Biological roles for SUMOylation of Arginyl tRNA synthetase

A thesis

submitted to
Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme
by

Aarti Kejriwal
20151174

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

March 2020
Supervisor: Dr. Girish Ratnaparkhi

All rights reserved

Certificate

This is to certify that this dissertation entitled '*Biological roles for SUMOylation of Arginyl tRNA synthetase*' towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Aarti Kejriwal at Indian Institute of Science Education and Research under the supervision of Dr. Girish Ratnaparkhi, Associate Professor, Department of Biology, during the academic year 2019-2020.

Dr. Girish Ratnaparkhi

Committee:

Dr. Girish Ratnaparkhi

Dr. Saikrishnan Kayarat

This thesis is dedicated to the thousands of flies that sacrificed their lives for me.

Declaration

I hereby declare that the matter embodied in the report entitled ‘Biological roles for SUMOylation of Arginyl tRNA synthetase’ are the results of the work carried out by me at the Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Girish Ratnaparkhi and the same has not been submitted elsewhere for any other degree.

Aarti Kejriwal

Date: 05/04/2020

Table of Contents

Declaration **4**

Abstract **8**

Acknowledgments **9**

Chapter 1 Introduction **10**

Objectives of MS research project **15**

Chapter 2 Generation of a SUMO Conjugation Resistant Arginyl tRNA Synthetase (SCR-RRS) **17**

- Summary 17
- Introduction 17
- Results 18
- Discussion 22
- Conclusions 22
- Contributions 22

Chapter 3 Using CRISPR/Cas9 based genome editing to generate a RRSSCR Drosophila transgenic **23**

- Summary 23
- Introduction 23
- Materials & Methods 26
- Results 29
- Discussion 34
- Conclusions 35
- Future Work 35
- Contributions 35

Chapter 4 Generation of a transgenic animal expressing SCR-RRS using the Null: Rescue methodology **36**

- Summary 36

Introduction 36
Materials & Methods 38
Results 41
Discussion 50
Conclusions 51
Future Work 51
Contributions 51

Future directions **.. 52**

Appendix **.. 54**

References **.. 58**

List of Figures

Fig. 1: Two-step reaction mechanism of aminoacylation by aaRSs (Kim 2014) 10
Fig. 2: Hypothetical MARS complex interactions in 2D 11
Fig. 3: Classic example of Non-canonical function of aaRS (Figure adapted from Arif et al. 2009.) 12
Fig. 4: Overview of non-canonical or moonlighting functions known of all the components of the MARS complex as described by *Park, Schimmel, and Kim 2008* 13
Fig. 5: The multi-step SUMOylation reaction in *Drosophila*. 15
Fig. 6: SUMO conjugation of tRNA synthetases(Figure adapted from Handu et al. 2015). 18
Fig. 7: RRS gets SUMOylated using *in-bacto* SUMOylation system 19
Fig. 8: Putative SUMO sites on RRS predicted by an online tool: JASSA. 20
Fig. 9: Expression of RRS and its predicted SUMO conjugation resistant (SCR) mutants (*In bacto*) 21
Fig. 10: RRS protein sequence alignment of multiple model organisms 21
Fig. 11: CRISPR-Cas9 technology directed gene-editing mechanisms(Cribbs and Perera 2017) 24
Fig. 12: pHD-ScarlessDsRed donor vector 25
Fig. 13: Site-directed mutagenesis for homology arms of scarless donor plasmid ... 29

Fig. 14: Strategy for creating SUMO-conjugation resistant RRS using scarless CRISPR-Cas9 technique 30
Fig. 15: RRS gRNA sequencing results. 30
Fig. 16: Amplification of vector fragments and homology regions with K→R mutations 31
Fig. 17: Gibson assembly for PCR amplified scarless fragements 32
Fig. 18: Colony PCR for transformants of NEBuilder HiFi DNA assembly kit (Gibson assembly with high fidelity polymerase)..33
Fig. 19: DNA digestion of plasmid extracted from a positive colony after Gibson assembly 34
Fig. 20: UAS-GAL4 driven overexpression of genes in *Drosophila melanogaster*....37
Fig. 21: Generation of RRS CRISPR null lines 42
Fig. 22: Homozygous RRS null is not embryonic lethal but dies in early 2nd instar stage 44
Fig. 23: RRS gene is haplosufficient in *Drosophila melanogaster* 45
Fig. 24: Genomic PCR for $RRS^{C\Delta}$ does show any changes in the amplicon size as compared to the wildtype (W^{1118})...46
Fig. 25: RRS null rescue with RRS wildtype and RRS-SCR. 47
Fig. 26: Mild increase in transcript levels of Met and Dros in RRS^{SCR}. 48
Fig. 27: Enhanced expression of *DiptA* and *AttD* in RRS^{SCR} 49
Fig. 28. A figure showing changes in sequences in the RRS null line 6B1 and 18B1. 57

List of Tables

Table 1: RRSCΔ is larval lethal 43
Table 2: Primers for PCR...54
Table 3. Drosophila lines generated 56

Abstract

Aminoacyl tRNA synthetases (aaRS), many of which are components of the 1.2 Megadalton Multiple aminoacyl tRNA complex (MARS), are SUMO-conjugated in response to infection (Handu et al. 2015) in *Drosophila*, suggesting that this post-translational modification (PTM) of AARS is part of host-defense response to pathogens. Small Ubiquitin like modifier (SUMO) is an important post translational (PTM) in the cell, with 10-30% of cellular proteins being SUMO conjugated.

We are studying the SUMOylation of Arginyl tRNA synthetase (RRS) and its role in the fly immune response. First, we validated the SUMOylation of RRS. Next, we used lysine mutagenesis to identify K147 and K383 as SUMO targets. The RRS (K147R, K383R) double mutant appears to be resistant to SUMO conjugation (RRS^{SCR}). In order to uncover roles for SUMOylation, we planned to compare the immune response of RRS^{wt} and RRS^{SCR}. To achieve this, our strategy was two-fold. One, we generated RRS^{null} animals using CRISPR/Cas9 based gene editing. The null animals were rescued with RRS^{wt} and RRS^{SCR} transgenes; using the UAS/Gal4 system. Two, we planned to use CRISPR/Cas9 editing to directly edit the RRS genomic locus and replace codons for K147 and K383 with those that code for Arginine.

The transgenic RRS^{SCR} animal, once generated through either one or both of the strategies listed above will be used to uncover biological roles for SUMO conjugation of RRS, in the context of the immune response.

Acknowledgments

I express my heartfelt gratitude to Dr. Girish Ratnaparkhi, Department of Biology, Indian Institute of Science Education and Research, Pune for the giving me the opportunity to work in his lab, and for his valuable guidance and support during my project. I also thank him for his constant support in pursuing my ideas. I also convey my sincere gratitude towards Dr. Saikrishnan Kayarat, IISER, Pune, for his comments during my evaluation and encouraging me to work towards by goal.

I am especially grateful to Prajna Nayak for her invaluable input, support and mentoring throughout my project. I would like to acknowledge the constant help, feedback and support from all the Ratnaparkhi lab members without which being here would have been far more challenging. I thank all of them for maintaining a conducive and friendly environment for conducting research.

IISER Pune fly stock and media facility's regular and efficient technical services as well as NCBS fly facility's help in the embryo injections and transgenic fly line generation are truly appreciated. I thank the Department of Science and Technology, Government of India, for granting the INSPIRE Scholarship for Higher Education to me. Lastly, I thank IISER Pune for providing me with an excellent infrastructure and all the opportunities for the last 5 years that have helped me not only in cultivating a scientific mind but have also transformed me as an individual.

Chapter 1

Introduction

Aminoacyl tRNA synthetases (aaRS) are enzymes that catalyze the charging of tRNAs with their cognate amino acids (Arnez and Moras 1997). Every amino acid has a unique aaRS to catalyse its charging (Ibba et al. 2000). Aminoacylation is a two-step reaction mechanism (Fig.1), which begins with the formation of the aminoacyl-adenylate by nucleophilic attack of the amino acid's α-carboxylate carbon on the α-phosphate of adenosine triphosphate (ATP) (Kim 2014). This is followed by the transfer of the aminoacyl-adenylate to the cognate tRNA and release of adenosine monophosphate (AMP).

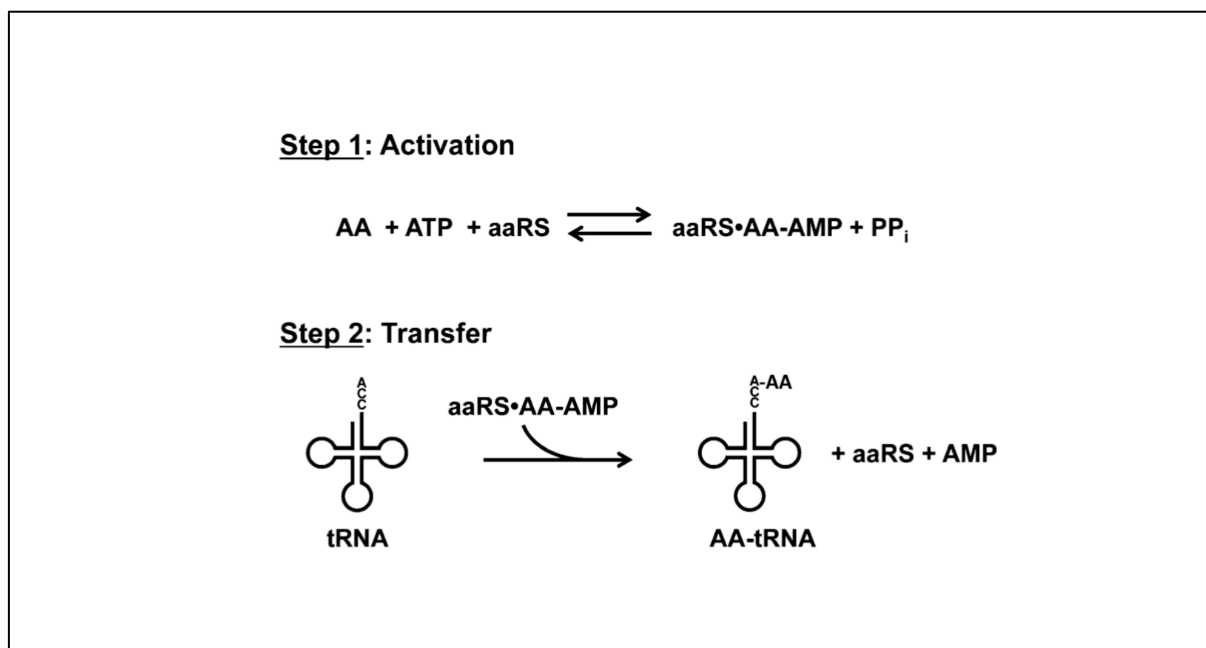


Fig. 1: Two-step reaction mechanism of aminoacylation by aaRSs (Kim 2014). Step 1: aaRS activates the amino acid by formation of aminoacyl-adenylate using ATP. Step 2: The aminoacyl-adenylate is then transferred to cognate tRNA and AMP is released.

Typically, all aaRSs have a catalytic domain and anticodon binding domain. Some aaRS have additional editing domains to increase the fidelity of protein synthesis (Schmidt and Schimmel 1994). In addition to facilitating aminoacylation, many synthetases show non-canonical functions. They have extra N- or C- terminal domains that help them bind to other proteins or RNA (Ray et al. 2011). For example, in eukaryotes, fused terminal domains are found in many aaRSs that help them form a complex known as the multiple aminoacyl tRNA synthetase (MARS) complex or multi-synthetase complex (MSC; Fig.2).

MARS Complex or MSC

The complex includes 9 different aaRSs- glutamyl prolyl tRNA-synthetase (EPRS), aspartyl-tRNA synthetase (DRS), isoleucyl-tRNA synthetase (IRS), leucyl-tRNA synthetase (LRS), methionyl tRNA-synthetase (MRS), glutaminyl-tRNA synthetase (QRS), arginyl tRNA synthetase (RRS), and lysyl-tRNA synthetase (KRS) and 3 non-synthetase proteins (AIMP1, AIMP2, and AIMP3). This complex structure is conserved from *Drosophila* (Fig. 2B) to humans (Fig.2A) (Kerjan et al. 1994). Over evolutionary timescales, new domains have been added to the ARSs keeping the catalytic domain conserved. These new domains are involved in performing non-canonical functions as well as forming complexes like MSC. Almost half the number of total aaRSs in the eukaryotic cell are a part of the MARS complex.

This complex is proposed to channel tRNAs for aminoacylation and further for translation and also serve as a reservoir for proteins which can be released upon some stress related cues (Hyeon et al. 2019). The complex size and density increases from yeast to mammals, suggesting that the MSC might play a role in the system development of complex organisms.

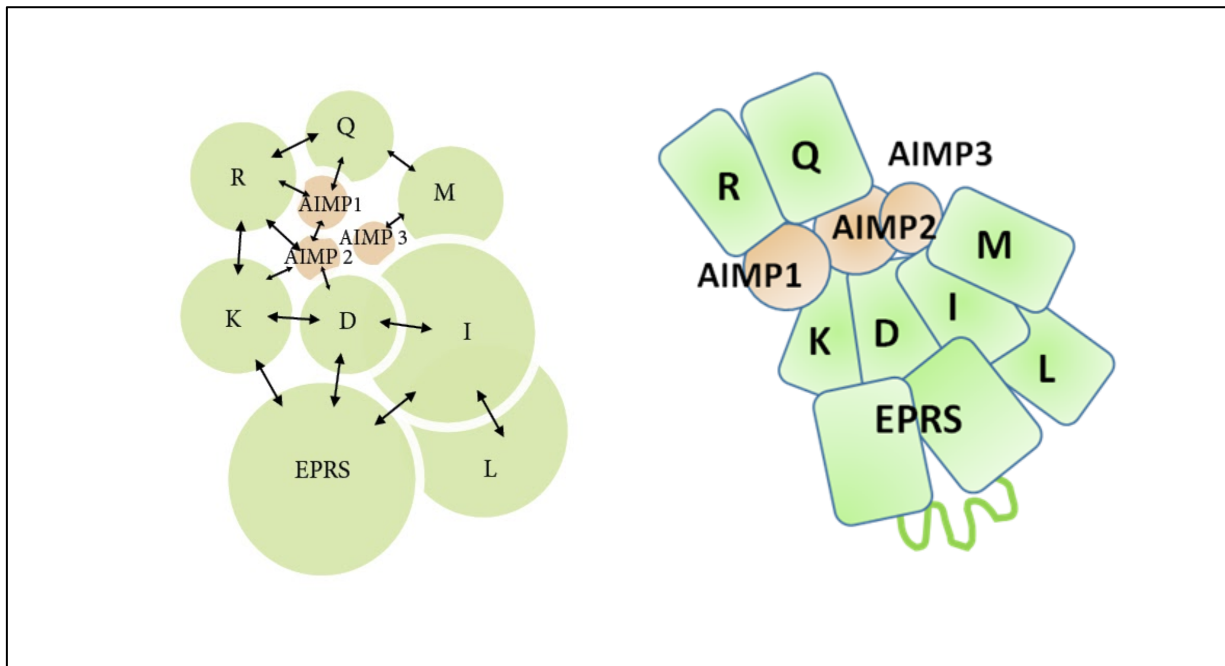


Fig. 2: Hypothetical MARS complex interactions in 2D. A. Schematic of Mammalian MARS complex showing all the predicted interactions. Arrows represent the interactions between different components of the complex. Green circles are all aaRS and brown are non-synthetase proteins of the complex. B. Schematic of Drosophila MARS complex. Green circles are all aaRS and brown are non-synthetase proteins of the complex.

Non-canonical functions of tRNA synthetases

We know that aaRSs are housekeeping enzymes and are involved in translation in all life forms. With evolution, many additional domains have been added to these genes which gives them potential to perform some non-translational functions upon various signalling cues. For example, QRS represses the activity of apoptosis signal regulating kinase-1(ASK-1) in the presence of glutamine (Fig.4; Ko et al. 2001). Post-translational modifications control many of these non-canonical functions like KRS gets phosphorylated upon immune challenge and is released from the MARS complex to regulate the transcriptional factor MITF (Fig.4) Similarly, EPRS also gets phosphorylated and released from MARS complex upon interferon-γ stimulation (Fig.3; Arif et al. 2009).

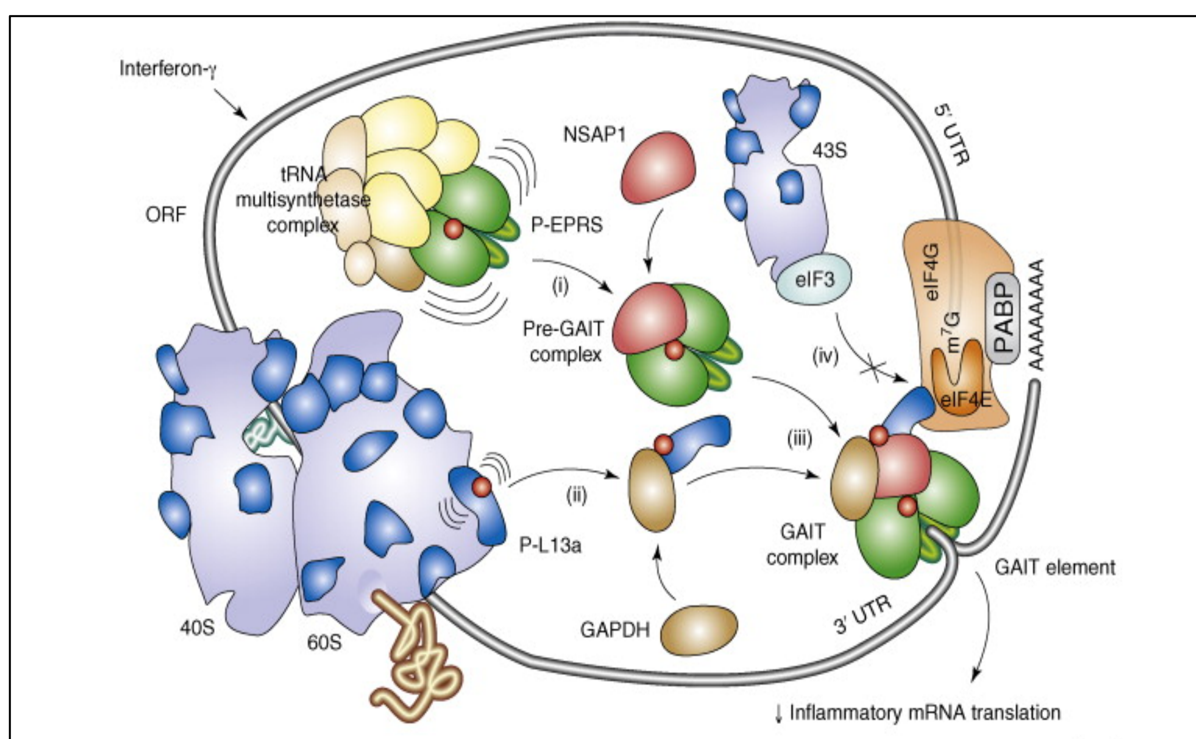


Fig. 3: Classic example of Non-canonical function of aaRS. Schematic representation of GAIT-mediated translational control. IFN-γ induces the phosphorylation of the EPRS leading to its detachment from the MARS complex. Phosphorylated EPRS (P-EPRS) becomes a part of the inactive pre-GAIT complex. Upon activation of the GAIT complex by phosphorylated L13a and GAPDH, it binds to the GAIT element in the 3'UTR of target transcripts to repress their translation (Arif et al. 2009). (Figure adapted from Arif et al. 2009.)

Just like there are glucose and fatty acid sensors, there should be sensors for determining the levels of amino acids in the cells as well. GCN2 senses the intracellular level of amino acids indirectly by binding to uncharged tRNAs and amino acid transporters sense the extracellular levels. To be true sensors, the protein must

bind to the specific amino acids. Since, aaRS directly bind to the amino acids, they can be true amino acid sensors. *In vitro* cellular and biochemical assays have shown that aaRSs have additional aminoacyl transferase activities which help in amino acid sensing (He et al. 2018). This is done by lysine aminoacylation (K-AA) of their specific substrates which is, modifying 3-amine of lysine in proteins with the reactive bond of aminoacyl-AMP. Formation of K-AA enables cell signalling through amino acid modification. These are reversible modifications which depend on the level of amino acids present in the cell and can be removed my deacetylases.

aaRSs are also involved in proofreading activities that selectively deacetylate non-cognate aa-tRNAs in cis or trans. aaRS proofreading ensures correct incorporation of cognate amino acid to a tRNA which maintains the fidelity of translation. Some aaRS have additional editing domains for specific aa-tRNAs misincorporations for example, AlaXP family proteins have editing domains for Ser-tRNA (Guo and Schimmel 2012).

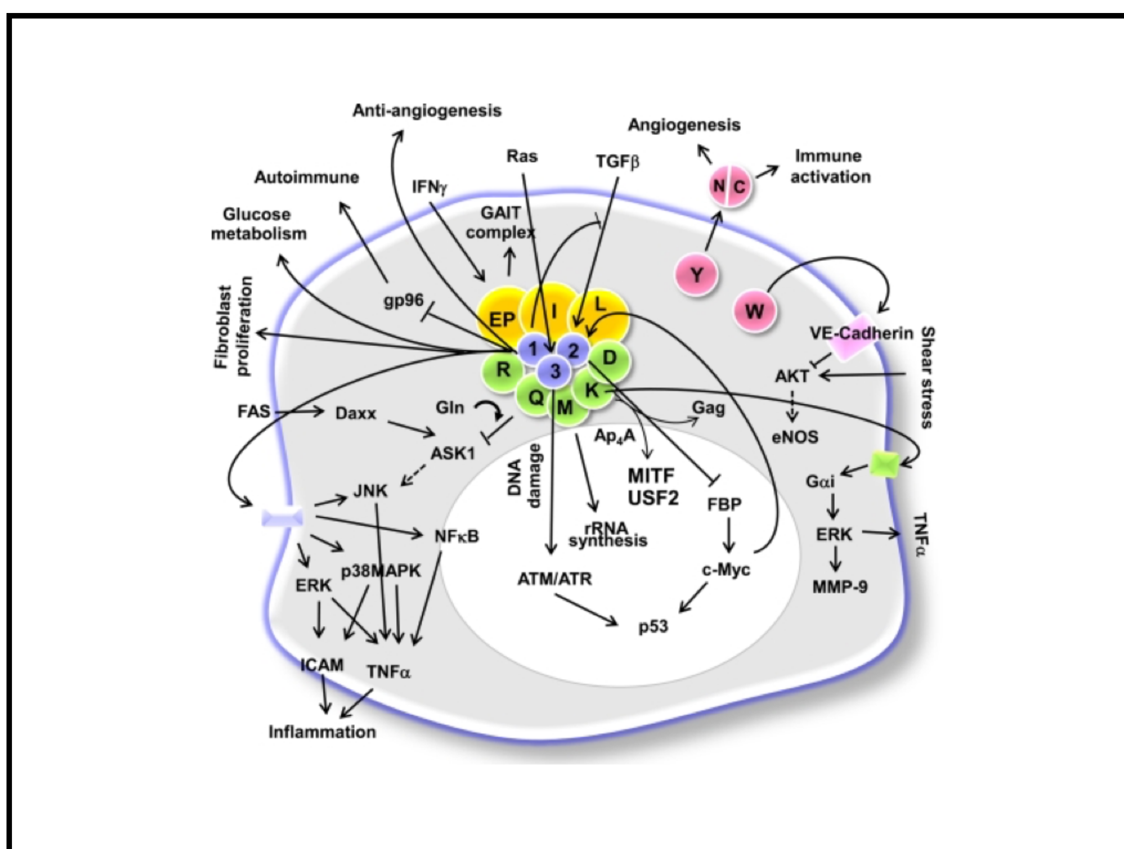


Fig. 4: Overview of non-canonical or moonlighting functions known of all the components of the MARS complex as described by *Park, Schimmel, and Kim 2008*. QRS regulates apoptosis in a glutamine dependent manner by binding to ASK-1. MRS stimulates rRNA synthesis. EPRS, upon IFN-γ stimulation forms leaves the MARS complex and forms

the GAIT complex. KRS activates microphthalmia-associated transcription factor (MITF) and upstream stimulatory factor (USF2) through Ap_4A to control transcriptionally target genes. AIMP1 is involved in inflammation, angiogenesis, wound healing, and glucose metabolism, and influences the autoimmune response via gp96. AIMP2 suppresses c-Myc via ubiquitin-dependent degradation of far upstream element binding protein (FBP). AIMP3 is released from the complex upon DNA damage to activate p53. (Figure adapted from Park, Schimmel, and Kim 2008)

tRNA synthetases in diseases

Many aaRS are involved in different diseases which are not related to their translation functions. They are found to be involved in neuronal, auto-immune diseases as well as in different cancers. Heritable mutations in GRS (Glycyl-tRNA synthetase) and YRS (Tyrosyl-tRNA synthetase) result in a motor and sensory neuropathy called Charcot-Marie-Tooth (CMT) (Park, Schimmel, and Kim 2008). The mutant proteins are shown to be fully active translationally in many cases. Mutations in the editing domains of aaRS result in ataxia without affecting it's aminoacylation activity. GRS is shown to be up-regulated in papillary thyroid carcinoma(PTC) (Scandurro et al. 2001) and Lysyl-tRNA synthetase (KRS) is overexpressed in breast cancer (Park et al. 2005). Angiogenesis is an important component of cancer development. Thus, proteins regulating angiogenesis have roles in tumorigenesis. Glutamyl-prolyl-tRNA synthetase (EPRS) negatively regulates angiogenesis by suppressing vascular endothelial growth factor (VEGF-A) (Ray and Fox 2007). YRS and WRS act as pro and antiangiogenic cytokines and regulate this process (Tzima et al. 2003; Liu 2004). There are many autoantibodies against different aaRS which cause autoimmune diseases called antisynthetase syndrome.

SUMOylation

Small Ubiquitin-like modifier (SUMO) proteins are small proteins that modify other proteins in the cell by covalently attaching to them. SUMO is conserved from yeast to mammals, although the number of gene copies varies in different organisms. *Drosophila melanogaster* has a single SUMO gene (smt3) (Huang et al. 1998) whereas humans have four (SUMO 1,2,3 & 4) (Talamillo et al. 2008).

SUMOylation is a post-translational modification that is involved in various cellular activities like growth, differentiation and stress response. It is a multi-step enzyme process (Fig.5) where SUMOs are translated as precursor proteins which are further processed by isopeptidases called sentrin-specific protease (SENPs). In *Drosophila*

there are ubiquitin-like protease (Ulp) family proteins which are similar to SENP family proteins in humans (Smith et al. 2012).

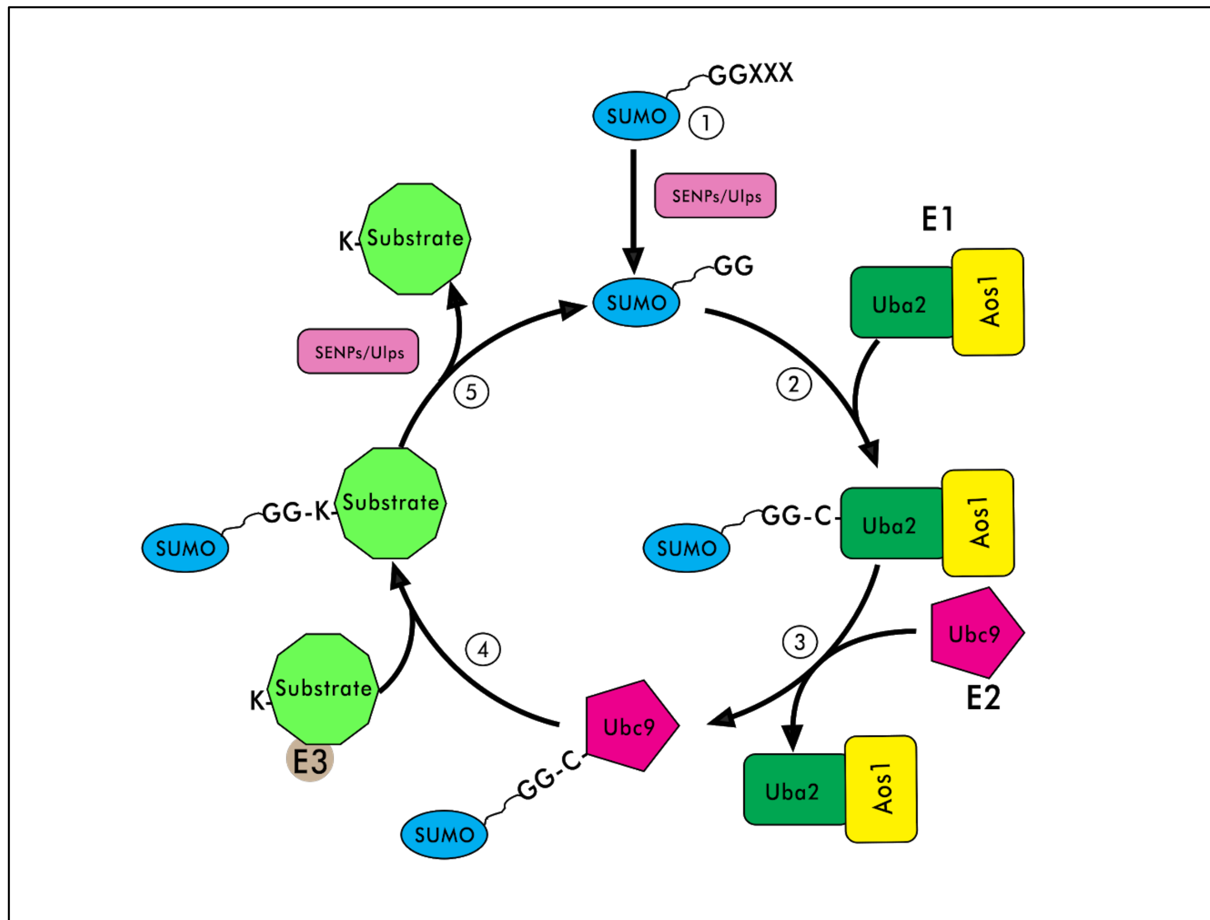


Fig. 5: The multi-step SUMOylation reaction in *Drosophila*. Precursor of SUMO is cleaved by ubiquitin-like protease family proteins into a mature form which is transferred to the target protein by a series of enzymes. E1 SUMO activating enzyme (Aos1/Uba2) activates SUMO, an E2 SUMO conjugation enzyme (Ubc9) helps in conjugation and an E3 SUMO ligase finally ligates it onto the target protein. SUMOylation is reversible and deconjugation is facilitated again by Ulp proteins.

They remove the C-terminal residues to generate mature SUMO with diglycine. Mature SUMO then interacts with SUMO activating enzyme(SAE1/2) to form a SUMO-SAE1/2 complex (Gareau and Lima 2010). The activating enzyme transfers SUMO to Ubc9 which is SUMO E2 conjugating enzyme. Ubc9 then transfers SUMO to its protein substrate where SUMO forms a covalent bond with its lysine residue. Since, post-translational modifications are reversible, SUMO can be deconjugated using SENPS or Ulps (Ulp1 and Ulp2).

Objectives of MS research project

Mass spectrometry data from the Ratnaparkhi laboratory (Handu et. al., 2015) suggests that *Drosophila* Arginyl tRNA Synthase (RRS) is SUMO conjugated in

response to infection. What is not understood is the biological reason for this SUMO conjugation and the role SUMOylated-RRS plays in the cell, during the immune response. The goal of my MS project is to generate a transgenic *Drosophila* where the RRS protein is engineered to be SUMO conjugation resistant (SCR-RRS or RRS^{SCR}). The first step towards this is the screening of different RRS Lysine→ Arginine mutants to identify a SCR-RRS (Chapter 2). This RRS^{SCR} variant was identified by Ms. Prajna Nayak, a graduate student in the Ratnaparkhi lab. Once the RRS^{SCR} variant was identified, the next step was to generate transgenic RRS^{SCR} flies, either through CRISPR/Cas9 based genome editing (Chapter 3), or by the null rescue approach (Chapter 4). The bulk of my research and thesis focuses on experiments to generate the RRS^{SCR} *Drosophila* lines, followed by experiments to compare and contrast the immune response of RRS^{WT} to RRS^{SCR}.

Chapter 2

Generation of a SUMO Conjugation Resistant Arginyl tRNA Synthetase (SCR-RRS)

Summary

SUMOylation of RRS is validated using in-bacto SUMOylation. Further, the same technique is used to pinpoint the two SUMOylation sites in RRS, and site directed mutagenesis is used to generate a double mutant, SCR-RRS allele (RRS^{SCR}).

Introduction

SUMOylation is a post translational modification. Small Ubiquitin like Modifier (SUMO) is a protein that conjugates and deconjugates to other proteins to modify them. SUMO attaches usually at a consensus site ψKXE (where ψ is a hydrophobic residue) (Desterro et al. 1999; Müller et al. 2001; Sampson, Wang, and Matunis 2001; Rodriguez, Dargemont, and Hay 2001; Handu et al. 2015) but it can also occur at non-consensus sequences. Thus, presence of a SUMO consensus motif is not a robust determinant of SUMOylation.

In a mass spectrometry based SUMO enrichment screen upon innate immune challenged performed by Ratnaparkhi lab (Handu et. Al. 2015) has shown MARS component proteins to be SUMOylated (Fig.6). Arginyl tRNA synthetase (RRS) is one of them. RRS is shown to be SUMOylated in *Drosophila melanogaster* S2 cells (Nie et al. 2009) and S2R+ cells (Pirone et al. 2017). The role of SUMOylation of RRS is not known. Thus, studying SUMO resistant RRS mutants and comparing it to wildtype RRS would help us uncover the biological role of SUMOylation of RRS in *Drosophila melanogaster*.

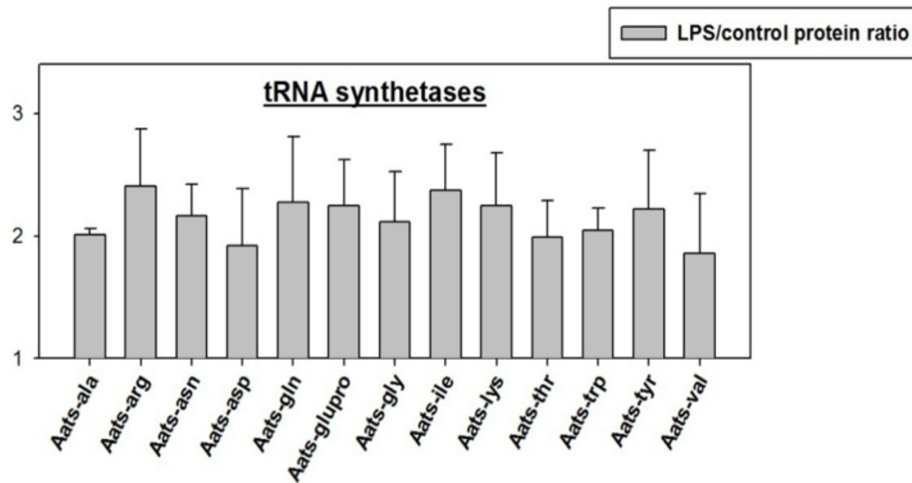


Fig. 6: SUMO conjugation of tRNA synthetases. The SUMO enriched proteome (Handu et. Al. 2015) suggests that many aaRS's are SUMO conjugated. This was subsequently tested by in-bacto SUMOylation and found to be correct. Figure shows fold increase in SUMO enriched tRNA synthetases in response to LPS treatment in 529SU cells, as measured by ITRAQ proteomics **(**Figure adapted from Handu et al. 2015).

Results

RRS is SUMOylated *in-bacto*. SUMOylation is not a natural modification in bacteria. Thus, *E. coli* BL21 was co-transformed with either 6Xhis-FLAG-tagged SUMO- ΔGG (non-functional form) or 6Xhis-FLAG-tagged SUMO- GG (mature form) and Q- SUMO plasmid (containing either functional or non-functional form of SUMO along with other necessary components of SUMO machine, Nie et. al. 2009). GST-tagged RRS wildtype and SCR construct were also transformed. Post-GST affinity pull down, in western blot probed with anti-GST antibody, RRS (75kDa) expression is seen in both the SUMO-GG and SUMO-ΔGG lanes. Only the SUMO-GG lane would show additional bands >15 kDa above the main band. In the western blot probed with anti-His antibody SUMO bands are seen in SUMO-GG lane but not in SUMO-ΔGG lane, confirming that RRS gets SUMO conjugated (Fig. 7). co-transformed with and pGEX 4T1 plasmid containing either wildtype or SCR construct of RRS.

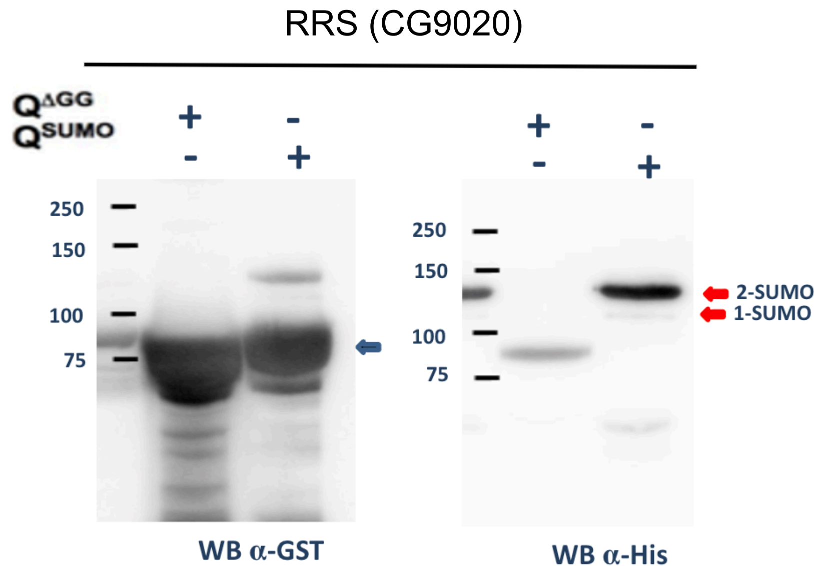


Fig. 7: RRS gets SUMOylated using *in-bacto* SUMOylation system. GST tagged *Drosophila* RRS CDS (CG9020) was subcloned into pGEX-4T1 vector (Promega) as plasmid construct for expression in bacteria. It was co-transformed with either 6Xhis-FLAG-SUMO-ΔGG (lane 1 from left in both blots), which could not conjugate to substrates, or 6Xhis-FLAG-SUMO-GG (lane 2 from left in both blots), a mature, active form of SUMO. The vector also expressed *Drosophila* SUMO activation enzymes (SAE1, SAE2) and the SUMO conjugase Ubc9 (E2). GST affinity pull down was performed on the transformed bacteria and western blot was performed using Mouse Anti-GST antibody and Mouse Anti-His. Molecular weight of RRS is around 75kDa, while SUMO is around 15kDa.

Putative SUMO sites for RRS predicted by bioinformatics tool. Joined Advanced Sumoylation Site and Sim Analyser (JASSA) was used to predict putative SUMO sites of RRS (Fig.8). This tool predicts SUMO sites based on sequence alignment and making a position frequency matrix based on experimental data collected. The output is the sequence of the SUMO motif along with its strength. It also says if it’s a direct or an inverted consensus. There were 20 predicted SUMO motifs amongst which K383 was in a direct consensus sequence and K147 and K579 were in inverted consensus sequences. Since this tool is not completely accurate, it needed experimental validation.

Results for putatifs SUMO site								
Position K	Sequence	Best PS	Consensus direct			Consensus Inverted		
			Type	PSd	DB Hit	Type	PSi	DB Hit
K9	--MSELNMELKKLRELELKTQ	Low	None	None		None	Low	3
K10	-MSELNMELKKLRELELKTQG	None	None	None		None	None	2
K17	ELKKLRELELKTQGLAARIQT	None	None	None		None	None	2
K46	VDLVQLQIENKKLKNRLFILK	Low	None	None		None	Low	
K47	DLVQLQIENKKLKNRLFILKK	None	None	None	1	None	None	
K74	TAAGGDVSKPKESSSITEHLE	None	None	None		None	None	2
K128	YQCNNAMGLSKKLKEKGINKA	None	None	None		None	None	1
K129	QCNNAMGLSKKLKEKGINKAP	None	None	None	1	None	None	
K131	NNAMGLSKKLKEKGINKAPRD	None	None	None		None	None	1
K147	KAPRDIATELKGHCPASPIIE	None	None	None		Weak consensus inv	None	2
K195	RNGVKPPEVIKKRVLVDFSSP	None	None	None		None	None	3
K281	SDLQLFYKESKKRFDEDEEFK	Low	None	None		None	Low	
K291	KKRFDEDEEFKKRAYSRVVSL	Low	None	None	1	None	Low	1
K310	SLQKGVPNSIKAWELICNVSR	None	None	None		None	None	1
K335	TIYERLDISVKERGESFYQSR	None	None	None		None	None	5
K383	TKTGIPLTIVKSDGGFTYDTS	High	Strong Consensus	High	2	None	None	1
K427	GQSTHFNTIFKAAERSAILNP	None	None	None		None	None	1
K579	NIVLDHEKEWKLAKTLLKLHD	Low	None	None		Strong consensus inv	Low	
K586	KEWKLAKTLLKLHDILIKCSK	None	None	None		None	None	1
K626	EFYDSCYCIEKNKQGDIIGVN	None	None	None		None	None	1

Fig. 8: Putative SUMO sites on RRS predicted by an online tool: JASSA. JASSA stands for Joined Advanced Sumoylation Site and Sim Analyser. It predicts SUMO sites based on sequence alignment and making a position frequency matrix based on experimental data collected. The figure shows the sequence of the SUMO motif along with the strength of the motif. It says if it's a direct or an inverted consensus sequence.

K147 and K383 are SUMO sites of RRS in *Drosophila melanogaster.*

Multiple predicted Lysines in RRS were mutated to Arginines and plasmid constructs were made to validate the SUMO site lysine of RRS *in bacto*. The constructs were co-transformed with the SUMO machinery. Single lysine mutants as well as all possible pairs were made in order to check for the loss of SUMOylation. RRS-wildtype and RRS- SCR were His and HA tagged. Denaturing His pulldown was performed and western blot was done with Anti-HA antibody (Fig.9). The lower band in the blot represents non-SUMOylated RRS whereas the upper band represents SUMOylated RRS. Loss of the upper band suggest loss of SUMOylaton *in bacto* was observed due to mutation of K147R and K383R together.

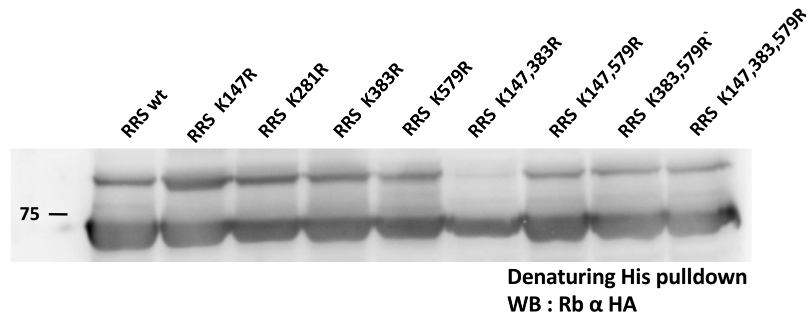


Fig. 9: Expression of RRS and its predicted SUMO conjugation resistant (SCR) mutants (*In bacto*). pET 45b His-RRS-HA and pET 45b His-RRS SUMO SCR-HA plasmids were co-transformed with the SUMO machinery plasmid. RRS SUMO SCR constructs had Lysine 147, 281, 383, 579 converted to Arginine as single mutations, in pairs of two and three. The western blot is a result of a denaturing His pulldown probed with Rabbit Anti-HA antibody. RRS is around 75kDa in weight. The lower band represents non-SUMOylated RRS and the upper band represents SUMOylated RRS. Data courtesy Prajna Nayak.

Predicted SUMO site lysines are conserved across species. RRS protein sequence of different species *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Xenopus laevis*, *Danio rerio*, *Gallus gallus*, *Mus musculus*, *Homo sapiens* were aligned using T-Coffee (Fig.10; Chang et. Al. 2012).

```
Saccharomyces     85  DLAVOWAEKFPC-GDFLEKVEANG  107
Drosophila       141  DIATELKGHCPA-SPIIEKLEIAG  163
Caenorhabditis   188  DVAKEIOAKLPTKIDFVEKIDVMP  211
Xenopus          136  EIADKIVKNIPT-NELVEKVDIAG  158
Danio            137  EIAEKIVONIPN-NDLIEKTEIAG  159
Gallus           137  EIAEKITKHIPA-NECIEKVEIAG  159
Mus              136  EIAENITKHLPN-NKYIDKVEIAG  158
Homo             136  EIAENITKHLPD-NECIEKVEIAG  158

cons             193  ::* :    : *     ::* :    216
```

```
Saccharomyces    321  AVLIDLTKF--NKKLGKAIVOKSD  342
Drosophila       365  REIMWPDDTKTGI---PLTIVKSD  385
Caenorhabditis   415  ROIMFPTGC--DI---PLTVVKSD  433
Xenopus          361  RKVVFPPGC--SV---PLTIVKSD  379
Danio            362  RKIVFAPGO--SI---PLTIVKSD  380
Gallus           362  RKIVFVPGF--SV---PLTIMKSD  380
Mus              361  RKIVFVPGC--SV---PLTIVKSD  379
Homo             361  RKIVFVPGC--SI---PLTIVKSD  379

cons             433   ::        .         : ***  456
```

```
Saccharomyces    524  AKLLIRLLGOYPDVLRNAIKTHEP  547
Drosophila       577  EWKLAKTLLKLHDILIKCSKELFL  600
Caenorhabditis   626  EFKLAKOLLKLSDCVLLVLDSLML  649
Xenopus          571  EWKLSKCILRFPEILOKILDDLLL  594
Danio            572  EWKLGKCILRFPEILOKITDDLLL  595
Gallus           572  EWKLGKCILRFPEILOKILDDLLL  595
Mus              571  EWKLGRCILRFPEILOKILDDLFL  594
Homo             571  EWKLGRCILRFPEILQKILDDLFL  594

cons             649     * : : :  : :     .     672
```

Fig. 10: RRS protein sequence alignment of multiple model organisms. Multiple sequence alignment was used to align the RRS protein sequence of species *Saccharomyces cerevisiae, Drosophila melanogaster, Caenorhabditis elegans, Xenopus laevis, Danio rerio, Gallus gallus, Mus musculus, Homo sapiens*. The three black arrows point to the three predicted SUMO sites lysines (K 147, K 383, K 579) in *Drosophila* RRS.

The result shows that K383 is fully conserved in all species. K579 is conserved in all except *Saccharomyces cerevisiae* where it is amino acid apart. K147 is conserved in all the species except in *Drosophila*, where it is one amino acid away.

Discussion

RRS was shown be SUMOylated in *Drosophila* S2 cells (Nie et. Al. 2009). To confirm this RRS SUMOylation was shown through an *in-bacto* assay. Many putative SUMO sites were predicted on RRS using a bioinformatics tool. The predicted Lysine residues on which SUMOylation takes place were converted to Arginine using site-directed mutagenesis. Using affinity pull-down and western blotting the mutant RRS were checked for the loss of SUMOylation *in bacto*. Conversion of K147 and K383 to R147 and R383 respectively together, showed a loss of SUMOylation *in-bacto*. Although, loss of K147, K383 and K579 together did not show a loss in SUMOylation *in-bacto*. Thus, we wanted to check the effect of conversion of K147R and K383R in the whole organism. The study in the following chapters talks about the efforts in SUMO-SCR RRS in *Drosophila melanogaster* using CRISPR-Cas9 and UAS-GAL4 technology.

Conclusions

The SUMO target lysine residues were found and validated *in-bacto*. In order to study the role of SUMOylation of RRS, we had to make *in-vivo* SUMO conjugation resistant RRS (SCR-RRS) by mutating K147 and K383 to R147 and R383 respectively. In order to achieve this, two strategies have been followed. First, to generate a genomic SCR-RRS using scarless CRISPR-Cas9 technology (discussed in Chapter 3). Second, to generate a genomic null of RRS using CRISPR-Cas9 technology and rescue it with wildtype RRS and SCR RRS using the UAS-Gal4 system (discussed in Chapter 4).

Contributions

Data for Figure 1 is taken from the Ph.D. thesis of Dr. Mithila Handu (Handu, 2015). Data for Figures 7 and 9 were contributed by Prajna Nayak and are part of her Ph.D. thesis.

Chapter 3

Using CRISPR/Cas9 based genome editing to generate a RRS^{SCR} *Drosophila* transgenic

Summary

This chapter discusses a strategy and efforts to create genomic SUMO conjugation resistant RRS (RRS^{SCR}) using CRISPR-Cas9 genome editing technology in *Drosophila melanogaster*. This strategy involves replacement of the lysine residues that are SUMO conjugated by Arginine, thus blocking SUMOylation.

Introduction

CRISPR-Cas9 which is originally the so-called bacterial immune system against viruses has become a powerful genome editing tool in the past few years (Barrangou 2015; Horvath and Barrangou 2010). CRISPR stands for clustered regularly interspaced short palindromic repeat DNA sequences and Cas-9 for CRISPR-associated protein 9. The CRISPR-Cas9 technology allows faster and more precise *in-vivo* genome editing. In this technology, the researches design a short and unique single-guide RNA (sgRNA) which recognizes a specific region of the genome as well as binds to the Cas-9 and directs it to that locus. Cas-9 is an endonuclease which makes a precise double stranded break (DSB) three bases 5' of the 3' end of the sgRNA target sequence (Jinek et al. 2012). sgRNA target site requires to have a protospacer-adjacent motif (PAM) immediate to the 3' of the target sequence (Jinek et al. 2012; Lamb, Walker, and Wittkopp 2017). DSB induces the cell's repair mechanism which is used by the researchers to either insert or delete or replace DNA segments according to their requirement (Fig.11). The cell's repair mechanism can either join the ends by non-homologous end-joining leading to small insertions or deletions (indels) or precisely close the gap by homology-directed repair (HDR) using the homologous chromosome or exogenous DNA molecule with homology arms, as template.

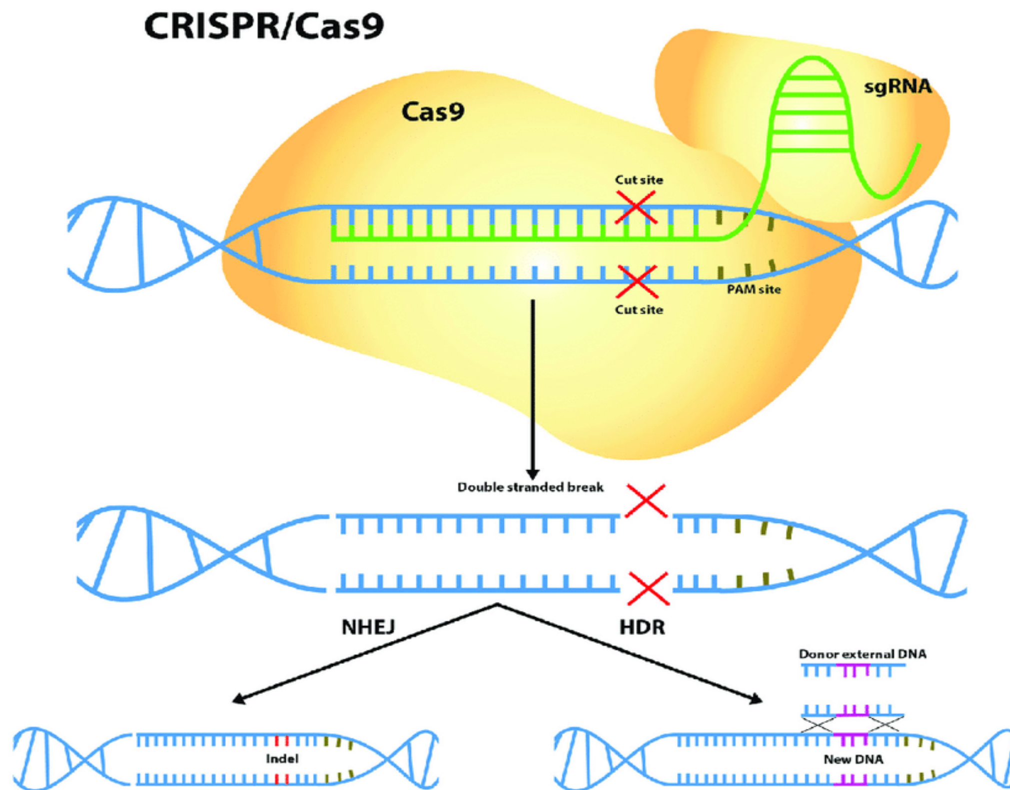


Fig. 11: CRISPR-Cas9 technology directed gene-editing mechanisms. A single gRNA specific to a region in the genome recognizes the region near a PAM site and guides the Cas9 endonuclease to make a precise double stranded break in the DNA. The double stranded break induces two kinds of repair mechanism which is manipulated to be used in the genome editing. (i) Non-homologous end joining (NHEJ) leads to indels in the locus helping in disrupting the gene, (ii) Homology directed repair (HDR) in the presence of a donor DNA, leads to addition of a required new stretch of DNA. (Cribbs and Perera 2017)

For making single or few nucleotide changes in the genome, researchers rely on HDR by providing an exogenous template. The exogenous template can be of two forms: a single stranded oligonucleotide donor (ssODN) or a double stranded DNA plasmid (Ran et al. 2013). The template has the desired modification as well as flanking homology arms for repair. After HDR the modification gets incorporated at the desired locus in the genome.

In the strategy employed to generate SCR-RRS we have used scarless CRISPR-Cas9 technique using pHD-ScarlessDsRed donor vector (Fig.12) ("Scarless Gene Editing – flyCRISPR" n.d.). This vector has 3xP3-DsRed marker cassette flanked by PBac transposon ends. The homology arms are inserted in either ends of the

transposon ends at a naturally occurring TTAA site of the target gene. This leads insertion of this transposon marker in between the target gene. The marker expresses (fluorescent red) in the larval tissue as well as in the adult eyes which makes the screening faster. This transgenic line is then crossed to a line expressing PBac tranposase for precise excision of the marker leaving behind the actual TTAA site of the gene. Since, the modification is made in the gene without disrupting any other nucleotides of the gene, it is called the scarless technique.

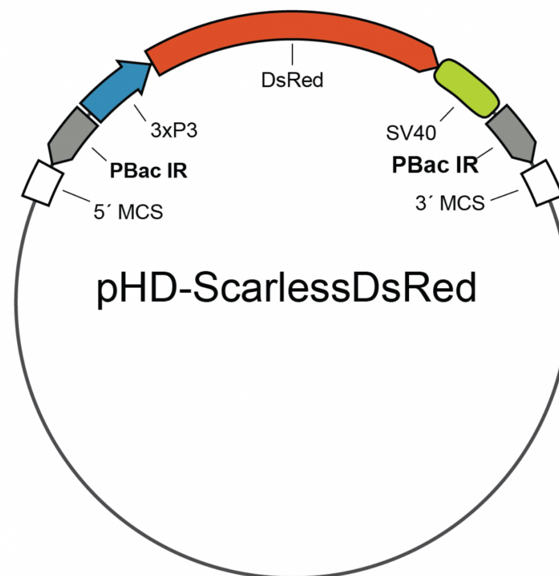


Fig. 12: pHD-ScarlessDsRed donor vector. Vector for generating dsDNA donors for homology-directed repair for CRISPR-Cas9 genome editing. It contains a visible marker 3xP3-DsRed and PBac transposon ends. (pHD-DsRed; Addgene plasmid 51434).

Gibson assembly is a molecular cloning technique for joining two or more than two DNA fragments with overlapping ends for the adjacent fragment(s) (Fig.17A). In Gibson assembly, the reaction mixture contains three enzymes, a 5’ exonuclease, a polymerase and a DNA ligase. The 5’ exonuclease generates overhangs which helps is annealing of the complimentary strands (homology arms of adjacent fragments). The polymerase polymerizes the single strand and fills the gap and the DNA ligase finally ligates the nicked end.

Materials & Methods

***Drosophila* Lines and Genetics.** Flies were maintained at 25 °C on 6% yeast cornmeal media. pBFv U6.2 vector was used for gRNA cloning for CRISPR-Cas9 approach. This method relies on the homology directed repair mechanism of the organism. Thus, a donor vector (pHD-Scarless DsRed) with mutations on the homology regions is injected.

Polymerase Chain Reaction (PCR). PCRs were done in Eppendorf Mastercycler X50s. Eppendorf 96-well plates or 200 µL microcentrifuge or PCR tubes. Most amplification reactions were of 50 µL and colony PCR screening reactions were 10 µL in volume. Invitrogen™ Platinum™ SuperFi™ DNA Polymerase with GC-enhancer was used for all cloning PCR reactions. HIMEDIA Taq polymerase with 10X Buffer S with 17.5mM $MgCl_2$ was used for colony PCR screening.

PCR amplifications using Invitrogen™ Platinum™ SuperFi™ DNA Polymerase were performed at: 98°C for 5 minutes, followed by 30 cycles at 98°C for 30 seconds, annealing temperature (specific to each primer pair) for 30 seconds and extension temperature 72°C for 30 seconds/kb, and finally 1 extra extension cycle at 72°C for 5-10 minutes.

PCR amplifications using HIMEDIA Taq polymerase were performed at: 98°C for 5 minutes, followed by 30 cycles at 98°C for 30 seconds, annealing temperature (specific to each primer pair) for 30 seconds and extension temperature 72°C for 1 minute/kb, and finally 1 extra extension cycle at 72°C for 5-10 minutes.

Agarose gel electrophoresis. A 0.8-1% agarose gel was prepared in 1X TAE and ethidium bromide (Final concentration: 200 ng/mL). DNA was added to 6X loading dye and loaded in the wells of the gel. HIMEDIA 100 bp or 1kb ladders were used as size markers. Gels were run at 110V for 30 minutes on an average unless more resolution was needed. Gels with DNA were visualized in a Gel Doc (Syngene G-box-Chemi 16) and images were captured and stored. For gel extraction, DNA bands were visualized under an UV transilluminator for gel extraction.

DNA gel extraction. Qiagen gel extraction kit was used for gel purifications. Desired DNA bands were resolved on a gel and visualized under an UV transilluminator.

Desired gel bands were cut out from the gel using a clean blade and the gel piece is stored in a centrifuge tube. Qiagen QG buffer equivalent to 3 times the volume of the gel was added to the gel piece and incubated at 50°C to dissolve the gel piece. Isopropanol, equivalent to the volume of the gel fragment was added next and mixed well. The solution was then transferred into a gel extraction column and spun at 13000rpm on a table top centrifuge and the flow-through was discarded. 500μL QG buffer was added to the column, spun and the flow-through was discarded. The column was the washed with 700μL PE wash buffer spun and the flow-through was discarded. It was spun again to remove any residual buffer and the column was transferred to a clean labelled 1.5mL microcentrifuge tube. After the column becomes dry and free of the ethanol smell, 20-50μL of elution buffer was added to the column. Column was incubated at room temperature for 5 minutes and spin to elute the purified PCR DNA fragment.

Transformation and plasmid extraction. 50uL of PPY (for fragment recombination) or DH5α (for plasmids) competent cells were used for individual transformations. The cells were thawed on ice. 1μL DNA plasmid was added for transformation. 1:3 molar ration of backbone to insert was used for fragment recombination (Not more than 5 μL of total DNA). DNA was mixed well by gently tapping the tube. The cells were incubated on ice for 10-15 minutes and then given a heat shock 42°C for 90 seconds. After the heats shock, the cells were chilled on ice for 2-5 minutes. 750μL fresh LB was added to the tube containing cells and was incubated at 37°C at 180rpm for 45 minutes to 1 hour. For plasmid transformation, 100μL of the sample was plated on a LB-Agarose plate. In case of fragment recombination, the cells were pelleted, and resuspended in 100μL of LB and plated on a LB-Agarose plate with appropriate antibiotic (in most cases Ampicillin). The plate was incubated at 37°C overnight and the transformed colonies were screened for required constructs using PCR. Positive colonies were inoculated in 5mL of LB with appropriate antibiotic and grown overnight at 37°C and 180rpm for plasmid extraction. Plasmid extraction was done using Qiagen Miniprep and Midiprep kits. The culture was pelleted at 5000-6000g, 4°C, for 10 minutes. The pellet was resuspended in P1 buffer and was lysed using lysis buffer P2 (for less than 5 minutes). For mini-prep, buffer N3 was used for neutralization and was then spun at maximum speed for 10 minutes. In all the following steps the column is spun at the same speed for 1 minute each. Supernatant was transferred carefully to a DNA binding column and spun down and

the flow through was discarded. The column was washed with PE buffer, the flow through was discarded and then DNA was eluted in 50µL elution buffer. For midi-prep, the bacterial lysate was neutralized with buffer P3 and then spun down at max speed. The column was activated using buffer QBT and was allowed to be emptied using gravity flow. The supernatant was added to the column and washed twice using buffer QC. DNA was eluted using 5mL of buffer QF followed by precipitation using chilled isopropanol. The precipitated DNA solution is spun for 30 minutes. The pellet was washed with 70% ethanol and air dried later. DNA is re-dissolved in 200µL of nuclease free water.

Colony screening PCR. In order to screen for positive transformants after transformation, bacterial colonies were screened using PCR with appropriate primers. A small part of the colony is picked using a pipette tip and mixed in 10µL of PCR mixture. A replica of this colony is made in another LB-agar plate by streaking. The PCR is performed as mentioned above. The product is loaded into a 1% agarose gel is run to check for the appropriate DNA band size.

Genomic DNA isolation. In order amplify RRS gene, single W^{1118} fly was collected and crushed in squishing buffer (10mM Tris-Cl pH 8, 1mM EDTA, 25mM NaCl, 200ug/mL freshly added proteinase K) to extract genomic DNA. The lysate was incubated at 37°C for 30 minutes, heat inactivated at 85°C for 5 minutes followed by a short spin to settle the debris and stored at -20°C.

Site-directed mutagenesis. For creating Lysine (AAA)—»Arginine (CGC) in RRS homology arms (5' and 3') in the scarless donor vector, 4 primers (forward, reverse, two for mutagenesis) were designed for each homology arm (Fig.). After doing genomic DNA isolation and amplifying the RRS gene from the W^{1118} fly, the PCR product was gel purified (Fig.13). The mutagenesis primers had AAA converted to CGC. First PCR was done with the forward primer and reverse mutagenesis primer with the isolated genomic RRS as template in one reaction and the reverse primer and forward mutagenesis primer with the isolated genomic RRS as template the other reaction. After the amplification and gel running, the correct DNA band was gel purified. In the final reaction, the product of both reactions were put as templates (equal molar ratio) along with the forward and reverse primers. The final product was

gel purified and amplified again to increase the yield. This procedure was followed for both 5' and 3' homology arms.

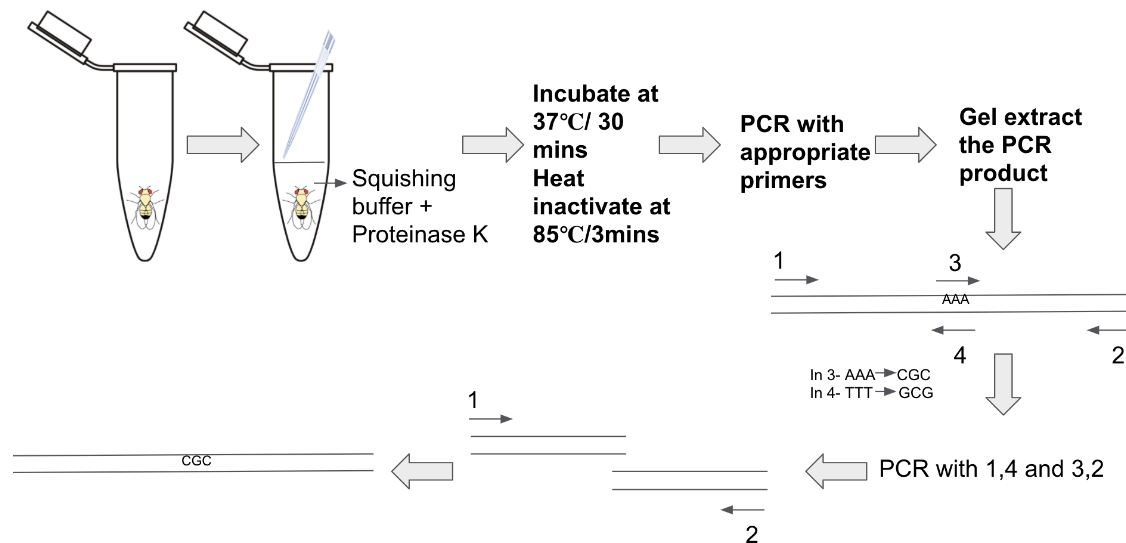


Fig. 13: Site-directed mutagenesis for homology arms of scarless donor plasmid. Wildtype W^{1118} fly was squished with a pipette tip using 50μL squishing buffer and 200μg/mL of freshly added proteinase K. The mixture was incubated at 37°C for 30 minutes, heat inactivated at 85°C for 3 minutes and spun to settle the debris. 1μL of the product was used as template along with appropriate primers for RRS for PCR. The product was gel extracted and used for further PCRs. Primer 1 and 2 are forward and reverse primers for the fragment respectively. Primer 3 and 4 are reverse and forward mutagenesis primers containing AAA—»CGC respectively. PCR with primers 1,4 and 3,2 were done simultaneously. This was followed by gel purification. Their products were used as templates (equal molar ration) for the next reaction with 1 and 2 as primers. The final product was gel purified and amplified again to increase the yield.

Restriction digestion. The assembled plasmid (100ng in each reaction) was digested with NEB XhoI, NdeI and NotI restriction enzymes. NEB Cutsmart buffer was used for all the restriction digestions. Each reaction was 10μL in volume and incubated at 37°C for 3 hours and heat inactivated at 65°C for 20 minutes. The entire reaction was loaded on a 0.8% agarose gel to visualize the DNA bands.

Results

gRNA transgenic generation for RRS SUMO mutant. gRNA was cloned in pBFv U6.2 vector. It was sequenced and injected in *Drosophila* embryos (Fig.15) (NCBS fly facility). The stable transgenic gRNA lines are made (Fig.14).

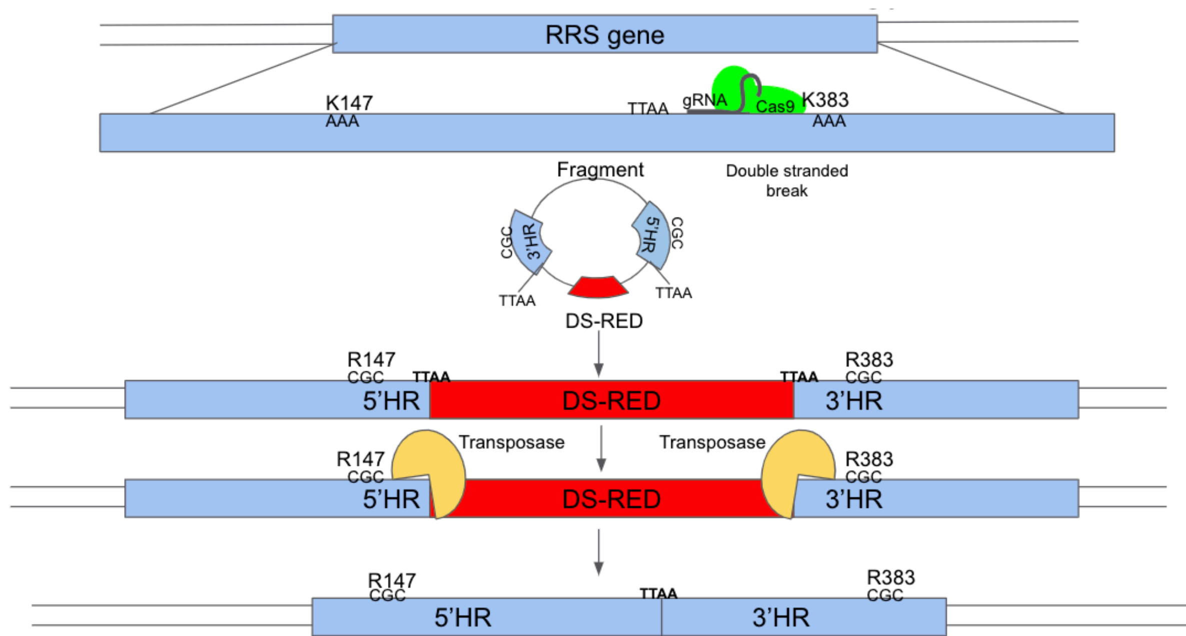


Fig. 14: Strategy for creating SUMO-conjugation resistant RRS using scarless CRISPR-Cas9 technique. The sgRNA-Cas9 complex binds to the RRS locus and makes a double-stranded DNA break. The donor vector containing 5' and 3' homology arms of RRS with specific mutations and a marker DsRed in between gets integrated by HDR. DsRed marker helps in screening the flies. Upon crossing this line to a PBac tranposase the DsRed transposon is precisely excised out.

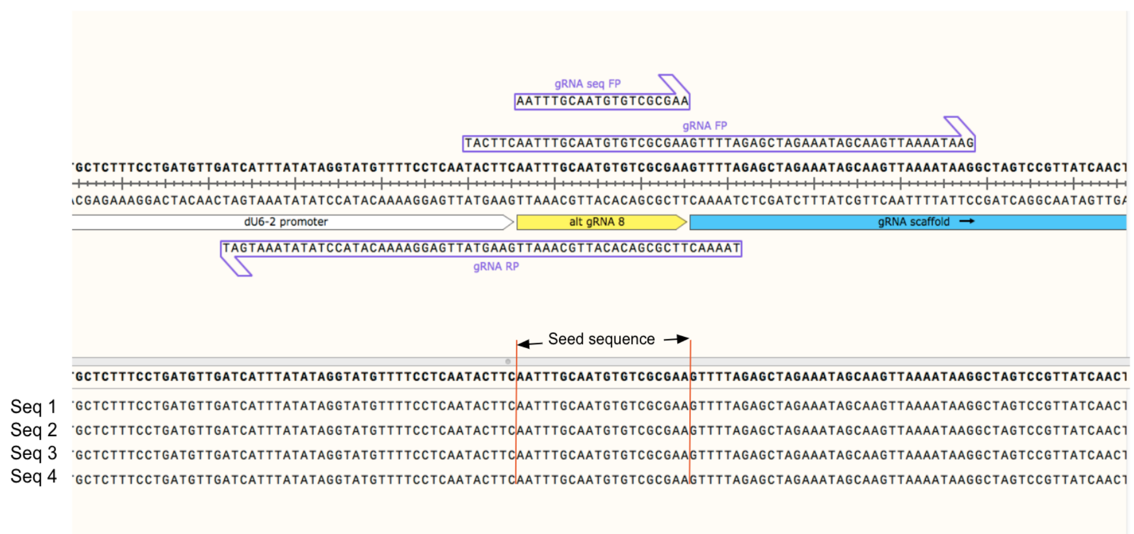


Fig. 15: RRS gRNA sequencing results. gRNA was cloned into pBFv U6.2 vector and was transformed into competent cells. Four positive clones were sequenced. All four sequences are aligned against the correct sequence file in *Snapgene*.

Donor vector fragments amplified and site-directed mutagenesis done for K147R and K383R. 5' and 3' homology regions were amplified from genomic DNA

of wild-type (W^{1118}) flies. Site directed mutagenesis was done using suitable primers to change K147 to R147 and K383 to R147 (Fig.16 B, C). The donor vector plasmid (pHD-Scarless DsRed **)** is also amplified in two parts (Fig.16A).

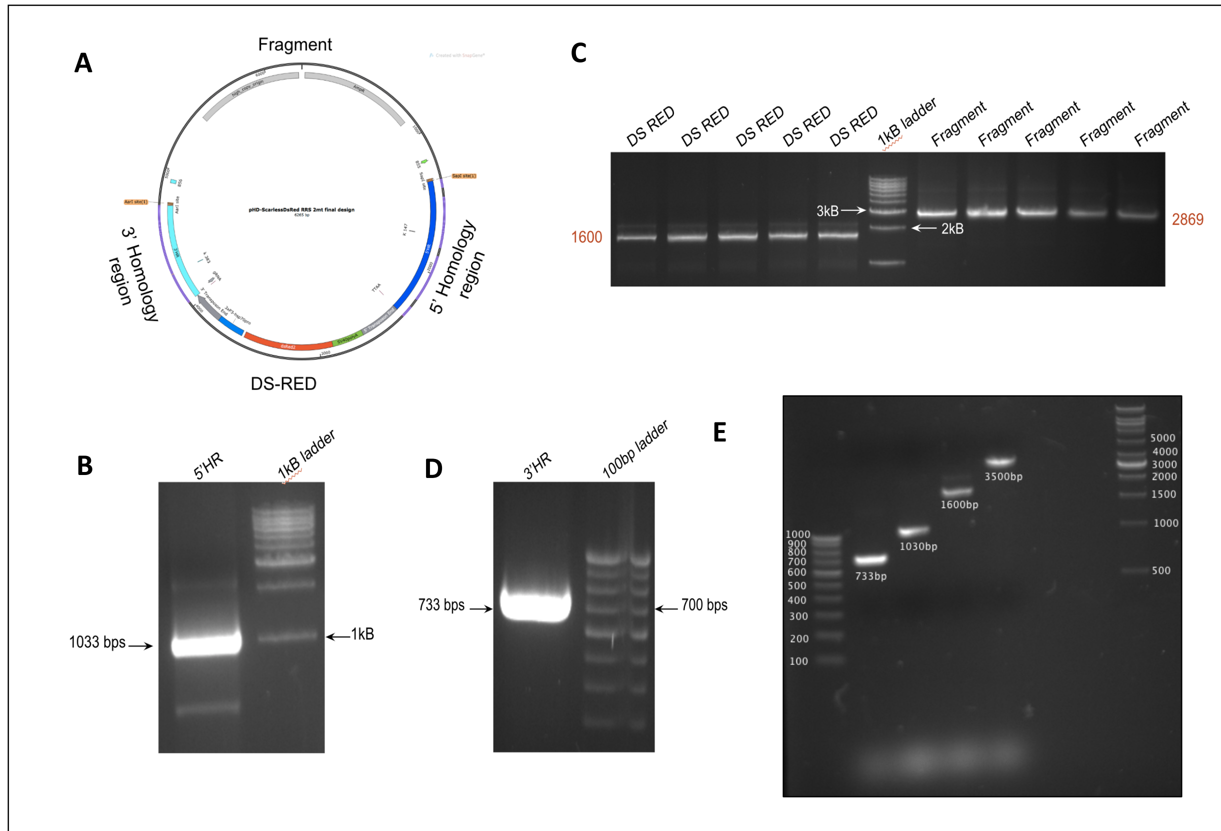


Fig. 16: Amplification of vector fragments and homology regions with K→R mutations. **A.** Design of the donor vector showing all the four parts that were PCR amplified separately. **B.** Gel image showing PCR amplification of pHD-Scarless DsRed vector fragments. Row1-5 shows DsRed amplification (part of pHD-Scarless DsRed plasmid) , Row 6 shows 1 kB DNA ladder by *Genei* and Row 7-11 show PCR amplification of another fragment of pHD-Scarless DsRed plasmid. **C.** Gel image showing PCR amplification of the 5' homology region with K147R. **D.** Gel image showing PCR amplification of the 3' homology region with K383R **E.** Gel image showing all four fragments after gel extraction.

Gibson assembly cloning for assembling the vector. All four fragments have overhangs which are homologous to one another. Thus, Gibson assembly was performed to join all the fragments to assemble them into a plasmid vector (Fig.17A). After the transformation of the vector into competent cells, 29 colonies were found (Fig.). All the colonies were screened for the presence of the correct plasmid but none of them turned out to be positive (Fig.17B). The process was repeated two more times and no positive colony could be found.

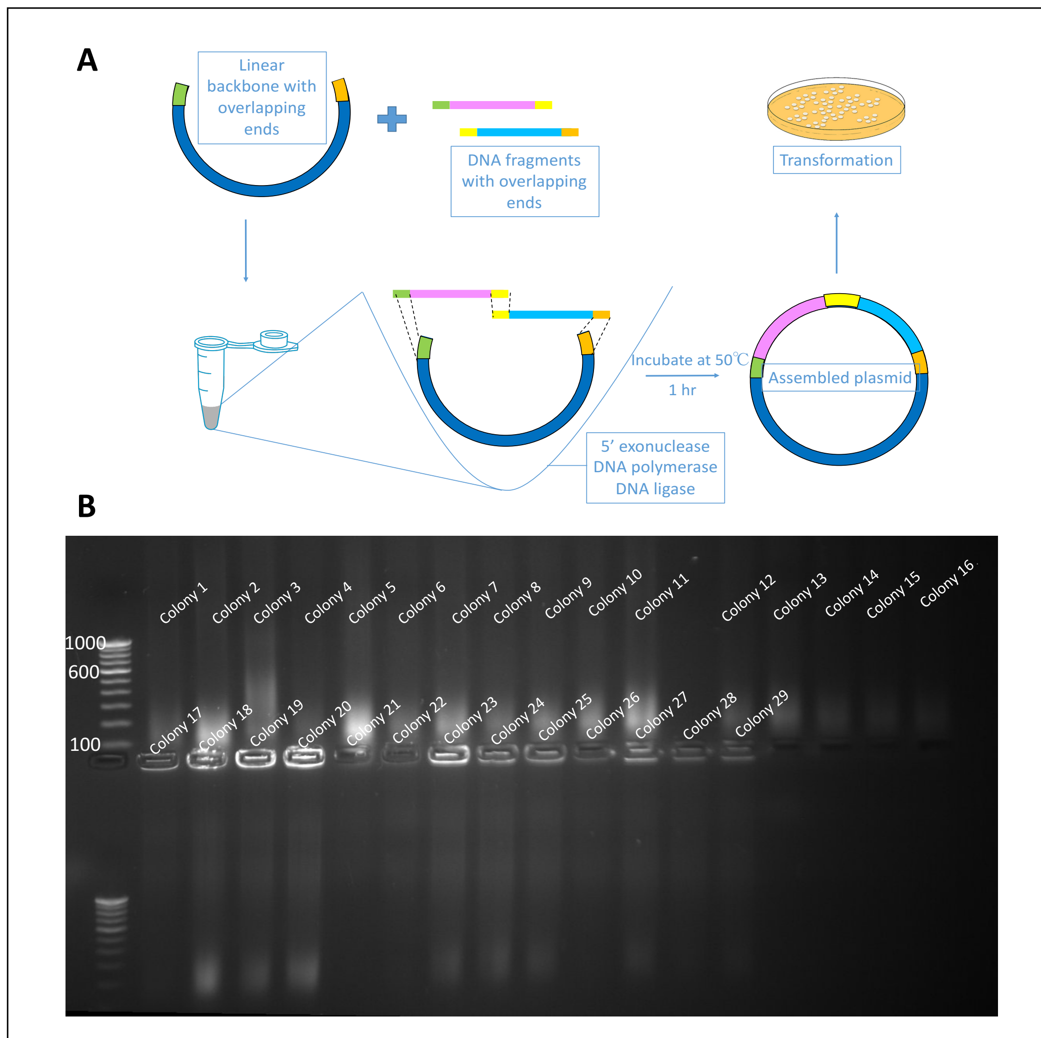


Fig. 17: Gibson assembly for PCR amplified scarless fragements. A. Diagrammatic representation of Gibson assembly workflow. The linear backbone and other DNA fragments with overlapping ends are mixed with the Gibson assembly enzyme mastermix containing a 5' exonuclease, DNA polymerase and a DNA ligase. The mixture is incubated at 50°C for 60 minutes and the assembled product is transformed into competent bacterial cells. **B.** Gel image showing colony PCR for 29 colonies obtained after transformation of Gibson assembly product.

Gibson assembly cloning with high fidelity polymerase (using NEBuilder HiFi DNA assembly kit) for assembling the vector. The transformation of this reaction resulted in more than 500 colonies. After screening around 230 colonies, 9 colonies were positive for one of the four fragments (3'HR) (Fig.18 A,B,C). Plasmids were extracted from all 9 colonies and 5'HR and other fragments were amplified. Only one colony (colony 48) showed a positive result for 3'HR. It was digested with NotI, XhoI and NdeI to confirm the presence of 3'HR and DS-Red but not 5'HR (Fig.19). After sequencing it was confirmed that 3'HR fragment and DS-Red were completely intact, whereas 5'HR was not present.

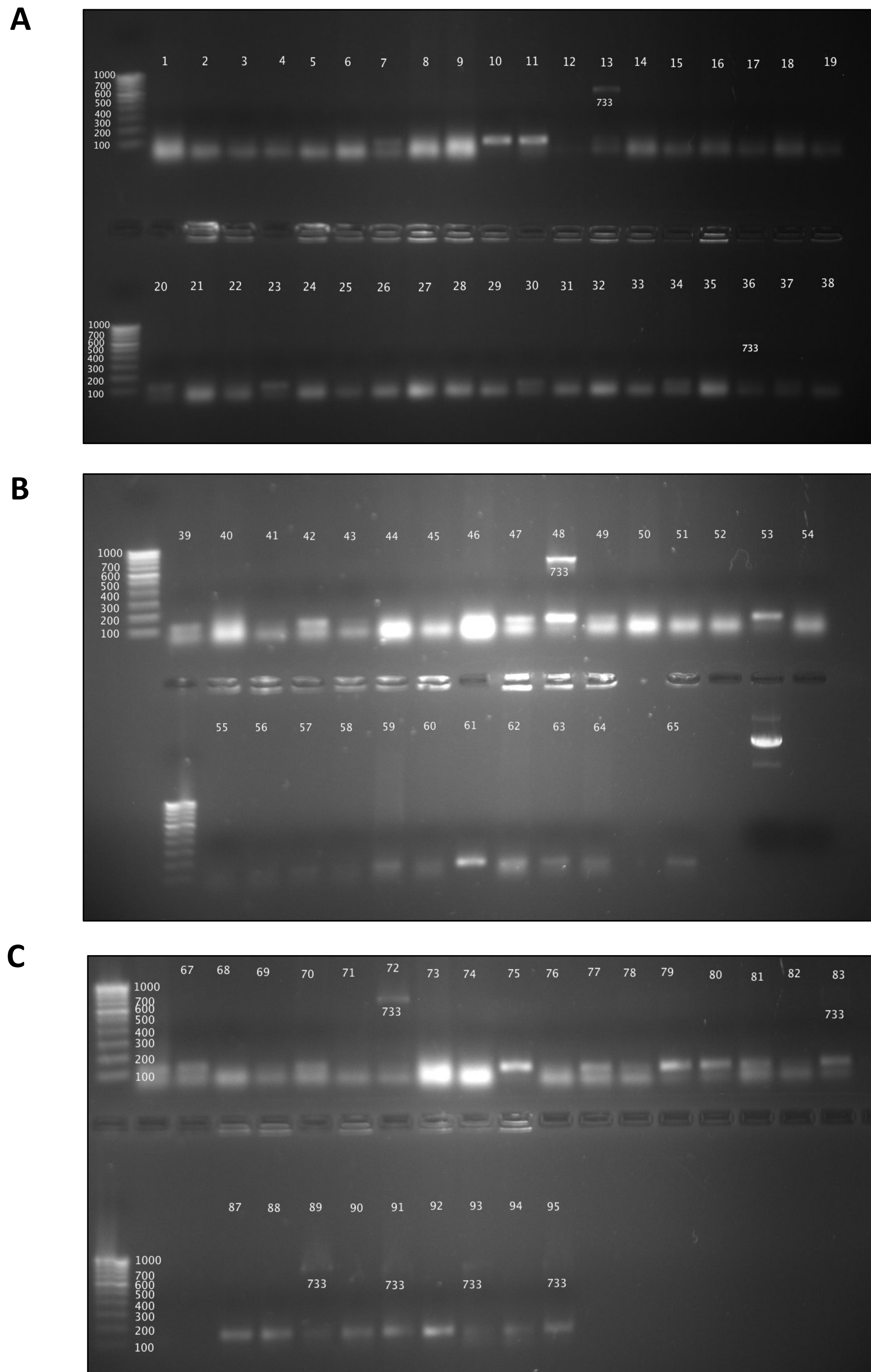


Fig. 18: Colony PCR for transformants of NEBuilder HiFi DNA assembly kit (Gibson assembly with high fidelity polymerase). A,B,C- Gel images of colony PCR (with 3'HR primers) of 95 colonies obtained after transformation. The correct size of the fragment is 733 base pairs.

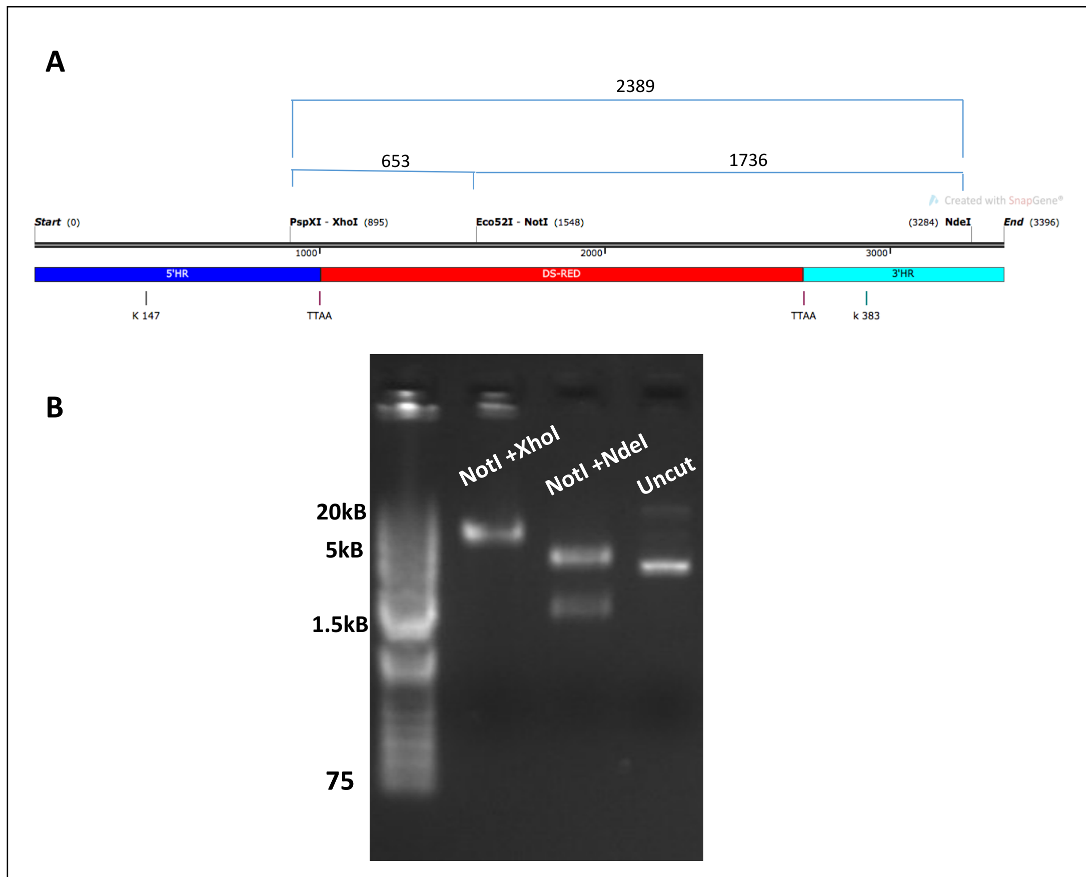


Fig. 19: DNA digestion of plasmid extracted from a positive colony after Gibson assembly. A. Map showing position of restriction sites XhoI, NotI, NdeI on 5' region, DsRed region and 3'region respectively. **B.** Gel image of the digested plasmid. Lane 1- ladder, lane-2 Digested with XhoI and NotI, lane-3 digested with NdeI and NotI, lane-4 undigested plasmid. The expected size of single cut is 6265 and dual cut with NotI and XhoI is 5612 and 653, NotI and Ndei is 4529 and 1736.

Discussion

We are using the scarless CRISPR-Cas9 technology to create RRS-SCR as it is an efficient way to precisely make changes in the genome with less screening time because of the presence of a visible marker. We have successfully cloned the sgRNA vector and generated the sgRNA transgenic line. The assembled scarless donor plasmid does not have 5' fragment as well as some part of the vector fragment which is strange because there is only one theoretical possibility of this assembly as the fragments have overlapping ends only for adjacent fragments. The other possibilities could be that some part of one of the fragments has homology/overlap with 3' or DsRed region which led the assembly of three fragments leaving the forth one out. There could also be some blunt end ligations which would have led to this result.

Conclusions

The sgRNA is cloned and transgenic line is made. All the fragments of the donor plasmid are PCR amplified with K147R and K383R mutations in 5’ and 3’ fragments respectively. The classical Gibson assembly cloning did not work for the assembly of the donor plasmid so a new assembly kit with high fidelity polymerase was used. This gave higher number of transformants with a few positives but none of them turned out to be the required clones.

Future Work

The donor plasmid still has to be assembled properly. The currently assembled plasmid has DsRed and 3’ corrected oriented in the plasmid with K383R mutation in 3’. This would be amplified from the plasmid and another round of Gibson assembly cloning would be done with DsRed-3’, vector fragment and 5’. Once the plasmid is cloned it would be injected in the fly embryos with sgRNA and Cas9.

Contributions

I thank Prajna Nayak for teaching me cloning methodology and CRISPR/Cas9 design principles.

Chapter 4

Generation of a transgenic animal expressing SCR-RRS using the Null: Rescue methodology

Summary

This chapter discusses the strategy and efforts to create a genomic RRS null using CRISPR-Cas9 technology. Since RRS is an essential housekeeping gene, flies stall in early developmental stages in its absence. This chapter discusses the use of UAS-GAL4 system to rescue this lethality by overexpression of RRS^{WT} as well as $RRS^{SCR.}$ This is an alternate strategy, in contrast to genome editing, of studying the role of SUMOylation of RRS in *Drosophila melanogaster*.

Introduction

The UAS-GAL4 system is a powerful tool in fly genetics for spatial and temporal control of targeted gene (Brand and Perrimon 1993). GAL4 is a yeast *Saccharomyces cerevisiae* DNA binding protein (Giniger and Ptashne 1988). It regulates the transcription of *GAL10* and *GAL1* genes by binding to four similar 17 base pair regions called the Upstream Activating Sequences (UAS) which act similar to an enhancer element for metazoans (Duffy 2002; Giniger, Varnum, and Ptashne 1985). This system has been adapted in *Drosophila* as well as other model organisms to express desired proteins or reporters under UAS by expressing GAL4 (Kakidani and Ptashne 1988; Webster et al. 1988). The *GAL4* gene is driven by a tissue-specific promoter and UAS is placed upstream of the gene of interest. GAL4 is not a natural *Drosophila* gene and does not have any deleterious effects upon expression in flies. This helps in its expression only in the specific tissue where it goes and binds to the UAS leading to the expression of the downstream gene without disrupting any other pathway. Fig.20 shows a simplistic schematic of the working of this system.

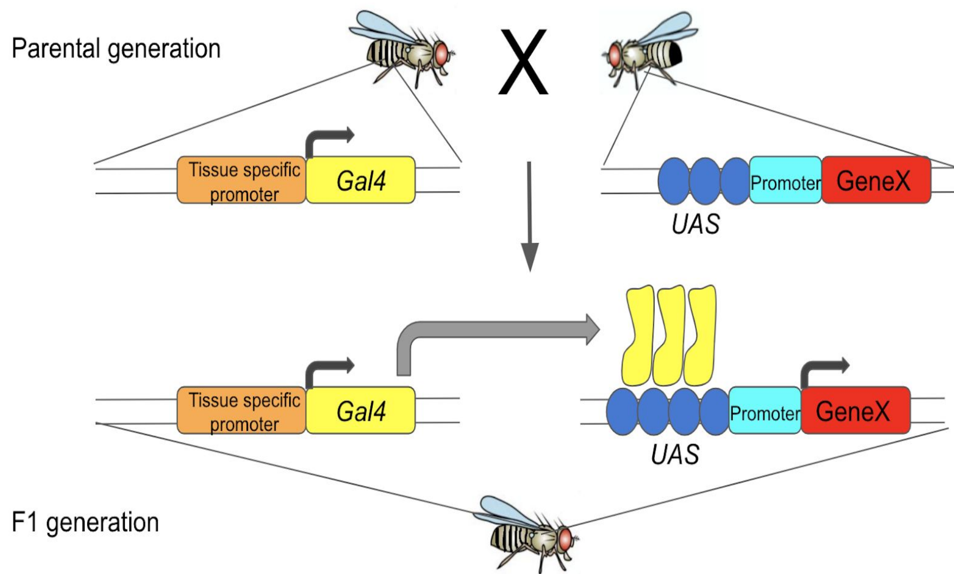


Fig. 20: UAS-GAL4 driven overexpression of genes in *Drosophila melanogaster*. Diagrammatic representation of the UAS-GAL4 system in flies. The parental generation flies Gal4 driven by tissue specific promoter and UAS upstream of the gene of interest in the and male respectively. When the male and female are crossed, the progeny (F1) contains both GAL4 binds to the UAS and drives the expression of the gene of interest in tissue specific manner.

We have used the UAS-GAL4 system to over-express RRS wildtype and RRS-SCR in the whole fly which genomic null for RRS ($RRS^{C\Delta}$). UAS-RRS (wildtype and SCR) is expressed with GAL4 driven by ubiquitous promoters such as Actin-Gal4.

Since many of the key signaling pathways are evolutionarily conserved between *Drosophila* and mammals, *Drosophila melanogaster* serves as an excellent model system for studying innate immune system. Upon parasitic and microbial invasion, just like other vertebrates, *Drosophila* elicit a range of innate immune responses for their defense. Three major categories of immune responses include: (i) epithelial barrier (cuticle, trachea, gut), clotting and melanisation (ii) humoral reactions (release of antimicrobial peptides or AMPs) and (iii) cellular reactions (hemocyte-dependent phagocytosis, encapsulation) (Govind 2008; Imler and Bulet 2005; Brennan and Anderson 2004; Hoffmann 2003).

The systemic immune response, in which antimicrobial peptides are secreted from cells of the fat body into the hemolymph is one of the best studied immune responses in *Drosophila*. This response is induced mostly by the microbial pathogens mostly fungal and bacterial. There are two key microbial-sensing immune pathways in *D. melanogaster*, the Toll/NF*k*B and immune deficiency (Imd)/NF*k*B pathways. Toll pathway is similar to myeloid differentiation primary response protein 88 (MYD88)-dependent Toll-like receptor (TLR) pathway (Lemaitre, -M. Reichhart, and Hoffmann 1997) in mammals whereas Imd pathway is similar to the tumour necrosis factor (TNF) pathway and TIR domain-containing adaptor protein inducing IFNβ (TRIF)-dependent TLR pathways (Buchon, Silverman, and Cherry 2014). Toll pathway is triggered mostly by Gram positive bacteria whereas Imd is triggered by Gram negative bacteria. Both result in NF-κB signalling pathway which work parallel to each other (Buchon, Silverman, and Cherry 2014). They induce the expression of an overlapping but distinct set of effector proteins, including AMPs (Kaneko et al. 2006).

In general, flies with impairment in the components of these pathways (Toll and Imd) are immune compromised and produce lower levels of AMPs upon microbial infections as compared to wild-type flies (Govind 2008).

Materials & Methods

***Drosophila* Lines and Genetics.** $RRS^{C\Delta}$ lines were generated using CRISPR-Cas9 system with pU6-BbsI-CG9020sg2-gRNA vector and *nanos*-Cas9. y'v';p {CG9020 sg2 dual v+} attP 40/Cyo line2 (Generated in NCBS) was the transgenic line. The lines that did not have any non-balancer males were considered as nulls and others as CRISPR control. Nulls were confirmed by rescuing the males by crossing the parents to RRS Myc overexpression line generated by Dr. Bhagyashree Khaduskar from Ratnaparkhi lab. pUASp-attB RRS wildtype and pUASp-attB RRS-SCR transgenic lines were generated using pUASp-attB vector cloning using PCR and PPY cells transformation. The plasmids were injected in the fly embryos at NCBS fly facility in Bangalore. The stable transgenic were screened using eye colour marker (mini-white gene). The flies were balanced over third chromosome balancers at NCBS and later balanced over both second and third chromosome balancer by me.

Embryonic Lethality Assay. 30 flies of each RRS$^{C\Delta}$ line and W^{III8} were collected and put in a fly egg collection cages. Small sucrose agar plates (3% w/v agar, 2.5% w/v sucrose) with some yeast paste were placed at the bottom of the cage. Flies were allowed to acclimatize in the cage for 24 hours at 25 °C. Plates were changed with fresh yeast paste every 12 hours. Eggs were collected after 3-6 hours of placing a fresh plate. 100 embryos were collected and were arranged in 10X10 grid on new plate (without yeast). The plates were incubated at 25°C for 24 hours. Number of eggs hatched were counted under a microscope after 24 hours.

Life span assay. RRS$^{C\Delta}$ (lines 6B1 and 18B1) were crossed to W^{1118} .100-120 virgin female flies of RRS$^{C\Delta}$/+ genotype were collected and stored with fresh 6% vials. 15-20 flies were kept in a single vial. 100-120 W^{1118} virgin females of the same age were also collected as wild-type control. The flies were maintained at 29°C and were transferred to new vials after every three days to provide fresh media. The number of flies surviving and dead in every vial was noted down every 24 hours till all the flies were dead. The 'percentage survival VS number of days' graph was plotted to analyse the result.

Null rescue. RRS$^{C\Delta}$ lines 6B1 and 18B1 balanced over X chromosome balancer FM7i were crossed to lines expressing ubiquitous Gal4 with UAS RRS wildtype or RRS SCR lines to get RRS rescued males. These crosses were set in 6% bottles at 29°C to increase the expression of Gal4 and thus increasing the number of rescue males. The males without FM7i (normal red eyes) were all the rescued males with genotype RRS$^{C\Delta}$ 6B1/Y ; Ubiquitous Gal4/+ ; UAS RRS wildtype/+ and RRS$^{C\Delta}$ 6B1/Y ; Ubiquitous Gal4/+ ; UAS RRS SCR/+ .

Infection assay. The 6-9 days old null rescue flies (RRS$^{C\Delta}$ 6B1/Y ; Actin Gal4/+ ; UAS RRS wildtype/+ and RRS$^{C\Delta}$ 6B1/Y ; Actin Gal4/+ ; UAS RRS SCR/+) along with controls (+/Y ; Actin Gal4/+ ; UAS RRS wildtype/+ , +/Y ; Actin Gal4/+ ; UAS RRS SCR/+ , RRS CRISPR control/Y ; Actin Gal4/+ ; UAS RRS wildtype/+ and RRS CRISPR control/Y; Actin Gal4/+ ; UAS RRS SCR/+) were infected with gram positive *Micrococcus luteus* (*M.lutues*) and gram negative *Erwinia carotovora carotovora 15* (Ecc15) by needle pricking. Pricking was done in the sternopleural plate of the thorax with the bacterial solution on the needle tip, avoiding the wings and the legs attachment sites. 5 uninfected same-aged flies of each genotype were crushed and

stored in TRIzol at -80°C. 5 flies of each genotype were infected and kept in a vial with fresh media for 12 hours in case of Ecc15 and 22 hours in case of *Micrococcus luteus* and then crushed and stored in TRIzol at -80°C. RNA was extracted from each sample and qRT-PCR was done for antimicrobial peptides (AMPs) of Toll and Imd pathway for *M.lutues* and *Ecc15* respectively. Drosomycin and Metchikowin were used for Toll pathway and Attacin D and Diptericin A were used for Imd pathway.

RNA extraction. The lysate was crushed and stored in TRIzol. It was spun and the supernant was stored in another microcentrifuge tube. Zymo research quick-RNA miniprep kit was used for RNA extraction. All centrifugation steps were performed at 13000g for 30 seconds. 1 volume of 95-100% ethanol was added to the lysate and mixed well. The mixture was added to a spin column, spun and the flow-through was discarded. To remove any trace of DNA, in column DNase treatment was done by first prewashing the column with 400µL RNA wash buffer and then adding DNase I with DNA digestion buffer and incubating the column for 15 minutes at room temperature. 400µL RNA prep buffer was added and spun followed by 700µL RNA wash buffer. Finally, the RNA was eluted in 30-50µL of nuclease free water. The RNA was stored at -80°C.

cDNA synthesis. Applied Biosystems High-Capacity cDNA Reverse Transcription Kit was used to synthesize cDNA from RNA. 2X reverse transcription master mix was prepared and equal amounts of RNA was added to each sample. Reverse transcription was performed in a thermal cycler with 3 steps. The first step was at 25°C for 10 minutes, followed by 120 minutes at 37°C and finally 5 minutes at 85°C. The cDNA was either directly used for doing Real-time PCR or stored at -20°C.

Quantitative Real Time-Polymerase Chain Reaction (qRT-PCR). KAPA SYBR® FAST qPCR Kit Master Mix (2X) Universal was used for performing real-time PCR. All reactions were 10µL in volume and carried out in Eppendorf white-back 96 well plate in Eppendorf Mastercycler Realplex. In the reaction master mix, 2X SYBR master mix was added along with the appropriate forward and reverse primers and 7µL was aliquoted in each well of the plate. 3µL of 1:3 or 1:5 time diluted cDNA was added to the mixture from top. The plate was tightly sealed with an adhesive transparent sheet and spun for a few seconds before placing it in the RT-PCR

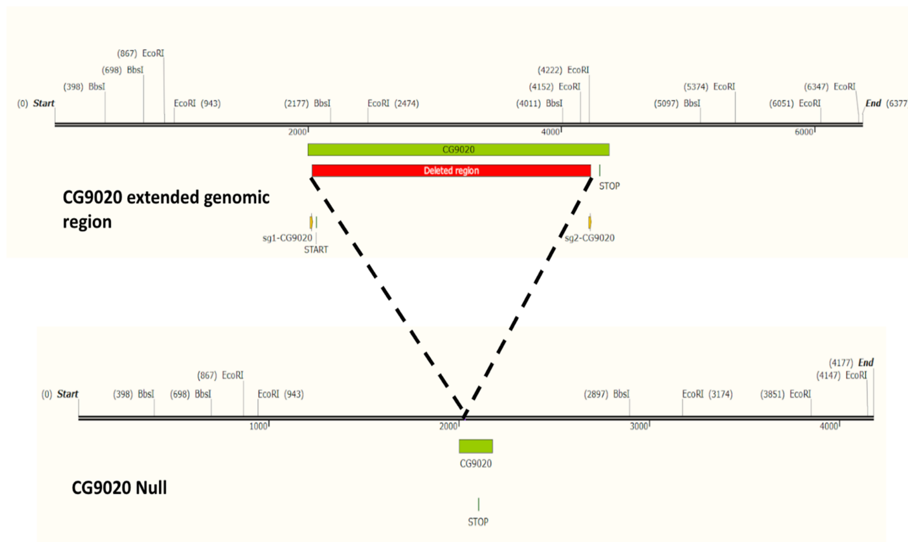
machine. A three step q-PCR cycle was performed. Enzyme activation was done at 95°C for 5 minutes followed by 40 cycles of denaturation at 95°C for 15 seconds, annealing at 60°C for 30 seconds and acquisition at 72°C for 30 seconds. Dissociation was done according to the melting curve set in the machine. The results were exported to excel sheet and analysed.

Bacterial growth and culture. *M. luteus* glycerol stock was streaked on a LB agar plate and a single colony was used to inoculate a 100mL culture (in LB broth). It was grown at 29°C for 18 hours and spun down at 4°C for 10 minutes at 6000rpm. The pellet was resuspended in 1X PBS to get an OD 70 which was used for infection. Ecc15 glycerol stock was streaked on a LB agar plate with 600µg/mL Ampicillin. A single colony was used to inoculate a 100mL culture of Ecc15 (in LB broth) which was grown at 29°C for 16 hours and spun down at 4°C for 10 minutes at 6000rpm. The pellet was resuspended in 1X PBS to get an OD 200 which was used for infection.

Results

Generation of RRS CRISPR Null ($RRS^{C\Delta}$). RRS nulls lines were generated using a dual gRNA method (Fig.21 A). The gRNA transgenic was generated and crossed to *nanos*- Cas9 transgenic flies (Fig.22 B). The progeny was crossed to FM7a (X chromosome balancer) line. The lines in which non-balancer males did not survive were chosen (Table 1) and crossed to UAS-RRS overexpression line (expressed with ubiquitous Gal4). The lines in which the male lethality was rescued were maintained. The rescued flies were morphologically normal and fertile.

A



B

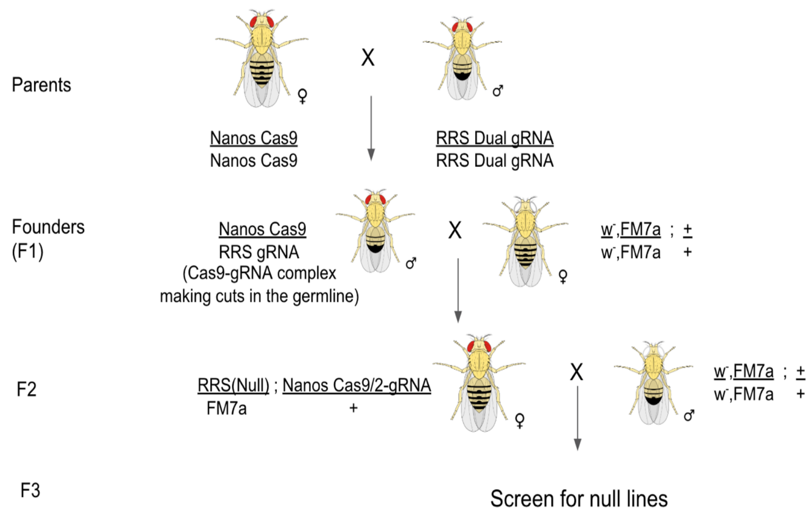


Fig. 21: Generation of RRS CRISPR null lines. A. Diagrammatic representation of RRS gene locus and gRNA target sites. The rectangle on the top shows the gene locus before the cut is made by Cas9 and the one below it shows the region after a predicted Cas9 cut and repair. **B.** The Homozygous nanos promoter (germline driver) driven Cas9 flies are crossed to homozygous dual gRNA expressing flies. The progeny (F1) are called founder flies as the gRNAs would make cuts in the germline of these flies. Each founder is maintained separately and crossed to X chromosome balancer (w^-, FM7a) as the target gene (RRS) is in the X chromosome. The progeny is again crosses to X chromosome balancer (w^-, FM7a) to clear either the dual gRNA or Cas9 in the background. The male of the resulting progeny is screened for normal red eyes.

Table 1: RRS$^{C\Delta}$ is larval lethal. The RRS$^{C\Delta}$/Y (males) and RRS$^{C\Delta}$ /RRS$^{C\Delta}$ (females) do not survive as RRS is a housekeeping gene. To test CRISPR-Cas9 null lines, the lethality was rescued with Myc-RRS wildtype; Daughterless Gal4 overexpression line. The table lists all the RRS null lines generated and rescued with UAS-GAL4 system and the phenotype showed.

RRS Null line (label)	Rescue (RRS Null; UAS-Myc-RRS wildtype; Daughterless Gal4)
3B5	Male lethality rescued by ectopic expression of RRS
5B4	Male lethality rescued by ectopic expression of RRS
6B1	Male lethality rescued by ectopic expression of RRS
12B1	Male lethality rescued by ectopic expression of RRS
18B1	Male lethality rescued by ectopic expression of RRS
18B7	Male lethality rescued by ectopic expression of RRS, rescued flies have shorter thoracic bristles
19B12	Male lethality rescued by ectopic expression of RRS
12A	Male lethality rescued by ectopic expression of RRS

RRS$^{C\Delta}$ is larval lethal. RRS is present on the X-chromosome of *Drosophila*. Homozygous RRS null female and RRS null male flies do not survive. Since homozygous RRS null female and RRS null males were lethal we could not confirm which portion is deleted with PCR or sequencing. In order to identify homozygous null larvae, we balanced the line with FM7i-Act GFP (Fig.22 A). The larvae without GFP were the true nulls and these animals were not embryonic lethal as they number of hatched embryos were similar in RRS null and W^{1118} (control) (Fig.22 B). They died in larval stages.

RRS$^{C\Delta}$ survives till early 2nd instar larval stage. Non-GFP larvae and wild-type W^{1118} (control) larvae of the same age were collected and kept separately in vials with media and on small sugar agar plates with yeast. They were observed every 12 hours for 4-5 days. The RRS$^{C\Delta}$ larvae did not grow beyond early second instar stage (Fig.22 C). Their growth was halted for 3 days and then died whereas the control larvae grew to become third instars.

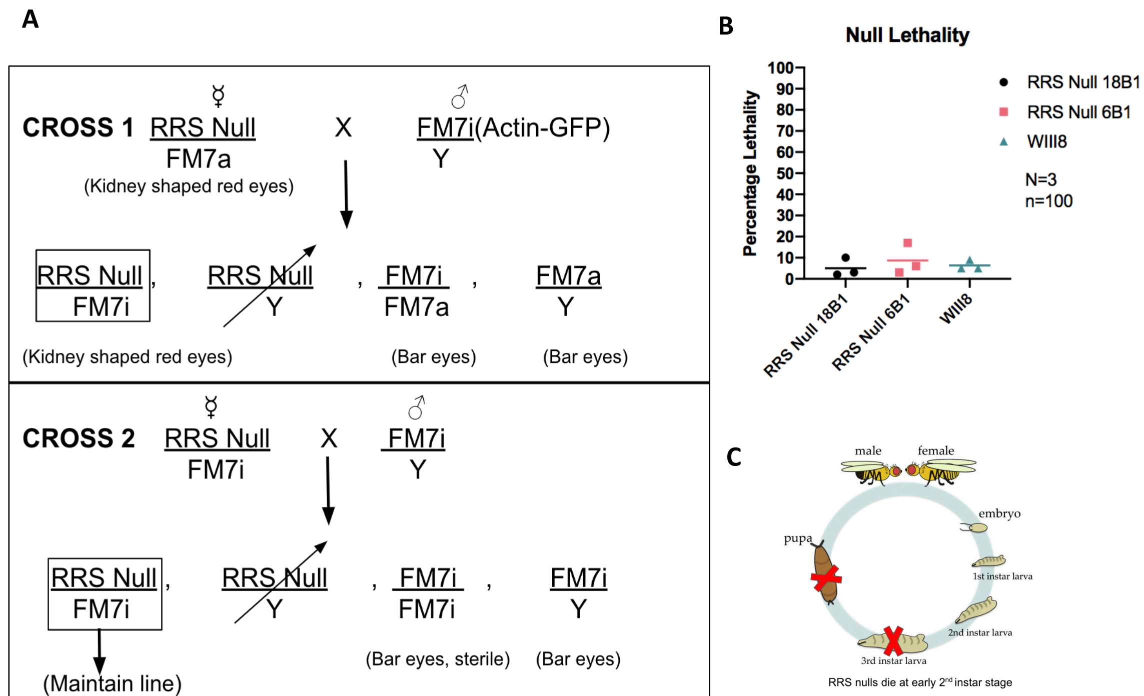


Fig. 22: Homozygous RRS null is not embryonic lethal but dies in early 2nd instar stage. A. Crosses to balance RRS null over a GFP containing balancer in order to distinguish homozygous RRS null from heterozygotes. Heterozygous RRS null females were crossed to FM7i males twice, sequentially. **B.** Lethality assay performed for 0-3 hours embryos with heterozygous RRS null lines (6B1 and 18B1) along with wildtype (W^{1118}). N=3, n=100 **C.** Diagrammatic representation of the *Drosophila* life cycle where homozygous RRS nulls do not survive beyond early 2nd instar stage.

Life span assay suggests haplosufficiency of RRS. Comparison of life span of two heterozygous $RRS^{C\Delta}$ lines (6B1 and 18B1) with W^{1118} flies at 29°C did not show any significant difference suggesting that one copy of RRS is sufficient for survival (Fig.23). Heterozygous $RRS^{C\Delta}$ lines do not show any morphological difference as well.

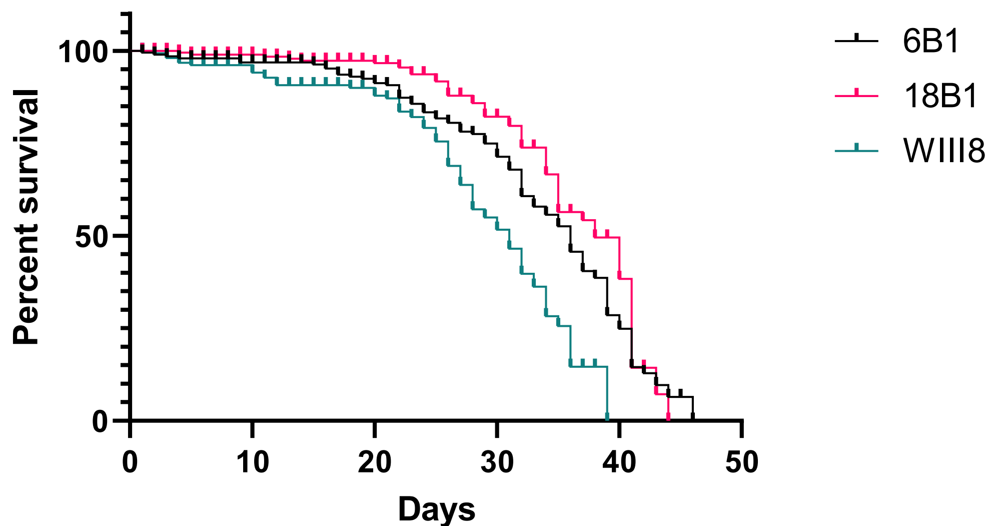


Fig. 23: RRS gene is haplosufficient in *Drosophila melanogaster*. The graph shows the life span of heterozygous $RRS^{C\Delta}$ flies (6B1 and 18B1) along with wildtype (W^{1118}) at 29°C. Here, N=3, n=~100, Log-rank (Mantel Cox) Test P<0.0001

Homozygous RRS mutants are functional nulls. The genomic DNA of the homozygous RRS null larvae were extracted and the RRS gene locus was amplified using suitable primers. The size of the null locus DNA was same as that of wildtype RRS locus (Fig 24). To confirm this result, the RRS locus of the homozygous RRS null larvae were sequenced. Few indels were found both in 3' and 5' UTR as well as the 3' coding region of RRS. To check if mRNA was produced, RT-PCR was performed. The results were inconclusive, thus the cDNA of homozygous RRS null larvae were amplified using RRS RT primers.

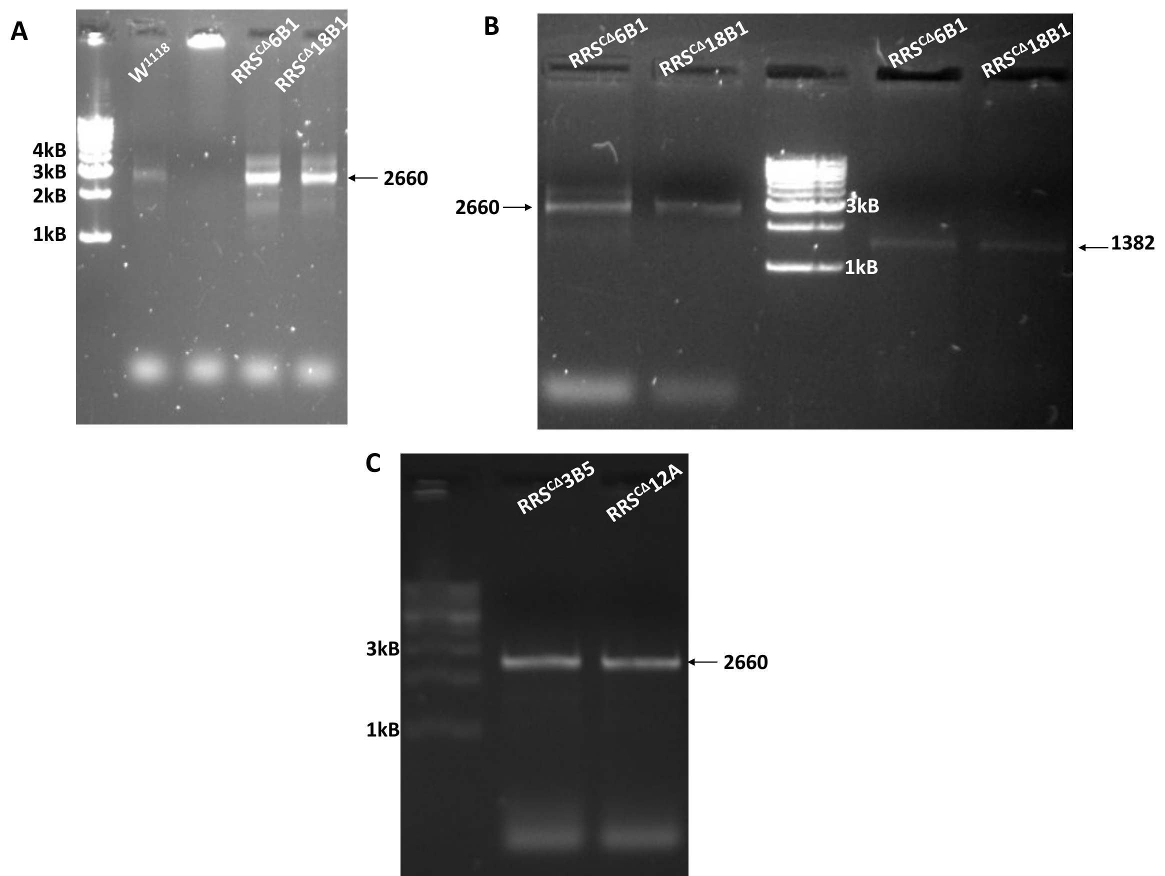


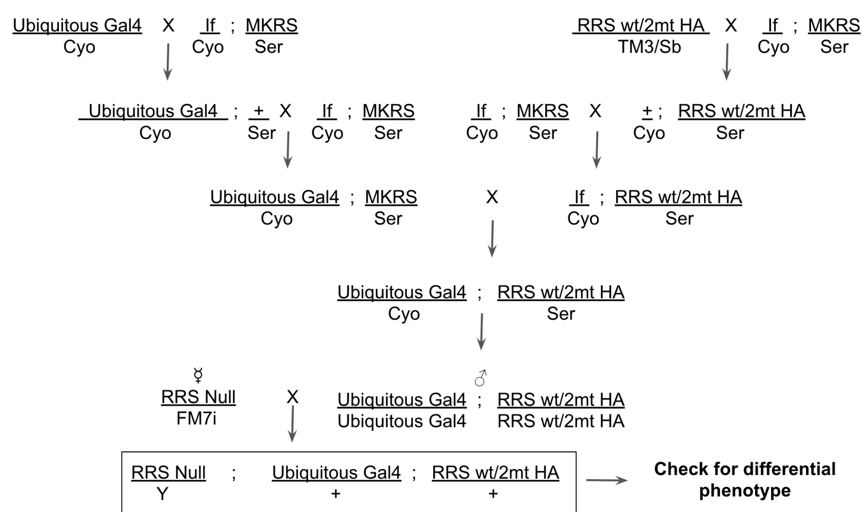
Fig. 24: Genomic PCR for RRS$^{C\Delta}$ does show any changes in the amplicon size as compared to the wildtype (W^{1118}). A. Gel image of PCR amplification of RRS extended genomic locus from early 2nd instar larvae of W^{1118} and RRS$^{C\Delta}$ 6B1, 18B1 using suitable primers (Appendix 1). The expected size of W^{1118} was 2660 base pairs and around 662 bp for RRS$^{C\Delta}$ 6B1, 18B1. **B.** Gel image of PCR amplification of RRS extended genomic locus from early 2nd instar larvae RRS$^{C\Delta}$ 6B1, 18B1 using suitable primers for the extended gene region and within the coding region. Lane-1,2 With primers upstream and downstream of the coding region, expected size in wildtype RRS-2660 bp, Lane-4,5 With primers within in the coding region, expected size in wildtype RRS-1382 bp **C.** Gel image of PCR amplification of RRS extended genomic locus from early 1st instar larvae RRS$^{C\Delta}$ 3B5, 12A using suitable primers for the extended gene region.

Sequencing of genomic region of RRS in RRS$^{C\Delta}$ lines. The sequences suggested that CRISPR/Cas9 editing did not give a large deletion as expected. However, the genomic sequences (Appendix Fig.28) indicated that deletions were made in the the RRS locus, which led to an absence of RRS function.

Cloning of pUASp-RRS and pUASp-SCR-RRS. RRS wildtype and RRS SCR were cloned into pUASp-AttP vector and sequenced. The plasmids were injected into *Drosophila* embryos and stable transgenic lines were made. These lines were balanced with second and third chromosome balancers and maintained.

Creating the null rescue system. Four different second chromosome GAL4 lines (Actin GAL4, Ubiquitin GAL4, Armadillo GAL4, Shibiri GAL4) were balanced with second and third chromosome balancers. UAS RRS wildtype and SCR lines were also balanced with second and third chromosome balancers. These balanced lines were crossed to to each other to create stable UAS-GAL4 lines (Fig.25 A). The UAS-GAL4 lines expressing wildtype and SCR RRS ubiquitously in flies were crossed to $RRS^{C\Delta}$ to check differential phenotype due to loss of SUMOylation of RRS (Fig.25 B).

A



B

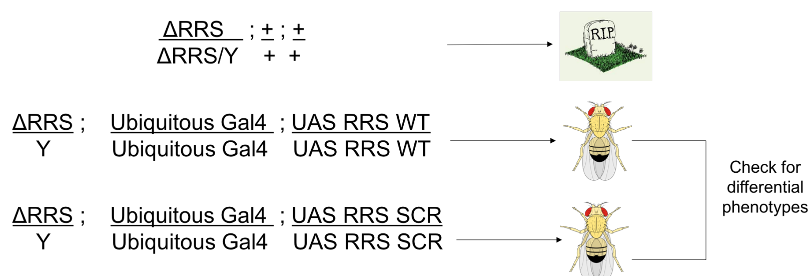


Fig. 25: RRS null rescue with RRS wildtype and RRS-SCR. A. A map of the series of crosses to obtain the null rescue males. **B.** A representative image showing the null rescue system. Homozygous RRS null flies do not survive which is rescued by overexpression of RRS-wildtype and SCR using UAS-GAL4 system. The rescued flies are checked for differential phenotypic expression or behaviour.

UAS-GAL4 rescue. The $RRS^{C\Delta}$ or RRS Null flies were rescued with UAS-RRS SCR and UAS- RRS WT as shown in Fig. 25 B. Four different genotypes (FM7i/Y; Actin GAL4/+; UAS-RRS SCR/WT, $RRS^{C\Delta}$ /Y; Actin GAL4/+; UAS-RRS SCR/WT, $RRS^{C\Delta}$ /X; Actin GAL4/+; UAS-RRS SCR/WT and FM7i /X; Actin GAL4/+; UAS-RRS SCR/WT) progeny were obtained from the both the crosses. The percentage of rescue males ($RRS^{C\Delta}$ /Y; Actin GAL4/+; UAS-RRS SCR/WT) was approximately 3% in both RRS-WT and RRS-SCR rescue of the total number of offspring. The percentage of male progeny in RRS-WT rescue was 13% whereas the RRS-SCR rescue was 24%.

Host response to pathogens. The UAS-GAL4 rescued $RRS^{C\Delta}$ males (5-9 days old) were infected with Gram-positive (*M. luteus*) and Gram-negative (Ecc-15) bacteria. qRT-PCR was performed for all the samples and the fold-change between amount of transcripts of AMPs such as *Metchnikowin*, *Drosomycin* (Fig. 26) and *DiptericinA, AttacinD* (Fig. 27) for Toll and Imd pathways respectively were plotted for both uninfected (0hr) and infected (12 or 22 hrs). The transcript levels of RRS were also compared before and after the infection.

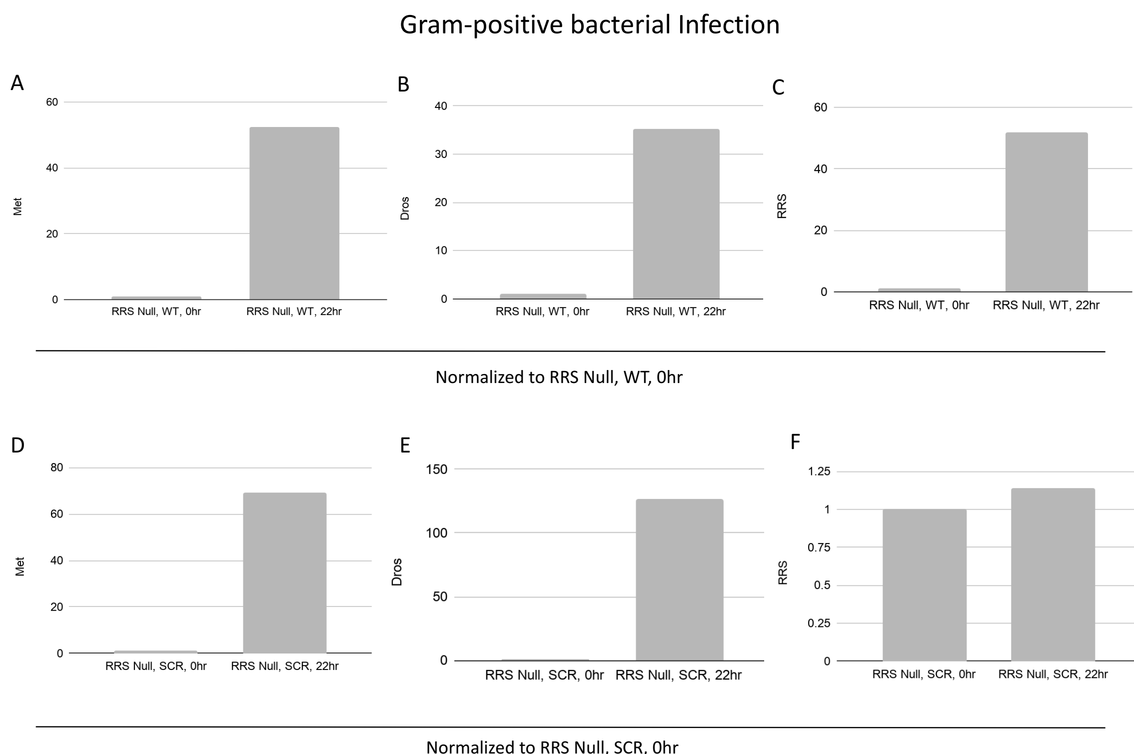


Fig. 26: Mild increase in transcript levels of Met and Dros in RRS^{SCR}. Fold-change in transcript levels of Toll pathway AMPs were measured upon Gram-positive bacterial infection

in RRS Null flies rescued with UAS-RRS-WT and UAS-RRS-SCR. The above plots show relative quantification of qRT-PCR data for *M. luteus* on RRS Null flies, analyzed by the Livak method (Reference gene- Rp49). **A,B,C** Fold-change in amount of transcripts of Metchnikowin, Drosomycin and RRS respectively in RRS Null rescued with UAS-RRS-WT uninfected (0hr) and UAS-RRS-WT infected (22hr) normalized to RRS Null rescued with UAS-RRS-WT at 0 hr. **D,E,F** Fold-change in amount of transcripts of Metchnikowin, Drosomycin and RRS respectively in RRS Null rescued with UAS-RRS-SCR uninfected (0hr) and UAS-RRS-SCR infected (22hr) normalized to RRS Null rescued with UAS-RRS-SCR at 0 hr.

For gram positive bacterial infection, where Toll/NFkB pathway is primary, the *Met* and *Dros* transcripts were higher by 1.3-fold and 3-fold respectively (Fig. 26 A, B, D, E), when comparing RRS^{SCR} to RRS^{WT}. The 26 C result appears to be an aberration and may be due to experimental error.

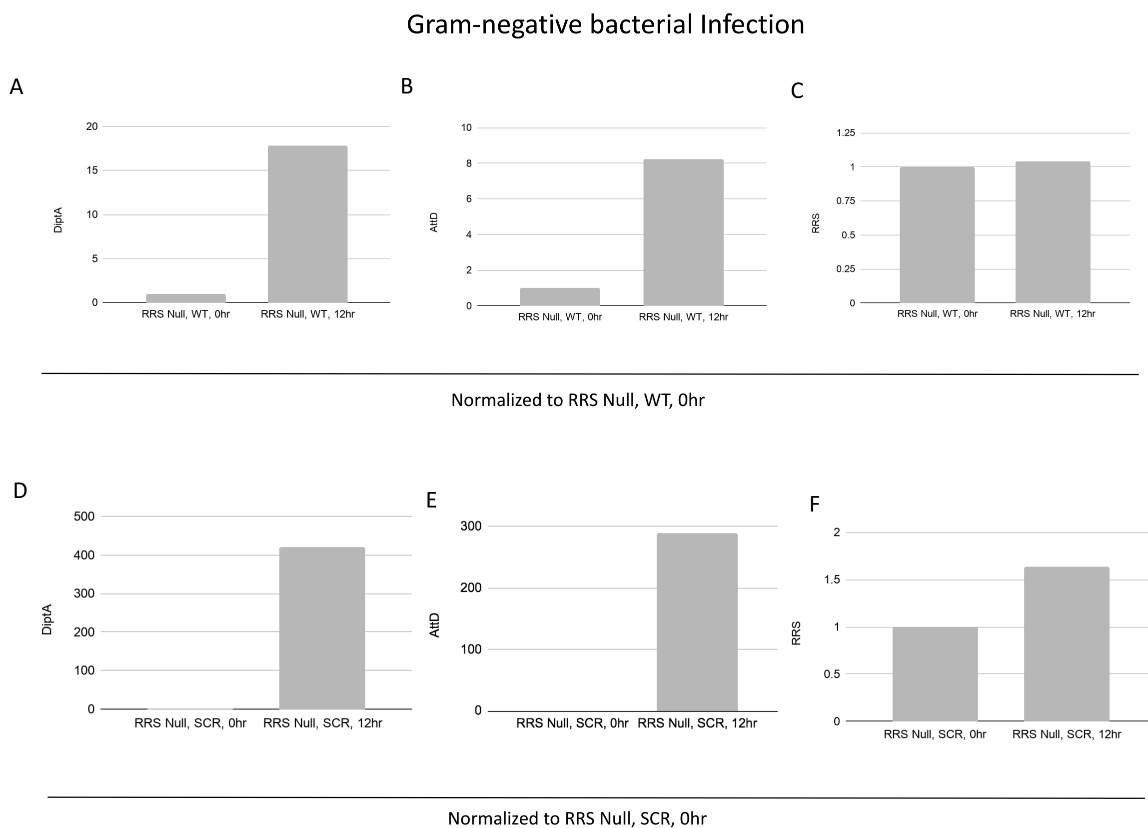


Fig. 27: Enhanced expression of *DiptA* and *AttD* in RRS^{SCR}. Fold-change in transcript levels of Imd pathway AMPs upon Gram-negative bacterial infection in RRS Null flies rescued with UAS-RRS-WT and UAS-RRS-SCR were measured**.** The above plots show relative quantification of qRT-PCR data for Gram-negative bacterial infection (Ecc15), analyzed by the Livak method (Reference gene- Rp49). **A,B,C** Fold-change in amount of transcripts of *Diptericin A*, *Attacin D* and *RRS* respectively in RRS Null rescued with UAS-RRS-WT uninfected (0hr) and UAS-RRS-WT infected (12hr) normalized to RRS Null rescued with UAS-RRS-WT at 0 hr. **D,E,F** Fold-change in amount of transcripts of Diptericin A, Attacin D and RRS respectively in RRS Null rescued with UAS-RRS-SCR uninfected (0hr) and UAS-RRS-SCR infected (12hr) normalized to RRS Null rescued with UAS-RRS-SCR at 0 hr.

For gram negative bacterial infection (Fig. 27), where IMD/NFkB pathway is primary, the *DiptA* and *AttD* transcripts were higher by 24-fold and 36-fold respectively (Fig.

27 A, B, D, E), when comparing RRS^{SCR} to RRS^{WT}. The RRS transcript level by itself is up by 1.5-fold.

Discussion

RRS gene is located on the X chromosome of *Drosophila melanogaster*. In *Drosophila* the dosage compensation of X chromosome happens by increasing the transcription of genes on single X chromosome of male equal to that of both the X chromosomes in females. Thus, loss of one copy of RRS in males is equivalent to homozygous RRS null females. While maintaining the RRS null lines, the males are always maintained as balancers since null males do not survive. RRS is a house-keeping enzyme so the absence of it is lethal for the flies. Surprisingly, RRS nulls are not embryonically lethal in *D. melanogaster*. This could be possibly due to the maternal deposition of RRS mRNA as well as protein. The maternal RNA gets degraded in the first few hours of development but most likely the residual protein allows the organism to survive till early 2nd instar stage.

The dual-sgRNA CRISPR technique should ideally lead to a complete deletion of the RRS locus but in most the null lines that we generated there are only some indels. We hypothesize that, RRS being a housekeeping gene and an important locus is more protected and has more efficient repair mechanism which makes it difficult to disrupt the locus. We amplified the cDNA of homozygous RRS null larvae using suitable primers for RRS coding region which suggest that the mRNA of RRS is produced in the null lines. Since, we do not have an antibody against RRS we could not confirm the presence or absence of the protein but lethality of RRS null *Drosophila* beyond 2nd instar stage suggests the absence of the functional enzyme. This lethality can be rescued by ectopic expression of RRS in the null organism. Thus, this suggests that the RRS CRISPR mutant flies are functional RRS nulls.

In order to test whether RRS is haplosufficient, we compared the life span of heretozygous RRS females with W^{1118} at 29°C. Since, the life spans were comparable, it suggests that the loss of one copy of RRS does not affect the physiology of flies.

Using the UAS-GAL4 system RRS null males which were lethal, were rescued by both RRS-wildtype and RRS-SCR mutant. The rescued flies were all morphologically

similar and normal along with being completely fertile. Thus, there were no visible differences between RRS-wildtype and RRS-SCR flies. To uncover the role of SUMOylation of RRS in *D. melanogaster*, an immunity assay was performed to check for differential immune activity. SUMOylation of RRS was shown to upregulated during innate immune challenge in *Drosophila* S2 cells, suggesting that RRS SUMOylation might have a role in innate immunity. To follow this result, we performed the infection assay with $RRS^{C\Delta}$ or RRS null flies rescued with UAS-RRS SCR and UAS- RRS WT. The rescued flies were infected with Gram-positive and Gram-negative bacteria, the transcript levels of AMPs as well as RRS were compared before and after the infection.

Conclusions

Functional null flies ($RRS^{C\Delta}$) were generated using CRISPR-Cas9 technology. Homozygous $RRS^{C\Delta}$ do not survive and can be rescued by ubiquitous expression of RRS. Homozygous $RRS^{C\Delta}$ do not die in the embryonic stages, rather survive up to early 2nd instar stage, supported by maternal RRS mRNA deposits. Heterozygous ($RRS^{C\Delta}$ /+) show no morphological or fertility defects and are equivalent to wildtype flies. UAS-RRS^{WT} and UAS-RRS^{SCR} transgenic fly lines were generated and used to rescue the $RRS^{C\Delta}$ flies. The rescued flies were morphologically normal as well as similar and had no fertility defects. Preliminary data indicates differences in the humoral response between the RRS^{WT} and RRS^{SCR} flies in response to bacterial infections.

Future Work

The Null: Rescue system for RRS gives very few rescued flies (~3% of total flies). Though not as ideal as CRISPR/Cas9 based genome editing of the RRS locus, the system allows us to address roles for SUMOylation of RRS in host defense. The data collected is preliminary, but the system is set-up for a host of experiments that will compare the immune reponse in RRS^{WT} and RRS^{SCR} flies.

Contributions

I thank the fly facility at NCBS for injection of the pUASp-AttP HA:RRS wildtype and pUASp-AttP HA:RRS^{SCR} plasmids in fly embryos and for creating the stable transgenic lines

Future directions

The major objective of this project was to create a SUMO conjugation resistant RRS mutant fly and check the difference in the phenotype compared to a wildtype fly in order to understand the biological role of SUMOylation of Arginyl tRNA synthetase. I have successfully created $RRS^{C\Delta}$ flies and rescued those with UAS-RRS-WT and UAS-RRS-SCR using ubiquitous GAL4. This system allowed us to study and compare the immune system functioning of the RRS^{WT} and RRS^{SCR} flies. The preliminary data suggests an increase in transcripts of anti-microbial peptides (AMPs) of primarily IMD/NFkB (*DiptA* and *AttD)* when the RRS^{SCR} is compared to RRS^{WT}. This result has to be validated by repeating the assay and generating more replicates to gain confidence in the findings. The system also allows us to perform various other assays and experiments to understand the biological role of SUMOylation of RRS. Life span of the RRS^{WT} and RRS^{SCR} can be compared upon heat stress and immune challenge. This will help us understand is SUMOylation of RRS plays a role in stress response and immunity. Since, RRS is a part of the Multi-aminoacyl tRNA synthetase complex, it would also be interesting to check if RRS SUMOylation affects its ability to interact in the complex. This could be done by co-immunoprecipitation assays. RRS is known to interact with EPRS and other AIMPs. Using affinity beads, RRS^{WT} and RRS^{SCR} can be pulled down and using western blotting and mass spectrometry the change in interactors (if any) can be found. Even though the Null-rescue strategy is an excellent system but it has a limitation that only males can be studied and only about 3% of total progeny are rescued male flies which makes it very difficult to collect enough number of flies to perform experiments. UAS-GAL4 system is also not very robust in expression and results might fluctuate based on the expression level of genes driven by it.

To overcome this problem, the CRISPR/Cas9 based genome editing of the RRS locus is very crucial, the sgRNA transgenic for which is already ready. The next in this project would be to clone the donor plasmid for homologous recombination and inject the plasmid in sgRNA-Cas9 expressing embryos. All the results obtained in the UAS-GAL4 rescue system would be reproduced in this for validation. To understand which pathways SUMOylated RRS is involved in, a quantitive RNA sequencing can be performed to check change in transcript levels of genomic RRS^{WT} and RRS^{SCR}.

Also, if the results suggest a non-canonical function of RRS upon SUMOylation or loss of SUMOylation then we would have to prove that aminoacylation is not affected in genomic RRS^{SCR} mutant flies. To validate this, *in-vitro* charging assays can be performed with both RRS^{WT} and RRS^{SCR}.

Appendix

Table 2: Primers for PCR. A list of primers used for my research. Primers were procured from 'Eurofins' and 'First Base' and concentration made up to 100µM using MilliQ water as stcoks and 10µM as working concentration.

Name	Sequence	Use
RRS(SCR) 5'HR FP	gggaagagcgacggctcttcccttgttcagcttc aaattgaaaataagaagctgaag	Scarless plasmid 5' fragment forward primer
RRS(SCR) 5'HR RP	cgcagactatctttctagggttaatggagttcgg cacacctttctg	Scarless plasmid 5' fragment reverse primer
RRS(SCR) DS RED FP	ggtgtgccgaactccattaaccctagaaagat agtctgcgtaaa	Scarless plasmid DSRED fragment forward primer
RRS(SCR) DSred RP	caaattagctcccacgccttaaccctagaaag ataatcatattgtgacgtac	Scarless plasmid DSRED fragment reverse primer
RRS(SCR) 3' FP	atgattatctttctagggttaaggcgtgggagcta atttg	Scarless plasmid 3' fragment forward primer
RRS(SCR) 3' RP	aggtggcagcacctgcgatctgtaaagtaagt acaccgcagtgttgc	Scarless plasmid 3' fragment reverse primer
RRS(SCR) Frag FP	gatcgcaggtgctgccacc	Scarless plasmid vector fragment forward primer
RRS(SCR) Frag RP	ggaagagccgtcgctcttcc	Scarless plasmid vector fragment reverse primer
K147R FP	attgcaaccgagttgcgcggacactgcccagc a	Mutagenesis primer for K147R forward
K147R RP	tgctgggcagtgtccgcgcaactcggttgcaat	Mutagenesis primer for K147R reverse
K383R FP	cctctgacaatcgtgcgctcggatggcggcttt	Mutagenesis primer for K383R forward
K383R RP	aaagccgccatccgagcgcacgattgtcaga gg	Mutagenesis primer for K383R reverse
CG9020 FP2	acgacatcctagaagtgactg	Null screening primer forward
CG9020 RP	cagctagtgtgaatgcgaac	Null screening primer reverse
RRS GENOMIC RP	aaactgtacgtggtcaacacgatggctgagtg gatttaggatagcag	Null screening primer reverse
RRS RT 1 FP	gccttgccgccagaataca	RRS RT primer forward
RRS RT 1 RP	ggcagttgattcctcagcaat	RRS RT primer reverse
RRS RT 2 FP	gcccaaggaatcctcttcgat	RRS RT primer forward
RRS RT 2 RP	tccgggaaagctgatgcaatc	RRS RT primer reverse
q_AttD-FP	cgg tca acg cca atg gtc at	Attacin D RT primer forward

q_AttD-RP	cat tca gag cgg cgt tat tg	Attacin D RT primer reverse
q-DptA-FP	gctgcgcaatcgcttctact	Diptericin A RT primer forward
q-DptA-RP	cat cgc cgc tct ggc cac	Diptericin A RT primer reverse
q_Drs-FP	ctg tcc gga aga tac aag gg	Drosomycin RT primer forward
q_Drs-RP	tcg cac cag cac ttc aga ct	Drosomycin RT primer reverse
Mtk FP	gctacatcagtgctggcaga	Metchnikowin RT primer forward
Mtk RP	ttaggattgaagggcgacgg	Metchnikowin RT primer reverse

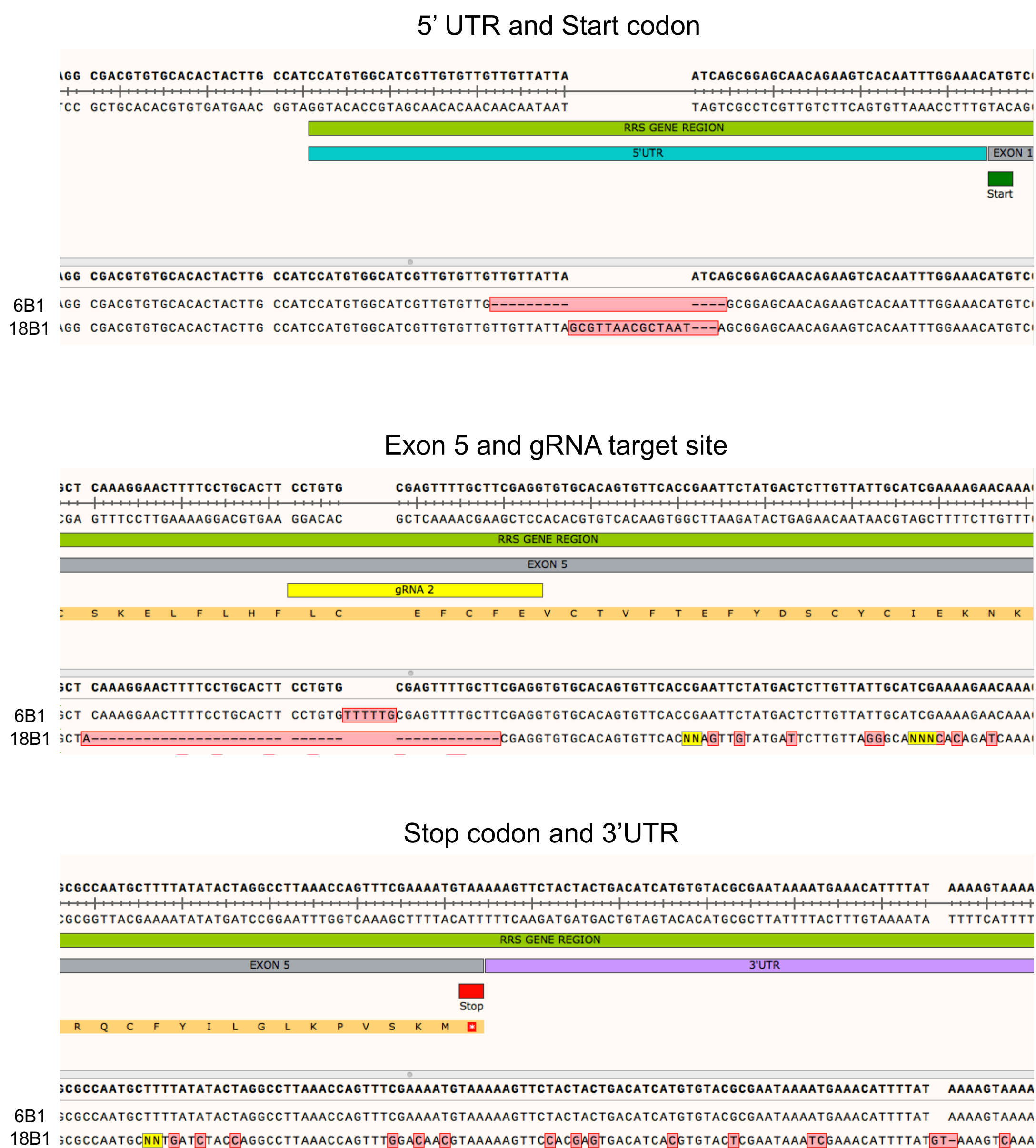
Table 3. Drosophila lines generated. A list of transgenic and mutant lines used for my experiments which were generated during the course of this project.

Sl.No.	Genotype	Description	Source
1	RRS 12a/ Fm7i	RRS Null line ($RRS^{C\Delta}$)	CRISPR/Cas9 transgenic
2	RRS CRISPR control/Fm7i	RRS Null CRISPR Control	CRISPR/Cas9 transgenic
3	Actin Gal4/Cyo; RRS wt/Ser	UAS-Gal4 line-wildtype for rescue	Actin Gal4/Cyo ; MKRS/Ser and RRS WT HA/TM3 Sb
4	Actin Gal4/Cyo; RRS 2mt/Ser	UAS-Gal4 line-SCR mutant for rescue	Actin Gal4/Cyo ; MKRS/Ser and RRS 2mt HA/TM3 Sb
5	If/Cyo; RRS wt (Line 1)/Ser	UAS-RRS:HA wildtype	RRS WT HA(Line 1) / TM3 Sb
6	If/Cyo; RRS wt (Line 2)/Ser	UAS-RRS:HA wildtype	RRS WT HA(Line 2) / TM3 Sb
7	If/Cyo; RRS 2mt (Line 1)/Ser	UAS-RRS:HA SCR mutant	RRS SCR HA(Line 1) / TM3 Sb
8	If/Cyo; RRS 2mt (Line 2)/Ser	UAS-RRS:HA SCR mutant	RRS SCR HA(Line 2) / TM3 Sb
9	RRS wt (Line 1)/TM3 Sb	UAS-RRS:HA wildtype	pUASp AttB transgenic
10	RRS wt (Line 2)/TM3 Sb	UAS-RRS:HA wildtype	pUASp AttB transgenic
11	RRS 2mt (Line 1)/TM3 Sb	UAS-RRS:HA SCR mutant	pUASp AttB transgenic
12	RRS 2mt (Line 2)/TM3 Sb	UAS-RRS:HA SCR mutant	pUASp AttB transgenic
13	Shibiri Gal4/Cyo ; MKRS/Ser	Chorosome II ubiquitous Gal4	Shibiri Gal4/Cyo
14	Ubiquitin Gal4/Cyo ; MKRS/Ser	Chorosome II ubiquitous Gal4	Ubiquitin Gal4/Cyo
15	Armadilo Gal4/Cyo ; MKRS/Ser	Chorosome II ubiquitous Gal4	Armadilo Gal4/Cyo
16	RRS 6B1/ Fm7i	RRS Null line	RRS 6B1/ Fm7a
17	RRS 18B1/Fm7i	RRS Null line	RRS 18B1/Fm7a
18	RRS 3B5/Fm7i	RRS Null line	RRS 3B5/Fm7a
19	RRS sg2	RRS SCR gRNA transgenic	CRISPR/Cas9 transgenic
20	Ubiquitin Gal4/Cyo ; RRS wt (Line1)/Ser	UAS-Gal4 line-wildtype for rescue	Ubiquitin Gal4/Cyo ; RRS WT HA(Line 1) / TM3 Sb
21	Ubiquitin Gal4/Cyo ; RRS 2mt /Ser	UAS-Gal4 line-SCR mutant for rescue	Ubiquitin Gal4/Cyo ; RRS SCR HA / TM3 Sb

Notes:

1. All lines have been deposited to the IISER Stock Centre (GR13, GR Lab)
2. The HA tagged lines are N-terminal fusions
3. C(1)DX, $y^1 f^1$ is the background for RRS Null lines
4. $y^1 w^1$; P{CaryP}attP2 is the background for RRS WT and SCR HA lines

Fig. 28. A figure showing changes in sequences in the RRS null line 6B1 and 18B1. This is a representative of all the RRS null sequencing result aligned with the original RRS sequence.Sequences of $RRS^{C\Delta}$ 6B1 and $RRS^{C\Delta}$ 18B1 lines are aligned along with the ideal sequence. The RRS locus was sequenced with primers in annealing in 5'UTR and 3'UTR (Approximately 200 bases upstream of the start and stop codons). $RRS^{C\Delta}$ 6B1 and $RRS^{C\Delta}$ 18B1 lines were majorly used for experiments.



References

1. 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.
2. Brand, A. H., and N. Perrimon. 1993. “Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes.” *Development* 118 (2): 401–15.
3. Brennan, Catherine A., and Kathryn V. Anderson. 2004. “Drosophila: The Genetics of Innate Immune Recognition and Response.” *Annual Review of Immunology*. https://doi.org/10.1146/annurev.immunol.22.012703.104626.
4. Buchon, Nicolas, Neal Silverman, and Sara Cherry. 2014. “Immunity in Drosophila Melanogaster — from Microbial Recognition to Whole-Organism Physiology.” *Nature Reviews Immunology*. https://doi.org/10.1038/nri3763.
5. Cribbs, Adam P., and Sumeth M. W. Perera. 2017. “Science and Bioethics of CRISPR-Cas9 Gene Editing: An Analysis Towards Separating Facts and Fiction.” *The Yale Journal of Biology and Medicine* 90 (4): 625–34.
6. Desterro, Joana M. P., Manuel S. Rodriguez, Graham D. Kemp, and Ronald T. Hay. 1999. “Identification of the Enzyme Required for Activation of the Small Ubiquitin-like Protein SUMO-1.” *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.274.15.10618.
7. Duffy, Joseph B. 2002. “GAL4 System Indrosophila: A Fly Geneticist’s Swiss Army Knife.” *Genesis*. https://doi.org/10.1002/gene.10150.
8. Giniger, E., and M. Ptashne. 1988. “Cooperative DNA Binding of the Yeast Transcriptional Activator GAL4.” *Proceedings of the National Academy of Sciences of the United States of America* 85 (2): 382–86.
9. Giniger, E., S. M. Varnum, and M. Ptashne. 1985. “Specific DNA Binding of GAL4, a Positive Regulatory Protein of Yeast.” *Cell* 40 (4): 767–74.
10. Govind, Shubha. 2008. “Innate Immunity in Drosophila: Pathogens and Pathways.” *Insect Science* 15 (1): 29–43.
11. Guo, Min, and Paul Schimmel. 2012. “Structural Analyses Clarify the Complex Control of Mistranslation by tRNA Synthetases.” *Current Opinion in Structural Biology* 22 (1): 119–26.
12. Handu, Mithila, Bhagyashree Kaduskar, Ramya Ravindranathan, Amarendranath Soory, Ritika Giri, Vijay Barathi Elango, Harsha Gowda, and Girish S. Ratnaparkhi. 2015. “SUMO-Enriched Proteome for Drosophila Innate Immune Response.” *G3* 5 (10): 2137–54.
13. He, Xia-Di, Wei Gong, Jia-Nong Zhang, Ji Nie, Cui-Fang Yao, Fu-Shen Guo, Yan Lin, et al. 2018. “Sensing and Transmitting Intracellular Amino Acid Signals through Reversible Lysine Aminoacylations.” *Cell Metabolism* 27 (1): 151–66.e6.
14. Hoffmann, Jules A. 2003. “The Immune Response of Drosophila.” *Nature*. https://doi.org/10.1038/nature02021.
15. Imler, Jean-Luc, and Philippe Bulet. 2005. “Antimicrobial Peptides in Drosophila: Structures,Activities and Gene Regulation.” *Chemical Immunology and Allergy*. https://doi.org/10.1159/000086648.

16. Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science* 337 (6096): 816–21.
17. Kakidani, H., and M. Ptashne. 1988. “GAL4 Activates Gene Expression in Mammalian Cells.” *Cell* 52 (2): 161–67.
18. Kaneko, Takashi, Tamaki Yano, Kamna Aggarwal, Jae-Hong Lim, Kazunori Ueda, Yoshiteru Oshima, Camilla Peach, et al. 2006. “PGRP-LC and PGRP-LE Have Essential yet Distinct Functions in the Drosophila Immune Response to Monomeric DAP-Type Peptidoglycan.” *Nature Immunology* 7 (7): 715–23.
19. Kim, Sunghoon. 2014. *Aminoacyl-tRNA Synthetases in Biology and Medicine*. Springer.
20. Ko, Y. G., E. Y. Kim, T. Kim, H. Park, H. S. Park, E. J. Choi, and S. Kim. 2001. “Glutamine-Dependent Antiapoptotic Interaction of Human Glutaminyl-tRNA Synthetase with Apoptosis Signal-Regulating Kinase 1.” *The Journal of Biological Chemistry* 276 (8): 6030–36.
21. Lamb, Abigail M., Elizabeth A. Walker, and Patricia J. Wittkopp. 2017. “Tools and Strategies for Scarless Allele Replacement in Drosophila Using CRISPR/Cas9.” *Fly*. https://doi.org/10.1080/19336934.2016.1220463.
22. Lemaitre, B., J. -M. Reichhart, and J. A. Hoffmann. 1997. “Drosophila Host Defense: Differential Induction of Antimicrobial Peptide Genes after Infection by Various Classes of Microorganisms.” *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.94.26.14614.
23. Liu, J. 2004. “A New -Interferon-Inducible Promoter and Splice Variants of an Anti-Angiogenic Human tRNA Synthetase.” *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkh240.
24. Müller, Stefan, Carsten Hoege, George Pyrowolakis, and Stefan Jentsch. 2001. “Sumo, Ubiquitin’s Mysterious Cousin.” *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/35056591.
25. 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.
26. Park, Sang Gyu, Paul Schimmel, and Sunghoon Kim. 2008. “Aminoacyl tRNA Synthetases and Their Connections to Disease.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (32): 11043–49.
27. Pirone, Lucia, Wendy Xolalpa, Jón Otti Sigurðsson, Juanma Ramirez, Coralia Pérez, Monika González, Ainara Ruiz de Sabando, et al. 2017. “A Comprehensive Platform for the Analysis of Ubiquitin-like Protein Modifications Using in Vivo Biotinylation.” *Scientific Reports* 7 (January): 40756.
28. Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, and Feng Zhang. 2013. “Genome Engineering Using the CRISPR-Cas9 System.” *Nature Protocols* 8 (11): 2281–2308.
29. Ray, Partho Sarothi, and Paul L. Fox. 2007. “A Post-Transcriptional Pathway Represses Monocyte VEGF-A Expression and Angiogenic Activity.” *The EMBO Journal*. https://doi.org/10.1038/sj.emboj.7601774.
30. 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.
31. Rodriguez, Manuel S., Catherine Dargemont, and Ronald T. Hay. 2001. “SUMO-1 Conjugationin VivoRequires Both a Consensus Modification Motif and Nuclear Targeting.” *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.m009476200.

32. Sampson, Deborah A., Min Wang, and Michael J. Matunis. 2001. "The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and Is Essential for SUMO-1 Modification." *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.m100006200.
33. Scandurro, A., C. Weldon, Y. Figueroa, J. Alam, and B. Beckman. 2001. "Gene Microarray Analysis Reveals a Novel Hypoxia Signal Transduction Pathway in Human Hepatocellular Carcinoma Cells." *International Journal of Oncology*. https://doi.org/10.3892/ijo.19.1.129.
34. "Scarless Gene Editing – flyCRISPR." n.d. Accessed March 26, 2020. https://flycrispr.org/scarless-gene-editing/.
35. Tzima, E., J. S. Reader, M. Irani-Tehrani, K. L. Ewalt, M. A. Schwartz, and P. Schimmel. 2003. "Biologically Active Fragment of a Human tRNA Synthetase Inhibits Fluid Shear Stress-Activated Responses of Endothelial Cells." *Proceedings of the National Academy of Sciences of the United States of America* 100 (25): 14903–7.
36. Webster, N., J. R. Jin, S. Green, M. Hollis, and P. Chambon. 1988. "The Yeast UASG Is a Transcriptional Enhancer in Human HeLa Cells in the Presence of the GAL4 Trans-Activator." *Cell* 52 (2): 169–78.
37. Arnez, J.G. and Moras, D. 1997. Structural and functional considerations of the aminoacylation reaction. *Trends in Biochemical Sciences* 22(6), pp. 211–216.
38. Gareau, J.R. and Lima, C.D. 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews. Molecular Cell Biology* 11(12), pp. 861–871.
39. Handu, M., Kaduskar, B., Ravindranathan, R., et al. 2015. SUMO-Enriched Proteome for Drosophila Innate Immune Response. *G3 (Bethesda, Md.)* 5(10), pp. 2137–2154.
40. Huang, H.W., Tsoi, S.C., Sun, Y.H. and Li, S.S. 1998. Identification and characterization of the SMT3 cDNA and gene encoding ubiquitin-like protein from Drosophila melanogaster. *Biochemistry and molecular biology international* 46(4), pp. 775–785.
41. Hyeon, D.Y., Kim, J.H., Ahn, T.J., Cho, Y., Hwang, D. and Kim, S. 2019. Evolution of the multi-tRNA synthetase complex and its role in cancer. *The Journal of Biological Chemistry* 294(14), pp. 5340–5351.
42. Ibba, M., Becker, H.D., Stathopoulos, C., Tumbula, D.L. and Söll, D. 2000. The adaptor hypothesis revisited. *Trends in Biochemical Sciences* 25(7), pp. 311–316.
43. Kerjan, P., Cerini, C., Sémériva, M. and Mirande, M. 1994. The multienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from Drosophila to mammals. *Biochimica et Biophysica Acta* 1199(3), pp. 293–297.
44. Park, S.G., Kim, H.J., Min, Y.H., et al. 2005. Human lysyl-tRNA synthetase is secreted to trigger proinflammatory response. *Proceedings of the National Academy of Sciences of the United States of America* 102(18), pp. 6356–6361.
45. Schmidt, E. and Schimmel, P. 1994. Mutational isolation of a sieve for editing in a transfer RNA synthetase. *Science* 264(5156), pp. 265–267.
46. Smith, M., Turki-Judeh, W. and Courey, A.J. 2012. Sumoylation in drosophila development. *Biomolecules* 2(3), pp. 331–349.
47. Talamillo, A., Sánchez, J., Cantera, R., et al. 2008. Smt3 is required for Drosophila melanogaster metamorphosis. *Development* 135(9), pp. 1659–1668.
48. Barrangou, R., 2015. The roles of CRISPR–Cas systems in adaptive immunity and beyond. *Current Opinion in Immunology*, 32, pp.36-41.
49. Horvath, P. and Barrangou, R., 2010. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327(5962), pp.167-170.