.96	2.35	1.40
	<i>Serine Protease Immune Response Integrator</i>	<i>spirit</i>	0.83	1.53	0.70
Serine Protease Inhibitor	-	<i>CG42259</i>	1.32	3.41	2.09
Tyrosine Hydroxylase	<i>pale</i>	<i>ple</i>	0.68	1.46	0.78

TYPE	NAME	SYMBOL	RRS WT 12 hours	RRS SCR 12 hours	Difference
Glycoside hydrolase	-	<i>CG9701</i>	0.67	1.94	1.27
GTP Cyclohydrolase	<i>Punch</i>	<i>Pu</i>	0.70	2.76	2.07
juvenile hormone binding	<i>takeout</i>	<i>to</i>	-0.64	-2.61	-1.97
Lipase	-	<i>CG4267</i>	0.70	1.67	0.98
lncRNA	<i>long non-coding RNA:CR32661</i>	<i>lncRNA:CR32661</i>	-0.71	1.64	2.34
	<i>long non-coding RNA:CR33942</i>	<i>lncRNA:CR33942</i>	0.71	1.82	1.10
methylenetetrahyd rofolate dehydrogenase	-	<i>CG4716</i>	0.82	-0.90	-1.71
Serine protease inhibitor	<i>Serpin 28F</i>	<i>Spn28F</i>	0.73	-1.52	-2.25
Serine-type endopeptidase	<i>Jonah 99Ci</i>	<i>Jon99Ci</i>	-0.70	-1.36	-0.66
	-	<i>CG17242</i>	-0.58	-1.35	-0.77
Sorbitol dehydrogenase	<i>Sorbitol dehydrogenase 1</i>	<i>Sodh-1</i>	-1.15	-2.42	-1.27
Sterol esterase	-	<i>CG31091</i>	-1.03	-2.49	-1.46
Un-annotated	-	<i>CG13641</i>	1.03	3.18	2.15
	-	<i>CG15461</i>	-0.64	1.20	1.84
	-	<i>CG43920</i>	5.10	6.16	1.07
	-	<i>CG9759</i>	0.99	1.97	0.98
	-	<i>CG17107</i>	1.14	2.06	0.92
	<i>Larval serum protein 1 beta</i>	<i>Lsp1beta</i>	-1.85	-2.46	-0.61
	<i>Larval serum protein 2</i>	<i>Lsp2</i>	-0.91	-2.99	-2.08

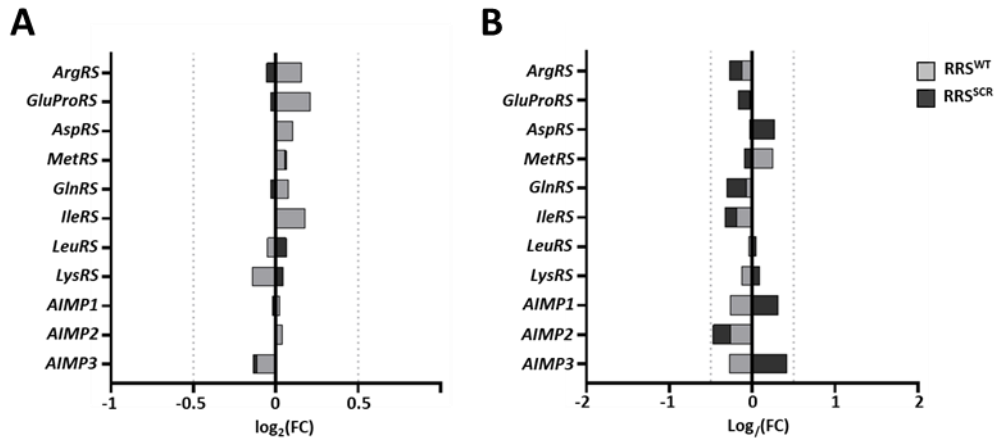


Figure 3.13: Transcriptome changes for MARS Complex genes. On infection by *M. luteus* (A) and *Ecc15* (B), the changes in transcript levels for genes that code for proteins in the MARS complex are well below the significance cut-off of 0.5 $\log_2(\text{FC})$.

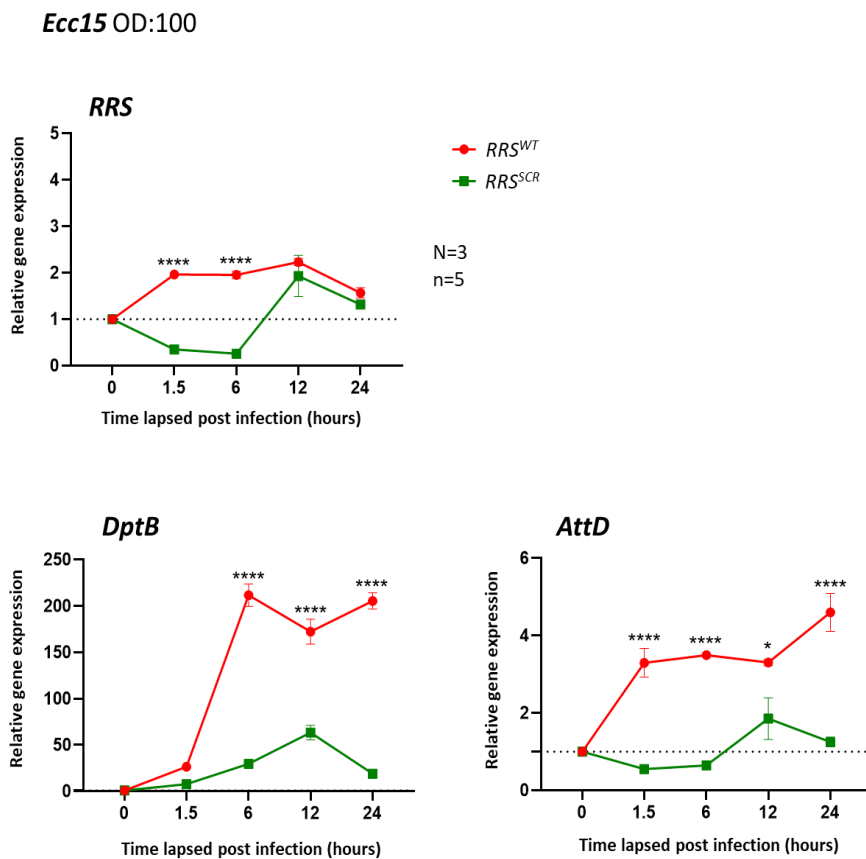


Figure 3.14: Expression of *RRS*, Toll pathway target gene *Drosomycin* and *Immune regulated catalase* (*Irc*) in *RRS*^{WT} and *RRS*^{SCR} upon *M. luteus* infection across 0-48 Hr. Values on the Y-axis depict the fold change normalised to the housekeeping gene *rp49*. Values shown are Mean \pm SEM. N=3, n=5. Statistical analysis by two-way ANOVA followed by Tukey's multiple comparison test. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

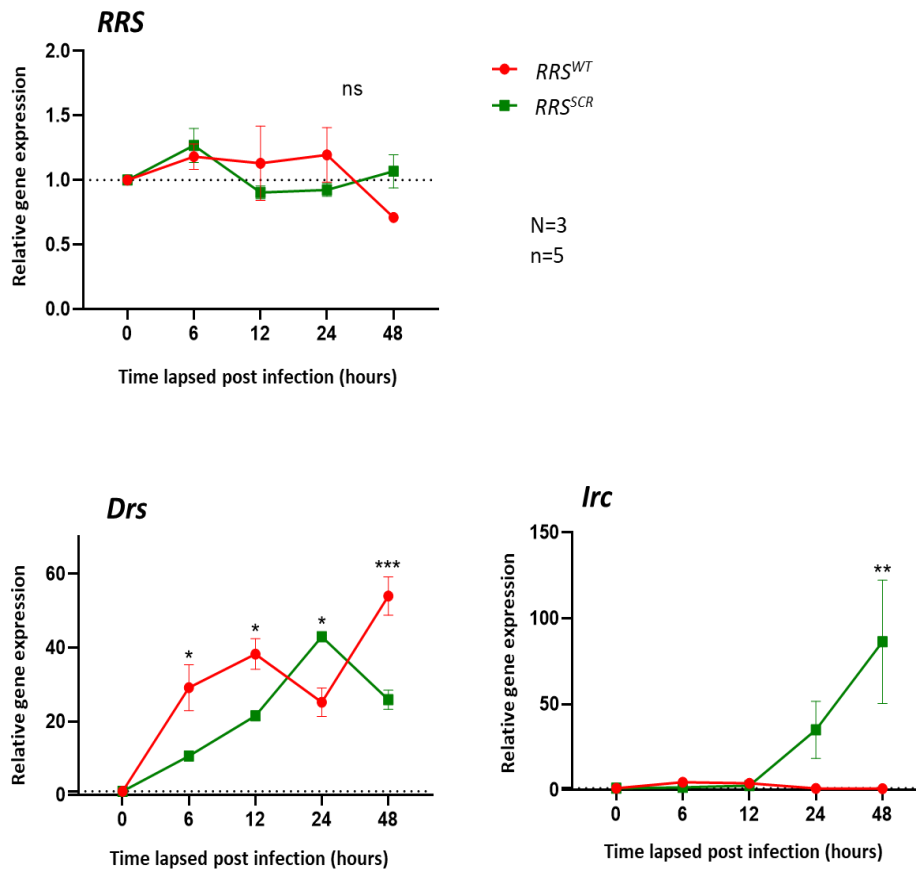


Figure 3.15: Expression of *RRS*, *Imd* pathway target genes *Diptericin B* and *Attacin D* in RRS^{WT} and RRS^{SCR} upon *Ecc15* infection across 0-24 Hr. Values on the Y-axis depict the fold change normalised to the housekeeping gene *rp49*. Values shown are Mean \pm SEM. N=3, n=5. Statistical analysis by two-way ANOVA followed by Tukey's multiple comparison test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

3.5 Modulation of the immune transcriptome in RRS^{SCR} transgenics. Next, we compared the change in immune transcriptome for RRS^{SCR} with reference to RRS^{WT} (Table 3.2A-B, Fig. A.3.4-6). In the case of *M. luteus* infection, a total of 22 immune-responsive genes, including AMPs, Bomanins, Serine hydrolases, and genes involved in ROS production, were significantly differentially upregulated (Table 3.2A, Fig. A.3.4A, Fig. A.3.5) in RRS^{SCR} . Both *Drosomycin* (*Drs*) and *Bomanin Bicipital 1* (*BomBc1*) are upregulated 5-6-fold in RRS^{SCR} , while other AMP genes (Table 3.2A) were not strongly or significantly upregulated.

In the case of *Ecc15* infection, the trends were stronger. A total of 28 genes showed enhanced upregulation, and 13 genes showed enhanced repression in RRS^{SCR} . Genes involved in metabolism, such as hydrolases, esterases, non-coding RNA, and AMP genes, were modulated. Amongst the strongly expressed genes (Table 3.2, Fig. A.3.4B, Fig. A.3.6) were immune-responsive genes involved in gram-negative bacterial recognition (*PGRP-LB*, *PGRP-*

LF, and *PGRP-SD*), and melanization (*Hayan*, *Pale* and *Punch*). Genes involved in oxidoreductase pathways like *Sodh 1* and Larval serum proteins *Lsp1* and *Lsp2* were repressed. For the AMP genes (Table 3.2B), *CecA1* and *AttD* were upregulated 2-4 fold, while *AttC*, *AttA*, and *DptB* downregulated 3-9 fold. We also looked at the transcriptional changes in the genes of the MARS complex. For both *M. luteus* and *Ecc15*, the transcriptional changes on infection were minimal, with none of the transcript levels crossing our cut-off of significance, $0.55 \log_2(\text{FC})$ (Fig. 3.13A-B).

Next, we validated the QuantSeq data by qRT-PCR for a few targets at time points ranging from 0-48 hours. For *M. luteus*, *RRS* levels did not change significantly from 0-48 hours (Fig. 3.14), while *Drs* levels, though significant, showed similar trends over 48 hrs (Fig. 3.14). *Irc* levels were distinctly higher in *RRS^{SCR}* animals at later time points (Fig. 3.14). For *Ecc15*, *RRS* transcript levels were different, showing a 5-fold decrease in *RRS^{SCR}* animals for the time points 1.5 and 6 hours (Fig. 3.15). *DptB* and *AttD* transcripts are significantly lower in the case of *RRS^{SCR}* animals (Fig. 3.15).

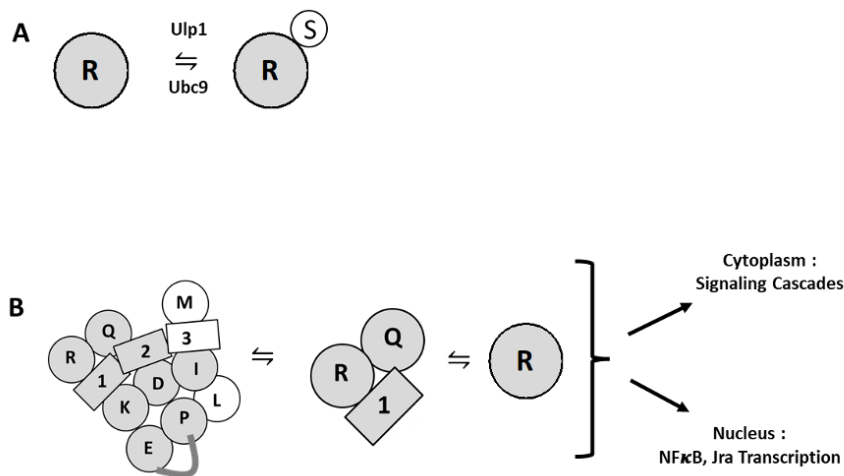


Figure 3.16: A. SUMO Conjugation of RRS. RRS is a target for the cellular SUMO conjugation machinery. A small fraction of RRS is SUMO conjugated and is in equilibrium with non-conjugated RRS. **B. Model for immune regulation by RRS.** RRS may influence the immune response either as part of the MARS complex or as part of the AIMP1:RRS: QRS complex, or as free RRS. In a SUMO-conjugated state, RRS may influence signalling cascades, interacting with partners containing SIM sites and modifying their function. These influences can be either cytoplasmic or nuclear.

4. Discussion

The MARS Complex has been implicated as a sensor and regulator of the immune response (Guo and Schimmel, 2013; Kim et al., 2014b; Arif et al., 2018; Nie et al., 2019). Mutations and misregulation of MARS function can lead to immune disease (Lee et al., 2018; Nie et al., 2019). In the best-studied mechanistic example, in response to infection and release of IFN- γ , EPRS dissociates from the MARS Complex (Sampath et al., 2004). The dissociation is triggered by phosphorylation of the WHEP domain. EPRS now associates with L13a, NSAP1, and GAPDH to form a ‘GAIT complex,’ which can now bind to a GAIT RNA element. The GAIT-RNA element (interferon-gamma-activated inhibitor of translation) (Sampath et al., 2003; Marquez-Jurado et al., 2015) is present in UTRs of mRNA transcripts and binding leads to a block of translation of the transcript.

Roles for RRS in the immune response are unknown. In terms of disease, mutations in RRS have been implicated in neuronal hypomyelination with severe spasticity and nystagmus (Antonellis and Green, 2008; Wolf et al., 2014). Autoantibodies against ARSs were found in anti-synthetase syndrome (ASSD), suggesting that ARSs are likely to be involved in the development and progression of autoimmune disease. In *Drosophila*, RRS is not studied in any physiological context.

How, then, does RRS modulate the transcription of defence genes? In mammals, the MARS complex itself is believed to be a cytoplasmic complex, though a few studies suggest nuclear localization (Wolfe et al., 2003; Cui et al., 2020). RRS could be available in at least three species, one as a free, unbound entity, the second as a complex with AIMP1 and QRS, and finally as part of the MARS Complex (Fig. 3.16B). Deletion of the RRS LZD leads to its dissociation from the MARS complex, but this does not affect charging (Cui et al., 2020). Interestingly, the nuclear fraction of MARS decreases when cells contain RRS (Δ LZD). In its dissociated state, RRS’s canonical functions are unaffected, but developmental genes such as homeobox and forkhead box genes are modulated (Cui et al., 2020).

Each of the RRS species could exist in a SUMO conjugated or unconjugated state (Fig. 3.16A). These species can ultimately regulate gene expression either by influencing signalling pathways in the cytoplasm or by affecting the transcription of the nuclear-localized NF κ Bs. SUMOylation could affect the stability or interaction with other proteins. RRS lacks a nuclear localization signal (NLS), as does *Drosophila* SUMO. Transport to the nucleus would require RRS to be part of a complex that includes an NLS, for example, the AIMP1:RRS: QRS complex, as AIMP1 may travel to the nucleus (Lee et al., 2008; Park et al., 2010). AIMP1 in mammals is a precursor for EMAP2, which can trigger an inflammatory response (Lee et al.,

2019), and a similar mechanism may exist in flies. Other possible mechanisms include modulation of NF κ B (Ko et al., 2001) by regulation of secretion of AIMP1 or by regulation of Jun signalling (Park et al., 2002), which in turn can regulate the immune response.

In Summary, RRS is SUMO conjugated, and SUMOylation appears to modulate, indirectly, the transcriptional host-defense response. The mechanisms underlying these phenomena are currently unknown.

5. Contributions

The research chapter is a collaboration with a BSMS student, Aarti Kejriwal (Kejriwal A, Thesis, 2021), who I mentored for her MS research project. Aarti worked under my supervision to set up stable lines for the UAS/Gal4 rescue experiments.

6. Acknowledgements

We thank Bloomington *Drosophila* Stock Center (BDSC), Indiana, supported by NIH grant P40OD018537, for fly stocks; Fly facility at the National Centre for Biological Sciences (NCBS), Bangalore, for embryonic injections. Dr. Deepti Trivedi for her input on the design of the dual gRNA construct for RRS.

References

- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166-169.
- Antonellis, A., and Green, E.D. (2008). The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu Rev Genomics Hum Genet* 9, 87-107.
- Arif, A., Jia, J., Halawani, D., and Fox, P.L. (2017). Experimental approaches for investigation of aminoacyl tRNA synthetase phosphorylation. *Methods* 113, 72-82.
- Arif, A., Yao, P., Terenzi, F., Jia, J., Ray, P.S., and Fox, P.L. (2018). The GAIT translational control system. *Wiley Interdiscip Rev RNA* 9.
- Beauclair, G., Bridier-Nahmias, A., Zagury, J.F., Saib, A., and Zamborlini, A. (2015). JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics* 31, 3483-3491.
- Curran, J.A., and Weiss, B. (2016). What Is the Impact of mRNA 5' TL Heterogeneity on Translational Start Site Selection and the Mammalian Cellular Phenotype? *Front Genet* 7, 156.
- Chang, J.M., Di Tommaso, P., Taly, J.F., and Notredame, C. (2012). Accurate multiple sequence alignment of transmembrane proteins with PSI-Coffee. *BMC Bioinformatics* 13 Suppl 4, S1.
- Cui, H., Kapur, M., Diedrich, J.K., Iii, J.R.Y., Ackerman, S.L., and Schimmel, P. (2020). Regulation of ex-translational activities is the primary function of the multi-tRNA synthetase complex. *Nucleic Acids Res.*
- Craig, D.B., and Dombkowski, A.A. (2013). Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* 14, 346.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.
- Fu, Y., Kim, Y., Jin, K.S., Kim, H.S., Kim, J.H., Wang, D., Park, M., Jo, C.H., Kwon, N.H., Kim, D., Kim, M.H., Jeon, Y.H., Hwang, K.Y., Kim, S., and Cho, Y. (2014). Structure of the ArgRS-GlnRS-AIMP1 complex and its implications for mammalian translation. *Proc Natl Acad Sci U S A* 111, 15084-15089.
- Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* 8, 947-956.
- Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. *Sci Signal* 2, ra24.
- Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M., and O'connor-Giles, K.M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* 196, 961-971.
- Guo, M., and Schimmel, P. (2013). Essential nontranslational functions of tRNA synthetases. *Nature Chemical Biology* 9, 145-153.
- Handu, M., Kaduskar, B., Ravindranathan, R., Soory, A., Giri, R., Elango, V.B., Gowda, H., and Ratnaparkhi, G.S. (2015). SUMO-Enriched Proteome for *Drosophila* Innate Immune Response. *G3-Genes Genomes Genetics* 5, 2137-2154.
- Havrylenko, S., and Mirande, M. (2015). Aminoacyl-tRNA synthetase complexes in evolution. *Int J Mol Sci* 16, 6571-6594.
- Hay, R.T. (2005). SUMO: a history of modification. *Mol Cell* 18, 1-12.
- Hendriks, I.A., and Vertegaal, A.C. (2016). A comprehensive compilation of SUMO proteomics. *Nat Rev Mol Cell Biol* 17, 581-595.
- Kerjan, P., Cerini, C., Semeriva, M., and Mirande, M. (1994). The multienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from *Drosophila* to mammals. *Biochim Biophys Acta* 1199, 293-297.
- Khan, K., Baleanu-Gogonea, C., Willard, B., Gogonea, V., and Fox, P.L. (2020). 3-Dimensional architecture of the human multi-tRNA synthetase complex. *Nucleic Acids Research* 48, 8740-8754.

- Kim, H.S., Cha, S.Y., Jo, C.H., Han, A., and Hwang, K.Y. (2014a). The crystal structure of arginyl-tRNA synthetase from *Homo sapiens*. *FEBS Lett* 588, 2328-2334.
- Kim, J.H., Han, J.M., and Kim, S. (2014b). Protein-protein interactions and multi-component complexes of aminoacyl-tRNA synthetases. *Top Curr Chem* 344, 119-144.
- Ko, Y.G., Park, H., Kim, T., Lee, J.W., Park, S.G., Seol, W., Kim, J.E., Lee, W.H., Kim, S.H., Park, J.E., and Kim, S. (2001). A cofactor of tRNA synthetase, p43, is secreted to upregulate proinflammatory genes. *J Biol Chem* 276, 23028-23033.
- Lee, D.D., Hochstetler, A., Murphy, C., Lowe, C.W., and Schwarz, M.A. (2019). A distinct transcriptional profile in response to endothelial monocyte activating polypeptide II is partially mediated by JAK-STAT3 in murine macrophages. *Am J Physiol Cell Physiol* 317, C449-C456.
- Lee, E.Y., Kim, S., and Kim, M.H. (2018). Aminoacyl-tRNA synthetases, therapeutic targets for infectious diseases. *Biochem Pharmacol* 154, 424-434.
- Lee, Y.S., Han, J.M., Son, S.H., Choi, J.W., Jeon, E.J., Bae, S.C., Park, Y.I., and Kim, S. (2008). AIMP1/p43 downregulates TGF-beta signaling via stabilization of smurf2. *Biochem Biophys Res Commun* 371, 395-400.
- Leppek, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* 19, 158-174.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.
- Lu, J.M., Marygold, S.J., Gharib, W.H., and Suter, B. (2015). The aminoacyl-tRNA synthetases of *Drosophila melanogaster*. *Fly* 9, 53-61.
- Marquez-Jurado, S., Nogales, A., Zuniga, S., Enjuanes, L., and Almazan, F. (2015). Identification of a gamma interferon-activated inhibitor of translation-like RNA motif at the 3' end of the transmissible gastroenteritis coronavirus genome modulating innate immune response. *mBio* 6, e00105.
- Nie, A.Z., Sun, B., Fu, Z.H., and Yu, D.S. (2019). Roles of aminoacyl-tRNA synthetases in immune regulation and immune diseases. *Cell Death & Disease* 10.
- Nie, M.H., Xie, Y.M., Loo, J.A., and Courey, A.J. (2009). Genetic and Proteomic Evidence for Roles of *Drosophila* SUMO in Cell Cycle Control, Ras Signaling, and Early Pattern Formation. *PLoS One* 4.
- Panse, V.G., Hardeland, U., Werner, T., Kuster, B., and Hurt, E. (2004). A proteome-wide approach identifies sumoylated substrate proteins in yeast. *J Biol Chem* 279, 41346-41351.
- Park, S.G., Choi, E.C., and Kim, S. (2010). Aminoacyl-tRNA Synthetase-Interacting Multifunctional Proteins (AIMPs): A Triad for Cellular Homeostasis. *Iubmb Life* 62, 296-302.
- Park, S.G., Kang, Y.S., Ahn, Y.H., Lee, S.H., Kim, K.R., Kim, K.W., Koh, G.Y., Ko, Y.G., and Kim, S. (2002). Dose-dependent biphasic activity of tRNA synthetase-associating factor, p43, in angiogenesis. *J Biol Chem* 277, 45243-45248.
- Pirone, L., Xolalpa, W., Sigurethsson, J.O., Ramirez, J., Perez, C., Gonzalez, M., De Sabando, A.R., Elortza, F., Rodriguez, M.S., Mayor, U., Olsen, J.V., Barrio, R., and Sutherland, J.D. (2017). A comprehensive platform for the analysis of ubiquitin-like protein modifications using in vivo biotinylation. *Sci Rep* 7, 40756.
- Rubio Gomez, M.A., and Ibba, M. (2020). Aminoacyl-tRNA synthetases. *RNA* 26, 910-936.
- Sampath, P., Mazumder, B., Seshadri, V., and Fox, P.L. (2003). Transcript-selective translational silencing by gamma interferon is directed by a novel structural element in the ceruloplasmin mRNA 3' untranslated region. *Mol Cell Biol* 23, 1509-1519.
- Sampath, P., Mazumder, B., Seshadri, V., Gerber, C.A., Chavatte, L., Kinter, M., Ting, S.M., Dignam, J.D., Kim, S., Driscoll, D.M., and Fox, P.L. (2004). Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell* 119, 195-208.
- Schimmel, P.R., and Soll, D. (1979). Aminoacyl-tRNA synthetases: general features and recognition of transfer RNAs. *Annu Rev Biochem* 48, 601-648.
- Talamillo, A., Sánchez, J., and Barrio, R. (2008). Functional analysis of the SUMOylation pathway in *Drosophila*. *Biochem Soc Trans* 36, 868-873.
- Tatham, M.H., Matic, I., Mann, M., and Hay, R.T. (2011). Comparative proteomic analysis identifies a role for SUMO in protein quality control. *Sci Signal* 4, rs4.

- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., De Beer, T.a.P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46, W296-W303.
- Wolf, N.I., Salomons, G.S., Rodenburg, R.J., Pouwels, P.J., Schieving, J.H., Derks, T.G., Fock, J.M., Rump, P., Van Beek, D.M., Van Der Knaap, M.S., and Waisfisz, Q. (2014). Mutations in RARS cause hypomyelination. *Ann Neurol* 76, 134-139.
- Wolfe, C.L., Warrington, J.A., Davis, S., Green, S., and Norcum, M.T. (2003). Isolation and characterization of human nuclear and cytosolic multisynthetase complexes and the intracellular distribution of p43/EMAPII. *Protein Sci* 12, 2282-2290.
- Yao, P., Poruri, K., Martinis, S.A., and Fox, P.L. (2014). Non-catalytic Regulation of Gene Expression by Aminoacyl-tRNA Synthetases. *Aminoacyl-Trna Synthetases in Biology and Medicine* 344, 167-187.
- Zhang, Y., Werling, U., and Edlmann, W. (2014). Seamless Ligation Cloning Extract (SLiCE) cloning method. *Methods Mol Biol* 1116, 235-244.

CHAPTER IV

Does SUMOylation of EPRS regulate innate immune response?**Summary**

Glutamyl Prolyl tRNA Synthetase (EPRS) has been demonstrated to be a target of SUMO conjugation. In this chapter, I utilise genome editing, using the CRISPR/Cas9 toolbox to generate an EPRS variant where all five SUMO conjugation sites have been modified to be SUMO conjugation-resistant EPRS^{5M-SCR}. Although I was successful in this exercise, the variant generated was not stable and was lost over time. Despite repeated attempts, I could not regenerate the mutant variant and thus, in contrast to RRS (Chapter 3), could not demonstrate the effect of loss of EPRS SUMOylation on the *Drosophila* innate immune response.

1. Introduction

The second Aminoacyl tRNA Synthetase we choose as a target is Glutamyl Prolyl tRNA Synthetase (EPRS), it again being an understudied target showing heavy and robust SUMOylation. Glutamyl Prolyl tRNA Synthetase (EPRS) is ubiquitous, cytoplasmic and the only known bifunctional tRNA synthetase. It has two synthetase catalytic cores, an N-terminal Glutamyl tRNA Synthetase (ERS) and a C-terminal Prolyl tRNA synthetase (PRS) in a single polypeptide chain. ERS and PRS ligate Glutamate and Proline, respectively. ERS is a Class I tRNA Synthetase, whereas PRS is a Class II tRNA Synthetase. The two enzymatic domains are separated by a non-catalytic, central linker region constituting of helix turn helix WHEP domains (named after AARSs first shown to contain these domains WRS, HRS and EPRS) (Cahuzac et al. 2000; Jeong et al. 2000). These domains are known to bind RNA. The presence of a linked tRNA Synthetase was detected as early as in a filasterean, *Capsaspora owczarzaki*, which has a linker with two WHEP domains. The bifunctionality of the enzyme is conserved from insects to mammals with the exception of a nematode, *Caenorhabditis elegans* and its closely related species (Cerini et al. 1997). The *Drosophila* variant has six WHEP domains (the maximum number of WHEP domains found in any organism), whereas its human counterpart has three such domains (Cerini et al. 1997; Rho et al. 1998). Interestingly in a Cnidarian *Nematostella vectensis* (stinging sea anemone), *EPRS* has three isoforms with a variable number and sequence of WHEP domains with different RNA binding capabilities by virtue of alternative splicing. Apart from appended WHEP domains in the central linker region EPRS

also has an N-terminus Glutathione-S-transferase (GST) like domain which assists in its assembly within the MARS complex (Fig. 4.1).

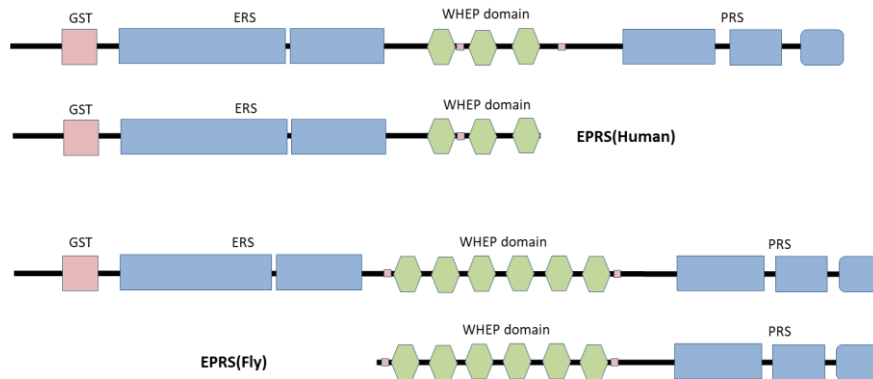


Figure 4.1: Schematic of EPRS. EPRS is a bi-functional multi-enzymatic Aminoacyl tRNA Synthetase. A. Human EPRS consists of two enzymatic(catalytic) domains; an N-terminal Glutamyl tRNA Synthetase (ERS) domain and a C-terminal Prolyl tRNA Synthetase (PRS) domain separated by a central, non-catalytic linker region bearing three RNA binding, WHEP domains. The N-terminal ERS domain is preceded by a GST-like domain (involved in MARS Complex assembly and structure). There are two protein-coding transcripts and two polypeptides associated with the gene in humans formed as a result of alternative splicing. The shorter isoform lacks the PRS domain. B. The *Drosophila* EPRS differs from its human counterpart and harbours six WHEP domains in the central linker region. In the case of *Drosophila*, the gene encodes two isoforms via alternative splicing. The shorter isoform lacks the ERS domain.

It is part of the sub-complex I within the MARS complex, wherein it is closely associated with Isoleucyl tRNA Synthetase and AIMP2 via its N-terminal GST-like domain (Quevillon et al. 1999; Norcum and Dignam 1999; Robinson et al. 2000). Biallelic mutations in EPRS is known to cause Hypomyelinating Dystrophy.

Noncanonical functions of EPRS

GAIT-dependent noncanonical function: Regulation of gene expression influences all aspects of cellular life. It can be modulated by either transcriptional or post-transcriptional/translational control mechanisms. The latter provides fine and localised control of protein production and accumulation. Post-transcriptional control can be either global, regulating the expression of most transcripts or transcript selective, modulating a defined set of mRNAs (a subset of functionally related genes). It is mediated by proteins and protein complexes binding to specific elements in 5'UTR and 3'UTR of target mRNAs. It is a co-ordinate post-transcriptional regulon.

The formation of GAIT complex upon interferon-gamma (a proinflammatory cytokine) stimulation in human myeloid cells is one such post-transcriptional control mechanism. In humans, the GAIT complex is heterotetrameric and constitutes of Glutamyl Prolyl tRNA Synthetase (EPRS), Syncrip (NSAP1), RPL13a and Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). A heterotrimeric GAIT complex exists in mice wherein it is devoid of Syncrip(NSAP1). The GAIT complex inhibits the translation of a subset of inflammation-responsive mRNAs by interacting with a stem-loop bipartite 29-33 nucleotide long GAIT element located in the 3'UTR of the target mRNAs. The incorporation of EPRS into the GAIT complex does not affect total protein synthesis.

Recruitment and assembly of the four components occur in two distinct, tightly regulated stages. In the first stage, 2-4 hours post Interferon Gamma (INF- γ) stimulation, hEPRS is phosphorylated by two kinase systems at Ser-886 and Ser-999. INF- γ triggers activation of Cdk5. Cdk5, along with ERK2 and its co-activator p35, phosphorylate EPRS at Ser-886. Parallely, INF- γ also activates the mTORC1 pathway resulting in the activation of downstream p70 ribosomal protein, S6K1, which phosphorylates EPRS on Ser-999. Sequential phosphorylation of EPRS leads to its release from the MARS complex and aids in its interaction with Syncrip via phosphorylated ser-886, resulting in the formation of an inactive pre-GAIT complex. Pre-GAIT complex is incapable of binding to GAIT elements stationed in the 3'UTRs of target mRNAs.

The second stage occurs 12-16 hours post-INF- γ stimulation. RPL13A, a component of 60S Ribosome, is phosphorylated at Ser-77 by DAPK-activated ZIPK and is released from its parental complex. It thereafter associates with GAPDH and the Pre-GAIT complex to form the Interferon Gamma Activated Inhibitor of Translation complex (GAIT complex). This interaction is facilitated by phosphorylated Ser-999 of EPRS. Formation of the GAIT complex brings about a conformational shift in EPRS, exposing its RNA binding sites, i.e. the WHEP domains in the central linker region rendering the complex active. EPRS is the sole GAIT complex constituent that binds to the 3'UTR GAIT element in target mRNAs. This phenomenon leads to translational repression of GAIT element-bearing transcripts.

Formation of the GAIT complex requires two independent and temporally distinct signalling pathways to come together, an early induction of EPRS phosphorylation and a delayed induction of RPL13A phosphorylation. All the GAIT complex components have canonical activities distinct from their function in the GAIT complex. EPRS and RPL13A are ubiquitous and are integral members of the translation machinery. GAPDH is involved in Glycolysis and energy production, but in the GAIT complex, it has a chaperone-like activity.

Ser-886 and Ser-999 in EPRS are conserved from opossums to primates. Ser-886 is absent in mice. Ser-886 is involved in binding with Syncrip upon phosphorylation. Mice have a heterotrimeric GAIT complex devoid of Syncrip. Ser-77 in RPL13A is evolutionarily conserved from yeast to humans. The existence of a functional GAIT complex is validated only in mice and humans. The targets of the GAIT complex show species-specific differences. Mice and human GAIT complexes recognise different structural elements. Thus, the success of bioinformatically detecting GAIT elements is limited by structural diversity.

EPRS has two isoforms in humans. The second isoform encodes for a polypeptide encompassing Glutamyl tRNA Synthetase, followed by the three WHEP domains. It is incapable of forming a GAIT complex as it lacks Ser-999. On the flip side, it competes with the full-length isoform for interaction with the GAIT element, shielding the proinflammatory mRNAs from the GAIT complex, thereby helping in the maintenance of the homeostatic basal level of expression of immune-responsive mRNAs.

GAIT in-dependent noncanonical function: EPRS is also involved in GAIT-independent moonlighting functions like adiposity, antiviral defence and perhaps in breast cancer progression.

Adiposity: In human adipocytes, insulin stimulation activates the mTORC1 pathway leading to the phosphorylation of EPRS at Ser-999 and its subsequent release from the MARS complex. Phosphorylated EPRS associates with the Solute carrier family 27 fatty acid transporter member 1 (FATP1) in the cytoplasm. This interaction promotes translocation and integration of FATP1 in the membrane leading to Long-chain fatty acid uptake.

Antiviral defence: Pattern recognition proteins RIG1 and MDA5 are viral genome sensors. Upon viral infection, RIG1 and MDA5 activate Mito-antiviral signalling proteins (MAVs) which in turn inhibits viral replication. Poly(rC) binding protein (PCBP2) ubiquitinates and degrades MAVs. As an independent event, viral infection triggers phosphorylation of EPRS at Ser990 and its subsequent release from the MARS complex. Phosphorylated EPRS interacts with PCBP2, thereby inhibiting the ubiquitination of MAVs protein.

Breast cancer progression: EPRS acts as a critical regulator of cell proliferation and estrogen signalling in ER-positive breast cancer.

2. Materials & Methods

2.1 SUMO conjugation assay: SUMOylation of EPRS was tested by expressing the target/substrate protein simultaneously with the *Drosophila* SUMO cycle components based on a published protocol (Nie et al., 2019). Target proteins from the *Drosophila* Gold cDNA collection, procured from the *Drosophila* Genome Resource Centre (DGRC), Bloomington, Indiana, were sub-cloned into *pGEX-4T1* (Promega) and *pET-45b* and subsequently sequenced for validation. For visualisation of SUMO conjugation, bacterial lysates were affinity purified using Glutathione-Agarose beads (Invitrogen) or Ni NTA-Agarose beads (Qiagen), run on an SDS-PAGE gel and monitored using mouse anti-GST antibody (sc53909, 1:5000; Santa-Cruz-Biotechnology), Rabbit anti-HA antibody (DW2, 1:3000; Millipore) and mouse anti-6X-His antibody (H1029, 1:1000; SIGMA) using Western blotting. The SUMO-conjugated forms appear as bands of a higher molecular weight.

2.2 SUMO-binding-motif and SIM-motif prediction: Putative SUMO acceptor lysines and SIM-motifs of EPRS were predicted *in-silico*, using Joined Advanced SUMOylation and Sim motif Analyzer (JASSA) tool with cut-off threshold criteria set at "high" (BEAUCLAIR *et al.* 2015).

2.3 Identification of evolutionarily conserved SUMO target lysine residues in-silico: FASTA sequences of EPRS for model organisms belonging to different eukaryotic groups were procured from the Uniport protein database. Multiple sequence alignment (MSA) was done based on homology extension using PSI-COFFEE (CHANG *et al.* 2012). SUMO acceptor lysines were compared across different representative organisms post-alignment.

2.4 Generation of EPRS^{SCR} using CRISPR Cas9 technology: CRISPR Cas9 technology was employed to generate SUMO conjugation-resistant EPRS^{SCR} fly lines. Single guide (sg)-RNA targeting the EPRS coding region in between DNA sequence coding for K1106 and K1198, close to a TTAA site, was designed using CRISPR Optimal Target Finder (COTF;(GRATZ *et al.* 2014)), a web tool for identifying CRISPR target sites and evaluating their specificity. The EPRS gene region was sequenced prior to designing the gRNA to account for SNPs at the sgRNA target sites. The sgRNA was cloned into the *pU6-BbsI-chiRNA* (Addgene # 45946) plasmid. The donor template was assembled using Gibson assembly. The transgenic sgRNA and the donor template were injected in the germline of 640 *Actin 5C Cas9 lig4* (BDSC #) embryos (the Cas9-sgRNA complex is formed in the germline) at the NCBS-CCAMP

transgenic facility, Bangalore, India. 544 adults eclosed, and 453 survived. All the 453 flies obtained were crossed to IIIrd chromosome balancer flies; *w-;*; *Sb/Tb*. 453 fly lines were maintained as separate lines over a *Sb/Tb* balancer since the genomic EPRS is located on the III chromosome. Putative *EPRS*^{SCR} lines were screened for the presence of DsRED in their eyes. Single-fly genomic PCR for the *EPRS* locus was performed on the putative *EPRS*^{SCR} fly lines, and the mutations were confirmed through sequencing.

2.5 Generation of Δ EPRS using CRISPR Cas9 technology: CRISPR Cas9 technology was employed to generate EPRS null fly lines. Single guide (sg)-RNAs targeting the EPRS coding region in the 5'end of Exon 3 and middle region of Exon 6 (adjacent to the region coding for the last WHEP domain in the series) were designed using CRISPR Optimal Target Finder (COTF;(GRATZ *et al.* 2014)), a web tool for identifying CRISPR target sites and evaluating their specificity. The EPRS gene region was sequenced prior to designing the gRNAs to account for SNPs at the sgRNA target sites. The sgRNAs were cloned into the *pU6-BbsI-chiRNA* (Addgene # 45946) plasmid, which was then docked into *y¹ v¹; P{CaryP}attP40* *Drosophila* line (BDSC 36304), by transgenic injections, at the NCBS-CCAMP transgenic facility, Bangalore, India. The transgenic dual sgRNA line was crossed to the *nanos-Cas9* (BDSC 54591) line. The founder male progenies obtained were crossed to *w-;*; *Sb/Tb* balancer females wherein the Cas9-sgRNA complex is formed in the germline. In the next generation, three heterozygous female progenies from each cross were maintained as a separate line over a *Sb/Tb* balancer. Since the genomic EPRS is located on the III chromosome, putative *EPRS* null lines were screened for lethality. Lines showing lethality were chosen for PCR-based confirmation of the deletion. Single-fly genomic PCR for the coding region of *EPRS* was performed on heterozygous animals, and the mutations were confirmed through sequencing.

3. Results

3.1 EPRS is a target for SUMO machinery: Prediction of SUMO conjugation sites (Beauclair *et al.*, 2015) in the EPRS sequence suggests that EPRS has numerous strong consensus SUMO conjugation motifs. Our experimental data and also data from the literature show that EPRS can show upto five SUMO conjugates (Fig. 4.3 A-B), and multiple rounds of mutagenesis followed by *in-bacto* SUMOylation shows that a mutant *EPRS*^{K957R, K1063R, K1083R, K1106R, K1198R} is SUMO conjugation resistant (*EPRS*^{SCR})(Fig.4.3 A). EPRS is part of subcomplex-I (Fig. 4.2 B) in the MARS complex, associating intimately with IRS and AIMP2(Fig 4.2 B). Analysis of the crystal structure of sub-complex-I suggests that the equivalent amino acids in the human

structure are not part of the protein: protein interface with either IRS or AIMP2(5A34,(J Biol Chem 290: 29313)). K957, K1106 and K1198 are part of high-scoring Synergy control (SC)-SUMO target motifs (PVKVKQEKNP, PAVVKKEASP and PVKKEP respectively), at intercalated zones between alpha-helical WHEP domains, in a region that is not conserved (Fig. 4.2C). K1063 and K1083 are part of low-scoring SUMO consensus motifs (LKSE and LKGEYKT, respectively) in a conserved alpha-helical WHEP domain. K1063 is part of a strong consensus motif, whereas K1083 is part of an extended phosphorylation-dependent SUMOylation (PDSM)-SUMO motif. Data from the literature confirms that EPRS^{K957R, K1063R, K1083R, K1106R, K1198R} is SUMO conjugation resistant (EPRS^{SCR}) in S2 cells (*Drosophila* Schneider Cells). In addition, our experimental data show that EPRS is SUMOylated *in-vivo*.

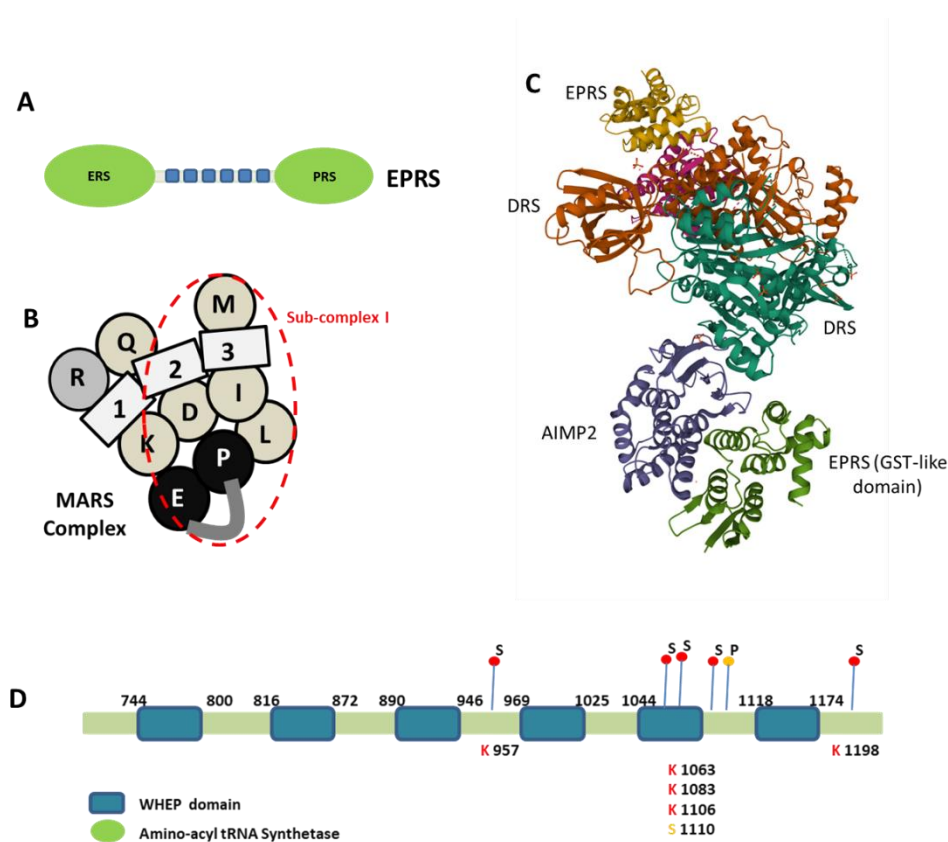


Figure 4.2: A. Schematic of EPRS. *Drosophila* EPRS is a bifunctional tRNA Synthetase. It constitutes an N-terminal Glutamyl tRNA Synthetase (ERS) domain and a C-terminal Prolyl tRNA Synthetase (PRS) domain separated by a central linker region having six tandem WHEP domains. **B. Schematic of the MARS complex.** In the MARS schematic, the ARSs are labelled with a single letter code, with the grey shading denoting mass spectrometric or *in-bacto* evidence for SUMO conjugation. EPRS is part of sub-complex I (marked with a red dashed line), interacting with IRS and AIMP2. **C. The structure for the human DRS:EPRS: AIMP2 complex (PDB-ID 6IY6).** The SUMOylation sites in the fly EPRS are present in the central linker region bearing WHEP domains. The SUMOylation sites were distant from the binding regions of both IRS and AIMP2 and did not appear to interact with any component of the MARS complex. **D.** Based on predictions of SUMO conjugation sites from JASSA (Beauclair et al., 2015) and literature survey.

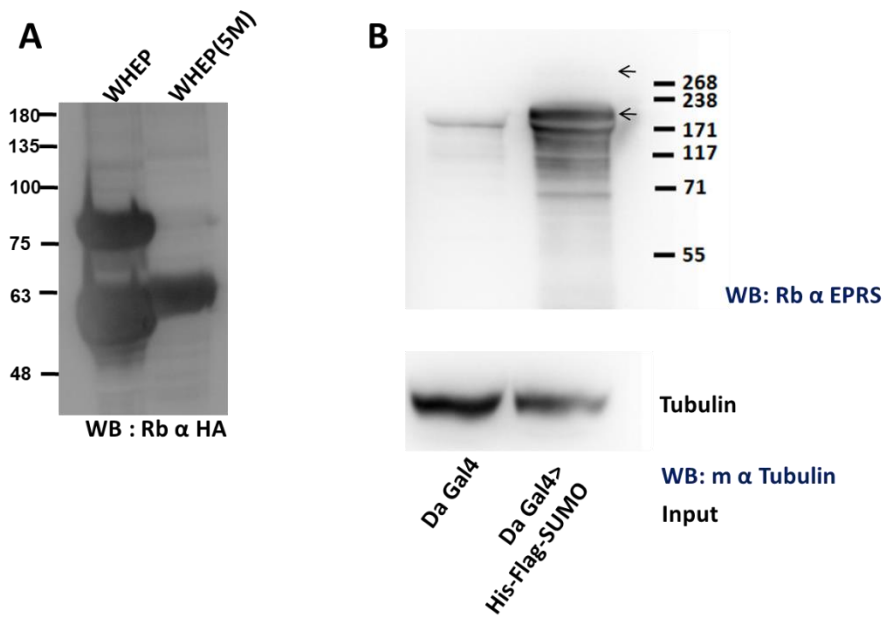


Figure 4.3: A. *EPRS* is SUMO conjugated at K957, K1063, K1083, K1106 and K1198. When co-expressed with *Drosophila* E1, E2 and 6X-His: SUMO, EPRS shows a single extra band running ~20 kDa higher than EPRS itself in Western blots. The band also cross-reacts with an anti-His antibody (data not shown), confirming that it represents a SUMO-conjugated species. The band is not seen when a SUMO(Δ GG) variant, which is unable to conjugate to a substrate, is used. Four more faint bands seen in overexposed Western blots suggest that EPRS may have five SUMOylation sites. Based on predictions of SUMO conjugation sites from JASSA (Beauclair et al., 2015) and a literature survey, mutagenesis of five lysines was carried out. None of the single mutants showed loss of SUMOylation. A quintuple mutant; EPRS^{K957R,K1063R,K1083R,K1106R,K1198R} was resistant to SUMO conjugation. **B. *EPRS* Is SUMO conjugated in-vivo.** *In-vivo*, when His and Flag-tagged SUMO is overexpressed using a ubiquitous Daughterless Gal4, EPRS shows three extra bands running ~20kDa higher or more than EPRS itself in Western blots upon immunoprecipitation of SUMOylated proteins using beads decorated with antibody against Flag tag, raised in rabbit. Tubulin is used as a control for input, and a line expressing Daughterless Gal4 devoid of expression of His-Flag tagged SUMO is used as a master control.

3.2 Generation of *EPRS*^{SCR} using CRISPR Cas9 genome editing: We have used CRISPR Cas9 genome editing to generate SUMO-conjugation resistant mutants of EPRS.

CRISPR, also known as 'Clustered regularly interspaced short palindromic repeats, is a genome engineering method that enables precise modification of any genome. In nature, CRISPR systems provide adaptive immunity in bacteria and archaea. In simple systems, RNAs containing sequences complementary to viral or plasmid DNA interact with DNA nuclease Cas9 to direct sequence-specific cleavage of invading DNA. This process has been harnessed for genomic engineering in several model systems, including *Drosophila*, where a gRNA or a combination of gRNAs can be targeted to the regions in the genome where mutations need to be incorporated during double-stranded breaks (DSBs). These DSBs induce DNA repair. Cells

employ two major pathways to repair double-stranded DNA breaks. It can either be repaired by non-homologous end joining (NHEJ), resulting in insertions and deletions at the breakpoint. This method is used to generate a loss of function alleles. In contrast, homology-directed repair (HDR) uses homologous DNA as a template for DNA synthesis to bridge the gap. By providing a DNA repair template, precise modifications can be induced.

We have used the second approach for generating SUMO conjugation-resistant mutants of EPRS. A transgenic gRNA was generated that would recognise a region within the DNA sequence where the mutations were supposed to be incorporated, and a donor template was provided as a repair template which harboured the desired mutations. The repair template also had a marker in the form of DsRED located downstream to a promoter PXP3 which drives its expression in the eye. DsRED, in turn, is present within a transposon that can be precisely transposed out by using a piggy-BAC transposase, thus resulting in a mutant allele in which EPRS cannot be SUMO-conjugated. This method is referred to as the Scarless mode of CRISPR Cas9 genome editing.

The mutations are distributed in 5' and 3' homology arms such that they are close to the site of double-stranded breaks inflicted by the gRNA, ensuring the successful incorporation of mutations. The four fragments, two parts of the vector and 5'HR and 3'HR are assembled using Gibson assembly (Fig 4.4 A).

The transgenic gRNA and the donor template were injected into the germline of embryos expressing nuclease cas9 under a ubiquitous actin promoter. Founder animals harbouring Cas9-gRNA complexes in their germline were stabilised by using a balancer on the III chromosome, w^+ ; TM3Sb/TM6Tb. In the next generation, only the fly lines positive for the presence of DsRED in their eyes were maintained and sequenced for the genomic locus of EPRS. Nine flies of a single lineage (453 lineages were screened) showed the presence of DsRED in the eyes (Fig. 4.4 B).

Seven out of nine lines showed the presence of all the desired mutations. Three of these nine lines showed aberrant mutations and insertions in addition to the desired mutations (Table 4.1, 4.2). To our dismay, these lines were highly unstable and reverted to wild type within three months. Hence, we could not explore the immune regulation of EPRS in EPRS^{SCR} flies. Hence, in order to study the role of SUMOylation of EPRS *in-vivo*, we opted for the Null-Rescue method wherein EPRS^{WT} and EPRS^{SCR} can be expressed in an EPRS-null (Δ EPRS) background and scored for differential phenotypes.

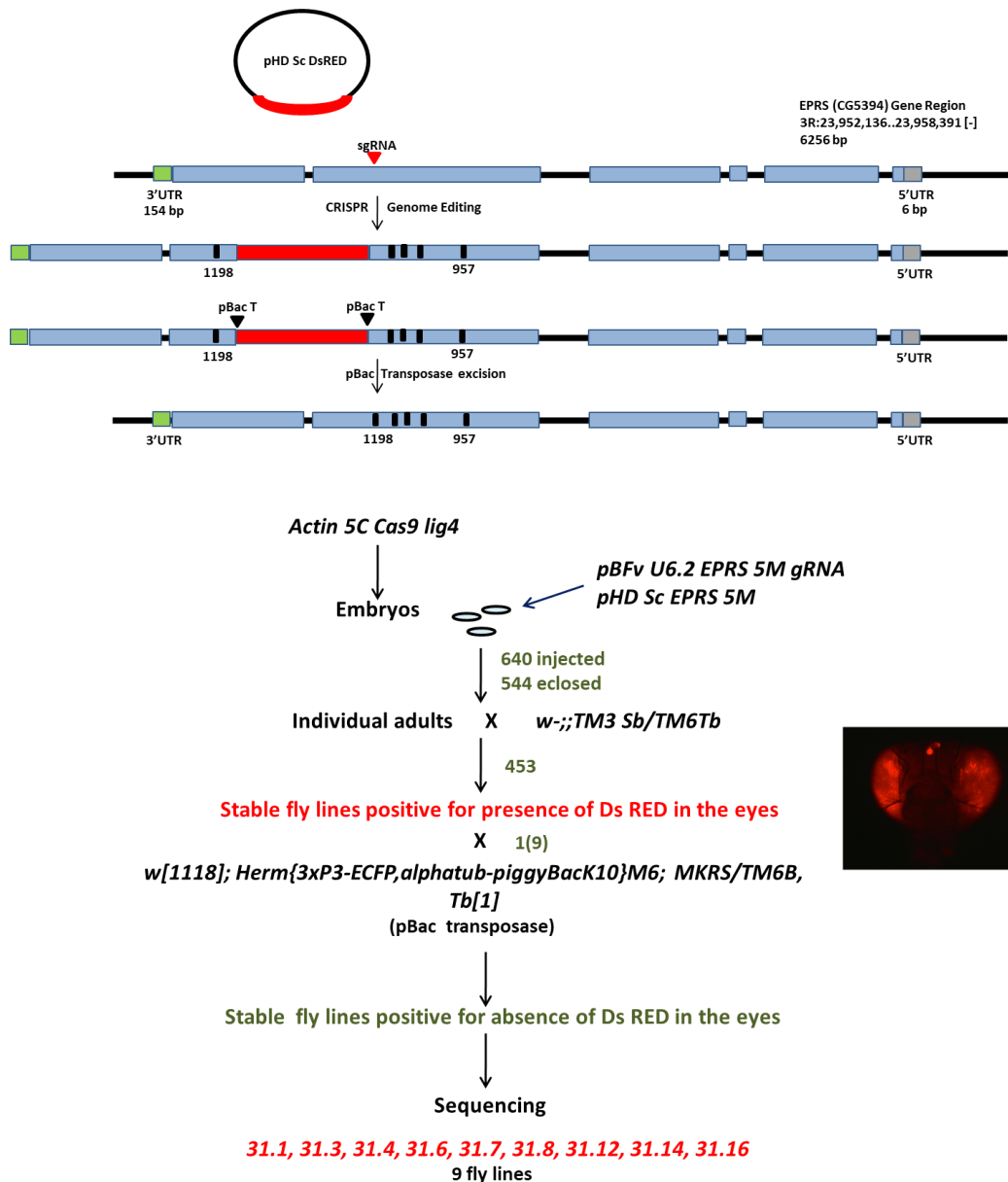


Figure 4.4: A. Design of single gRNA for generation of EPRSSCR. The EPRS gene is located on the IIIrd chromosome. It has six exons, which code for two annotated transcripts via alternative splicing. The full-length transcript spans 5404 bp, and the shorter splice variant spans 3151 bp. A single gRNA is designed in the sixth WHEP domain in the series in-between K1106 and K1198; the last two SUMOylation motifs are adjacent to a TTAA site (pBac transposon insertion site). The goal of the experiment is to replace the wild-type DNA sequence with a SUMO-resistant variant **EPRSSCR**^{K957, K1063, K1083, K1106, K1198->R}. **B. Strategy for generation of EPRSSCR:** Embryos from *Actin 5C cas9 lig4* flies (express nuclease cas9 under a ubiquitous promoter actin 5C) were injected with *pBFv U6.2 EPRS^{5M} gRNA* (codes for a single gRNA which targets a region in the 6th WHEP domain in the series) and a donor template; *pHDSc EPRS^{5M}* (harbours all the mutations necessary to render EPRS SUMO-conjugation resistant). A total of 640 embryos were injected, of which 544 eclosed. The individual adults were crossed to flies having a IIIrd chromosome balancer *w⁻; TM3 Sb/TM6 Tb*. In the next generation, fly lines positive for the presence of DsRED in their eyes were selected. Only nine animals allied with a single lineage (of the 453 lineages screened) were positive for the presence of DsRED in the eye. Single-fly genomic PCR was conducted on these lines to probe for the presence of desired mutations.

Table 4.1: Multiple Sequence Alignment of *Drosophila* EPRS and EPRS^{SCR} lines. The genomic regions of EPRS^{SCR} lines 31.1,31.3,31.4,31.6,31.8,31.12,31.14 and 31.16 were sequenced and the DNA sequence aligned to that of wild-type EPRS.

31.8	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	675
31.16	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	669
EPRS	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	3960
31.3	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	671
31.6	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	671
31.1	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAA <u>T</u> GGAACCAGGTACTG	672
31.14	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	671
31.4	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	671
31.12	TGCTGGCTTTGAAAACGGACTACAAATCTCTAACTGGTCAAGAGTGGAACCAGGTACTG	671

31.8	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	795
31.16	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	789
EPRS	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAAAT	4080
31.3	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	791
31.6	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAG <u>A</u> TTGAGAT	791
31.1	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	792
31.14	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	791
31.4	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	791
31.12	ACGTAGGCAGTGATTGAGCAAGATTCAGGCCCAAGGTGATAAGATCAGGAAATTGAGAT	791

		K1063R
31.8	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTAAGACTCTGCTTGCTTTAGAG	855
31.16	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	849
EPRS	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAAG	4140
31.3	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	851
31.6	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCNGCTTGCTTTAGAG	851
31.1	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	852
31.14	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	851
31.4	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	851
31.12	CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTTTAGAG	851

31.8	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	914
31.16	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	908
EPRS	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	4199
31.3	GAGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	910
31.6	GTGAATATAAAACGCTAAGCGGTAAGGGATTGGACGCCAGACGCTAAATCTGAACCAGCT	911
31.1	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	911
31.14	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	910
31.4	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	910
31.12	GTGAATATAAAACGCTAAGCGGTAAG - GATTGGACGCCAGACGCTAAATCTGAACCAGCT	910
	* *****	
31.8	GTAGTAGAAA - GGAAGCTAGTCCCCTTCGATGGGATCGCCGGCTAAGATGAAATCCCC	973
31.16	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	968
EPRS	GTAGTAAAAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	4259
31.3	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGG - NCGCCAGCTAAGGATGAACTCACC	969
31.6	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATC - - - - -	949
31.1	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	971
31.14	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	970
31.4	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	970
31.12	GTAGTAAGAAAGGAAGCTAGTCCCCTTCGATGGGATCGCCAGCTAAGGATGAACTCACC	970

```

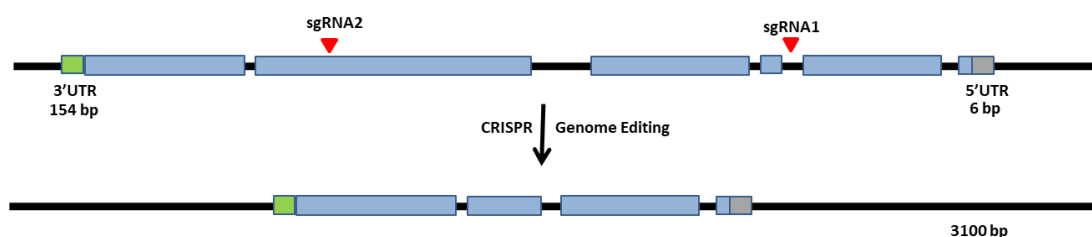
31.8      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 795
31.16     ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 789
EPRS      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAAAT 4080
31.3      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 791
31.6      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAGATTGAGAT 791 AAA(Lys) AGA(Arg)
31.1      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 792
31.14     ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 791
31.4      ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 791
31.12     ACGTAGGCAGTGTATTGAGCAAGATTCAGGCCAAGGTGATAAGATCAGGAAATTGAGAT 791
*****
K1063R
31.8      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGGTAAGACTCTGCTTGCTCTTAGAG 855 GTT(Val) GGT(Gly)
31.16     CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 849
EPRS      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 4140
31.3      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 851 GAA/GAG – Glu
31.6      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCNGCTTGCTCTTAGAG 851
31.1      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 852
31.14     CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 851
31.4      CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 851
31.12     CAGAGAAGGCAGCCAAGAACGTAATCGATCCTGAGGTTAAGACTCTGCTTGCTCTTAGAG 851
*****
K1083R
31.8      GTGAATATAAAACGCTAAGCGGCAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 914 GGT/GGC – Gly
31.16     GTGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 908
EPRS      GTGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 4199
31.3      GAGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 910 GGT/GGA – Gly
31.6      GTGAATATAAAACGCTAAGCGGTAAGGGATTGGACGCCAGACGCTAAATCTGAACCAAGCT 911 Extra Nucleotide
31.1      GTGAATATAAAACGCTAAGCGGCAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 911 GGT/GGC – Gly
31.14     GTGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 910
31.4      GTGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 910
31.12     GTGAATATAAAACGCTAAGCGGTAAG-GATTGGACGCCAGACGCTAAATCTGAACCAAGCT 910
* *****
31.8      GTAGTAGAAA-GGAAGCTAGTCCCGTTTCGATGGGATCGCCGGCTAAGATGAAATCCCC 973 AAA(Lys) AGA(Arg) GAA(Glu)
31.16     GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 968
EPRS      GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 4259
31.3      GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGC-NCGCCAGCTAAGGATGAACTCACC 969
31.6      GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGGATC----- 949
31.1      GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 971
31.14     GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 970
31.4      GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 970
31.12     GTAGTAAGAAAGGAAGCTAGTCCCGTTTCGATGGCATCGCCAGCTAAGGATGAACTCACC 970
*****
K1106R

```

Table 4.2: Tabulation of mutations confirmed upon sequencing, listed against the corresponding lines positive for the presence of DsRED in their eyes.

Line No.	Allelic mutations
31.1	957 1063 1083 1106 1198
31.3	957 1063 1083 1106 1198
31.4	957 1063 1083 1106 1198
31.6	957 1063 1083 1106 1198 with additional mutations
31.7	957 1063 with additional mutations and insertions
31.8	957 1063 1083 1106 with additional mutations
31.12	957 1063 1083 1106 1198
31.14	957 1063 1083 1106 1198
31.16	957 1063 1083 1106 1198

3.3 Generation of a $\Delta EPRS$ line using CRISPR Cas9 genome editing. The UAS-Gal4 system is an ideal system to express EPRS^{WT} and EPRS^{SCR} in an EPRS-null ($\Delta EPRS$) background. Since the $\Delta EPRS$ line is not available, as a first step, we used CRISPR Cas9 genome editing to generate the same. A transgenic dual-guide RNA line (*U6.2b-EPRS^{dual-gRNA}*) was created (See Materials and Methods) to express gRNA that would recognise the 5' end of exon-3 and the central region of exon-6 (adjacent to the region coding for the last WHEP domain in the series) of the EPRS gene (inverted red triangles, Fig. 4.4). Our goal was to remove a major portion of the coding region to create a $\Delta EPRS$ animal. The DNA sequence coding for the EPRS gene also codes for an antisense RNA asRNA: CR46092 in an opposite reading frame (encoded by exons 1-2). The *U6.2b-EPRS^{dual-gRNA}* line was crossed to a *nos-Cas9* animal (Fig. 4.4), and five hundred twenty-three lines were stabilised by balancing the putative nulls over a III chromosome balancer, TM6-Tb (TM6B). Of these lines, one (402c/TM6-Tb) is homozygous lethal, which is indicative of a successful excision of the *EPRS* locus since the absence of the *EPRS* on the III chromosome would lead to lethality. Twelve other lines show melanotic tumours but are homozygous viable, indicating the gRNA activity caused minor changes to the sequence of the wild-type genome in the sites targeted by gRNA, and these modifications presumably led to the generation of mutants of the EPRS gene. Single-fly genomic PCR was conducted on the homozygous lethal line; the genomic PCR product showed the expected 3.1 kb deletion that would be a consequence of the removal of the desired EPRS genomic region. To corroborate the observed lethality, we sequenced the genomic region of the line 402c/TM6-Tb. We found that the coding region was deleted in 402c/TM6-Tb, and this modification presumably led to the generation of a functional EPRS-null ($\Delta EPRS$). $\Delta EPRS$ die during IInd instar larval stages, with embryonic survival till 1st Instar presumably driven by maternal EPRS (Fig. 4.5).



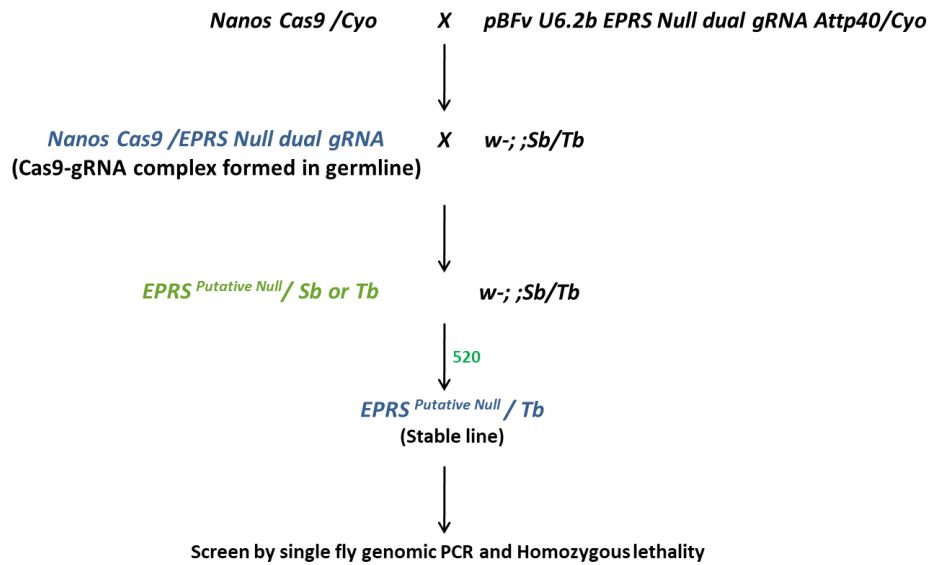


Figure 4.5: An *EPRS*-null ($\Delta EPRS$) line generated using CRISPR Cas9 genome editing. A. Design of the dual guide-RNA for excision of the *EPRS* locus. The *EPRS* gene is located on the III chromosome. It has six exons, which code for two annotated transcripts via alternative splicing. The longer isoform spans 5405 bp, and the shorter one spans 3151 bp. Two gRNAs (inverted red triangles) were designed in the 5' end of exon 3 and the central region of exon 6 (adjacent to the region coding for the last WHEP domain in the series) of the *EPRS* gene. Our goal was to excise most of the coding region and generate a *EPRS* null line. **B. Excision of the *EPRS* locus to generate a $\Delta EPRS$ line.** A *U6.2b-EPRS*^{dual-gRNA} line was generated (Materials & Methods) and crossed to *nos-Cas9* animals. Five hundred twenty-three lines were balanced over a third chromosome balancer and screened for homozygous lethality. One line (402c/TM6-Tb) was homozygous lethal and had the expected deletion in the *EPRS* locus, based on PCR.

4. Discussion

The goal of this project, as executed in the previous chapter for RRS, was to generate a *Drosophila* line where EPRS could not be SUMO conjugated. Once successful, the *EPRS^{SCR}* line would be a fantastic reagent to address specific roles for SUMO conjugation of EPRS in animal development and in the immune response. Unfortunately, despite my successful creation of transgenic constructs as well as the detection of Ds Red positive transgenics, our experiment failed after ~2 years of effort. The reasons for the failure remain elusive. At the point of selection of the mutants, the DsRed positive flies clearly had the designed mutations incorporated, indicating that we had been successful in our attempts. This was confirmed by DNA sequencing of the *EPRS* locus. At this stage in the experiment, the flies were balanced against the TM3Sb or TM6Tb balancer chromosomes, both being used effectively in fly biology for over 70 years. In spite of this, each and every line lost the mutations with time, and by the step where we conducted the DsRed excision experiment, we were unable to detect the mutations both in the excised lines as well as the seemingly balanced stocks that we had maintained.

The failure to generate *EPRS^{SCR}* was a major disappointment and a significant loss of time. We attempted to repeat the steps for the generation of *EPRS^{SCR}* by redoing injections and screens, but these did not lead to any Ds Red positive transgenics. We then attempted the alternative experiment, as carried out successfully for RRS, which was a generation of a *EPRS* null animal, again using CRISPR/Cas9 genome editing. In this step, we were successful, and this reagent would allow us to execute a null: rescue experiment sometime in the near future. We plan to clone *EPRS* and *EPRS^{SCR}* constructs under the UAS promoter and use the UAS/Gal4 system to generate flies where EPRS is SUMO conjugation resistant.

5. Acknowledgements

I thank the Fly facility at the National Centre for Biological Sciences (NCBS), Bangalore, for embryonic injections. Dr. Deepti Trivedi for her input on the design of the CRISPR/Cas9 genomic constructs for EPRS.

References

- Arif, Abul, Fulvia Terenzi, Alka A. Potdar, Jie Jia, Jessica Sacks, Arnab China, Dalia Halawani, et al. 2017. "EPRS Is a Critical mTORC1-S6K1 Effector That Influences Adiposity in Mice." *Nature* 542 (7641): 357–61.
- Barrangou, Rodolphe, and John van der Oost. 2012. *CRISPR-Cas Systems: RNA-Mediated Adaptive Immunity in Bacteria and Archaea*. Springer Science & Business Media.
- Beauclair, Guillaume, Antoine Bridier-Nahmias, Jean-François Zagury, Ali Saïb, and Alessia Zamborlini. 2015. "JASSA: A Comprehensive Tool for Prediction of SUMOylation Sites and SIMs." *Bioinformatics* 31 (21): 3483–91.
- Cahuzac, B. 2000. "A Recurrent RNA-Binding Domain Is Appended to Eukaryotic Aminoacyl-tRNA Synthetases." *The EMBO Journal*. <https://doi.org/10.1093/emboj/19.3.445>.
- Cerini, C., M. Semeriva, and D. Gratecos. 1997. "Evolution of the Aminoacyl-tRNA Synthetase Family and the Organization of the Drosophila Glutamyl-Prolyl-tRNA Synthetase Gene. Intron/exon Structure of the Gene, Control of Expression of the Two mRNAs, Selective Advantage of the Multienzyme Complex." *European Journal of Biochemistry / FEBS* 244 (1): 176–85.
- Chang, Jia-Ming, Paolo Di Tommaso, Jean-François Taly, and Cedric Notredame. 2012. "Accurate Multiple Sequence Alignment of Transmembrane Proteins with PSI-Coffee." *BMC Bioinformatics* 13 Suppl 4 (March): S1.
- Cho, Ha Yeon, Seo Jin Maeng, Hyo Je Cho, Yoon Seo Choi, Jeong Min Chung, Sangmin Lee, Hoi Kyoung Kim, et al. 2015. "Assembly of Multi-tRNA Synthetase Complex via Heterotetrameric Glutathione Transferase-Homology Domains." *The Journal of Biological Chemistry* 290 (49): 29313–28.
- Gratz, Scott J., Fiona P. Ukken, C. Dustin Rubinstein, Gene Thiede, Laura K. Donohue, Alexander M. Cummings, and Kate M. O'Connor-Giles. 2014. "Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila." *Genetics* 196 (4): 961–71.
- Jeong, E. J., G. S. Hwang, K. H. Kim, M. J. Kim, S. Kim, and K. S. Kim. 2000. "Structural Analysis of Multifunctional Peptide Motifs in Human Bifunctional tRNA Synthetase: Identification of RNA-Binding Residues and Functional Implications for Tandem Repeats." *Biochemistry* 39 (51): 15775–82.
- Jia, Jie, Abul Arif, Partho S. Ray, and Paul L. Fox. 2008. "WHEP Domains Direct Noncanonical Function of Glutamyl-Prolyl tRNA Synthetase in Translational Control of Gene Expression." *Molecular Cell* 29 (6): 679–90.
- Kang, J., T. Kim, Y. G. Ko, S. B. Rho, S. G. Park, M. J. Kim, H. J. Kwon, and S. Kim. 2000. "Heat Shock Protein 90 Mediates Protein-Protein Interactions between Human Aminoacyl-tRNA Synthetases." *The Journal of Biological Chemistry* 275 (41): 31682–88.
- Katsy, Igor, Minghui Wang, Won Min Song, Xianxiao Zhou, Yongzhong Zhao, Sun Park, Jun Zhu, Bin Zhang, and Hanna Y. Irie. 2016. "EPRS Is a Critical Regulator of Cell Proliferation and Estrogen Signaling in ER+ Breast Cancer." *Oncotarget* 7 (43): 69592–605.
- Kim, Jong Hyun, Jung Min Han, and Sunghoon Kim. 2014. "Protein-Protein Interactions and Multi-Component Complexes of Aminoacyl-tRNA Synthetases." *Topics in Current Chemistry* 344: 119–44.
- Lamb, Abigail M., Elizabeth A. Walker, and Patricia J. Wittkopp. 2017. "Tools and Strategies for Scarless Allele Replacement in Drosophila Using CRISPR/Cas9." *Fly* 11 (1): 53–64.
- Lee, Eun-Young, Hyun-Cheol Lee, Hyun-Kwan Kim, Song Yee Jang, Seong-Jun Park, Yong-Hoon Kim, Jong Hwan Kim, et al. 2016. "Infection-Specific Phosphorylation of Glutamyl-Prolyl tRNA Synthetase Induces Antiviral Immunity." *Nature Immunology* 17 (11): 1252–62.
- Lee, Yi-Hsueh, Chia-Pei Chang, Yu-Ju Cheng, Yi-Yi Kuo, Yeong-Shin Lin, and Chien-Chia Wang. 2017. "Evolutionary Gain of Highly Divergent tRNA Specificities by Two Isoforms of Human Histidyl-tRNA Synthetase." *Cellular and Molecular Life Sciences: CMLS* 74 (14): 2663–77.
- Mendes, Marisa I., Mariana Gutierrez Salazar, Kether Guerrero, Isabelle Thiffault, Gajja S. Salomons, Laurence Gauquelin, Luan T. Tran, et al. 2018. "Bi-Allelic Mutations in EPRS,

- Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy.” *American Journal of Human Genetics* 102 (4): 676–84.
- Nie, Minghua, Yongming Xie, Joseph A. Loo, and Albert J. Courey. 2009. “Genetic and Proteomic Evidence for Roles of Drosophila SUMO in Cell Cycle Control, Ras Signaling, and Early Pattern Formation.” *PLoS One* 4 (6): e5905.
- Norcum, Mona Trempe, and J. David Dignam. 1999. “Localisation Of Glutamyl/Prolyl-Trna Synthetase Within The Eukaryotic Multienzyme Complex By Immunoelectron Microscopy.” *Microscopy and Microanalysis*. <https://doi.org/10.1017/s1431927600019796>.
- Quevillon, S., J. C. Robinson, E. Berthonneau, M. Siatecka, and M. Mirande. 1999. “Macromolecular Assemblage of Aminoacyl-tRNA Synthetases: Identification of Protein-Protein Interactions and Characterization of a Core Protein.” *Journal of Molecular Biology* 285 (1): 183–95.
- Ray, Partho Sarothi, and Paul L. Fox. 2014. “Origin and Evolution of Glutamyl-Prolyl tRNA Synthetase WHEP Domains Reveal Evolutionary Relationships within Holozoa.” *PLoS One* 9 (6): e98493.
- Ray, Partho Sarothi, James C. Sullivan, Jie Jia, John Francis, John R. Finnerty, and Paul L. Fox. 2011. “Evolution of Function of a Fused Metazoan tRNA Synthetase.” *Molecular Biology and Evolution* 28 (1): 437–47.
- Rho, S. B., J. S. Lee, E. J. Jeong, K. S. Kim, Y. G. Kim, and S. Kim. 1998. “A Multifunctional Repeated Motif Is Present in Human Bifunctional tRNA Synthetase.” *The Journal of Biological Chemistry* 273 (18): 11267–73.
- Robinson, J. C., P. Kerjan, and M. Mirande. 2000. “Macromolecular Assemblage of Aminoacyl-tRNA Synthetases: Quantitative Analysis of Protein-Protein Interactions and Mechanism of Complex Assembly.” *Journal of Molecular Biology* 304 (5): 983–94.
- Shin, Chinho, Geum-Sook Hwang, Hee-Chul Ahn, Sunghoon Kim, and Key-Sun Kim. 2015. “(1)H, (13)C and (15)N Resonance Assignment of WHEP Domains of Human Glutamyl-Prolyl tRNA Synthetase.” *Biomolecular NMR Assignments* 9 (1): 25–30.
- Singh, Vijai, and Pawan K. Dhar. 2020. *Genome Engineering via CRISPR-Cas9 System*. Academic Press.
- Smith, Matthew, Vinay Bhaskar, Joseph Fernandez, and Albert J. Courey. 2004. “Drosophila Ulp1, a Nuclear Pore-Associated SUMO Protease, Prevents Accumulation of Cytoplasmic SUMO Conjugates.” *The Journal of Biological Chemistry* 279 (42): 43805–14.
- Cahuzac, B. 2000. “A Recurrent RNA-Binding Domain Is Appended to Eukaryotic Aminoacyl-tRNA Synthetases.” *The EMBO Journal*. <https://doi.org/10.1093/emboj/19.3.445>.
- Cerini, C., M. Semeriva, and D. Gratecos. 1997. “Evolution of the Aminoacyl-tRNA Synthetase Family and the Organization of the Drosophila Glutamyl-Prolyl-tRNA Synthetase Gene. Intron/exon Structure of the Gene, Control of Expression of the Two mRNAs, Selective Advantage of the Multienzyme Complex.” *European Journal of Biochemistry / FEBS* 244 (1): 176–85.
- Jeong, E. J., G. S. Hwang, K. H. Kim, M. J. Kim, S. Kim, and K. S. Kim. 2000. “Structural Analysis of Multifunctional Peptide Motifs in Human Bifunctional tRNA Synthetase: Identification of RNA-Binding Residues and Functional Implications for Tandem Repeats.” *Biochemistry* 39 (51): 15775–82.
- Jia, Jie, Abul Arif, Partho S. Ray, and Paul L. Fox. 2008. “WHEP Domains Direct Noncanonical Function of Glutamyl-Prolyl tRNA Synthetase in Translational Control of Gene Expression.” *Molecular Cell* 29 (6): 679–90.
- Lee, Yi-Hsueh, Chia-Pei Chang, Yu-Ju Cheng, Yi-Yi Kuo, Yeong-Shin Lin, and Chien-Chia Wang. 2017. “Evolutionary Gain of Highly Divergent tRNA Specificities by Two Isoforms of Human Histidyl-tRNA Synthetase.” *Cellular and Molecular Life Sciences: CMLS* 74 (14): 2663–77.
- Mendes, Marisa I., Mariana Gutierrez Salazar, Kether Guerrero, Isabelle Thiffault, Gajja S. Salomons, Laurence Gauquelin, Luan T. Tran, et al. 2018. “Bi-Allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy.” *American Journal of Human Genetics* 102 (4): 676–84.

- Norcum, Mona Trempe, and J. David Dignam. 1999. "Localisation Of Glutamyl/Prolyl-Trna Synthetase Within The Eukaryotic Multienzyme Complex By Immunoelectron Microscopy." *Microscopy and Microanalysis*. <https://doi.org/10.1017/s1431927600019796>.
- Quevillon, S., J. C. Robinson, E. Berthonneau, M. Siatecka, and M. Mirande. 1999. "Macromolecular Assemblage of Aminoacyl-tRNA Synthetases: Identification of Protein-Protein Interactions and Characterization of a Core Protein." *Journal of Molecular Biology* 285 (1): 183–95.
- Ray, Partho Sarothi, and Paul L. Fox. 2014. "Origin and Evolution of Glutamyl-Prolyl tRNA Synthetase WHEP Domains Reveal Evolutionary Relationships within Holozoa." *PloS One* 9 (6): e98493.
- Ray, Partho Sarothi, James C. Sullivan, Jie Jia, John Francis, John R. Finnerty, and Paul L. Fox. 2011. "Evolution of Function of a Fused Metazoan tRNA Synthetase." *Molecular Biology and Evolution* 28 (1): 437–47.
- Rho, S. B., J. S. Lee, E. J. Jeong, K. S. Kim, Y. G. Kim, and S. Kim. 1998. "A Multifunctional Repeated Motif Is Present in Human Bifunctional tRNA Synthetase." *The Journal of Biological Chemistry* 273 (18): 11267–73.
- Robinson, J. C., P. Kerjan, and M. Mirande. 2000. "Macromolecular Assemblage of Aminoacyl-tRNA Synthetases: Quantitative Analysis of Protein-Protein Interactions and Mechanism of Complex Assembly." *Journal of Molecular Biology* 304 (5): 983–94.
- Shin, Chinho, Geum-Sook Hwang, Hee-Chul Ahn, Sunghoon Kim, and Key-Sun Kim. 2015. "(1)H, (13)C and (15)N Resonance Assignment of WHEP Domains of Human Glutamyl-Prolyl tRNA Synthetase." *Biomolecular NMR Assignments* 9 (1): 25–30.

Appendix I

Fig A.3.1 Multiple Sequence Alignment of *Drosophila* RRS and Δ RRS lines. The genomic regions of Δ RRS lines 6B1 and 18B1 were sequenced and the DNA sequence aligned to that of wild-type RRS.

A. Forward and reverse sequencing of line 6B1.

Forward Primer (+ strand; RRS gene region) **LINE 6B1**

```

RRS      -----CCATGTGGCATCGTTGTGTTGTTGTTATTAATCAGCGGAGCAACAGAAGTCA 52
6B1      ACTTGCCACCATGTGGCATCGTTGTGTTG-----GCGGAGCAACAGAAGTCA
          *****
RRS      CAATTTGGAAACATGTCCGAGCTAAATATGGAGCTGAAAAAACTCAGGGAGCTGGTAAGC 112
6B1      CAATTTGGAAACATGTCCGAGCTAAATATGGAGCTGAAAAAACTCAGGGAGCTGGTAAGC
          *****
RRS      GAATAGCTGTTACTTTACGCCAACATTTTGACATGACATGCCGGCTTTAAAGGAAGTCAA 172
6B1      GAATAGCTGTTACTTTACGCCAACATTTTGACATGACATGCCGGCTTTAAAGGAAGTCAA
          *****
RRS      GACCCAAGGCCTTGCCGCCAGAATACAACTGCCAAAAGCGGTGAACAGTTGGACGTCGA 232
6B1      GACCCAAGGCCTTGCCGCCAGAATACAACTGCCAAAAGCGGTGAACAGTTGGACGTCGA
          *****
RRS      TCTTGTTTCAGCTTCAAATTGAAAATAAGAAGCTGAAGAACCGCCTGTTTATCCTAAAGAA 292
6B1      TCTTGTTTCAGCTTCAAATTGAAAATAAGAAGCTGAAGAACCGCCTGTTTATCCTAAAGAA
          *****
RRS      GGTGAGTTTATGACCCCAATGTGTTTAAGCTTAAGAATTTTCATGTCTTTTATTTCCCTG 352
6B1      GGTGAGTTTATGACCCCAATGTGTTTAAGCTTAAGAATTTTCATGTCTTTTATTTCCCTG
          *****
RRS      TTAGTCCATTGCTGAGGAATCAACTGCCGCCGGCGGCGACGTTTCGAAGCCCAAGGAATC 412
6B1      TTAGTCCATTGCTGAGGAATCAACTGCCGCCGGCGGCGACGTTTCGAAGCCCAAGGAATC
          *****
RRS      CTCTTCGATCACCGAACACCTGGAAAGCGTCTTTCGCCAGGCGATTGCATCAGCTTTCCC 472
6B1      CTCTTCGATCACCGAACACCTGGAAAGCGTCTTTCGCCAGGCGATTGCATCAGCTTTCCC
          *****
RRS      GGAATTCAGAGATACGCCTGTTATAATTGCACCAGTTAATAGTACGTCTGCGAAATTCGG 532
6B1      GGAATTCAGAGATACGCCTGTTATAATTGCACCAGTTAATAGTACGTCTGCGAAATTCGG
          *****
RRS      CGACTATCAGTGCAACAATGCCATGGGATTGTCCAAGAAGCTGAAAGAGAAGGGCATTA 592
6B1      CGACTATCAGTGCAACAATGCCATGGGATTGTCCAAGAAGCTGAAAGAGAAGGGCATTA
          *****
RRS      TAAAGCACCACGTGATATTGCAACCGAGTTGAAAGGACACTGCCAGCATCACCAATCAT 652
6B1      TAAAGCACCACGTGATATTGCAACCGAGTTGAAAGGACACTGCCAGCATCACCAATCAT
          *****

```

Reverse Primer (+ strand; RRS gene region) LINE 6B1

```

RRS      ACCAATTTACCCGAAATACTGAAGAAGACCAACATTGTGTTGGACCACGAAAAGGAATGG 2040
6B1      ACCAATTTACCCGAAATACTGAAGAAGACCAACATTGTGTTGGACCACGAAAAGGAATGG
          *****

RRS      AAGCTGGCTAAGACTCTACTGAAACTCCACGACATACTCATCAAGTGCTCAAAGGAACTT 2100
6B1      AAGCTGGCTAAGACTCTACTGAAACTCCACGACATACTCATCAAGTGCTCAAAGGAACTT
          *****

RRS      TTCCTGCACTTCCTG-----TGCGAGTTTTGCTTCGAGGTGTGCACAGTGTTCACCGAA 2154
6B1      TTCCTGCACTTCCCTGTGTTTTTGCGAGTTTTGCTTCGAGGTTGTGCACAGTGTTCACCGAA
          *****

RRS      TTCTATGACTCTTGTTATTGCATCGAAAAGAACAACAAGGCGATATTATTGGGGTCAAT 2214
6B1      TTCTATGACTCTTGTTATTGCATCGAAAAGAACAACAAGGCGATATTATCGGGGTCAAT
          *****

RRS      CATAGCCGAATTCTATTGTGCGAGGCAACTGCGGCTGTGTTGCGCCAATGCTTTTATATA 2274
6B1      CATAGCCGAATTCTATTGTGCGAGGCAACTGCGGCTGTGTTGCGCCAATGCTTTTATATA
          *****

RRS      CTAGGCCTTAAACCAGTTTCGAAAATGTAAAAAGTTCTACTACTGACATCATGTGTACGC 2334
6B1      CTAGGCCTTAAACCAGTTTCGAAAATGTAAAAAGTTCTACTACTGACATCATGTGTACGC
          *****

```

B. Forward and reverse sequencing of line 18B1.

Forward Primer (+ strand; RRS gene region) **LINE 18B1**

```

RRS      -----CCATGTGG 8
18B1     GGANCCAGCTGTTATCGTTATCGATAGGCGACGTGTGCACACTACTTGCCATCCATGTGG
          *****

RRS      CATCGTTGTGTTGTGTTATTAA-----TCAGCGGAGCAACAGAAGTCACAATT 57
18B1     CATCGTTGTGTTGTGTTATTAGCGTTAACGCTAATAGCGGAGCAACAGAAGTCACAATT
          *****

RRS      TGGAAACATGTCGAGCTAAATATGGAGCTGAAAAAACTCAGGGAGCTGGTAAGCGAATA 117
18B1     TGGAAACATGTCGAGCTAAATATGGAGCTGAAAAAACTCAGGGAGCTGGTAAGCGAATA
          *****

RRS      GCTGTTACTTTACGCCAACATTTTGACATGACATGCCGGCTTTAAAGGAACGAAGACCC 177
18B1     GCTGTTACTTTACGCCAACATTTTGACATGACATGCCGGCTTTAAAGGAACGAAGACCC
          *****

RRS      AAGGCCTTGCCGCCAGAATACAACTGCCAAAAGCGGTGAACAGTTGGACGTCGATCTTG 237
18B1     AAGGCCTTGCCGCCAGAATACAACTGCCAAAAGCGGTGAACAGTTGGACGTCGATCTTG
          *****

RRS      TTCAGCTTCAAATGAAAATAAGAAGCTGAAGAACCCTGTTTATCCTAAAGAAGGTGA 297
18B1     TTCAGCTTCAAATGAAAATAAGAAGCTGAAGAACCCTGTTTATCCTAAAGAAGGTGA
          *****

RRS      GTTTATGACCCCCAATGTGTTTAAGCTTAAGAATTTTCATGTCTTTTATTTCCCTGTTAGT 357
18B1     GTTTATGACCCCCAATGTGTTTAAGCTTAAGAATTTTCATGTCTTTTATTTCCCTGTTAGT
          *****

RRS      CCATTGCTGAGGAATCAACTGCCGCCGGCGGCGACGTTTCGAAGCCCAAGGAATCCTCTT 417
18B1     CCATTGCTGAGGAATCAACTGCCGCCGGCGGCGACGTTTCGAAGCCCAAGGAATCCTCTT
          *****

RRS      CGATCACCGAACACCTGGAAGCGTCTTTCGCCAGGCGATTGCATCAGCTTTCCCGGAAT 477
18B1     CGATCACCGAACACCTGGAAGCGTCTTTCGCCAGGCGATTGCATCAGCTTTCCCGGAAT
          *****

RRS      TCAGAGATACGCTGTTATAATTCACCAGTTAATAGTACGTCTGCGAAATTCGGCGACT 537
18B1     TCAGAGATACGCTGTTATAATTCACCAGTTAATAGTACGTCTGCGAAATTCGGCGACT
          *****

RRS      ATCAGTGCAACAATGCCATGGGATTGTCCAAGAAGCTGAAAGAGAAGGGCATTAAATAAAG 597
18B1     ATCAGTGCAACAATGCCATGGGATTGTCCAAGAAGCTGAAAGAGAAGGGCATTAAATAAAG
          *****

RRS      CACCACGTGATATTGCAACCGAGTTGAAAGGACACTGCCAGCATCACCAATCATTGAAA 657
18B1     CACCACGTGATATTGCAACCGAGTTGAAAGGACACTGCCAGCATCACCAATCATTGAAA
          *****

RRS      AGCTGGAAATTGCCGGAGCTGGCTTCGTAAACGTGTTTCCTTAGCAAGTGAGTGGCATAAG 717
18B1     AGCTGGAAATTGCCGGAGCTGGCTTCGTAAACGTGTTTCCTTAGCAAGTGAGTGGCATAAG
          *****

RRS      TACACAAAGTCCATTATAAACATTAATAATTTTGCAGAGATTATGCATCTTTA 777
18B1     TACACAAAGTCCATTATAAACATTAATAATTTTGCAGAGATTATGCATCTTTA
          *****

```

ATC, 6 bp
upstream of PAM
site of gRNA
sequence is
replaced by a 14
bp insertion in 5'
UTR.

Reverse Primer (Sequencing Run 1) (- strand; RRS gene region) **LINE 18B1**

```

RRS      -----ATTCGCGTA 2329
18B1a    ACCAATTCCTTTCAAGGTATATAATTTACTTTTATAAAATGTTTCATTTTATTCGCGTA
          *****
          Translational Stop
RRS      CACATGATGTCAGTAGTAGAACTTTTTACATTTTCGAAACTGGTTTAAGGCCCTAGTATAT 2269
18B1a    CACATGATGTCAGTAGTAGAACTTTTTACATTTTCGAAACTGGTTTAAGGCCCTAGTATAT
          *****
RRS      AAAAGCATTGGCGCAACACAGCCGCGAGTTGCCCTCGCACAAATAGAATTCGGCTATGATTGA 2209
18B1a    AAAAGCATTGGCGCAACACAGCCGCGAGTTGCCCTCGCACAAATAGAATTCGGCTATGATTGA
          *****
RRS      CCCCAATAATATCGCCTTGTTTGTCTTTTCGATGCAATAACAAGAGTCATAGAATTCGG 2149
18B1a    CCCCGATAATATCGCCTTGTTTGTCTTTTCGATGCAATAACAAGAGTCATAGAATTCGG
          *****
          sgRNA1
RRS      TGAACACTGTGCACA CCTCGAAGCAAACTCGCACAGGAAGTGCAGGAAAAGTTCCTTTG 2089
18B1a    TGAACACTGTGCACCTCGAAGCAAACTCGCACAGAAGTGCAGGAAAAGTTCCTTTG
          *****
          * * * * *
RRS      AGCACTTGATGAGTATGTCGTGGAGTTTCAGTAGAGTCTTAGCCAGCTTCCATTCCTTTT 2029
18B1a    ACTTCTTGATCATTATGACTATGACTTGCAGTATAGTCTTAGCCATCTTCCATTCCTTTT
          * * * * *
RRS      CGTGGTCCAACACAATGTTGGTCTTCTTCAGTATTT----- 1969
18B1a    CGTGTCCAGCACCATGTTGGTGTTCCTTCATCTCTATTACATTGATGATGCTTAAAT
          **** * * * * *

```

Sequencing fails
after sgRNA1
(attempt 1)Reverse Primer (Sequencing Run 2) (- strand; RRS gene region) **LINE 18B1**

```

RRS      -----TTATTCGCGTACACATGATGTCAGTAGTAGAACTTTTTACATTTTCGAAACTGG 2083
18B1b    TTCATTTTATTCGCGTACACATGATGTCAGTAGTAGAACTTTTTACATTTTCGAAACTGG
          *****
          Translational Stop
RRS      TTTAAGGCCCTAGTATATAAAAGCATTGGCGCAACACAGCCGCGAGTTGCCCTCGCACAAATAG 2023
18B1b    TTTAAGGCCCTAGTATATAAAAGCATTGGCGCAACACAGCCGCGAGTTGCCCTCGCACAAATAG
          *****
RRS      AATTCGGCTATGATTGACCCCAATAATATCGCCTTGTTTGTCTTTTCGATGCAATAACA 1963
18B1b    AATTCGGCTATGATTGACCCCGATAATATCGCCTTGTTTGTCTTTTCGATGCAATAACA
          *****
          sgRNA1
RRS      AGAGTCATAGAATTCGGTGAACACTGTGCACA CCTCGAAGCAAACTCGCACAGGAAGTG 1903
18B1b    AGAGTCATAGAATTCGGTGAACACTGTGCACACCTCGAAGCAAACTCGCACAGAAGTGC
          *****
          * * * * *
RRS      CAGGAAAAGTTCCTTTGAGCACTTGTGAGTATGTCGTGGAGTTTCAGTAGAGTCTTAGC 1843
18B1b    GAAGAGCAGGACCTTTGACCTCTAGATCATGATGACTAGGACTTGCAGTATAGTCATAGC
          * * * * *
RRS      CAGCTTCCATTCCTTTTCGTGGTCCAACACAATGTTGGTCTTCTTCAGTATTTCCGGGTAA 1783
18B1b    CTTATTCATTCCTTTTCCTGGTCCAGCACCATGTCGGTGTGGTCTTCATTTCTATTAC
          * * * * *
RRS      ATTGGTGAAATCTTCGCCAGAGTTT-----C 380
18B1b    ATTGATGATGTCGACATCATAGCCTCATGCTCTGGA 396
          **** * * * * *

```

Sequencing fails
after sgRNA1
(attempt 2)

Table A.3.1: Gene expression levels for RRS WT and RRS SCR upon *M.luteus* infection were measured using the counts generated by HTSeq-count v 0.6.0. The differential expression analysis was performed on all the genetic features listed in the annotation file resulting in ~16,000 coding and non-coding genes with non-zero counts in at least one of the samples. The gene expression counts were normalized for all samples together and the biological conditions were compared pairwise using DESeq2. The differential expression is represented as base mean and log₂ (Fold Change) with its corresponding False discovery rate (FDR). A. List of genes significantly differentially expressed in RRS SCR as compared to RRS WT 0 hours post *M.luteus* infection.

<i>Up-regulated</i>			
Gene ID	log ₂ (Fold Change)	Annotation Symbol	Name
FBgn0263412	1.06	CR43458	long non-coding RNA:CR43458
FBgn0039189	1.01	CG18528	-
<i>Down-regulated</i>			
Gene ID	log ₂ (Fold Change)	Annotation Symbol	Name
FBgn0032768	-1.01	CG17564	-
FBgn0266588	-1.01	CR45113	long non-coding RNA:CR45113
FBgn0264086	-1.01	CG43755	-
FBgn0263257	-1.02	CG43395	Cyclic nucleotide-gated ion channel-like
FBgn0033864	-1.03	CG18368	-
FBgn0030429	-1.03	CG4661	-
FBgn0030764	-1.07	CG9777	-
FBgn0037040	-1.07	CG12983	-
FBgn0053017	-1.07	CG33017	-
FBgn0264255	-1.08		
FBgn0051988	-1.08	CG31988	-
FBgn0262102	-1.10	CG42855	-
FBgn0029728	-1.11	CG2861	-
FBgn0013700	-1.13	CR34062	mitochondrial transfer RNA:Methionine-CAT
FBgn0030975	-1.14	CG7349	Succinate dehydrogenase, subunit B (iron-sulfur)-like
FBgn0264606	-1.23	CG43955	Fife
FBgn0052106	-1.28	CG32106	-

B. List of genes significantly differentially expressed in RRS SCR as compared to RRS WT 0 hours post Ecc15 infection

<i>Up-regulated</i>			
Gene ID	log ₂ (Fold Change)	Annotation Symbol	Name
FBgn0263763	1.42	CG43680	-
FBgn0038095	1.17	CG7241	Cytochrome P450 304a1
FBgn0038914	0.95	CG17820	female-specific independent of transformer
FBgn0001233	0.85	CG1242	Heat shock protein 83
<i>Down-regulated</i>			
Gene ID	log ₂ (Fold Change)	Annotation Symbol	Name
FBgn0035348	-0.82	CG16758	-
FBgn0034997	-0.91	CG3376	-
FBgn0038983	-1.06	CG5326	-
FBgn0030262	-1.07	CG2081	Vago
FBgn0036619	-1.10	CG4784	Cuticular protein 72Ec
FBgn0039298	-1.18	CG11853	takeout
FBgn0029990	-1.24	CG2233	-
FBgn0039798	-1.52	CG11313	-
FBgn0038067	-1.60	CG11598	-
FBgn0005614	-1.72	CG18345	transient receptor potential-like

Table A.3.2: Gene expression levels for RRS WT and RRS SCR upon Ecc15 infection were measured using the counts generated by HTSeq-count v 0.6.0. The differential expression analysis was performed on all the genetic features listed in the annotation file resulting in ~16,000 coding and non-coding genes with non-zero counts in atleast one of the samples. The gene expression counts were normalized for all samples together and the biological conditions were compared pairwise using DESeq2. The differential expression is represented as base mean and log₂(Fold Change) with its corresponding False discovery rate (FDR).

A. List of genes uniquely expressed for RRS WT at 12 hours post Ecc15 infection compared to its baseline

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0010385	CG1385	Defensin	Def	1.00
FBgn0039023	CG4723	Neprilysin-like 14	Nep14	0.98
FBgn0085325	CG34296	-	CG34296	0.83
FBgn0035022	CG11413	-	CG11413	0.94
FBgn0002565	CG6806	Larval serum protein 2	Lsp2	-2.26

B. List of genes uniquely expressed for RRS SCR at 12 hours post Ecc15 infection compared to its baseline

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0012042	CG10146	Attacin-A	AttA	2.34
FBgn0053462	CG33462	-	CG33462	2.12
FBgn0039102	CG16705	Spatzle-Processing Enzyme	SPE	1.38
FBgn0000047	CG5178	Actin 88F	Act88F	0.10
FBgn0031973	CG7219	Serpin 28Dc	Spn28Dc	0.95
FBgn0000276	CG1365	Cecropin A1	CecA1	0.93
FBgn0034162	CG6426	-	CG6426	0.99
FBgn0000281	CR32934	Cecropin 2	Cec2	0.88
FBgn0032973	CG6675	-	CG6675	1.00
FBgn0263321	CG43402	-	CG43402	0.95
FBgn0031560	CG16713	-	CG16713	0.74
FBgn0027563	CG9631	-	CG9631	0.86
FBgn0031561	CG16712	Immune induced molecule 33	IM33	0.80
FBgn0052185	CG32185	elevated during infection	edin	0.86
FBgn0034741	CG4269	-	CG4269	0.79
FBgn0051769	CG31769	-	CG31769	0.75
FBgn0259973	CG42483	Seminal fluid protein 79	Sfp79B	0.83
FBgn0263766	CR43683	antisense RNA:CR43683	asRNA:CR43683	0.86
FBgn0050154	CG30154	-	CG30154	0.80
FBgn0000279	CG1373	Cecropin C	CecC	0.84
FBgn0263083	CG43351	-	CG43351	0.74
FBgn0041182	CG7052	Thioester-containing protein 2	Tep2	0.77
FBgn0030051	CG2056	Serine Protease Immune Response Integrator	spirit	0.80
FBgn0043791	CG8147	phurba tashi	phu	-0.81
FBgn0031689	CG10833	Cytochrome P450 28d1	Cyp28d1	-0.85
FBgn0029990	CG2233	-	CG2233	-0.88
FBgn0043783	CG32444	-	CG32444	-0.89
FBgn0028853	CG15263	-	CG15263	-0.97
FBgn0000406	CG13279	Cytochrome b5-related	Cyt-b5-r	-1.02

C. List of genes uniquely expressed for RRS WT at 12 hours post *M.luteus* infection compared to its baseline

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0052185	CG32185	elevated during infection	edin	4.69
FBgn0040653	CG15231	Daisho1	Dso1	2.34
FBgn0000047	CG5178	Actin 88F	Act88F	1.40
FBgn0261989	CG42807	-	CG42807	1.39
FBgn0039023	CG4723	Neprilysin-like 14	Nep14	1.29
FBgn0083938	CG34102	BG642163	BG642163	1.27
FBgn0039564	CG5527	Neprilysin 7	Nep7	1.24
FBgn0037915	CG6790	-	CG6790	1.11
FBgn0038610	CG7675	-	CG7675	1.07
FBgn0039629	CG11842	-	CG11842	1.06
FBgn0023129	CG3705	astray	aay	1.06
FBgn0032286	CG7300	-	CG7300	1.05
FBgn0262717	CG43161	Skeletor	Skeletor	1.03
FBgn0028538	CG7578	Secretory 71	Sec71	1.02
FBgn0025583	CG18106	Bomanin Short 2	BomS2	1.02
FBgn0264077	CG9906	Calnexin 14D	Cnx14D	1.02
FBgn0020416	CG4472	Imaginal disc growth factor 1	ldgf1	1.01
FBgn0052110	CG32110	-	CG32110	1.01
FBgn0031626	CG15631	-	CG15631	1.01
FBgn0034143	CG8303	-	CG8303	0.98
FBgn0085454	CG34425	-	CG34425	0.97
FBgn0052694	CG32694	-	CG32694	0.96
FBgn0085244	CG34215	-	CG34215	0.96
FBgn0035505	CG15004	tipE homolog 2	Teh2	0.93
FBgn0036925	CG17736	schumacher-levy	schuy	0.92
FBgn0030307	CG33235	-	CG33235	0.92
FBgn0011559	CG7157	Accessory gland protein 36DE	Acp36DE	0.91
FBgn0028984	CG18525	Serpin 88Ea	Spn88Ea	0.89
FBgn0085256	CG34227	-	CG34227	0.89

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0262035	CG42846	-	CG42846	0.89
FBgn0039495	CG5909	-	CG5909	0.89
FBgn0283427	CG3523	Fatty acid synthase 1	FASN1	0.89
FBgn0263118	CG5441	taxi	tx	0.89
FBgn0265191	CG44244	Glycogenin	Gyg	0.88
FBgn0038250	CG3505	-	CG3505	0.87
FBgn0014395	CG14620	touch insensitive larva B	tilB	0.87
FBgn0034290	CG5773	-	CG5773	0.87
FBgn0015586	CG3801	Accessory gland protein 76A	Acp76A	0.86
FBgn0035957	CG5144	-	CG5144	0.86
FBgn0026415	CG1780	Imaginal disc growth factor 4	ldgf4	0.86
FBgn0031617	CG15635	-	CG15635	0.85
FBgn0039238	CG7016	-	CG7016	0.84
FBgn0035144	CG17181	Kahuli	Kah	0.84
FBgn0037534	CG2781	ELOVL fatty acid elongase 7	Elovl7	0.84
FBgn0031251	CG4213	-	CG4213	0.83
FBgn0265180	CG44245	-	CG44245	0.82
FBgn0261928	CG42795	-	CG42795	0.81
FBgn0052072	CG32072	Elongase 68alpha	Elo68alpha	0.81
FBgn0027949	CG10364	msb1l	msb1l	0.81
FBgn0262583	CG43123	-	CG43123	0.81
FBgn0027376	CG11908	rha	rha	0.80
FBgn0033031	CG8245	-	CG8245	0.80
FBgn0038130	CG8630	-	CG8630	0.80
FBgn0035343	CG16762	-	CG16762	0.80
FBgn0051320	CG31320	HEAT repeat containing 2	HEATR2	0.79
FBgn0038607	CG7669	hemingway	hmw	0.79
FBgn0034827	CG12192	Kinesin-like protein at 59D	Klp59D	0.79
FBgn0036106	CG6409	-	CG6409	0.79

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0033603	CG13214	Cuticular protein 47Ef	Cpr47Ef	0.79
FBgn0037810	CG12819	slender lobes	sle	0.79
FBgn0034515	CG13428	-	CG13428	0.78
FBgn0032520	CG10859	-	CG10859	0.78
FBgn0262894	CG43249	-	CG43249	0.78
FBgn0035638	CG10541	Tektin C	Tektin-C	0.77
FBgn0034121	CG6262	-	CG6262	0.77
FBgn0032495	CG16820	-	CG16820	0.77
FBgn0039152	CG6129	Rootletin	Root	0.77
FBgn0051921	CG31921	-	CG31921	0.76
FBgn0036311	CG17666	-	CG17666	0.76
FBgn0037064	CG9389	-	CG9389	0.76
FBgn0014073	CG7525	Tie-like receptor tyrosine kinase	Tie	0.76
FBgn0031279	CG3544	-	CG3544	0.76
FBgn0038179	CG9312	-	CG9312	0.75
FBgn0054002	CG34002	-	CG34002	0.75
FBgn0040759	CG13177	-	CG13177	0.75
FBgn0027843	CG6906	Carbonic anhydrase 2	CAH2	0.74
FBgn0085337	CG34308	-	CG34308	0.74
FBgn0003249	CG10888	Rhodopsin 3	Rh3	0.74
FBgn0036835	CG14075	-	CG14075	0.73
FBgn0029167	CG7002	Hemolectin	Hml	0.73
FBgn0032471	CG5122	-	CG5122	0.73
FBgn0050156	CG30156	-	CG30156	0.73
FBgn0053110	CG33110	-	CG33110	0.73
FBgn0033919	CG8547	-	CG8547	0.73
FBgn0038345	CG5213	-	CG5213	0.73
FBgn0033149	CG11060	-	CG11060	0.73
FBgn0038181	CG9297	-	CG9297	0.73
FBgn0263321	CG43402	-	CG43402	0.73

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0025388	CG12179	Alstrom syndrome 1a	Alms1a	0.73
FBgn0037796	CG12814	-	CG12814	0.72
FBgn0032472	CG9928	-	CG9928	0.72
FBgn0011695	CG11390	Ejaculatory bulb protein III	EbpIII	0.71
FBgn0051099	CG31099	-	CG31099	0.71
FBgn0032771	CG17349	-	CG17349	0.71
FBgn0028986	CG9334	Serpin 38F	Spn38F	0.70
FBgn0046763	CG17278	-	CG17278	0.70
FBgn0051025	CG31025	Protein phosphatase 1c interacting protein 1	Ppi1	0.70
FBgn0261816	CG42758	-	CG42758	0.70
FBgn0260474	CG30002	-	CG30002	0.69
FBgn0267253	CG32700	-	CG32700	0.69
FBgn0086906	CG1915	Sallimus	sls	0.69
FBgn0010482	CG9432	lethal (2) 01289	l(2)01289	0.68
FBgn0052568	CG32568	-	CG32568	0.68
FBgn0265264	CG17097	-	CG17097	0.68
FBgn0036415	CG7768	-	CG7768	0.68
FBgn0030814	CG4955	-	CG4955	0.68
FBgn0035346	CG1146	-	CG1146	0.68
FBgn0040735	CG16836	Bomanin Tailed 2	BomT2	0.67
FBgn0038690	CG11703	-	CG11703	0.67
FBgn0061197	CG13164	Salto	salto	0.67
FBgn0037339	CG2929	Phosphatidylinositol 4-kinase II alpha	Pi4KIIalpha	0.66
FBgn0022355	CG6186	Transferrin 1	Tsf1	0.66
FBgn0263048	CG43343	Glycerol-3-phosphate dehydrogenase 3	Gpdh3	0.66
FBgn0266801	CG45263	-	CG45263	0.66
FBgn0039521	CG5402	-	CG5402	0.66
FBgn0031690	CG7742	-	CG7742	0.66
FBgn0039114	CG10374	Lipid storage droplet-1	Lsd-1	0.66
FBgn0052425	CG32425	-	CG32425	0.66

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0036157	CG7560	-	CG7560	0.65
FBgn0003149	CG5939	Paramyosin	Prm	0.65
FBgn0004414	CG14560	male-specific opa containing gene	msopa	0.64
FBgn0020415	CG4475	Imaginal disc growth factor 2	ldgf2	0.64
FBgn0050271	CG30271	-	CG30271	0.64
FBgn0035380	CG9970	-	CG9970	0.64
FBgn0028870	CG4691	-	CG4691	0.64
FBgn0004882	CG10868	oo18 RNA-binding protein	orb	0.64
FBgn0028997	CG8362	nmdyn-D7	nmdyn-D7	0.63
FBgn0262579	CG43119	Sterile alpha and Armadillo motif	Sarm	0.63
FBgn0011725	CG31137	twin	twin	0.63
FBgn0259736	CG42390	-	CG42390	0.63
FBgn0260429	CG42524	-	CG42524	0.63
FBgn0005659	CG5583	Ets at 98B	Ets98B	0.62
FBgn0050069	CG30069	-	CG30069	0.62
FBgn0037819	CG14688	Phytanoyl-CoA dioxygenase domain containing 1	Phyhd1	0.61
FBgn0036931	CG14183	-	CG14183	0.61
FBgn0038952	CG7069	-	CG7069	0.61
FBgn0011230	CG14472	purity of essence	poe	0.61
FBgn0086359	CG3953	Invadolysin	Invadolysin	0.61
FBgn0260745	CG3359	midline fasciclin	mfas	0.61
FBgn0039419	CG12290	-	CG12290	0.60
FBgn0002865	CG11719	Male-specific RNA 98Ca	Mst98Ca	0.60
FBgn0034133				0.59
FBgn0033174	CG11125	-	CG11125	0.59
FBgn0031347	CG10869	no individualized sperm	nis	0.58
FBgn0032773	CG15825	fondue	fon	0.58
FBgn0031037	CG14207	-	CG14207	0.58
FBgn0038248	CG7886	-	CG7886	0.57
FBgn0034435	CG9975	wurstfest	fest	0.57
FBgn0031878	CG9188	septin interacting protein 2	sip2	0.57

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0036969	CG6663	Serpin 77Bb	Spn77Bb	0.56
FBgn0261258	CG6014	regeneration	rgn	0.56
FBgn0051092	CG31092	Lipophorin receptor 2	LpR2	0.56
FBgn0005666	CG32019	bent	bt	0.55
FBgn0033954	CG12860	-	CG12860	0.55
FBgn0265739	CR44546	long non-coding RNA:CR44546	lncRNA:CR44546	-0.55
FBgn0086355	CG2171	Triose phosphate isomerase	Tpi	-0.55
FBgn0267327	CG6555	Accessory gland-specific peptide 33A	Acp33A	-0.55
FBgn0015584	CG8622	Accessory gland protein 53Ea	Acp53Ea	-0.56
FBgn0002579	CG7622	Ribosomal protein L36	RpL36	-0.56
FBgn0038407	CG6126	-	CG6126	-0.56
FBgn0039761	CG18404	-	CG18404	-0.56
FBgn0040575	CG15922	-	CG15922	-0.57
FBgn0053977	CG33977	Dolichyl-phosphate mannosyltransferase subunit 3	Dpm3	-0.57
FBgn0038723	CG6195	-	CG6195	-0.57
FBgn0031538	CG3246	-	CG3246	-0.57
FBgn0034152	CG8626	Accessory gland protein 53C14a	Acp53C14a	-0.57
FBgn0261602	CG1263	Ribosomal protein L8	RpL8	-0.57
FBgn0032518	CG9282	Ribosomal protein L24	RpL24	-0.57
FBgn0037788	CG3940	Carbonic anhydrase 7	CAH7	-0.57
FBgn0010053	CG15101	Juvenile hormone epoxide hydrolase 1	Jheh1	-0.58
FBgn0030518	CG11134	-	CG11134	-0.58
FBgn0030999	CG7874	Mucin related 18B	Mur18B	-0.58
FBgn0083983	CG34147	mitochondrial ribosomal protein L34	mRpL34	-0.58
FBgn0036381	CG8745	-	CG8745	-0.59
FBgn0033093	CG3270	-	CG3270	-0.59
FBgn0030773	CG9676	-	CG9676	-0.59
FBgn0033428	CG1818	Uroporphyrinogen decarboxylase	Urod	-0.59
FBgn0028920	CG8997	-	CG8997	-0.59
FBgn0061198	CG30173	Haematopoietic stem/progenitor cell protein 300	HSPC300	-0.59
FBgn0039697	CG7834	-	CG7834	-0.60

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0010078	CG3661	Ribosomal protein L23	RpL23	-0.60
FBgn0039118	CG10208	-	CG10208	-0.60
FBgn0038516	CG5840	Pyrroline-5-carboxylate reductase-like 2	P5cr-2	-0.60
FBgn0030572	CG14413	mitochondrial ribosomal protein S25	mRpS25	-0.60
FBgn0019624	CG14724	Cytochrome c oxidase subunit 5A	COX5A	-0.60
FBgn0035722	CG10075	-	CG10075	-0.60
FBgn0038820	CG4000	-	CG4000	-0.60
FBgn0024293	CG1865	Serpin 43Ab	Spn43Ab	-0.60
FBgn0037612	CG8112	-	CG8112	-0.60
FBgn0250831	CG34103	BG642167	BG642167	-0.61
FBgn0028336				-0.61
FBgn0031653	CG8871	Jonah 25Biii	Jon25Biii	-0.61
FBgn0262571	CG43111	-	CG43111	-0.61
FBgn0051226	CG31226	-	CG31226	-0.61
FBgn0031654	CG8869	Jonah 25Bii	Jon25Bii	-0.62
FBgn0042119	CG18778	Cuticular protein 65Au	Cpr65Au	-0.62
FBgn0037913	CG6783	fatty acid binding protein	fabp	-0.62
FBgn0053002	CG33002	mitochondrial ribosomal protein L27	mRpL27	-0.62
FBgn0044511	CG32854	mitochondrial ribosomal protein S21	mRpS21	-0.63
FBgn0034645	CG10320	NADH dehydrogenase (ubiquinone) B12 subunit	ND-B12	-0.63
FBgn0050410	CG30410	Ribose-5-phosphate isomerase	Rpi	-0.63
FBgn0044030	CG32531	mitochondrial ribosomal protein S14	mRpS14	-0.63
FBgn0003274	CG4918	Ribosomal protein LP2	RpLP2	-0.63
FBgn0003034	CG17673	Sex Peptide	SP	-0.63
FBgn0031417	CG3597	-	CG3597	-0.63
FBgn0004087	CG14887	Dihydrofolate reductase	Dhfr	-0.63
FBgn0261055	CG42603	Seminal fluid protein 26Ad	Sfp26Ad	-0.64
FBgn0259971	CG42481	-	CG42481	-0.64
FBgn0083953	CG34117	-	CG34117	-0.65
FBgn0036135	CG7636	mitochondrial ribosomal protein L2	mRpL2	-0.65
FBgn0050373	CG30373	-	CG30373	-0.65

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0264513	CR43912	long non-coding RNA:CR43912	lncRNA:CR43912	-0.65
FBgn0013680	CG34063	mitochondrial NADH-ubiquinone oxidoreductase chain 2	mt:ND2	-0.66
FBgn0032066	CG9463	Lysosomal alpha-mannosidase III	LManIII	-0.66
FBgn0261534	CG7532	lethal (2) 34Fc	l(2)34Fc	-0.66
FBgn0263597	CG43619	Accessory gland protein 98AB	Acp98AB	-0.66
FBgn0013683	CG34086	mitochondrial NADH-ubiquinone oxidoreductase chain 4L	mt:ND4L	-0.66
FBgn0034354	CG5224	Glutathione S transferase E11	GstE11	-0.66
FBgn0040001				-0.66
FBgn0260444	CG18616	CCR4-NOT transcription complex subunit 10	Not10	-0.67
FBgn0031108	CG15459	-	CG15459	-0.67
FBgn0035042	CG3640	-	CG3640	-0.67
FBgn0265538	CG44388	-	CG44388	-0.67
FBgn0259701	CG42355	-	CG42355	-0.67
FBgn0036023	CG18179	-	CG18179	-0.68
FBgn0263082	CG43350	-	CG43350	-0.69
FBgn0040931	CG9034	-	CG9034	-0.69
FBgn0037723	CG8327	Spermidine Synthase	SpdS	-0.69
FBgn0261061	CG42609	Seminal fluid protein 96F	Sfp96F	-0.69
FBgn0040349	CG3699	-	CG3699	-0.69
FBgn0020906	CG8867	Jonah 25Bi	Jon25Bi	-0.70
FBgn0263830	CG40486	-	CG40486	-0.70
FBgn0031418	CG3609	-	CG3609	-0.70
FBgn0036759	CG5577	-	CG5577	-0.70
FBgn0261060	CG42608	Seminal fluid protein 93F	Sfp93F	-0.71
FBgn0036015	CG3088	-	CG3088	-0.71
FBgn0034001	CG12954	mitochondrial ribosomal protein L41	mRpL41	-0.71
FBgn0259968	CG42478	Seminal fluid protein 60F	Sfp60F	-0.71
FBgn0051788	CG31788	-	CG31788	-0.72
FBgn0262896	CG43251	-	CG43251	-0.72
FBgn0031505	CG12400	NADH dehydrogenase (ubiquinone) B14.5 B subunit	ND-B14.5B	-0.73
FBgn0030984	CG7440	twiggy	tgy	-0.73

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0262099	CG42852	-	CG42852	-0.73
FBgn0030251	CG2145	-	CG2145	-0.73
FBgn0038718	CG17752	-	CG17752	-0.74
FBgn0003076	CG5165	Phosphoglucose mutase 1	Pgm1	-0.74
FBgn0053511	CG33511	-	CG33511	-0.74
FBgn0052643	CG32643	-	CG32643	-0.74
FBgn0058469	CR40469	long non-coding RNA:CR40469	lncRNA:CR40469	-0.74
FBgn0262574	CG43114	-	CG43114	-0.75
FBgn0039802	CG1349	dj-1beta	dj-1beta	-0.75
FBgn0051779	CG31779	Accessory gland protein 24A4	Acp24A4	-0.75
FBgn0035695	CG10226	-	CG10226	-0.78
FBgn0052205	CR32205	hairpin RNA:CR32205	hpRNA:CR32205	-0.78
FBgn0039682	CG7584	Odorant-binding protein 99c	Obp99c	-0.78
FBgn0040705	CG15434	NADH dehydrogenase (ubiquinone) B8 subunit	ND-B8	-0.78
FBgn0031373	CG15358	-	CG15358	-0.78
FBgn0033367	CG8193	Prophenoloxidase 2	PPO2	-0.79
FBgn0031277	CG13947	-	CG13947	-0.80
FBgn0085453	CG34424	Methenyltetrahydrofolate synthetase	Mthfs	-0.80
FBgn0262601	CG5352	Small ribonucleoprotein particle protein SmB	SmB	-0.80
FBgn0022359	CG4649	Sorbitol dehydrogenase-2	Sodh-2	-0.81
FBgn0259957	CG42467	-	CG42467	-0.82
FBgn0023477	CG2827	Transaldolase	Taldo	-0.82
FBgn0263763	CG43680	-	CG43680	-0.83
FBgn0051485	CR31485	transfer RNA:Lysine-CTT 2-1 pseudogene	tRNA:Lys-CTT-2-1Psi	-0.85
FBgn0259715	CG42369	-	CG42369	-0.85
FBgn0022774	CG8782	Ornithine aminotransferase precursor	Oat	-0.87
FBgn0030968	CG7322	-	CG7322	-0.87
FBgn0262948	CG43267	-	CG43267	-0.88
FBgn0033723	CG13155	-	CG13155	-0.89
FBgn0030593	CG9512	-	CG9512	-0.90
FBgn0035484	CG11594	-	CG11594	-0.90

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0050042	CG30042	Cuticular protein 49Ab	Cpr49Ab	-0.91
FBgn0260393	CG17147	-	CG17147	-0.93
FBgn0026721	CG6953	fat-spondin	fat-spondin	-0.94
FBgn0085201	CG34172	Cytochrome c oxidase subunit 7A-like 2	COX7AL2	-0.94
FBgn0015035	CG4105	Cytochrome P450 4e3	Cyp4e3	-0.96
FBgn0030737				-0.98
FBgn0259951	CG42461	Seminal fluid protein 24Ba	Sfp24Ba	-0.98
FBgn0030362	CG1803	regucalcin	regucalcin	-0.99
FBgn0019928	CG4812	Ser8	Ser8	-1.01
FBgn0085195	CG34166	-	CG34166	-1.05
FBgn0040211	CG4779	homogentisate 1,2-dioxygenase	hgo	-1.07
FBgn0036992	CG11796	4-hydroxyphenylpyruvate dioxygenase	Hpd	-1.09
FBgn0283437	CG42639	Prophenoloxidase 1	PPO1	-1.21
FBgn0053319	CR33319	-	CR33319	-1.34
FBgn0039685	CG7592	Odorant-binding protein 99b	Obp99b	-2.41

D. List of genes uniquely expressed for RRS SCR at 12 hours post *M.luteus* infection compared to its baseline

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0032638	CG6639	Serine protease homolog 93	SPH93	2.07
FBgn0034094	CG3666	Transferrin 3	Tsf3	2.01
FBgn0262004	CG42822	-	CG42822	1.96
FBgn0263479	CR43594	small Cajal body-specific RNA : PsiU2-35.45	scaRNA:PsiU2-35.45	1.93
FBgn0030927	CG15046	-	CG15046	1.89
FBgn0031701	CG14027	Turandot M	TotM	1.85
FBgn0030981	CG14191	-	CG14191	1.85
FBgn0038631	CG7695	-	CG7695	1.76
FBgn0036877	CG9452	-	CG9452	1.72
FBgn0263766	CR43683	antisense RNA:CR43683	asRNA:CR43683	1.66
FBgn0029639	CG14419	-	CG14419	1.62
FBgn0029728	CG2861	-	CG2861	1.62
FBgn0013700	CR34062	mitochondrial transfer RNA:Methionine-CAT	mt:tRNA:Met-CAT	1.61
FBgn0086677	CG30040	jelly belly	jeb	1.59
FBgn0263839	CR43701	small non-messenger RNA:CR43701	snmRNA:CR43701	1.56
FBgn0013703	CR34061	mitochondrial transfer RNA:Glutamine-TTG	mt:tRNA:Gln-TTG	1.55
FBgn0003430	CG16738	sloppy paired 1	slp1	1.54
FBgn0035623	CG17795	methuselah-like 2	mthl2	1.54
FBgn0035023	CG13586	lon transport peptide	ITP	1.53
FBgn0033949	CG10131	beta Hydroxy acid dehydrogenase 2	Had2	1.52
FBgn0036951	CG7017	-	CG7017	1.52
FBgn0032943	CG8666	Tetraspanin 39D	Tsp39D	1.51
FBgn0036287	CG10663	-	CG10663	1.50
FBgn0013694	CR34075	mitochondrial transfer RNA:Glycine-TCC	mt:tRNA:Gly-TCC	1.50
FBgn0065099	CR33925	small nuclear RNA 7SK	snRNA:7SK	1.47
FBgn0000422	CG10697	Dopa decarboxylase	Ddc	1.45
FBgn0036101	CG6449	Ninjurin A	NijA	1.44
FBgn0030975	CG7349	Succinate dehydrogenase, subunit B (iron-sulfur)-like	SdhBL	1.44
FBgn0250871	CG2467	papillote	pot	1.43

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0033538	CG11883	-	CG11883	1.43
FBgn0065076	CR33930	snoRNA:185	snoRNA:185	1.43
FBgn0262107	CR42860	antisense RNA:CR42860	asRNA:CR42860	1.42
FBgn0034082	CG10734	-	CG10734	1.41
FBgn0263847	CR43708	small nuclear RNA Like U	snRNA:LU	1.41
FBgn0264503	CG43902	-	CG43902	1.40
FBgn0265064	CR44175	long non-coding RNA:CR44175	lncRNA:CR44175	1.39
FBgn0085426	CG34397	Rad, Gem/Kir family member 3	Rgk3	1.39
FBgn0013704	CR34078	mitochondrial transfer RNA:Arginine-TCG	mt:tRNA:Arg-TCG	1.37
FBgn0013685	CG34089	mitochondrial NADH-ubiquinone oxidoreductase chain 6	mt:ND6	1.37
FBgn0027865	CG6120	Tetraspanin 96F	Tsp96F	1.36
FBgn0033388	CG8046	-	CG8046	1.36
FBgn0085387	CG34358	shaking B	shakB	1.35
FBgn0029736	CG4041	-	CG4041	1.35
FBgn0036544	CG6114	sugar-free frosting	sff	1.35
FBgn0004191	CR31850	small nuclear RNA U2 at 34AB a	snRNA:U2:34ABa	1.33
FBgn0032123	CG3811	Organic anion transporting polypeptide 30B	Oatp30B	1.32
FBgn0083972	CG34136	-	CG34136	1.32
FBgn0264255				1.32
FBgn0039183	CG6413	Dis3	Dis3	1.32
FBgn0039453	CG6403	-	CG6403	1.31
FBgn0019828	CG1980	don juan	dj	1.30
FBgn0041150	CG12787	hoepel1	hoe1	1.29
FBgn0052373	CG32373	-	CG32373	1.29
FBgn0268063	CG32372	larval translucida	ltl	1.29
FBgn0052161	CG32161	-	CG32161	1.28
FBgn0013690	CR34065	mitochondrial transfer RNA:Cysteine-GCA	mt:tRNA:Cys-GCA	1.28
FBgn0054003	CG34003	Nimrod B3	NimB3	1.28
FBgn0030277	CG1394	-	CG1394	1.28

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0052412	CG32412	Glutaminy cyclase	QC	1.28
FBgn0054054	CG34054	-	CG34054	1.27
FBgn0037515	CG3066	Serine protease 7	Sp7	1.27
FBgn0033250	CG14762	-	CG14762	1.27
FBgn0034046	CG8253	tungus	tun	1.26
FBgn0002917	CG1517	narrow abdomen	na	1.26
FBgn0037396	CG11459	-	CG11459	1.26
FBgn0029003	CG4746	mab-21	mab-21	1.26
FBgn0040582	CG5791	Bomanin Bicipital 3	BomBc3	1.25
FBgn0266588	CR45113	long non-coding RNA:CR45113	lncRNA:CR45113	1.25
FBgn0011296	CG4533	lethal (2) essential for life	l(2)efl	1.25
FBgn0000108	CG7727	beta amyloid protein precursor-like	Appl	1.25
FBgn0039041	CG13838	-	CG13838	1.24
FBgn0030452	CG4330	Major Facilitator Superfamily Transporter 10	MFS10	1.24
FBgn0013697	CR34070	mitochondrial transfer RNA:Lysine-CTT	mt:tRNA:Lys-CTT	1.24
FBgn0025743	CG18582	mushroom bodies tiny	mbt	1.22
FBgn0040905	CG15578	-	CG15578	1.22
FBgn0034364	CG5493	-	CG5493	1.20
FBgn0003930	CR31625	small nuclear RNA U4 at 39B	snRNA:U4:39B	1.20
FBgn0266758	CG6643	Extended synaptotagmin-like protein 2	Esyt2	1.19
FBgn0040020	CG17397	Mediator complex subunit 21	MED21	1.19
FBgn0037167	CG11425	-	CG11425	1.19
FBgn0260400	CG4262	embryonic lethal abnormal vision	elav	1.18
FBgn0028496	CG30116	-	CG30116	1.18
FBgn0265042	CG44159	Inwardly rectifying potassium channel 1	Irk1	1.17
FBgn0038704	CG5316	-	CG5316	1.17
FBgn0263647	CG43638	-	CG43638	1.16
FBgn0263415	CR43461	long non-coding RNA:CR43461	lncRNA:CR43461	1.16
FBgn0029761	CG10706	small conductance calcium-activated potassium channel	SK	1.16
FBgn0263111	CG43368	cacophony	cac	1.16

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0053136	CG33136	-	CG33136	1.15
FBgn0041713	CG4182	yellow-c	yellow-c	1.14
FBgn0086915	CG3354	Male-specific transcript 77F	Mst77F	1.14
FBgn0030764	CG9777	-	CG9777	1.14
FBgn0052450	CG32450	-	CG32450	1.12
FBgn0001230	CG5436	Heat shock protein 68	Hsp68	1.12
FBgn0053329	CG33329	Serine-peptidase 212	Sp212	1.12
FBgn0040291	CG16988	Regulator of cullins 1b	Roc1b	1.11
FBgn0016926	CG4710	Pinocchio	Pino	1.10
FBgn0026059	CG31045	Myosin heavy chain-like	Mhcl	1.10
FBgn0086675	CG4396	found in neurons	fne	1.10
FBgn0038281	CG3843	Ribosomal protein L10Aa	RpL10Aa	1.10
FBgn0264819	CR44027	antisense RNA:CR44027	asRNA:CR44027	1.08
FBgn0051988	CG31988	-	CG31988	1.08
FBgn0039647	CG14509	julius seizure	jus	1.08
FBgn0083228	CG5907	Frequenin 2	Frq2	1.08
FBgn0020440	CG10023	Focal adhesion kinase	Fak	1.07
FBgn0004606	CG1322	Zn finger homeodomain 1	zfh1	1.06
FBgn0019929	CG2045	Ser7	Ser7	1.05
FBgn0262102	CG42855	-	CG42855	1.03
FBgn0031306	CG4577	-	CG4577	1.03
FBgn0013334	CG8884	Synapse-associated protein 47kD	Sap47	1.02
FBgn0051361	CG31361	defective proboscis extension response 17	dpr17	1.01
FBgn0050026	CG30026	-	CG30026	1.00
FBgn0031560	CG16713	-	CG16713	1.00
FBgn0013305	CG3798	N-methyl-D-aspartate receptor-associated protein	Nmda1	1.00
FBgn0022768	CG2984	Protein phosphatase 2C	Pp2C1	0.99
FBgn0033680	CG13186	-	CG13186	0.99
FBgn0032382	CG14935	Maltase B2	Mal-B2	0.97
FBgn0052675	CG32675	Transport and Golgi organization 5	Tango5	0.97

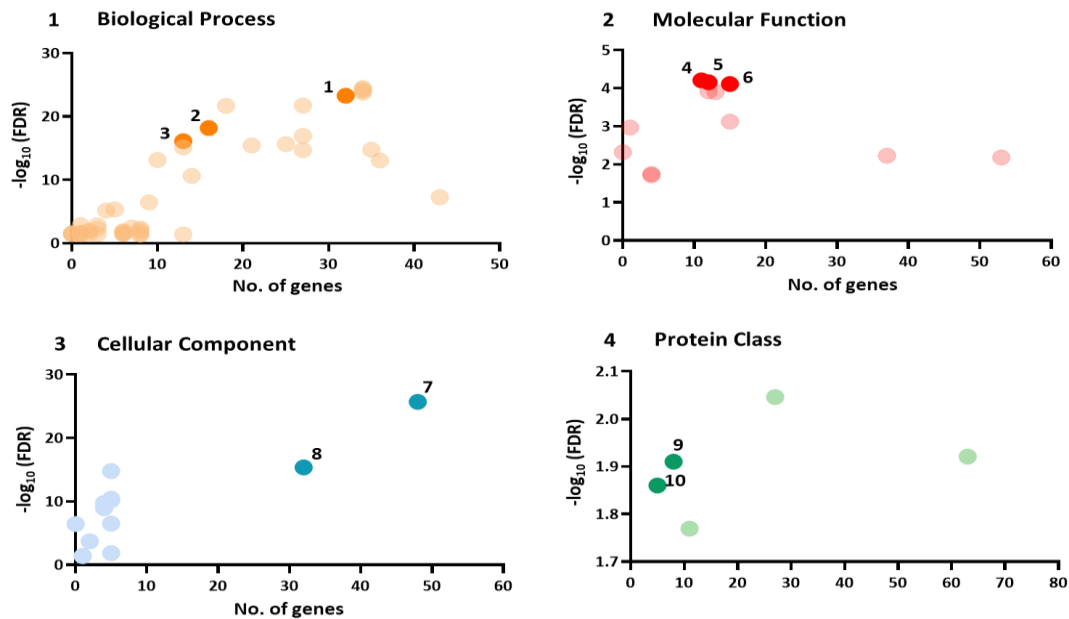
Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0027932	CG13388	A kinase anchor protein 200	Akap200	0.96
FBgn0035964	CG4665	Dihydropteridine reductase	Dhpr	0.96
FBgn0034145	CG5065	-	CG5065	0.96
FBgn0033128	CG12142	Tetraspanin 42Eg	Tsp42Eg	0.95
FBgn0261686	CG42733	-	CG42733	0.95
FBgn0034689	CG2921	-	CG2921	0.93
FBgn0263106	CG10578	DnaJ-like-1	DnaJ-1	0.92
FBgn0028990	CG11331	Serpin 27A	Spn27A	0.92
FBgn0003429	CG10693	slowpoke	slo	0.92
FBgn0052190	CG32190	NUCB1	NUCB1	0.91
FBgn0015320	CG6720	Ubiquitin conjugating enzyme 2	Ubc2	0.89
FBgn0020414	CG4559	Imaginal disc growth factor 3	ldgf3	0.88
FBgn0026206	CG12218	meiotic P26	mei-P26	0.86
FBgn0033875	CG6357	-	CG6357	0.81
FBgn0014028	CG3283	Succinate dehydrogenase, subunit B (iron-sulfur)	SdhB	-0.75
FBgn0040813				-0.83
FBgn0010611	CG4311	HMG Coenzyme A synthase	Hmgs	-0.83
FBgn0035779	CG8562	-	CG8562	-0.83
FBgn0035228	CG12091	-	CG12091	-0.86
FBgn0015585	CG10852	Accessory gland protein 63F	Acp63F	-0.88
FBgn0027073	CG4302	UDP-glycosyltransferase family 49 member B1	Ugt49B1	-0.91
FBgn0035189	CG9119	-	CG9119	-0.92
FBgn0039324	CG10553	-	CG10553	-0.95
FBgn0027586	CG5867	-	CG5867	-0.97
FBgn0039325	CG10560	-	CG10560	-0.98
FBgn0027560	CG4104	Trehalose-6-phosphate synthase 1	Tps1	-0.99
FBgn0032068	CG9466	Lysosomal alpha-mannosidase V	LManV	-1.00
FBgn0025885	CG11143	myo-inositol-1-phosphate synthase	Inos	-1.02
FBgn0001248	CG7176	Isocitrate dehydrogenase	ldh	-1.05
FBgn0040832	CG8012	-	CG8012	-1.07

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0267914	CR46195	antisense RNA:CR46195	asRNA:CR46195	-1.08
FBgn0026428	CG6170	Histone deacetylase 6	HDAC6	-1.10
FBgn0263024	CG43319	-	CG43319	-1.10
FBgn0031693	CG14032	Cytochrome P450 4ac1	Cyp4ac1	-1.10
FBgn0038257	CG7390	Senescence marker protein-30	smp-30	-1.11
FBgn0259952	CG42462	Seminal fluid protein 24Bb	Sfp24Bb	-1.11
FBgn0032864	CG2493	-	CG2493	-1.12
FBgn0036289	CG10657	-	CG10657	-1.17
FBgn0004181	CG2668	Ejaculatory bulb protein	Ebp	-1.17
FBgn0038463	CG3534	-	CG3534	-1.18
FBgn0042201	CG13061	Neuropeptide-like precursor 3	Nplp3	-1.18
FBgn0035978	CG4347	UDP-glucose pyrophosphorylase	UGP	-1.18
FBgn0001208	CG7399	Henna	Hn	-1.20
FBgn0015575	CG1112	alpha-Esterase-7	alpha-Est7	-1.21
FBgn0026314	CG6649	UDP-glycosyltransferase family 35 member B1	Ugt35B1	-1.23
FBgn0034394	CG15096	-	CG15096	-1.24
FBgn0031942	CG7203	-	CG7203	-1.26
FBgn0040606	CG6503	-	CG6503	-1.26
FBgn0030607	CG5560	doppelganger von brummer	dob	-1.27
FBgn0041194	CG10078	Phosphoribosylamidotransferase 2	Prat2	-1.28
FBgn0027657	CG9734	globin 1	glob1	-1.28
FBgn0263199	CG5288	Galactokinase	Galk	-1.28
FBgn0020385	CG4067	pugilist	pug	-1.32
FBgn0039612	CG14523	Neprilysin-like 19	Nep19	-1.32
FBgn0085360	CG34331	-	CG34331	-1.34
FBgn0033170	CG11124	secretory Phospholipase A2	sPLA2	-1.35
FBgn0033633	CG7759	SET and MYND domain containing, class 4, member 3	Smyd4-3	-1.36
FBgn0030558	CG1461	Tyrosine aminotransferase	Tat	-1.36
FBgn0031305	CG4715	Iris	Iris	-1.36
FBgn0011293	CG6642	antennal protein 10	a10	-1.37

Gene ID	ANNOTATION SYMBOL	NAME	SYMBOL	Log ₂ (FC)
FBgn0038702	CG3739	-	CG3739	-1.39
FBgn0040718	CG15353	-	CG15353	-1.43
FBgn0263868	CR43715	long non-coding RNA:CR43715	lncRNA:CR43715	-1.45
FBgn0011694	CG2665	Ejaculatory bulb protein II	EbplII	-1.49
FBgn0039801	CG11315	Niemann-Pick type C-2h	Npc2h	-1.52
FBgn0034415	CG15116	-	CG15116	-1.52
FBgn0038914	CG17820	female-specific independent of transformer	fit	-1.55
FBgn0043783	CG32444	-	CG32444	-1.56
FBgn0038291	CG3984	-	CG3984	-1.60
FBgn0043791	CG8147	phurba tashi	phu	-1.62
FBgn0032265	CG18301	-	CG18301	-1.69
FBgn0036262	CG6910	-	CG6910	-1.70
FBgn0039611	CG14528	Neprilysin-like 18	Nepl18	-1.88
FBgn0005614	CG18345	transient receptor potential-like	trpl	-2.01
FBgn0035076	CG10142	Ance-5	Ance-5	-2.03

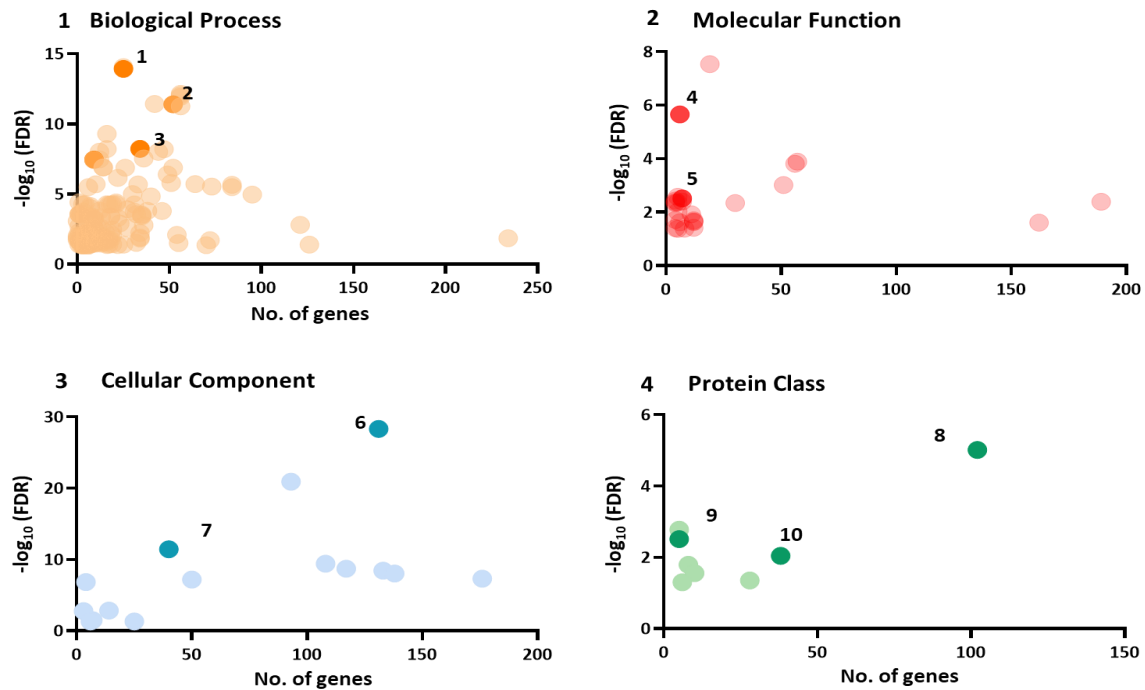
Figure A.3.2: Gene Ontology Enrichment Analysis (GOEA) of the significantly differentially expressed genes in RRS^{WT} and RRS^{SCR} post-infection with *M. luteus* and *Ecc15*. Both RRS^{WT} and RRS^{SCR} show differential expression of immune-responsive genes post-infection.

GOEA is done for 4 different categories (1) Biological Process, (2) Molecular Function, (3) Cellular Component, and (4) Protein Class for RRS^{WT} and RRS^{SCR} post-infection with *M. luteus*. The number of genes enriched in each sub-category is plotted against its corresponding $\{-\log_{10}(\text{FDR})\}$ value. A few key GO terms are highlighted and listed (5).



5	GO Term
1	Defense response
2	Defense response to Gram-positive bacterium
3	Antibacterial humoral response
4	Binding
5	Serine-type endopeptidase activity
6	Endopeptidase activity
7	Extracellular region
8	Extracellular space
9	Serine protease
10	Protease inhibitor

Figure A.3.3: GOEA is done for 4 different categories (B1) Biological Process, (B2) Molecular Function, (B3) Cellular Component, and (B4) Protein Class for RRS^{WT} and RRS^{SCR} post-infection with *Ecc15*. The number of genes enriched in each sub-category is plotted against its corresponding $-\log_{10}(\text{FDR})$ value. A few key GO terms are highlighted and listed (B5).



5
1. Humoral immune response
2. Defense response to bacterium
3. Regulation of transcription, DNA-templated
4. DNA binding
5. Transcription regulator activity
6. Extracellular region
7. Nucleus
8. Metabolite interconversion enzyme
9. DNA-binding transcription factor
10. Oxidoreductase

Figure A.3.4: Gene Ontology (GO) Analysis for significantly differentially expressed genes for RRS^{SCR/WT} post-infection with *M. luteus* (A) and *Ecc15* (B). Number of genes enriched in each category is plotted against their Panther Gene Ontology terms.

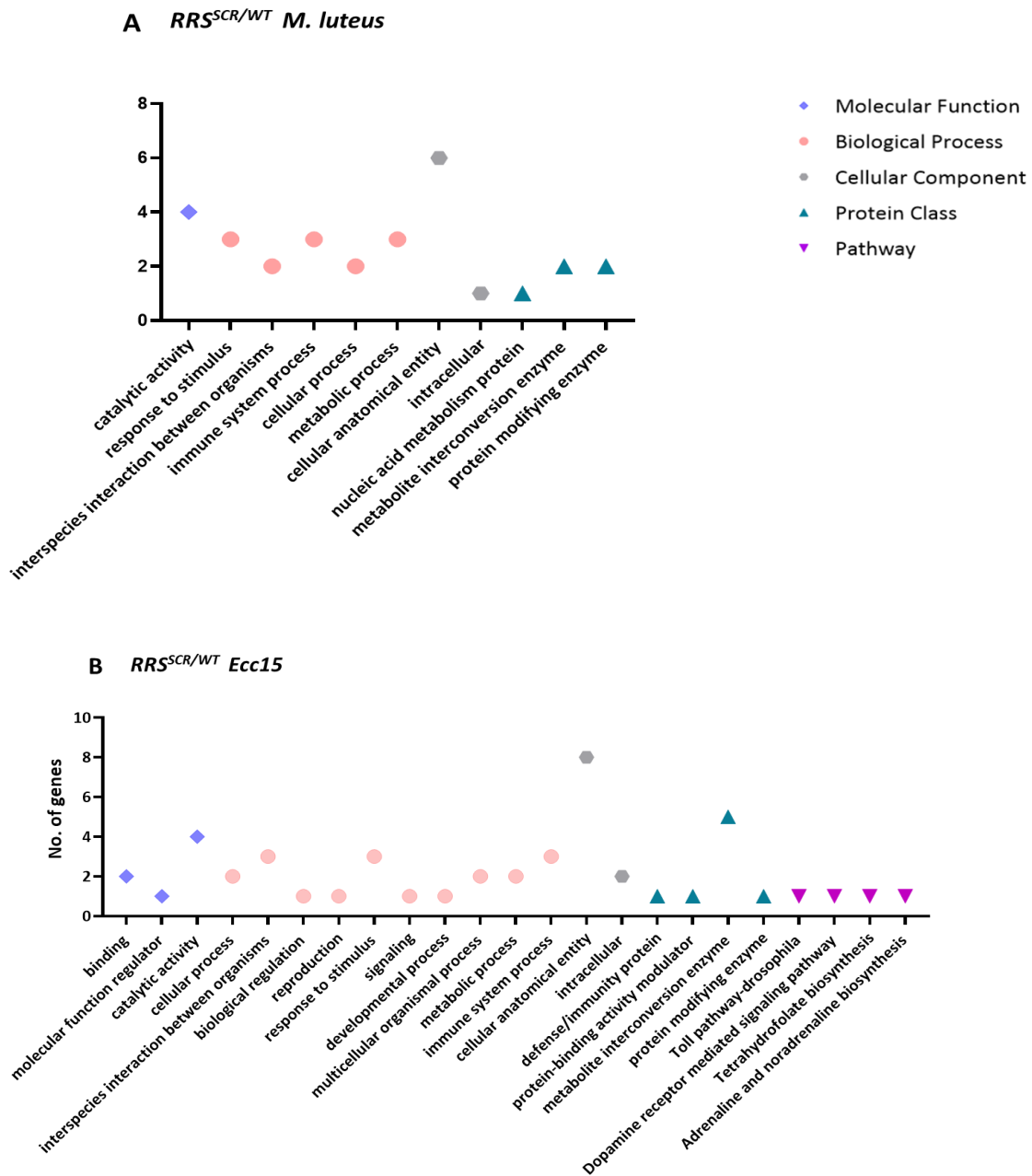
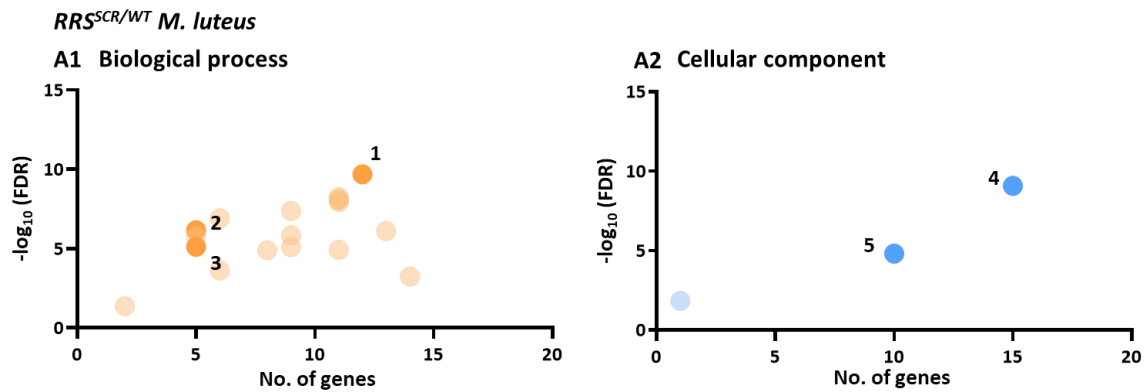


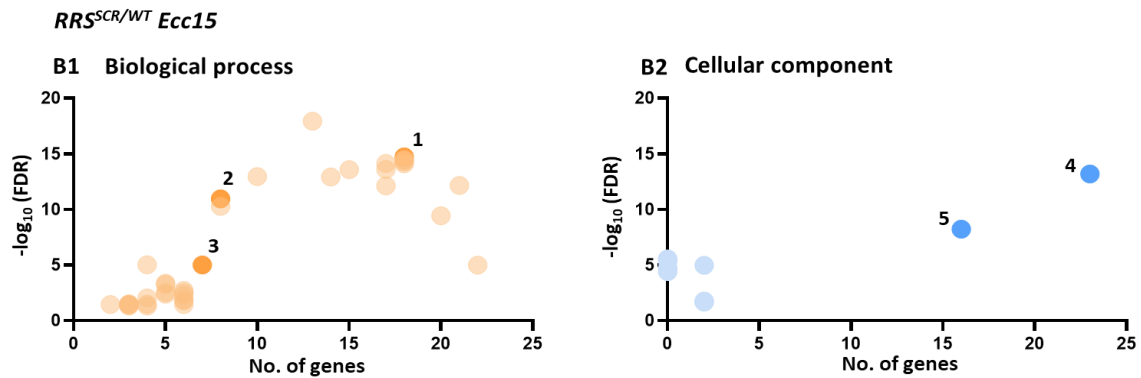
Figure A.3.5: Gene Ontology Enrichment Analysis (GOEA) for significantly differentially expressed genes for $RRS^{SCR/WT}$ post-infection with *M. luteus* and *Ecc15*.

GOEA is done for 2 different categories (A1) Biological process and (A2) Cellular Component and $RRS^{SCR/WT}$ post-infection with *M. luteus*. The number of genes enriched in each sub-category is plotted against its corresponding $\{-\log_{10}(\text{FDR})\}$ value. A few key GO terms are highlighted in the expanded list (A3).



Biological Process		
Category	No.of genes	$-\log_{10}(\text{FDR})$
Defense response (1)	12	9.681936665
Response to biotic stimulus	11	8.248720896
Response to other organism	11	8.123782159
Biological process involved in interspecies interaction between organisms	11	8.049635146
Response to external biotic stimulus	11	7.946921557
Response to bacterium	9	7.374687549
Humoral immune response	6	6.913640169
Antibacterial humoral response (2)	5	6.158015195
Response to stress	13	6.107905397
Immune response	9	5.835647144
Antimicrobial humoral response	5	5.777283529
Defense response to Gram-positive bacterium (3)	5	5.107348966
Immune system process	9	5.083019953
Response to external stimulus	11	4.924453039
Defense response to other organism	8	4.896196279
Defense response to bacterium	6	3.614393726
Response to stimulus	14	3.234331445
Defense response to insect	2	1.370590401
Cellular Component		
Category	No.of genes	$-\log_{10}(\text{FDR})$
Extracellular region (4)	15	9.07109231
Extracellular space (5)	10	4.806875402
Intracellular anatomical structure	1	1.838631998

Figure A.3.6: GOEA is done for 2 different categories (B1) Biological process and (B2) Cellular Component and $RRS^{SCR/WT}$ post-infection with *Ecc15*. The number of genes enriched in each sub-category is plotted against its corresponding $\{-\log_{10}(\text{FDR})\}$ values. A few key GO terms are highlighted in the expanded list (B3).



Biological Process		
Category	No.of genes	$-\log_{10}(\text{FDR})$
Defense response to Gram-positive bacterium	13	17.91721463
Defense response (1)	18	14.73282827
Response to biotic stimulus	18	14.49485002
Response to other organism	18	14.3990271
Response to external biotic stimulus	18	14.27327279
Defense response to other organism	17	14.1266794
Biological process involved in interspecies interaction between organisms	18	14.09691001
Immune response	17	13.58838029
Response to bacterium	15	13.57839607
Humoral immune response	10	12.94309515
Defense response to bacterium	14	12.91009489
Response to external stimulus	21	12.14691047
Immune system process	17	12.1214782
Antibacterial humoral response (2)	8	10.95467702
Antimicrobial humoral response	8	10.26760624
Response to stress	20	9.42945706
Defense response to insect	4	5.014573526
Defense response to Gram-negative bacterium (3)	7	5.005682847
Response to stimulus	22	4.987162775
Cellular process	5	3.359518563
Regulation of defense response to bacterium	5	3.238824187
Regulation of defense response	6	2.723538196
Response to wounding	5	2.536107011
Regulation of response to biotic stimulus	6	2.511449283
Innate immune response	5	2.380906669

Biological Process		
Category	No.of genes	$-\log_{10}(\text{FDR})$
Innate immune response	5	2.380906669
Regulation of immune response	6	2.283996656
Response to fungus	4	2.045757491
Regulation of response to external stimulus	6	1.850780887
Regulation of immune system process	6	1.790484985
Regulation of peptidoglycan recognition protein signaling pathway	3	1.512861625
Positive regulation of defense response	4	1.484126156
Regulation of response to stress	6	1.454692884
Detection of biotic stimulus	2	1.452225295
Positive regulation of Toll signaling pathway	3	1.448550002
Regulation of antibacterial peptide biosynthetic process	3	1.441291429
Regulation of antibacterial peptide production	3	1.390405591
Regulation of antimicrobial peptide biosynthetic process	3	1.312471039
Positive regulation of response to biotic stimulus	4	1.304518324
Cellular Component		
Category	No.of genes	$-\log_{10}(\text{FDR})$
Extracellular region (4)	23	13.18708664
Extracellular space (5)	16	8.214670165
Organelle	0	5.55129368
Intracellular organelle	0	5.36251027
Intracellular anatomical structure	2	4.970616222
Membrane-bounded organelle	0	4.721246399
Intracellular membrane-bounded organelle	0	4.416801226
Larval serum protein complex	2	1.759450752
Cytoplasm	2	1.671620397

Resource Table

Reagent Type (Species/Resource)	Designation	Source/Resource	Identifiers	Additional Information
Genetic reagent	Nos-Cas9/Cyo	Bloomington Drosophila Stock Center, Indiana	BDSC:78781	Ren et. al., / Perrimon Lab Expresses Cas9 in the germline under the control of nos.
Genetic reagent	CG9020-UASgRNA/Cyo	This Study		Transgene dual gRNA inserted at Attp40 site
Genetic reagent	RRS CRISPR control/Fm7i, Actin-GFP	This Study		RRS CRISPR Null Control
Genetic reagent	RRS_C_Δ_6B1/Fm7i, Actin-GFP	This Study		RRS CRISPR Null line
Genetic reagent	RRS_C_Δ_18B1/Fm7i, Actin-GFP	This Study		RRS CRISPR Null line
Genetic reagent	Actin Gal4/Cyo; HA-RRS WT/Ser	This Study		UAS-Gal4 line for RRS wildtype rescue
Genetic reagent	Actin Gal4/Cyo; HA-RRS SCR/Ser	This Study		UAS-Gal4 line for RRS SCR mutant rescue
Genetic reagent	Ubiquitin Gal4/Cyo; HA-RRS WT/Ser	This Study		UAS-Gal4 line for RRS wildtype rescue
Genetic reagent	Ubiquitin Gal4/Cyo; HA-RRS SCR/Ser	This Study		UAS-Gal4 line for RRS SCR mutant rescue
cDNA construct	Aats-Arg (RRS)	BDGP DGC clones	RE02962	AU.46, 35, pFlc-1
cDNA construct	Aats-GluPro (EPRS)	BDGP DGC clones	LD42739	AU.45,25, pOT2
cDNA construct	Aats-Isoleu (IRS)	BDGP DGC clones	LD27166	AU.44,17, pOT2
cDNA construct	Aats-Asp (DRS)	BDGP DGC clones	GM14334	AU.24,41, pOT2
cDNA construct	Aats-Leu (LRS)	BDGP DGC clones	LD44376	AU.44,32, pOT2
cDNA construct	Aats-Gln (QRS)	BDGP DGC clones	GH11673	AU.18,47, pOT2
cDNA construct	Aats-Lys (KRS)	BDGP DGC clones	LD41976	AU.132,86, pOT2
cDNA construct	Aats-Met (MRS)	BDGP DGC clones	GH13807	AU.27,26, pOT2
cDNA construct	AIMP1	cDNA (S2 cells)		
cDNA construct	AIMP2	BDGP DGC clones	LD25772	AU.79,13, pOT2
cDNA construct	AIMP3	BDGP DGC clones	RH48588	AU.57,60, pFlc-1

Reagent Type (Species/ Resource)	Designation	Source/ Resource	Identifiers	Additional Information
Sequence-based reagent	pGEX-4T1 FP	This study (Sigma-Aldrich)	Linearizing pGEX-4T1 for cloning	CCGGAATCCCGGGTCGA
Sequence-based reagent	pGEX-4T1 RP	This study (Sigma-Aldrich)	Linearizing pGEX-4T1 for cloning	TCCTCCGGATCCACGCGG
Sequence-based reagent	pGEX4T1 RRS FP	This study (Sigma-Aldrich)	Cloning RRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTC CGAGCTAAATATGGAGCTG
Sequence-based reagent	pGEX RRS RP	This study (Sigma-Aldrich)	Cloning RRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTACA TTTTCGAAACTGGTTAAGGCCTAG
Sequence-based reagent	pGEX DRS FP	This study (Sigma-Aldrich)	Cloning DRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGG TCGAGGACAAAGAGCAAG
Sequence-based reagent	pGEX DRS RP	This study (Sigma-Aldrich)	Cloning DRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAA GGCGTAAATCGTTTGGGATCG
Sequence-based reagent	pGEX EPRS-WHEP FP	This study (Sigma-Aldrich)	Cloning EPRS-WHEP coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGACAAG GCGATCTGGTCCGAG
Sequence-based reagent	pGEX EPRS-WHEP RP	This study (Sigma-Aldrich)	Cloning EPRS-WHEP coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAA GGTTCTTCTTACCCGGCT
Sequence-based reagent	pGEX QRS FP	This study (Sigma-Aldrich)	Cloning QRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTC AATAAAGCTCAAAGCGAACCC
Sequence-based reagent	pGEX QRS RP	This study (Sigma-Aldrich)	Cloning QRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAAT AGCTTCGTCGGAAGAGCGTGTA
Sequence-based reagent	pGEX IRS FP	This study (Sigma-Aldrich)	Cloning IRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTC AATAAAGCTCAAAGCGAACCTT
Sequence-based reagent	pGEX IRS RP	This study (Sigma-Aldrich)	Cloning IRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAAT AGCTTCGTCGGAAGAGCGTGTA
Sequence-based reagent	pGEX LRS FP	This study (Sigma-Aldrich)	Cloning LRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTC AATAAAGCTCAAAGCGAACCT
Sequence-based reagent	pGEX LRS RP	This study (Sigma-Aldrich)	Cloning LRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAAT AGCTTCGTCGGAAGAGCGG
Sequence-based reagent	pGEX KRS FP	This study (Sigma-Aldrich)	Cloning KRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTC AATAAAGCTCAAAGCGAACCTT
Sequence-based reagent	pGEX KRS RP	This study (Sigma-Aldrich)	Cloning KRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAAT AGCTTCGTCGGAAGAGCGTGTA
Sequence-based reagent	pGEX MRS FP	This study (Sigma-Aldrich)	Cloning MRS coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGAT AATCTACCAATGATGGCAACC
Sequence-based reagent	pGEX MRS RP	This study (Sigma-Aldrich)	Cloning MRS coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTACT TCTTCTTTTGCCCTTGCCCT

Reagent Type (Species/ Resource)	Designation	Source/ Resource	Identifiers	Additional Information
Sequence-based reagent	pGEX AIMP1 FP	This study (Sigma-Aldrich)	Cloning AIMP1 coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGCTGCG TAGAGCTGCCA
Sequence-based reagent	pGEX AIMP1 RP	This study (Sigma-Aldrich)	Cloning AIMP1 coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTACTTCAC ATTGACGTTCTTCAGGG
Sequence-based reagent	pGEX AIMP2 FP	This study (Sigma-Aldrich)	Cloning AIMP2 coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTACGA ACTGAAGACTCTGCTGC
Sequence-based reagent	pGEX AIMP2 RP	This study (Sigma-Aldrich)	Cloning AIMP2 coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAAATTTG AACCCAGAGCTTGCC
Sequence-based reagent	pGEX AIMP3 FP	This study (Sigma-Aldrich)	Cloning AIMP3 coding region into pGEX-4T1	TTCCGCGTGGATCCGGAGGAATGTGTGA CGTTGCAACA
Sequence-based reagent	pGEX AIMP3 RP	This study (Sigma-Aldrich)	Cloning AIMP3 coding region into pGEX-4T1	AGTCGACCCGGGAATCCGGTTAGATGTG GGTGCCTGTGGC
Sequence-based reagent	pET45b-C-HA FP	This study (Sigma-Aldrich)	Linearizing pET45b for cloning	GGAGGATACCCATACGATGTTCCAGATTA CGCTTAAGCGGCCGCACTCGAG
Sequence-based reagent	pET45b-N-His RP	This study (Sigma-Aldrich)	Linearizing pET45b for cloning	TCCTCCGTGATGGTGGTGGTATGTGCC
Sequence-based reagent	pET45b His-RRS-HA FP	This study (Sigma-Aldrich)	Cloning RRS coding region into pET45b with N-terminal His-tag and C-terminal HA tag	CATCACCACCACATCAGGAGGAATGTC CGAGCTAAATATGGAGCTG
Sequence-based reagent	pET45b His-RRS-HA RP	This study (Sigma-Aldrich)	Cloning RRS coding region into pET45b with N-terminal His-tag and C-terminal HA tag	ACATCGTATGGGTATCCTCCTTACATTTTC GAAACTGGTTAAGGCCTAG
Sequence-based reagent	K147R FP	This study (Sigma-Aldrich)	Mutagenesis primer for K147R forward	ATTGCAACCGAGTTGCGCGGACACTGCCC AGCA
Sequence-based reagent	K147R RP	This study (Sigma-Aldrich)	Mutagenesis primer for K147R reverse	TGCTGGGCAGTGTCCGCGCAACTCGGTT GCAAT
Sequence-based reagent	K383R FP	This study (Sigma-Aldrich)	Mutagenesis primer for K383R forward	CCTCTGACAATCGTGCGCTCGGATGGCGG CTTT
Sequence-based reagent	K383R RP	This study (Sigma-Aldrich)	Mutagenesis primer for K383R reverse	AAAGCCGCCATCCGAGCGCACGATTGTCA GAGG
Sequence-based reagent	CG9020 FP2	This study (Sigma-Aldrich)	Null screening primer forward	ACGACATCCTAGAAGTGACTG
Sequence-based reagent	CG9020 RP	This study (Sigma-Aldrich)	Null screening primer reverse	CAGCTAGTGTGAATGCCAAC
Sequence-based reagent	RRS GENOMIC RP	This study (Sigma-Aldrich)	Null screening primer reverse	CCAGATTACGCTGGCGGCATGTCCGAGCT AAATATGGAGCTG

Reagent Type (Species/ Resource)	Designation	Source/ Resource	Identifiers	Additional Information
Sequence-based reagent	pUASp_AttB HA FP	This study (Sigma-Aldrich)	Cloning RRS coding region into pUASp AttB with N-terminal HA-tag	ATAGGCCACTAGTGGATCTGTACCCAT ACGATGTTCCAGATTACGCTGGCGGC
Sequence-based reagent	pUASp_AttB HA RRS FP	This study (Sigma-Aldrich)	Cloning RRS coding region into pUASp AttB with N-terminal HA-tag	ATAGGCCACTAGTGGATCTGTACCCAT ACGATGTTCCAGATTACGCTGGCGGC
Sequence-based reagent	pUASp_AttB HA RRS RP	This study (Sigma-Aldrich)	Cloning RRS coding region into pUASp AttB with N-terminal HA-tag	GAGGTCGACTCTAGAGGATCTTACAT TTTCGAACTGGTTTAAAGCCCTAG
Sequence-based reagent	pUASp AttB RRS Seq FP	This study (Sigma-Aldrich)		ATAGGCCACTAGTGGATCTG
Sequence-based reagent	pUASp AttB RRS Seq RP	This study (Sigma-Aldrich)		GAGGTCGACTCTAGAGGATC
Sequence-based reagent	RRS RT FP	This study (Eurofins)	RT-PCR primer	GCCTTGCCGCCAGAATACA
Sequence-based reagent	RRS RT RP	This study (Eurofins)	RT-PCR primer	GGCAGTTGATTCTCAGCAAT
Sequence-based reagent	rp49 FP	This study (Sigma-Aldrich)	RT-PCR primer	GACGCTCAAGGGACAGTATC
Sequence-based reagent	rp49 RP	This study (Sigma-Aldrich)	RT-PCR primer	AAACGCGGTTCTGCATGAG
Sequence-based reagent	AttD FP	This study (Sigma-Aldrich)	RT-PCR primer	CGGTCAACGCCAATGGTCAT
Sequence-based reagent	AttD RP	This study (Sigma-Aldrich)	RT-PCR primer	CATTAGAGCGGCGTTATTG
Sequence-based reagent	DptB FP	This study (Sigma-Aldrich)	RT-PCR primer	ACCGCAGTACCACTCAATC
Sequence-based reagent	DptB RP	This study (Sigma-Aldrich)	RT-PCR primer	GGTCCACACCTTCTGGTGAC
Sequence-based reagent	Drs FP	This study (Sigma-Aldrich)	RT-PCR primer	CTGTCCGGAAGATACAAGGG
Sequence-based reagent	Drs RP	This study (Sigma-Aldrich)	RT-PCR primer	TCGCACCAGCACTTCAGACT
Sequence-based reagent	Irc FP	This study (Sigma-Aldrich)	RT-PCR primer	TAGGCAAAGCGACTGGAGGACA
Sequence-based reagent	Irc RP	This study (Sigma-Aldrich)	RT-PCR primer	GCAAGCTGGACTTAAGGATCTTC
Antibody	Anti-HA (Rabbit monoclonal)	Millipore	DW2, Millipore	WB (1:3000)
Antibody	Anti-His (Mouse monoclonal)	SIGMA	H1029, SIGMA	WB (1:1000)
Antibody	Anti-GST (Mouse monoclonal)	sc53909, Santa-Cruz-Biotechnology		WB (1:5000)
Software, Algorithm	GraphPad Prism 8.0.2	Prism		
Software, Algorithm	ImageJ, Fiji	National Institutes of Health		

Appendix II

Generation of $\Delta brwl$ by CRISPR Cas9-based Genome Editing

MADF-BESS family of genes is a family of sixteen transcription factors having a DNA binding N-terminal **MADF** (**Myb Adf-1**) domain and a protein-protein interaction C-terminal domain **BESS** (**BEAF-1**, **Stonewall**, **su (var)** domain). Adf and Dlp3 function as transactivators (England, Admon, and Tjian 1992; Cutler, Perry, and Tjian 1998; Bhaskar and Courey 2002; Ratnaparkhi, Duong, and Courey 2008). Coop, Hng1, Hng2, and Hng3 act as repressors for wing hinge development (Song et al. 2010; Shukla et al. 2014). *stwl* (*stonewall*) is required for germ cell development (Maines et al. 2007). It is required for the maturation of the cytoblast into nurse cells and the specification of the oocyte (Clark and McKearin 1996). A study in our lab has shown that *brwl* (*brickwall*) like *stwl* is required for germ cell maintenance, cyst formation, and oocyte specification. *brwl* mutants; *brwl*^{KG00824} (BDGP Gene Disruption project collection) and *brwl*^{MI054561} {transposon Minos-mediated integration cassette (MiMic) collection} wherein P-element is inserted in the N-terminus and C-terminus of the gene respectively; show age-dependent ovary phenotypes. Over-expression of *stwl* using a germ cell driver rescues the defects, indicating *brwl* is a genetic interactor and positive regulator of *stwl* in ovary development (Shukla et al. 2018).

A complete loss of function (lof) variant was not available, thus we used CRISPR Cas-9 to generate the same. A transgenic dual guide RNA line (*U6-brwl*^{dual gRNA}) was created (refer to section Materials and Methods) to express *gRNA* that could recognize the beginning of exon1 and the end of exon3 region of the *brwl* gene (inverted red triangles). Our goal was to remove the coding region to create $\Delta brwl$ animal. The *U6-brwl*^{dual gRNA} lines were crossed to *nos-Cas9* animal and 80 lines from each cross were stabilized by balancing the putative lofs' over a second chromosome balancer, *Cyo*; where the balancer chromosome expresses GFP (Fig. 1A-B). Since the genomic *brwl* is located on the second chromosome, putative *brwl* lof lines were screened by PCR for confirmation of the deletion. Single fly genomic PCR for the extended gene region of *brwl* was performed on heterozygous females where the primers used, flanked the ends of the coding region of the gene, and the mutations were confirmed through visualization of migration of the bands upon gel electrophoresis (Fig. 2A-D) followed by sequencing. Of these lines, four lines having precise (namely as *brwl*^{1a/1b/1c/1d}) and two lines having imprecise deletion (namely as *brwl*^{2/3}) were maintained. Three lines wherein the *brwl*

gene remained intact were maintained as CRISPR wildtype controls (namely as *brwl*^{C1/C2/C3}) (Fig. 2E).

Materials and Methods

Generation of *brwl* lof lines using CRISPR Cas9 technology

CRISPR Cas9 technology was employed to generate *brwl* loss of function (lof) fly lines. Single guide (sg)-RNAs targeting the *brwl* coding region at the beginning of exon 1 and the end of exon3 were designed using CRISPR Optimal Target Finder {COTF;(Gratz et al. 2014)}, a web tool for identifying CRISPR target sites and evaluating their specificity. The *brwl* gene region was sequenced prior to the experimental design to account for SNPs at the sgRNA target sites. The sg-RNAs (sg1 and sg2) were cloned into the *pU6-BbsI-chiRNA* (Addgene #45946) plasmid, which was then docked using *y¹v¹*, *P{CaryP}attP2* Drosophila line (BDSC-36303) by transgenic injections, at the NCBS-CCAMP transgenic facility, Bengaluru, India. The transgenic dual sg-RNA lines were crossed to the *nanos-Cas9* (BDSC-54591) line. The founder male progenies were crossed to *w⁺; Tft/Cyo* balancer females wherein Cas9-gRNA complex is formed in the germline. In the next generation, three heterozygous female progenies from each cross (160 lines, each labelled; L1-1 to L1-80; L2-1 to L2-80) were maintained as a separate line over a *Cyo* balancer (Fig. 1A-B). Since the genomic *brwl* is located on the second chromosome, putative *brwl* lof lines were screened by PCR-based confirmation of the deletion. Single-fly PCR for the extended gene region of *brwl* was performed on heterozygous females and the mutations were confirmed through sequencing (Fig. 2A-D).

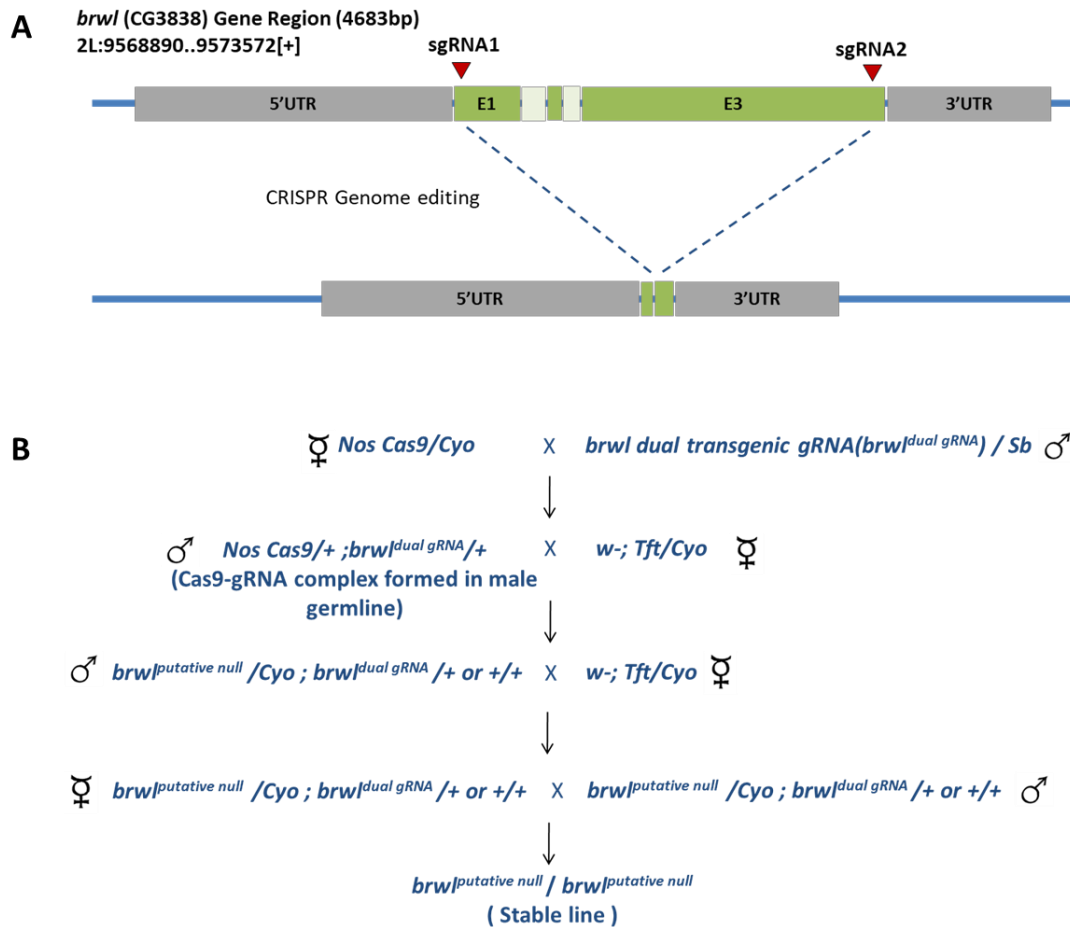
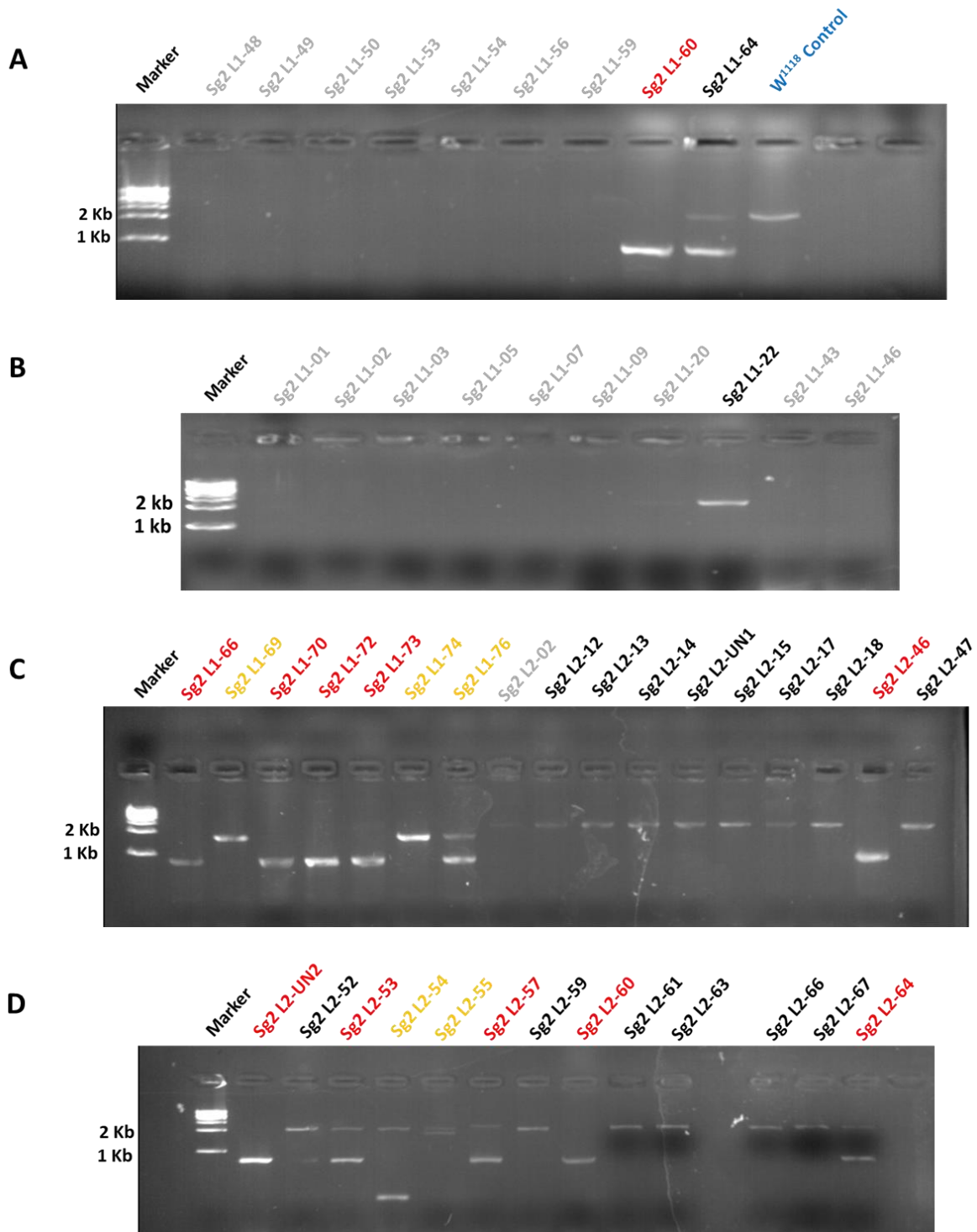


Figure 1: A *brwl* loss of function line was generated using CRISPR Cas9 genome editing. (A) Design of the dual guide-RNA for excision of *brwl* locus. The *brwl* gene is located on the second chromosome. It has five exons, which code for two annotated transcripts; the longer one spanning 3091 bp and the shorter one spanning 2256 bp. Two gRNAs (inverted red triangles) were designed at beginning of exon1 and the end of exon3. Our goal was to excise the coding region and generate a *brwl* loss of function line. (B) Excision of *brwl* locus to generate a *brwl-lof* line. A *U6-brwl*^{dual gRNA} line was generated (see section “Materials and Methods”) and crossed to *nos-Cas9* animals. 80 lines were balanced over a second chromosome balancer and screened for deletions via PCR.



E	Re-named as:
Sg2 L1-60	<i>brwl</i> ^{1a}
Sg2 L1-66	<i>brwl</i> ^{1b}
Sg2 L2-46	<i>brwl</i> ^{1c}
Sg2 L2-UN2	<i>brwl</i> ^{1d}
Sg2 L1-74	<i>brwl</i> ²
Sg2 L2-54	<i>brwl</i> ³
Sg2 L2-13	<i>brwl</i> ^{C1}
Sg2 L1-22	<i>brwl</i> ^{C2}
Sg2 L2-14	<i>brwl</i> ^{C3}

Figure 2: PCR based screening to detect a successful *brwl* gene knockout after a CRISPR genome editing experiment. A deletion spanning ~1kb in the *brwl* gene was generated using dual gRNA strategy, creating a lof mutant allele. PCR primers flanking the expected deletion were designed for detecting the lof allele in a population of flies such that the homozygous mutant fly would generate an ~1kb amplicon, while in unmodified wildtype fly, a PCR product of 2 kb is generated, as expected for the intact gene. The mutant or the balancer heterozygous fly yields both PCR products. The PCR products were run on 1% Agarose gel and stained with Ethidium Bromide for visualization. Lines labelled in ■, ■, ■, ■ denote no amplification, no deletion detected, precise deletion, and imprecise deletion respectively.

References

- Bhaskar, Vinay, and Albert J. Courey. 2002. "The MADF-BESS Domain Factor Dip3 Potentiates Synergistic Activation by Dorsal and Twist." *Gene* 299 (1-2): 173–84.
- Clark, K. A., and D. M. McKearin. 1996. "The Drosophila Stonewall Gene Encodes a Putative Transcription Factor Essential for Germ Cell Development." *Development* 122 (3): 937–50.
- Cutler, Gene, Kathleen M. Perry, and Robert Tjian. 1998. "Adf-1 Is a Nonmodular Transcription Factor That Contains a TAF-Binding Myb-Like Motif." *Molecular and Cellular Biology*. <https://doi.org/10.1128/mcb.18.4.2252>.
- England, B. P., A. Admon, and R. Tjian. 1992. "Cloning of Drosophila Transcription Factor Adf-1 Reveals Homology to Myb Oncoproteins." *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.89.2.683>.
- Gratz, Scott J., Fiona P. Ukken, C. Dustin Rubinstein, Gene Thiede, Laura K. Donohue, Alexander M. Cummings, and Kate M. O'Connor-Giles. 2014. "Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila." *Genetics* 196 (4): 961–71.

- Maines, Jean Z., Joseph K. Park, Meredith Williams, and Dennis M. McKearin. 2007. "Stonewalling *Drosophila* Stem Cell Differentiation by Epigenetic Controls." *Development* 134 (8): 1471–79.
- Ratnaparkhi, Girish S., Hao A. Duong, and Albert J. Courey. 2008. "Dorsal Interacting Protein 3 Potentiates Activation by *Drosophila* Rel Homology Domain Proteins." *Developmental and Comparative Immunology* 32 (11): 1290–1300.
- Shukla, Vallari, Neena Dhiman, Prajna Nayak, Neelesh Dahanukar, Girish Deshpande, and Girish S. Ratnaparkhi. 2018. "Stonewall and Brickwall: Two Partially Redundant Determinants Required for the Maintenance of Female Germline in." *G3* 8 (6): 2027–41.
- Shukla, Vallari, Farhat Habib, Apurv Kulkarni, and Girish S. Ratnaparkhi. 2014. "Gene Duplication, Lineage-Specific Expansion, and Subfunctionalization in the MADF-BESS Family Patterns the *Drosophila* Wing Hinge." *Genetics* 196 (2): 481–96.
- Song, Haiyun, Sandra Goetze, Johannes Bischof, Chloe Spichiger-Haeusermann, Marco Kuster, Erich Brunner, and Konrad Basler. 2010. "Coop Functions as a Corepressor of Pangolin and Antagonizes Wingless Signaling." *Genes & Development*. <https://doi.org/10.1101/gad.561310>.