ANALYSIS OF
ALTERNATIVE SPLICING REGULATION IN THE
HYPERVARIABLE RECEPTOR DSCAM

By
YASH RAMESH HEMANI

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham
September 2012
Abstract

The pattern recognition receptor Dscam is a key molecule mediating innate immunity and wiring of the nervous system in Drosophila. Intriguingly, massive molecular diversity is generated by alternative splicing in three exon clusters of Dscam.

Upon pathogen exposure in Anopheles gambiae, the AgDscam splicing pattern changes to express isoforms that bind pathogens with higher affinity. In order to test the generality of Dscam splicing regulation in Drosophila, similar experiments involving microbial exposure were carried out, which also showed changes in Dscam splicing pattern. Mutants in RNA regulatory pathways and in the RNA binding protein ELAV were analyzed due to their similar mutant phenotypes in nervous system development as Dscam. In each of these mutants, alterations of Dscam alternative splicing in a cluster specific manner were observed, eluding a unique mechanism for any of the analyzed pathways. In ELAV mutants, one of the three clusters of alternatively spliced exons is dramatically mis-regulated.

Since no ELAV binding site is present in this cluster, genes downstream of ELAV could mediate mis-regulation of alternative splicing. From the analysis of mutants in ELAV differentially regulated genes it was concluded that Dscam alternative splicing is most prominently affected by chromatin remodeling factors, along with RNA binding proteins, DNA binding proteins and small-RNA processing factors. A heterologous transgene for expression of Dscam pre-mRNA in Drosophila was also developed to characterize the role of the chromatin state in alternative splicing.
Dedicated to

Pujiyashri Gurudev

Baa

Mom and Dad

and the entire Hemani family
Acknowledgements

I would like to sincerely thank my PhD supervisor, Dr. Matthias Soller for guiding me throughout the course of my PhD. This project would not have been possible without your remarkable insights and technical prowess. Thank you for instilling in me the importance of efficient time management, multi-tasking and proaction. I earnestly thank The Darwin Trust of Edinburgh for funding my PhD.

I am very grateful to all members of the Soller lab namely Irmi, Emma, Min and Pinar for their support and cooperation without which this experience would have been far more difficult. Thanks to Irmi for her expert advice and useful discussions. Thanks to Emma for teaching me the basics of fly genetics and for being a great companion for all these years. Thanks to Min for teaching me fly transgenesis and for being a great listener in times of stress. I would like to extend my gratitude to all the past and present members of the Brogna, Hidalgo and Dove labs for making the ‘6th floor’ such a fun place to work in.

I sincerely thank Mr. Suresh Patel, Mrs. Nirmala Patel and the entire Patel family for your support and blessings. All of you have been so kind and loving towards me. Thank you for making my stay in the UK such a memorable experience. Thanks to Dr. Apporva Bhatt and Dr. Kiran Bhatt for your advice and invaluable inputs during my PhD.

Lastly, I would like to thank Nawsheen Damree and Riddhi Shah for being fantastic housemates with whom I shared my daily experiences. Thank you, Nawsheen for being my pillar of strength during this entire journey. Thank you for always believing in me and providing me with solace in times of anxiety and pressure. Thank you, Riddhi for being a great friend, a confidante and my cooking buddy!
# Table of Contents

## CHAPTER 1: INTRODUCTION

1.1. Pre-mRNA processing .......................... 2
1.2. Pre-mRNA splicing ............................ 3
1.3. Alternative splicing ............................ 6
   1.3.1. Modes of alternative splicing .......... 8
   1.3.2. Regulation of alternative splicing .... 10
      1.3.2.1. Splice site strength regulates alternative splicing .... 11
      1.3.2.2. Splicing regulatory elements and RNA binding proteins regulate alternative splicing .... 11
   1.3.2.3. Cellular signaling regulates alternative splicing .... 12
      1.3.2.4. Tissue-specific regulatory proteins affect alternative splicing regulation .......... 13
      1.3.2.5. Co-transcriptional regulation of alternative splicing .......... 13
      1.3.2.6. Promoter type can regulate alternative splicing .......... 14
      1.3.2.7. RNA pol II processivity can regulate alternative splicing .......... 15
      1.3.2.8. Nucleosome occupancy suggests a role in alternative splicing regulation .......... 17
      1.3.2.9. Histone modifications and chromatin structure can regulate alternative splicing .......... 20
   1.3.2.10. Mechanisms of mutually exclusive splicing .... 23
1.5. The Dscam gene ................................ 26
   1.5.1. Gene organization and protein structure .... 26
   1.5.2. Dscam mutually exclusive splicing and its regulation .... 28
      1.5.2.1. RNA secondary structures mediate Dscam variable exon selection .......... 28
      1.5.2.2. Dscam variable exon clusters are in a repressed state .......... 32
      1.5.2.3. RNA binding proteins regulate inclusion of Dscam variable exons .......... 35
1.6. Biological role of Dscam in invertebrates .... 37
   1.6.1. Role of Dscam in Drosophila nervous system .... 37
      1.6.1.1. Development of Drosophila nervous system .... 37
      1.6.1.2. Dscam isoforms exhibit homophilic binding .......... 38
      1.6.1.3. Role of Dscam in axon guidance and dendritic field organization .......... 41
      1.6.1.4. Diversity of Dscam intracellular domain is important for neuronal development .......... 45
   1.6.2. Overview of the immune system .......... 47
      1.6.2.1. Insect immune system .......... 48
      1.6.2.2. Evidence for ‘trained immunity’ in invertebrates .......... 49
      1.6.2.3. Dscam splicing regulation provides ‘trained immunity’ in invertebrates .......... 51
1.7. Model for Dscam splicing regulation .......... 53
1.8. Aims .......................................... 54
1.9. Objectives .................................... 54
CHAPTER 2: MATERIALS AND METHODS

2.1 *Drosophila* husbandry and genetics

2.1.1. Fly food
2.1.2. Fly maintenance
2.1.3. Embryo collection
2.1.4. Fly transgenesis with *UAS Dscam 9L*

2.2. Molecular biology

2.2.1. RNA isolation
2.2.2. Reverse transcription (RT)
2.2.3. Polymerase chain reaction (PCR)
2.2.3.1. Single fly PCR
2.2.4. Agarose gel electrophoresis

2.3. Radioactivity used to label primers

2.3.1. Primer radiolabeling
2.3.2. Restriction enzyme digests of PCR products from

* Dscam variable regions
2.3.3. Denaturing polyacrylamide gel electrophoresis
2.3.4 Quantification of bands

2.4. Molecular cloning

2.4.1. Media preparation
2.4.2. Preparation of competent cells
2.4.2.1. Chemically competent cells
2.4.2.2. Electro-competent cells
2.4.3. Primer phosphorylation
2.4.4. Phenol-chloroform extraction of PCR products
2.4.5. Restriction enzyme digests of PCR products
2.4.6. Gel extraction of digested DNA fragments
2.4.7. Ligation
2.4.8. Bacterial transformation
2.4.8.1. Heat shock transformation
2.4.8.2. Electro-transformation
2.4.9. Recombineering
2.4.9.1. Generation of pRed/ET electro-competent cells
2.4.9.2. Generation of Dscam BAC/pRed/ET electro-competent cells
2.4.9.3. Preparation of Dscam BAC/pRed/ET host for DNA retrieval
2.4.10. Plasmid DNA mini prep
2.4.11. Plasmid DNA midi prep
2.4.12. DNA sequencing
2.4.13. Cloning of *pAc5.1A Dscam Mut Exons 4-9*

2.5. Cell culture

2.5.1. Pathogen infection of *Drosophila* haemocytes
2.5.2. Treatment of *Drosophila* haemocytes with drugs affecting transcription and RNA pol II processivity
2.5.3. Transfection of *pAc5.1A Dscam Mut Exons 4-9* in *Drosophila* haemocytes

2.6. Western Blotting

2.6.1. Sample preparation
2.6.2. Sodiumdodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) 90
2.6.3. Transfer 91
2.6.4. Blocking 91
2.6.5. Blotting 91
2.6.6. Development 92

CHAPTER 3: RESULTS 93
3.1. Dscam variable exons in clusters 4, 6 and 9 have similar sizes 93
3.2. Separation of Dscam variable exons based on sequence variation 93
3.3. Dscam splicing pattern changes on exposure to pathogens in S2 cells 101
3.4. Variation in Dscam splicing between different developmental stages, strains and sexes 106
3.5. Analysis of Dscam splicing in mutants of genes involved in small RNA and mRNA processing 115
3.6. Analysis of ago1 and rrp6 maternal mutants and redundancy between Agos and rrp6 in Dscam splicing regulation 118
3.7. Overexpression of a single Dscam isoform does not reinforce selection of the same variable exons 126
3.8. Dscam intracellular signaling is not involved in regulating its own splicing pattern 128
3.9. Analysis of mRNA methylation for a role in Dscam splicing regulation 131
3.10. Analysis of Dscam exon 9 splicing regulation in mutants of genes differentially regulated in elav mutants 134
3.11. RNA pol II processivity does not influence splicing of Dscam 138
3.12. Dscam diversity is not generated by differential stability of isoforms 140
3.13. Development of a Dscam exon 9 reporter transgene recapitulating endogenous exon 9 splicing 140

CHAPTER 4: DISCUSSION 144
4.1. Future work 148

Appendices 151
A1. Separation of Dscam 4 variants 151
A2. Separation of Dscam 9 variants 152
A3. Separation of Dscam 6 variants 153
A4. Crossing scheme showing generation of zygotic ago1 mutants 155
A5. Crossing scheme showing generation of zygotic rrp6 mutants 156
A6. Crossing scheme showing generation of zygotic elav mutants 161
A7. Crossing scheme showing generation of ago1/Df; ago2/ago2 double mutants 158
A8. Crossing scheme showing generation of ago1/Df; rrp6/Df double mutants 159
A9. Crossing scheme showing generation
of ago1 zygotic and maternal mutants 160
A10. Crossing scheme showing generation of rrp6 zygotic and maternal mutants 162
A11. Analysis of Dscam exon 9 splicing in mutants of genes that are differentially regulated in elav mutants 164
A12. List of primers 172
A13. List of fly stocks 173

List of references 175
List of figures

Figure 1: Spliceosome assembly and pre-mRNA splicing 4
Figure 2: Transesterification reactions involved in pre-mRNA splicing 7
Figure 3: Modes of alternative splicing 9
Figure 4: Mechanisms of mutually exclusive splicing 25
Figure 5: Dscam gene organization and protein structure 27
Figure 6: Complementarity between the docking site and selector sequences consensus 30
Figure 7: Proposed model for Dscam exon 4 mutually exclusive alternative splicing 31
Figure 8: Proposed model for Dscam exon 17 mutually exclusive alternative splicing 33
Figure 9: Proposed model for Dscam exon 6 mutually exclusive alternative splicing 34
Figure 10: Regulation of Dscam exon 4 mutually exclusive alternative splicing by the iStem 36
Figure 11: Molecular mechanism of homophilic interactions between identical Dscam isoforms 40
Figure 12: Role of Dscam diversity in mushroom body development 43
Figure 13: Role of Dscam diversity in dendritic field organization of dendritic arborization (da) neurons 44
Figure 14: A schematic diagram of post-scuteellar (pSc) mechanosensory neuron projections into the adult fly thoracic central nervous system (CNS) 46
Figure 15: Specification of Dscam isoform expression 55
Figure 16: Diagrammatic representation of Dscam Ex 7-11 in pUC 3GLA 2.5 HA UAS 59
Figure 17: Diagrammatic representation of the initial cloning steps to generate pAc5.1A Dscam Mut Exons 4-9 82
Figure 18: Diagrammatic representation of the final cloning step to generate pAc5.1A Dscam Mut Exons 4-9 83
Figure 19: Amplification of Dscam variable exons by RT-PCR 94
Figure 20: Diagrammatic representation showing the separation of Dscam exon 4 variants 96
Figure 21: Resolution of Dscam exon 4 cluster 97
Figure 22: Resolution of Dscam exon 9 cluster 98
Figure 23: Resolution of Dscam exon 6 cluster 100
Figure 24 (A,B): Dscam exon 4 splicing pattern changes on exposure to heat inactivated pathogens in S2 cells 103
Figure 24 (C,D): Dscam exon 4 splicing pattern does not change in naive S2 cells over a period of 12 hr 104
Figure 24 (E,F): Dscam exon 9 splicing pattern changes on exposure to heat inactivated pathogens in S2 cells 105
Figure 25: Variation in Dscam exon 4 splicing pattern between different developmental stages, strains, sexes and tissues 107-110
Figure 26: Variation in Dscam exon 9 splicing pattern between different developmental stages, strains, sexes 112-114
Figure 27: ago1 and elav show similar Dscam mutant phenotype 116
Figure 28: Analysis of Dscam exon 4 splicing in mutants of genes involved in small RNA and mRNA processing 119
Figure 29: Analysis of Dscam exon 9 splicing in mutants of genes involved in small RNA and mRNA processing 120
Figure 30: Analysis of Dscam exon 6 splicing in mutants of genes involved in small RNA and mRNA processing 121
Figure 31: Schematic representation of generating germ line clones using the dominant female sterile (DFS) technique 123
Figure 32: PCR validation showing recombined FRT sites on the same chromosome as the (A) ago1 and (B) rrp6 lethal mutation 124
Figure 33: Analysis of ago1 and rrp6 maternal mutants and redundancy between Agos and rrp6 in Dscam splicing regulation 125
Figure 34: Diagrammatic representation of distinguishing the endogenous Dscam exon 4 variants from the exogenously overexpressed Dscam single isoform, containing exon variant 4.6 127
Figure 35: Overexpression of a single Dscam isoform does not reinforce selection of the same variable exons 129
Figure 36: Dscam intracellular signaling is not involved in regulating its own splicing pattern 130
Figure 37: Analysis of mRNA methylation for a role in Dscam splicing regulation 132-133
Figure 38: Graphical representation of proportions of genes differentially regulated in elav mutants 135
Figure 39: Heat map representation of Dscam exon 9 splicing analysis in mutants of genes that are differentially regulated in elav mutants 136
Figure 40: RNA pol II processivity does not influence splicing of Dscam 139
Figure 41: Dscam diversity is not generated by differential stability of isoforms 141
Figure 42: Development of a Dscam exon 9 reporter transgene recapitulating endogenous exon 9 splicing 143
List of tables

Table 1: Denaturing polyacrylamide gel composition. 68
Table 2: List of antibiotics used during the study. 71
Table 3: Estimation of DNA amounts corresponding to size
for DNA sequencing (Beckman Coulter). 81
Table 4: List of pathogens used to challenge S2 cells. 87
Table 5: List of drugs affecting transcription and RNA pol II processivity
used during the study. 89
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ActD</td>
<td>Actinomycin D</td>
</tr>
<tr>
<td>ago1</td>
<td>Argonaute 1</td>
</tr>
<tr>
<td>ago2</td>
<td>Argonaute 2</td>
</tr>
<tr>
<td>AMPs</td>
<td>Antimicrobial peptides</td>
</tr>
<tr>
<td>attB</td>
<td>Bacterial attachment site</td>
</tr>
<tr>
<td>attP</td>
<td>Phage attachment landing site</td>
</tr>
<tr>
<td>BAC</td>
<td>Bacterial artificial chromosome</td>
</tr>
<tr>
<td>BBP</td>
<td>Branch point binding protein</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CAMs</td>
<td>Cell adhesion molecules</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CELF4</td>
<td>CUG-BP and ETR3-like factor 4</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>Ci</td>
<td>Ciuries</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>Cpm</td>
<td>Counts per minute</td>
</tr>
<tr>
<td>CPSF</td>
<td>Cleavage and polyadenylation specificity factor</td>
</tr>
<tr>
<td>CstF</td>
<td>Cleavage stimulating factor</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC-binding factor</td>
</tr>
<tr>
<td>CTD</td>
<td>Carboxy terminal domain</td>
</tr>
<tr>
<td>CUG-BP</td>
<td>CUG-repeat-binding protein</td>
</tr>
<tr>
<td>da</td>
<td>Dendritic arborization</td>
</tr>
<tr>
<td>DDAB</td>
<td>Didecyldimethylammonium bromide</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Double distilled water</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethylpyrocarbonate</td>
</tr>
<tr>
<td>DFS</td>
<td>Dominant female sterile</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxyribonucleotides</td>
</tr>
<tr>
<td>dpm</td>
<td>Disintegrations per minute</td>
</tr>
<tr>
<td>dps</td>
<td>Disintegrations per second</td>
</tr>
<tr>
<td>DRB</td>
<td>Dichlororibofuranosylbenzimidazole</td>
</tr>
<tr>
<td>DSCAM</td>
<td>Down Syndrome Cell Adhesion Molecule</td>
</tr>
<tr>
<td>DSE</td>
<td>Downstream sequence element</td>
</tr>
<tr>
<td>Dsx</td>
<td>Doublesex</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothretil</td>
</tr>
<tr>
<td>E</td>
<td>Counting efficiency</td>
</tr>
<tr>
<td>EDI</td>
<td>Extra domain I</td>
</tr>
<tr>
<td>EDII</td>
<td>Extra domain II</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ELAV</td>
<td>Embryonic lethal abnormal visual system</td>
</tr>
<tr>
<td>ESEs</td>
<td>Exonic splicing enhancers</td>
</tr>
<tr>
<td>ESSs</td>
<td>Exonic splicing silencers</td>
</tr>
<tr>
<td>ETR3</td>
<td>Elav-type RNA-binding protein 3</td>
</tr>
<tr>
<td>Ewg</td>
<td>Erect wing</td>
</tr>
<tr>
<td>FBS</td>
<td>Foetal bovine serum</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>FGFR2</td>
<td>Fibroblast growth factor receptor 2</td>
</tr>
<tr>
<td>FN III</td>
<td>Fibronectin type III</td>
</tr>
<tr>
<td>FRT</td>
<td>FLP-recombinase target</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GMCs</td>
<td>Ganglion mother cells</td>
</tr>
<tr>
<td>hnRNP s</td>
<td>Heterogeneous nuclear ribonucleoproteins</td>
</tr>
<tr>
<td>HP1α</td>
<td>Heterochromatin-associated protein 1α</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>ISEs</td>
<td>Intronic splicing enhancers</td>
</tr>
<tr>
<td>ISSs</td>
<td>Intronic splicing silencers</td>
</tr>
<tr>
<td>iStem</td>
<td>Inclusion stem</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-morpholinopropane-1-sulfonic acid</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>msl-2</td>
<td>Male-specific-lethal-2</td>
</tr>
<tr>
<td>Na-valproate</td>
<td>Sodium valproate</td>
</tr>
<tr>
<td>NM</td>
<td>Non-muscle</td>
</tr>
<tr>
<td>NMD</td>
<td>Nonsense-mediated mRNA decay</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotides</td>
</tr>
<tr>
<td>OD 600</td>
<td>Optical density at 600 nm</td>
</tr>
<tr>
<td>PAMPs</td>
<td>Pathogen-associated molecular patterns</td>
</tr>
<tr>
<td>PAP</td>
<td>Poly A polymerase</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PGNs</td>
<td>Peptidoglycans</td>
</tr>
<tr>
<td>PP1</td>
<td>Protein phosphatase 1</td>
</tr>
<tr>
<td>pre-mRNA</td>
<td>Precursor messenger RNA</td>
</tr>
<tr>
<td>proPO</td>
<td>Prophenoloxidase</td>
</tr>
<tr>
<td>PRRs</td>
<td>Pattern recognition receptors</td>
</tr>
<tr>
<td>pSc</td>
<td>Post-scutellar</td>
</tr>
<tr>
<td>PTB</td>
<td>Polypyrimidine tract-binding protein</td>
</tr>
<tr>
<td>PTC</td>
<td>Premature termination codon</td>
</tr>
<tr>
<td>(Py)n</td>
<td>Polypyrimidine tract</td>
</tr>
<tr>
<td>RBP s</td>
<td>RNA binding proteins</td>
</tr>
<tr>
<td>RNA pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>RNA i</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNAsin</td>
<td>Ribonuclease inhibitor</td>
</tr>
<tr>
<td>RNS</td>
<td>Reactive nitrogen species</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodiumdodecylsulphate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SF2</td>
<td>Splicing factor 2</td>
</tr>
<tr>
<td>siRNAs</td>
<td>Small interfering RNAs</td>
</tr>
<tr>
<td>SM</td>
<td>Smooth muscle</td>
</tr>
<tr>
<td>snRNP s</td>
<td>Small nuclear ribonucleoproteins</td>
</tr>
<tr>
<td>SOC</td>
<td>Super optimal broth</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SR</td>
<td>Serine/arginine</td>
</tr>
<tr>
<td>SREs</td>
<td>Splicing regulatory elements</td>
</tr>
<tr>
<td>SRPK1</td>
<td>Serine-arginine protein kinase 1</td>
</tr>
<tr>
<td>ss</td>
<td>Splice sites</td>
</tr>
<tr>
<td>STET</td>
<td>Sucrose-Tris-EDTA-Triton-X 100</td>
</tr>
<tr>
<td>sxl</td>
<td>Sex-lethal</td>
</tr>
<tr>
<td>T4-PNK</td>
<td>T4 polyunucleotide kinase</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate-EDTA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate-EDTA</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>TBST</td>
<td>Tris-buffered saline-Tween 20</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>tra</td>
<td>Transformer</td>
</tr>
<tr>
<td>TRAMP</td>
<td>Trf4/Air2/Mtr4p polyadenylation</td>
</tr>
<tr>
<td>U2AF</td>
<td>U2 auxiliary factor</td>
</tr>
<tr>
<td>ubx</td>
<td>Ultrabithorax</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
<tr>
<td>VSP</td>
<td>Variant-specific surface protein</td>
</tr>
<tr>
<td>WSSV</td>
<td>White spot syndrome virus</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Sections 1.4 and 1.5 have been published as ‘Mechanisms of Drosophila Dscam mutually exclusive splicing regulation’ in the Biochemical Society Transactions with myself as first author and Dr. Matthias Soller as corresponding author. Both of us planned the topics and the layout of the manuscript. I wrote the manuscript and drew the figures. Dr. Soller proofread the manuscript prior to submission. For the full article please see (Hemani and Soller, 2012).
1.1. Pre-mRNA processing

Eukaryotic genes are characterized by the presence of protein-coding sequences called exons and non-coding sequences called introns. However, not all exons are located in the coding regions. Transcription of a gene by RNA polymerase II (RNA pol II) results in the formation of a single stranded precursor messenger RNA (pre-mRNA) molecule, which undergoes many complex modifications in the nucleus to form a mature messenger RNA (mRNA). The pre-mRNA processing reactions include capping, editing, splicing and 3’ end processing. The resulting mRNA is then transported from the nucleus to the cytoplasm where it serves as a template for protein synthesis via translation (Soller, 2006).

The first pre-mRNA processing reaction is capping which involves the addition of an N$^7$-methyl GMP at the 5’ end of the pre-mRNA by an atypical 5’-5’ triphosphate linkage (McCracken et al., 1997). Nucleotides adjacent to the 5’ cap are also methylated to various degrees at the 2’-hydroxyl positions of their ribose sugars (Bisaillon and Lemay, 1997). Capping is important for protecting the mRNA from 5’-3’ exonucleases, enhancing splicing of the first intron and 3’ end processing, and for translation initiation (Flaherty et al., 1997; Fortes et al., 2000; Izaurralde et al., 1994; Lewis et al., 1996).

RNA editing is a process of altering the sequence of nucleotides within an RNA transcript of a gene by either modification or insertion/deletion of bases, to a sequence, which does not correspond to the sequence of that gene. Editing by modification of bases is largely achieved by deamination of adenosine to inosine (A to I) or cytidine to uridine (C to U) by specialized deaminases. Editing by insertion involves a guide RNA, which is partially complementary to the pre-mRNA transcript;
interrupted by additional adenosine residues. The guide RNA acts as a template for the synthesis of an edited transcript with additional uridines as observed in the mitochondria of trypanosomes. Only a few transcripts are edited in most eukaryotes (Smith et al., 1997).

Splicing is a process whereby introns, interspaced between the exons, are removed by the spliceosome, a multimeric RNA-protein complex, to produce an mRNA containing contiguous exons (Black, 2003). The mechanism of splicing and its regulation is explained in more detail from 1.2. to 1.4.

The final step of pre-mRNA processing involves cleavage and polyadenylation at the 3’ end. Conserved sequence elements, which determine 3’ processing comprise an AAUAAA hexamer, a CA dinucleotide located upstream of the cleavage site and a U- or GU-rich downstream sequence element (DSE). These sequence elements are identified by two multiprotein complexes; cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulating factor (CstF) that trigger cleavage by cleavage factors I and II (CFI and II) and polyadenylation by poly A polymerase (PAP) together with CPSF (~200 adenosines). All RNA pol II transcripts have a poly-A tail except histone RNAs (Proudfoot, 2004; Venkataraman et al., 2005; Wahle and Ruegsegger, 1999; Zhao et al., 1999).

1.2. Pre-mRNA splicing

Pre-mRNA splicing comprises a series of reactions catalyzed by a multimeric RNA-protein complex called the spliceosome (Figure 1). The spliceosome is composed of small nuclear ribonucleoprotein (snRNP) subunits, which consist of five core structural RNAs namely U1, U2, U4, U5 and U6 and over 150 accessory proteins
Figure 1: Spliceosome assembly and pre-mRNA splicing. Stepwise recruitment of snRNPs (U1, U2, U4, U5 and U6), branch point binding protein (BBP/SF1) and U2 auxiliary factor (U2AF) on an intron results in spliceosome assembly which catalyses the excision of the intron and joining of the flanking exons. Exons flanking the excised intron are shown as blue boxes. Adapted from (Zaharieva et al., 2012).
orderly binding and release of these U snRNPs and auxiliary proteins results in excision of introns and joining of adjacent exons (Staley and Guthrie, 1998). Specific sequence elements in the pre-mRNA at the exon-intron junctions called splice sites (ss) are recognized by particular snRNPs and splicing factors. The 5¢ ss (splice donor site) and 3¢ ss (splice acceptor site) consensus sequences are defined by AG-guragu and yag-N (-designates the ss), respectively, however, only the first G of the 5¢ ss and the AG of the 3¢ ss are strictly conserved (Hertel, 2008; Smith et al., 1989). Additional sequence elements that mediate RNA splicing are the polypyrrimidine tract ((Py)_n) immediately before the 3’ss and a branch point (yuryrac), containing a conserved adenine nucleotide, located upstream of the (Py)_n (Black, 2003).

The splicing reaction commences with the recruitment of U1 snRNP to the 5¢ ss, followed by the attachment of the branch point binding protein (BBP/SF1) to the branch point sequence to form the E complex. U2 auxiliary factor (U2AF), a heterodimeric splicing factor, binds to both the (Py)_n via its larger subunit (65-kDa) and the 3¢ ss via its smaller subunit (35-kDa). This binding triggers the recruitment of U2 snRNP, which binds to the branch point after the release of BBP/SF1, to form the A complex. The U4/U6.U5 heterotrimeric complex binds to the A complex to form the B complex (Soller, 2006).

At this point, a conformational rearrangement occurs within the B complex which results in the release of U1 and U4 snRNPs to eventually form the catalytic C complex. This complex catalyzes the removal of the intervening intron and joining of the adjacent exons via two transesterification steps. The first transesterification reaction occurs with the cleavage of the 5¢ end of the intron from the proximal exon.
and its consequent ligation to the 2’ hydroxyl group of the adenine in the branch point to form an unusual 2’-5’ phosphodiester bond. At this point, there are two intermediates – the detached 5’ exon and the intron-3’ exon in the form of a lariat. The second transesterification reaction occurs with the nucleophilic attack on the phosphate at the 3’ end of the intron by the 3’ hydroxyl group of the separated exon. This results in the release of the intron, in the form of a lariat, and ligation of the two exons. The U2, U5 and U6 snRNPs release from the lariat, which is eventually linearized on hydrolysis of the 2’-5’ phosphodiester bond by the lariat debranching enzyme (Black, 2003; Soller, 2006) (Figure 2). This mechanism underlies the function of the major class of spliceosomes, which removes introns flanked by conserved elements, ‘gu’ at the 5’ end and ‘ag’ at the 3’ end (Patel and Steitz, 2003).

There is however, a minor class of spliceosome, which is involved in removal of introns with ‘au’ at the 5’ end and ‘ac’ at the 3’ end. Apart from recognizing different splicing signals, the minor spliceosome has different yet functionally equivalent snRNPs for U1, U2, U4 and U6, which are respectively called U11, U12, U4atac and U6atac (Patel and Steitz, 2003).

1.3. Alternative splicing

Alternative splicing is a process by which exons in the same pre-mRNA are differentially spliced to produce more than one variety of mature mRNA generating different protein isoforms from the same gene (Soller, 2006). Humans are estimated to have 22,000 genes, only a fraction more than the 20,000 genes postulated for the simple nematode, *Caenorhabditis elegans*, whereas grapevine (*Vitis vinifera*) has
Figure 2: Transesterification reactions involved in pre-mRNA splicing. The catalytic C complex performs two transesterification reactions to result in efficient splicing of introns and subsequent ligation of flanking exons. The first transesterification step results in two reaction intermediates: the upstream exon and an intron/downstream exon fragment in a lariat conformation. The second transesterification step joins the two exons and excises the intron lariat. Exons flanking the excised intron are shown as blue boxes. Adapted from (Black, 2003).
~30,000 genes suggesting that gene number alone does not entirely define organismal complexity (Claverie, 2001; Zaharieva et al., 2012).

The variability in splicing pattern via alternative splicing is a major mechanism to generate molecular diversity (transcriptomic and proteomic) and organismal complexity from the limited number of genes present in higher eukaryotes. Alternative splicing occurs in ~25% of C. elegans genes, ~60% of Drosophila melanogaster genes and ~95% of human genes, which accounts for the disparity between the estimated 22,000 genes in the human genome and the proposed 100,000 proteins synthesized from them. Thus, organisms with more cell and tissue type complexity exhibit more alternative splicing (McManus and Graveley, 2011).

1.3.1. Modes of alternative splicing

Alternative splicing events occur in a number of different ways (Keren et al., 2010):

- **Use of cassette exons**: A cassette exon can either be included or excluded in the final mRNA transcript (Figure 3A). When the exon is excised from the pre-mRNA, the splicing event is termed exon skipping which constitutes nearly 40% of all alternative splicing events in higher eukaryotes.

- **Alternative 3’ ss selection**: Using an alternative splice acceptor site changes the 5’ end of the downstream exon (Figure 3B). Alternative 3’ ss selection accounts for 18.4% of alternative splicing events in higher eukaryotes.

- **Alternative 5’ ss selection**: Using an alternative splice donor site changes the 3’ end of the upstream exon (Figure 3C). Alternative 5’ ss selection accounts for 7.9% of all alternative splicing events in higher eukaryotes.
Figure 3: Modes of alternative splicing. (A) A cassette exon can either be included in or be excised from an mRNA. (B and C) Alternative 3’ ss or 5’ ss selection alters the 5’ or 3’ end of the downstream or upstream exon, respectively. (D) Intron retention can either result in exclusion or inclusion of an intron in the mRNA. (E) Mutually exclusive splicing of a group of exons allows the splicing of only one exon at a time. (F) Alternative poly A sites alter the 3’ most exons in an mRNA transcript. Constitutive exons are shown as blue boxes and alternative exons as pink or yellow boxes. Solid lines represent splicing options. Adapted from (Keren et al., 2010).
- **Intron retention:** The intron can either be excluded or can be retained in the mRNA transcript (Figure 3D). Intron retention is the rarest form of alternative splicing in vertebrates and invertebrates constituting less than 5% of known alternative splicing events but is the most preferred mode of alternative splicing in plants, fungi and protozoa.

- **Mutually exclusive splicing:** Two or more cassette exons are spliced in a mutually exclusive manner such that only one exon will be included from the whole variable cluster (Figure 3E).

- **Alternative terminal exons:** Frequently, genes have alternative terminal exons, which is a different form of mutually exclusive exon usage. Here, regulation can occur at the level of splicing, 3’ end processing or both. Use of alternative 3’ ss and/or alternative polyadenylation sites result in a switch in the 3’ most exon of the mRNA transcript (Figure 3F).

### 1.3.2. Regulation of alternative splicing

Recognition of an intron by the spliceosomal components and their consequent assembly around its ss determines the splicing of that intron. These reactions are under the control of combinatorial interactions between *cis*-acting elements and *trans*-acting factors. Disruption of these interactions affects alternative splicing, which has been associated with many human disease conditions such as cancer and neurodegeneration. Hence, developing a better understanding of mechanisms which regulate alternative splicing might help rectify erroneous splicing, potentially leading to novel molecular therapies (Garcia-Blanco et al., 2004; House and Lynch, 2008; Zaharieva et al., 2012).
1.3.2.1. Splice site strength regulates alternative splicing

The strength of 5’ and 3’ ss is proportional to the degree of complementarity with U1 small nuclear RNA (snRNA) and the extent of polypyrimidine tract, respectively. Accordingly, an intron having stronger ss is removed more readily than an intron with weaker ss. Conversely, an alternative cassette exon flanked by stronger ss is included more frequently than an adjacent exon flanked by weaker ss (Hertel, 2008).

1.3.2.2. Splicing regulatory elements and RNA binding proteins regulate alternative splicing

Auxiliary sequences called splicing regulatory elements (SREs) regulate recognition and consequent splicing of nearby ss (Voelker et al., 2012). These cis-acting elements occur within both exonic and intronic regions and bind directly or indirectly to trans-acting RNA binding proteins (RBPs) that function as splicing activators or repressors (House and Lynch, 2008).

SREs are divided into four categories based on location and function: exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs), exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs). ESEs characteristically bind members of the serine/arginine-rich (SR) family of splicing factors and promote spliceosome assembly and exon inclusion (Garcia-Blanco et al., 2004). Sex-specific inclusion of Drosophila doublesex (dsx) exon 4 is mediated by the presence of an ESE in the exon, which binds RBP1 (SR protein) and transformer (Tra, an RNA binding protein), amongst other proteins, promoting spliceosome assembly and exon inclusion in females (Demir and Dickson, 2005; Manoli et al., 2005; Usui-Aoki et al., 2000). ESSs, however, typically bind heterogeneous nuclear ribonucleoproteins (hnRNPs)
and block spliceosome assembly and suppress exon inclusion (Garcia-Blanco et al., 2004). Variable exon 4 of the human CD45 gene contains an ESS, which binds to hnRNP L and results in exon skipping (Rothrock et al., 2005). SR proteins (splicing activators) and hnRNPs (splicing repressors) function antagonistically but exist in a balance and a change in concentration and/or activity of one factor affects the splicing outcome (Long and Caceres, 2009). SR proteins and hnRNPs are functionally flexible and in some cases are also associated with exon skipping and inclusion, respectively (Han et al., 2011; Hofmann and Wirth, 2002; Konig et al., 2011). ISEs and ISSs function similarly to their exonic counterparts and are known to regulate splicing of exon 8 splicing in human fibroblast growth factor receptor 2 (FGFR2) pre-mRNA (Black, 2003; Wagner and Garcia-Blanco, 2002).

1.3.2.3. Cellular signaling regulates alternative splicing

Post-translational modifications such as methylation or phosphorylation in RBPs are important for determining RNA binding specificities. The most abundant class of phosphorylated RBPs is the one that includes SR proteins. SR proteins are active upon phosphorylation, which is important for alternative splicing regulation. Global shut down of splicing during mitosis or heat shock due to dephosphorylation of SRSF10 exemplifies such a splicing regulation where 5’ ss recognition and subsequent spliceosome assembly is hampered (Shin et al., 2004). Inhibition of SRPK1 (serine-arginine protein kinase 1) causes hypophosphorylation of SF2 (splicing factor 2) which regulates alternative splicing of VEGF (vascular endothelial growth factor) from pro- to anti-angiogenic spliceforms beneficial in inhibiting progression of diabetic nephropathy and tumour growth (Oltean et al., 2012). C6
pyridinium ceramide, a potential anti-cancer drug, inhibits dephosphorylation of various splicing regulatory proteins (including, SF2/ASF and Tra2-beta1) by binding to the protein phosphatase 1 (PP1) catalytic subunit thereby changing alternative splicing patterns of several endogenous genes such as drf1, tau and syk (Sumanasekera et al., 2012).

1.3.2.4. Tissue-specific regulatory proteins affect alternative splicing regulation
Regulation of alternative splicing can also be mediated by tissue-specific factors, which are present in one cell type but not in the other. Embryonic lethal abnormal visual system (ELAV), a neuron-specific RBP binds to and regulates alternative splicing of Drosophila erect wing (ewg) and neuroglian (nrg) genes (Lisbin et al., 2001; Soller and White, 2003). Sex-lethal (Sxl), an RBP expressed only in females, is a key regulator of the Drosophila sex determination pathway. Sxl regulates the splicing of Drosophila tra exon 2 and autoregulates its own exon 3 splicing. In both cases, binding of Sxl blocks 3' ss recognition and suppresses exon inclusion to produce active Tra and Sxl proteins in females but not in males, where exon inclusion produces truncated and inactive proteins. Sxl also mediates X chromosome dosage compensation by blocking splicing of the first intron in Drosophila male-specific-lethal-2 (msl-2), and also translation of msl-2 to prevent transcriptional upregulation on the X chromosome as observed in males (Black, 2003; Lopez, 1998; Maniatis and Tasic, 2002; Schutt and Nothiger, 2000).

1.3.2.5. Co-transcriptional regulation of alternative splicing
Regulation of alternative splicing not only depends on the interaction of splicing
factors with their target pre-mRNA sequences but also transcription. Transcription and pre-mRNA processing were thought to be independent processes. However, studies in the Drosophila chorion genes (s36-1 and s38-1) showed that transcripts appear to be shorter while still being attached to the chromatin via RNA pol II suggesting co-transcriptional pre-mRNA splicing (Osheim et al., 1985). Moreover, nascent RNA sequencing data in Drosophila revealed that eighty-seven percent of the introns manifest >50% co-transcriptional splicing (Khodor et al., 2011). Key to coupling of transcription and splicing is the carboxy terminal domain (CTD) of RNA pol II which comprises 52 tandem repeats of the 7 amino acid consensus sequence YSPTSPS and is subject to post-translational phosphorylation of serine residues, which affects co-transcriptional alternative splicing by recruiting splicing factors. CTD promotes skipping of human fibronectin extra domain I (EDI) exon, independent of RNA elongation rate, by recruiting SR protein SRp20 demonstrating that alternative splicing regulation does not solely depend on regulatory element-specific trans-acting factors (de la Mata and Kornblihtt, 2006).

1.3.2.6. Promoter type can regulate alternative splicing

Different RNA pol II promoters alter splicing outcome of the same exon provided very strong evidence towards coupling of transcription and alternative splicing. Transcription under the α-globin promoter results in a ten-fold reduction in the inclusion of human fibronectin EDI exon than when transcription is driven by the cytomegalovirus promoters suggesting that splicing can be influenced by RNA pol II transcription (Cramer et al., 1997). A similar regulatory function of promoters has also been observed in splicing of human CD44 and cystic fibrosis transmembrane
regulator genes (Auboeuf et al., 2002; Pagani et al., 2003). These effects are not a result of differences in promoter strength but due to qualitative properties attributed to the RNA processing machinery. This observation is in line with microarray data showing that transcription based on promoter strength acts independently of alternative splicing on different sets of genes to result in tissue-specific expression profiles (Pan et al., 2004).

Promoter type affecting splicing outcome could function as a result of differences in promoter occupancy by factors varying in activation moieties and mechanistic properties (Kornblihtt, 2005). A possible mechanism that could explain such an effect is that the promoter itself recruits factors with functional domains for both transcription and splicing to the transcription site, via transcription factors that physically associate with the promoter or transcriptional enhancer elements within the promoter. The splicing of human fibronectin extra domain II (EDII) exon exemplifies such a mechanism where the promoter with a DR-1 transcriptional enhancer element binds transcription factor PPARγ, which subsequently recruits the transcriptional co-activator PGC-1. Due to its dual functionality, PGC-1 interacts with RNA pol II and the splicing factor SRp40, which results in EDII exon skipping. Conversely, when PGC-1 is not a part of a promoter-binding complex, EDII exon is included (Monsalve et al., 2000).

1.3.2.7. RNA pol II processivity can regulate alternative splicing

RNA pol II elongation rates or processivity can regulate alternative splicing outcomes. A putative model suggests that for an alternative exon, flanked by a weaker upstream and a stronger downstream ss, low RNA pol II elongation rates or
internal stalling sites, commonly called pause sites, would favour exon inclusion, whereas a high RNA pol II elongation rate or absence of pause sites would result in exon skipping (Nogues et al., 2003). A lower elongation rate between the two ss would allow more time for spliceosomal components to assemble on the weaker ss and delay the synthesis of the stronger downstream ss, thus favouring excision of the upstream intron and exon inclusion. A higher elongation rate would expose both the ss simultaneously before the splicing machinery, which would choose the stronger 3’ ss resulting in exon skipping. In the case of two constitutive strong ss, as observed in constitutive splicing, RNA pol II elongation rates do not affect splicing (Kornblihtt et al., 2004).

Altering RNA pol II elongation rates influences alternative splicing of human fibronectin EDI exon. Cells treated with dichlororibofuranosylbenzimidazole (DRB), a potent inhibitor of RNA pol II elongation showed a three-fold elevation in EDI inclusion (Nogues et al., 2002; Price, 2000). Also, the C4 RNA pol II mutant, which has a lower elongation rate due to a single amino acid substitution (741Arg-His) in its largest subunit, provided a direct evidence for the kinetic coupling of transcription and splicing. This slow polymerase was shown to cause a three-fold increase in fibronectin EDI exon inclusion in human cells. This mutation also affected the splicing of adenovirus E1a by favouring the inclusion of the most upstream exon out of the three alternative 5’ donor sites (de la Mata et al., 2003).

The Drosophila Hox gene Ultrabithorax (ubx), responsible for wing and leg formation in adult flies, can give rise to six different spliceforms by differential inclusion of its three variable exons: the B element and two microexons, mI and mII that are separated by very large introns. Joining of either mI or mII to the 5’ constitutive exon
regenerates a consensus 5’ ss which is subsequently recognized by the splicing machinery to result in exon skipping and shorter isoforms by a mechanism of recursive splicing (Burnette et al., 2005).

In C4 embryos, Ubx pre-mRNA splicing shows an enhancement in shorter isoforms, which translates as a mutant phenotype termed ‘Ubx effect’. This increase in recursive splicing is attributed to the lowered RNA pol II elongation rate which allows more time for the newly formed 5’ ss to be recognized by the spliceosome before the downstream 5’ ss can be synthesized. Contrary to EDI and E1a, ubx shows exon skipping under the influence of a slow polymerase. Splicing of other alternative exons, such as exon 7B of the Drosophila hnRNPA1 gene is not affected by the slow polymerase (de la Mata et al., 2003). Inhibition of RNA pol II elongation in Jurkat T cells revealed that RNA pol II occupancy was enhanced on introns flanking variable exons that became more included. This observation, at a genome-wide level, supports the RNA pol II elongation model (Ip et al., 2011).

1.3.2.8. Nucleosome occupancy suggests a role in alternative splicing regulation

DNA is packaged together with an octamer of highly evolutionarily conserved proteins called histones (two each of histones H2A, H2B, H3 and H4) to form the fundamental repeating unit of chromatin called nucleosome. Early sequence data from human and mouse genes suggested a relationship between nucleosome positioning and exon-intron architecture based on regular distribution of ss. Since splicing and transcription were regarded as independent processes then, no conclusion was derived from this study (Beckmann and Trifonov, 1991).
Deep-sequencing data from human and *C. elegans* genomes and computationally based predictions of nucleosome positioning in humans, *D. melanogaster* and *C. elegans* reveal that sequences favouring nucleosome positioning are located in exons and disfavouring sequences are a part of 50 nt intronic regions immediately before and after exons as if nucleosome-free regions delimit exon-intron boundaries (Schwartz et al., 2009). Studies in *Arabidopsis thaliana* and *Oryzias latipes* (Japanese killifish) have shown the enrichment of nucleosomes on internal exons (Andersson et al., 2009; Chodavarapu et al., 2010; Nahkuri et al., 2009). A nucleosome accommodates 147 bp of DNA around it, which interestingly is very similar to the average length of exons in higher eukaryotes (Venter et al., 2001). These findings might explain the mean exon length of 140-150 nt as it would facilitate wrapping around a nucleosome and possibly improve exon recognition (Beckmann and Trifonov, 1991). Nucleosomes have also been reported to cover ss junctions, which may protect them against mutations thereby facilitating efficient splicing (Hapala and Trifonov, 2011; Kogan and Trifonov, 2005).

Exons are known to contain a slightly higher GC content than introns. In addition, nucleosomes preferentially occupy GC-rich sequences. However, intronic regions with a much higher GC content than exons show lower levels of nucleosome occupancy and true exons show higher levels of nucleosome positioning, despite having the same level of GC content as pseudo-exons; intronic sequences flanked by strong ss but not included in the mRNA. These findings eliminate a possible bias of nucleosomes towards GC-rich sequences (Tilgner et al., 2009). Tilgner et al., (2009) also showed that nucleosomes are particularly enriched over exons flanked by weak ss. In contrast, exons with strong ss show reduced nucleosome occupancy but have
an extended region of nucleosomes upstream of the acceptor site. This inverse correlation between nucleosome occupancy and ss strength suggests an interplay between chromatin architecture and ss choice during pre-mRNA splicing (Tilgner et al., 2009). Schwartz et al. suggested a positive correlation between nucleosome occupancy and exon inclusion when analyzing splicing of three categories of exons: alternative exons with less than 50% inclusion, those with more than 50% inclusion and constitutive exons (Schwartz et al., 2009). In parallel, exon inclusion is disfavoured in pseudo-exons despite the presence of strong ss, possibly due to low nucleosome occupancy (Tilgner et al., 2009).

The link between chromatin architecture and ss choice is the CTD – a key component of transcription and splicing. *In vivo* data shows enhanced occupancy of RNA pol II over exons compared to introns (Brodsky et al., 2005; Dhami et al., 2010; Spies et al., 2009). These observations suggest a cross-talk between nucleosome positioning, RNA pol II occupancy and general/alternative splicing regulation. Nucleosomes positioned on alternative exons might act as natural hurdles or ‘speed-bumps’ for the transcribing RNA pol II and lower its elongation rate. This might delay the synthesis of competing downstream ss and consequently allow more time for spliceosome assembly and exon recognition eventually promoting exon inclusion (Churchman and Weissman, 2011; Hodges et al., 2009). In contrast, depletion of nucleosomes from alternative exons would result in a faster RNA pol II and hence, faster synthesis and recognition of the stronger downstream competing ss by the nucleosome resulting in exclusion of alternative exons (Carrillo Oesterreich et al., 2011).
1.3.2.9. Histone modifications and chromatin structure can regulate alternative splicing

Histones are post-translationally modified at many positions, particularly in the N-terminal tails that extend out of the nucleosomal core, by enzymes that write a ‘histone code’ by either adding or removing chemical groups. These histone marks primarily include methylation, acetylation and phosphorylation. A change in the degree of these modifications has been shown to regulate alternative splicing events (Hnilicova and Stanek, 2011).

Enrichment of specific histone modifications on nucleosomes helps distinguish exons from introns to facilitate efficient splicing. In order to analyse chromatin structure and function in C. elegans, a genome-wide map of histone H3 tail methylations, using chromatin immunoprecipitation (ChIP) -on-chip and ChIP sequence (ChIP-Seq) data, have shown that exons are preferentially marked with H3K36me3 relative to introns. Moreover, constitutive exons have higher H3K36me3 signal than alternative exons. Other modifications such as H3K4me3 and H3K9me3 do not exhibit this difference in levels (Kolasinska-Zwierz et al., 2009). H3K36me3 occurrence on alternative exons correlates with exon inclusion on a genome wide scale in humans suggesting interplay between H3K36me3 marking and alternative splicing regulation (Hon et al., 2009). Other histone methylations such as H3K27me1, H3K27me2, H4K20me1, H3K79me1 and H2BK5me1 have also been reported to favourably mark exons suggesting a possible role of chromatin structure in exon definition and regulation of alternative splicing (Andersson et al., 2009; Dhami et al., 2010; Huff et al., 2010; Spies et al., 2009).
RNA pol II serves as a link between the transcribed pre-mRNA and chromatin structure and promotes cross-talk between them such that chromatin remodeling can alter RNA elongation rate and consequently regulate alternative splicing. Membrane depolarization causes hyperacetylation of H3K9 and relaxation of chromatin around the mouse NCAM alternative exon 18 favouring exon skipping (Schor et al., 2009). Conversely, inhibiting histone deacetylases with trichostatin A results in enhanced skipping of exon 18. Moreover, a slow RNA pol II mutant promotes exon 18 inclusion, suggesting the role of histone acetylation in increasing RNA pol II elongation rate and regulating alternative splicing in mouse NCAM. Similar effects are observed in HeLa cells where inhibiting deacetylase activity results in skipping of human fibronectin EDII exon (Hnilicova et al., 2011).

On the contrary, small interfering RNAs (siRNAs) complimentary binding an intron, downstream of the alternative human fibronectin EDI exon, enhances local H3K9me2 and H3K27me3 marks on the exon, which promotes heterochromatin formation via recruitment of heterochromatin-associated protein 1α (HP1α), reduces RNA pol II elongation rate and as a consequence enhances inclusion of EDI exon (Allo et al., 2009). These examples suggest a correlation between chromatin structure and alternative splicing regulation where histone acetylation results in more open chromatin architecture around an alternative exon, allowing faster RNA pol II processivity and consequential skipping of the alternative exon. Histone methylation, on the other hand, results in a more compact chromatin structure, which slows down RNA pol II elongation rate and hence allows more time for spliceosome assembly around the alternative exon resulting in an increase in its inclusion.
Specific histone modifications are suggested to recruit certain splicing factors, which mediate alternative splicing regulation. Indeed, MRG15, an adaptor protein, recognizes and binds to H3K36me3 modification on human FGFR2 exon IIIb and recruits the polypyrimidine tract-binding protein (PTB), a splicing repressor. PTB binds to its intronic splicing silencer sites and causes exon skipping. Conversely, knockdown of MRG15 and reduction in H3K36me3 results in enhanced exon inclusion (Luco et al., 2011).

In human CD44, H3K9 trimethylation on variant exons recruits more HP1γ, a transcriptional repressor, in the phosphorylated form (HP1γS83p), which binds to the CD44 mRNA and slows down RNA pol II. A decrease in elongation rate in turn recruits splicing factors such as U2AF65 and PRP8 that facilitate inclusion of CD44 variant exons. Apart from regulating CD44 splicing, HP1γ affects alternative splicing of a few other genes such as PNK2, TAF4B, GLS, BRCA1 and DSN (Saint-Andre et al., 2011).

Chromatin and RNA binding factors, independent of recruitment by histone modifications, regulate alternative splicing of variable exons. CCCTC-binding factor (CTCF) binds to alternative human CD45 exon 5 and promotes inclusion by local pausing of RNA pol II at the variable exon. DNA methylation on exon 5, however, inhibits CTCF binding and favours exon skipping (Shukla et al., 2011). Splicing regulators ELAV/Hu proteins, bind to RNA pol II and their target pre-mRNA sequences around alternative exons of mouse FAS1 and Nf1 genes and stimulate local histone hyperacetylation. This modification increases RNA pol II elongation rate which promotes variant exon skipping. Moreover, Hu proteins also directly bind and inhibit histone deacetylase 2 activity, which favours local histone acetylation. These
observations suggest a ‘reach back’ mechanism wherein splicing regulators remodel chromatin structure when recruited to their target pre-RNA sequences cotranscriptionally to ensure efficient regulation of alternative splicing (Zhou et al., 2011).

1.4. Mechanisms of mutually exclusive splicing

Mutually exclusive splicing results in inclusion of only one exon from two or more variable exons (Keren et al., 2010). Mechanisms explaining mutually exclusive splicing involve steric hindrance due to overlapping signals required for splicing, incompatibility of splicing signals for the major and the minor spliceosome, regulation by trans-acting factors and removal of transcripts with premature stop codons by nonsense-mediated mRNA decay (NMD).

In mammalian introns, a minimum of 50 nt are required between the 5’ ss and the branch point for splicing. A shorter distance will not allow spliceosome assembly due to steric hindrance. Between exons 2 and 3 of human α-tropomyosin gene, the distance between the 5’ ss and the branch point is only 41nt resulting in inclusion of either exon 2 or 3 (Smith, 2005; Smith and Nadal-Ginard, 1989) (Figure 4A).

Although most introns are spliced by the major spliceosome using GU/AG splicing signals, a minor spliceosome is present where U1 and U2 snRNPs are substituted by U11 and U12 snRNPs, which use AU/AC as splicing signals. The ss for major and minor spliceosomes are incompatible and cannot be spliced together to include adjacent exons. The intron between mutually exclusive exons 6a and 6b in human JNK2 contains a U12-type 5’ ss and a U2-type 3’ ss and accordingly these exons cannot be spliced together (Chang et al., 2007) (Figure 4B).
In some exon pairs, tissue-specific expression or activity of trans-acting factors can favour inclusion or force exclusion of only one from two variable exons, which under certain conditions could also be spliced together. Such a situation is found in the rat α-actinin1 gene where the smooth muscle (SM) exon is preferentially included and the non-muscle (NM) exon is preferentially excluded in smooth muscles. Here, CUG-repeat-binding protein (CUG-BP) and elav-type RNA-binding protein 3 (ETR3) activate the SM exon but together with CUG-BP and ETR3-like factor 4 (CELF4) they also suppress the NM exon. Inclusion of the SM exon also involves removal of PTB to relieve repression from the SM exon indicating that combinatorial interaction of multiple RBPs is required to regulate this complex alternative splicing event (Gromak et al., 2003) (Figure 4C).

In some genes, alternative splicing generates isoforms, which contain a premature termination codon (PTC) due to a frame shift. Such PTC containing isoforms can be efficiently removed by NMD thus disguising the absence of an isoform as mutually exclusive splicing. This situation is observed in mammalian FGFR2 gene, where transcripts containing both mutually exclusive exons IIIb and IIIc, harbor a PTC and are therefore degraded (Jones et al., 2001) (Figure 4D).
Figure 4: Mechanisms of mutually exclusive splicing. (A) Steric hindrance in spliceosome assembly is imposed by insufficient distance between 5' ss and the branch point resulting in mutually exclusive splicing of adjacent exons 2 and 3 in the α-tropomyosin gene. (B) The GU/AG splicing signals used by major spliceosomal U1/U2 snRNPs are incompatible with the AU/AC splicing signals used by minor spliceosomal U11/U12 snRNPs. An intron with a U12-type 5' ss and a U2-type 3' ss cannot be spliced due to spliceosomal incompatibility as observed between exons 6a and 6b of JNK2 gene. (C) Trans-acting factors can regulate mutually exclusive splicing by acting as both promoters and inhibitors of mutually exclusive exon pairs to prevent them from being spliced together. CELF proteins regulate splicing of SM and NM exons of α-actinin gene resulting in mutually exclusive splicing. Disfavoured, but observed splicing is shown as dotted lines. (D) Inclusion of multiple mutually exclusive exons can result in generation of a PTC due to a frame shift. Such aberrant transcripts are removed by NMD as observed in FGFR2 transcripts containing both IIIb and IIIC exons. Constitutive exons are shown as blue boxes, and pairs of mutually exclusive exons are shown as pink and yellow boxes. Adapted from (Hemani and Soller, 2012).
1.5. The *Dscam* gene

*Dscam*, a *Drosophila* homolog of human Down Syndrome Cell Adhesion Molecule (DSCAM) was first isolated by its affinity to Dock, an SH3/SH2 adaptor protein required for axon guidance (Schmucker et al., 2000). *Drosophila Dscam* is expressed in the developing nervous system and in immune cells. *Dscam* is important for its role in wiring of the nervous system and pattern recognition in the immune system (Hattori et al., 2007; Watson et al., 2005).

1.5.1. Gene organization and protein structure

*Dscam* structure resembles that of an immunoglobulin (Ig) receptor and is related to other cell adhesion molecules (CAMs). The Ig superfamily is known to be the largest group of related cell surface receptors (Harpaz and Chothia, 1994; Williams and Barclay, 1988). They contain ten Ig domains and six fibronectin type III (FNIII) domains. Nine out of the ten Ig domains are arranged in tandem whereas the tenth domain is located between FN4 and FN5. It also contains a transmembrane domain and a unique cytoplasmic domain with multiple tyrosine phosphorylation sites (Schmucker et al., 2000; Yamakawa et al., 1998) (Figure 5).

The *Dscam* gene comprises 115 exons of which 95 exons are alternatively spliced. These 95 variable exons are organized into four exon clusters namely exon 4 (encoding the N-terminal half of Ig2 domain), exon 6 (encoding the N-terminal half of Ig3 domain), exon 9 (encoding the entire Ig7 domain) and exon 17 (encoding transmembrane domains) clusters. Alternative splicing of individual exons takes place in a mutually exclusive and combinatorial manner resulting in tens of thousands of *Dscam* isoforms. Each of the exon clusters can generate 12, 48, 33 and
Figure 5: Dscam gene organization and protein structure. The Dscam gene comprises four exon clusters spliced in a mutually exclusive manner generating a repertoire of up to 38,016 isoforms. Variable exon clusters are shown in colour: exon 4 cluster (red, 12 variables), exon 6 cluster (blue, 48 variables), exon 9 cluster (green, 33 variables), exon 17 cluster (yellow, 2 variables). The Dscam protein comprises ten Ig domains and six FN type III domains. Nine out of the ten Ig domains are arranged in tandem and the tenth Ig domain is present between FN4 and FN5 domains. Exons 4 and 6, code for the N-terminal halves of Ig2 and Ig3, respectively. Exons 9 and 17, code for the entire Ig7 and transmembrane domain, respectively. Adapted from (Wojtowicz et al., 2004).
two alternative forms, respectively. If all combinations were allowed, 38,016 different isoforms could be generated by inclusion of a single exon in each of the variable regions (Schmucker et al., 2000).

1.5.2. Dscam mutually exclusive splicing and its regulation

Alternative splicing of Dscam is extraordinary considering not only the resultant protein diversity but also its unique mechanism. Each variable exon cluster can give rise to multiple alternative exons, however; only one variant is included in the functional Dscam isoform.

The mechanisms of mutually exclusive splicing discussed previously do not seem to apply to splicing of the Dscam variable exons comprising the extracellular domain. Introns in the variable clusters are longer than the minimal size of about 59 nt in Drosophila and contain splicing signals for the major spliceosome (Graveley, 2005). Although splicing together of exons from the exon 6 cluster would result in a premature termination codon, removal of such isoforms by nonsense mediated mRNA decay would seem a highly inefficient way to remove isoforms with multiple inclusions due to the large number of possibilities. Hence, a mechanism is postulated whereby RNA secondary structure mediates the inclusion of a specific exon by relieving the variable cluster from its default repressed state (Graveley, 2005; Olson et al., 2007).

1.5.2.1. RNA secondary structures mediate Dscam variable exon selection

Extensive sequence analysis and phylogenomics have initially identified evolutionary conserved intronic sequences in the exon 6 cluster consisting of a 66 nt sequence in
the intron after exon 5 termed the “docking site” and a partially complementary shorter sequence called “selector sequence” in front of each exon 6 variant (Figure 6). Base pairing between the docking site and the selector site is then postulated to bring a specific variable exon into the proximity of the preceding constant exon resulting in splicing of the intron by the release of a repressor (e.g. hrp36 for exon 6, explained in 1.5.2.2.), subsequently triggering splicing to the distal constant exon. Since the selector sequences overlap with one another to a certain extent, the docking site is thought to interact with only one selector sequence at a time (Graveley, 2005).

The function of the docking site is evolutionary conserved and is required for exon 6 splicing. Deletions in the docking site increase exon 6 skipping and alter inclusion frequencies of exon 6 variables. Also, swapping selector sequences from rarely to frequently encoded variable exons, and vice versa, alters inclusion accordingly (May et al., 2011). Further support for the involvement of RNA secondary structure has recently come from sequence analysis of the variable clusters 4 and 9, where the docking site is present in the intron between the entire variable cluster and the downstream constitutive exon and the selector sequences are located in the introns after each exon variant (Yang et al., 2011) (Figure 7).

Most introns in *Drosophila* are spliced co-transcriptionally (Khodor et al., 2011). Although the variable exon 4, 6 and 9 clusters extend over 6, 12 and 15 kb respectively, a “polar effect” leading to preferential inclusion of variable exons closer to the proximal constant exon is not observed arguing for dedicated mechanisms regulating the inclusion of a single variable exon.
Figure 6: Complementarity between the docking site and selector sequences consensus. The most frequently occurring nucleotide at each position of the central 28 nucleotides of the selector sequence consensus is complementary to the docking site consensus. Adapted from (Graveley, 2005).
Figure 7: Proposed model for Dscam exon 4 mutually exclusive alternative splicing. Once an exon is chosen for inclusion, an RNA stem loop is formed by complimentary base pairing between the docking site (green box) and the selector sequence (red box) downstream of the chosen exon. For simplicity, only five exon 4 variables are shown. Constant exons are shown in blue and variable exons in a gradient from pink to yellow. Adapted from (Yang et al., 2011).
RNA secondary structure is also postulated to govern mutually exclusive splicing of the transmembrane domain encoded by the two exon 17 variables (Anastassiou et al., 2006; McManus and Graveley, 2011). Here, four conserved sequences that can form two competing RNA stem-loop structures are present in the intron distal of constitutive exon 16 (Figure 8). Mutually exclusive splicing of variable exon 17 is functionally important to localize Dscam isoforms either to dendrites (exon 17.1) or axons (exon 17.2) of mushroom bodies (Wang et al., 2004).

1.5.2.2. Dscam variable exon clusters are in a repressed state

Adjacent exons in each variable cluster have ss close to the consensus sequence; however they are not spliced together. Through RNA interference (RNAi) mediated knockdown of RBPs in Drosophila S2 cells (cell-line of haematopoietic origin), an hnRNP called hrp36 has been identified to specifically repress the joining of multiple exons throughout the exon 6 cluster but not the other clusters (Figure 9). An RT-PCR flanking the exon 6 cluster in cells depleted of hrp36 show multiple bands in contrast to a single band in the presence of hrp36. Although hrp36 also binds to the exon 4 and 17 cluster, no joining of multiple exons is observed when RNAi reduces hrp36 levels. Thus, hrp36 regulates mutually exclusive splicing of exon 6 in a cluster-specific and cluster-wide manner.

Other hnRNPs such as hrp38, hrp40 and hrp48 do not regulate splicing fidelity of exon 6. Moreover, hrp36 acts antagonistically to SR proteins to prevent ectopic inclusion of multiple exon variants. These observations suggest a mechanism, which keeps the variable exon cluster in a repressed state until one exon is chosen for inclusion. Though very compelling, the role of hrp36 in repressing splicing together of
Figure 8: Proposed model for Dscam exon 17 mutually exclusive alternative splicing. Four conserved intronic sequence elements are located upstream of Dscam exon 17 cluster, which have the capacity to form RNA secondary structures by complimentary base pairing. These stem-loop structures allow the inclusion of only one of the two exon 17 variable exons. Adapted from (McManus and Graveley, 2011).
Figure 9: Proposed model for Dscam exon 6 mutually exclusive alternative splicing. Once an exon is chosen for inclusion, an RNA stem loop is formed by complimentary base pairing between the docking site and the selector sequence preceding the chosen exon, the repressor hrp36 is released to allow for its inclusion. Only the most frequently occurring nucleotides of the core consensus docking site (CUG) and selector sequences (CAG) are shown. For simplicity, only five exon 6 variables are shown. Constant exons are shown in blue and variable exons in a gradient from pink to yellow. Adapted from (Hemani and Soller, 2012).
multiple exons in the exon 6 cluster seems not to extrapolate to the whole organism. No splicing repressor has been associated with other variable exon clusters (Olson et al., 2007). Recent work in shrimp (Litopenaeus vannamei) demonstrated the role of Lvhrp36 as a splicing repressor of Ig2, Ig3 and Ig7 variable regions (Lee et al., 2012). The repression of the exon 4 cluster is governed by an RNA secondary structure, wherein a 27 nt stem with a 2 nt bulge and a 275 nt loop structure, called inclusion stem (iStem), located in the intron 18 nt after exon 3 is functionally relevant for inclusion of one variable exon (Figure 10). Deletion or mutation of the double stranded part of the iStem results in skipping of the entire exon 4 cluster (in cell culture) but the length and sequence of the loop in the iStem seem not to play a role in regulating splicing in the exon 4 cluster. It is also unknown how this loop interacts with the spliceosome. The iStem plays a role in repression but is not involved in exon choice. Genome sequence comparison showed that the iStem is evolutionarily conserved in all eleven Drosophila species known. This study suggests that a fine balance exists between choosing a variant exon and repression of the entire exon 4 cluster (Kreahling and Graveley, 2005).

1.5.2.3. RNA binding proteins regulate inclusion of Dscam variable exons

Regulation of Dscam diversity is required to provide unique identities to individual neurons in mushroom bodies and to increase levels of high affinity binding isoforms that fight against pathogens in the immune system (Dong et al., 2006; Zhan et al., 2004). Prime candidates to regulate inclusion of individual variants are RBPs. RNAi knock down of 250 RBPs in S2 cells reveals 36 proteins that have an effect on the splicing of exon 4 cluster, but only eight proteins that affect alternative splicing of
Figure 10: Regulation of Dscam exon 4 mutually exclusive alternative splicing by the iStem. The iStem allows selection of only one of the 12 exon 4 variables. Deletion of the iStem results in complete skipping of the whole exon cluster. For simplicity, only four exon 4 variables are shown. Constant exons are shown in blue and variable exons in a gradient from pink to yellow. Adapted from (Kreahling and Graveley, 2005).
exon 17 cluster. All eight proteins involved in splicing of exon 17 cluster also affect splicing of exon 4 cluster (Park et al., 2004).

These results suggest combinatorial control in the regulation of Dscam mutually exclusive variable exon splicing by RBPs, but surprisingly those RBPs identified mostly affect inclusion of the same small set of variable exons implying that a large proportion of Dscam diversity is unutilized. Currently, it is not clear how specific exons are included over others to establish neuronal patterning and how a few exons are included more frequently on pathogen exposure.

1.6. Biological role of Dscam in invertebrates

1.6.1. Role of Dscam in Drosophila nervous system

1.6.1.1. Development of Drosophila nervous system

The insect central nervous system (CNS) comprises the ventral nerve cord and the brain. After gastrulation, neurogenic regions of the ectoderm are determined to form the neuroectoderm, a sheet of undifferentiated epithelial cells. The ventral neurogenic region gives rise to the ventral nerve cord and the procephalic neurogenic region gives rise to the brain. In each hemisegment, a few selected cells delaminate from the neuroectoderm as CNS progenitor stem cells, also called neuroblasts that divide asymmetrically by restoring themselves and producing a chain of secondary precursor cells called ganglion mother cells (GMCs). A single GMC divides to give rise to two progeny cells that differentiate into neurons and/or glia. Each hemisegment has ~30 neuroblasts that generate a total of 350 embryonic progeny cells which include ~290 interneurons, 30 motorneurons and 30 glial cells (Technau et al., 2006; Urbach and Technau, 2004). Subsequently, the axons and
dendrites branches from these neurons must find their target cells via axon pathfinding and dendritic field organization, respectively (Kaprielian et al., 2000; Nichols, 2006). The adult fly brain comprises highly organized neuropile structures such as mushroom bodies, optic lobes, antennal lobes and the central complex. The ventral nerve cord is relatively simpler, composed of a series of uniform segments called neuromeres (8 abdominal, 3 thoracic and 3 gnathal) (Technau et al., 2006).

Neural patterning in the nervous system comprises many different neurons that are interconnected by a complex network of synaptic connections. Precise assembly of these neural circuits depends on interactions between cell surface recognition proteins expressed on projecting neurons and specific guidance signals from target neurons (Zipursky et al., 2006). In Drosophila there are ~250,000 different neurons and only ~15,000 genes which clearly suggest that each cell recognition event cannot be governed by a different gene product. Despite the limited number of genes, several mechanisms facilitate complex neuronal wiring, including combinatorial use of guidance cues such as attractants, repellants or modifiers, diffusible gradient labels and coordinated firing activity between neurons (Schmucker and Flanagan, 2004). Alternative splicing of Drosophila Dscam serves as a mechanism to generate thousands of neuronal cell surface proteins with distinct recognition properties from a single gene, which could be used to establish connections within the fly neural circuitry (Schmucker et al., 2000; Shapiro, 2007).

1.6.1.2. Dscam isoforms exhibit homophilic binding

It would be interesting to know how these isoforms, expressed on the surface of one neuron, are recognized and interpreted by neighboring neurons. In vitro biochemical
binding assays reveal that 95% of Dscam isoforms, like many Ig superfamily proteins, exhibit isoform-specific homophilic binding (Wojtowicz et al., 2007). The first eight N-terminal Ig domains, including all three variable domains, form a symmetric S shaped bend and determine binding specificity of Dscam isoforms. Strong homophilic interactions require matching of all the three variable Ig domains. Isoforms with even a single variable Ig domain mismatch fail to engage in heterophilic binding (Wojtowicz et al., 2004) (Figure 11).

The Ig 1-4 domains form a horseshoe shaped ‘head’ wherein Ig2 and Ig3 comprise two distinct binding elements on opposite sides of the ectodomain - epitope I and epitope II. Dimerization and consequently homophilic interactions involve symmetric antiparallel pairing of epitope I but not epitope II (Sawaya et al., 2008). These striking observations suggest that the vast repertoire of Dscam isoforms mediates cell-surface recognition important for neural circuit assembly in Drosophila nervous system development.

Although isoform-specific homophilic interactions might appear to mediate adhesion forces to establish neuronal connectivity, it has been established that binding of identical Dscam isoforms on isoneuronal dendritic branches triggers cytoplasmic signaling events that result in repulsion forces between sister dendrites (Hughes et al., 2007; Soba et al., 2007). Accordingly, deletion of the cytoplasmic domain results in adhesion rather than repulsion, promoting ectopic self-crossing and formation of stable bridges between dendritic branches. Such contact dependent repulsion also underlies axonal branch segregation in mushroom body neurons. The signaling mechanisms downstream of Dscam however remain to be elucidated (Matthews et al., 2007).
Figure 11: Molecular mechanism of homophilic interactions between identical Dscam isoforms. (A) Dscam protein structure showing nine of ten Ig domains and one half of FN1 domain. Ig2, Ig3 and Ig7 domains are encoded by variable exon clusters 4, 6 and 9, respectively. (B) Ig2, Ig3 and Ig7 domains determine the binding specificity between Dscam isoforms. Ig 1-4 domains comprise a 'horseshoe' conformation and a strong skew between Ig 4-5 domains results in an S-shaped conformation important for dimer formation. Matching of all three variable Ig domains is necessary for strong homophilic interactions. Adapted from (Schmucker and Chen, 2009)
1.6.1.3. Role of \textit{Dscam} in axon guidance and dendritic field organization

\textit{Dscam} loss-of-function mutants show axon guidance defects in the embryonic central nervous system, disruption of axonal extension and sister branch segregation in mushroom body neurons, interrupted axon arborization within the ellipsoid body and altered olfactory receptor neuron target specificity and dendritic patterning of projection neurons in the antennal lobe (Hummel et al., 2003; Schmucker et al., 2000; Wang et al., 2002; Zhu et al., 2006). Although, these findings show the involvement of Dscam in \textit{Drosophila} nervous system development, they do not address the question of whether isoform diversity is essential for Dscam function. Indeed, individual cells express a stochastic yet biased repertoire of Dscam isoforms as observed in photoreceptor neurons and S2 cells (14-50 isoforms) and mushroom body neurons (8-30 isoforms) (Du Pasquier, 2005; Neves et al., 2004; Zhan et al., 2004). Hence, Dscam diversity provides each neuron with a unique cell surface ‘tag’, which allows it to distinguish between ‘self’ and ‘non-self’ essential for intra-neuronal avoidance and inter-neuronal recognition (Zipursky et al., 2006). Different mushroom body neurons express several isoforms in varying combinations such that no two neurons share the exact Dscam repertoire. Dscam null mushroom body neurons exhibit defects in axon extension. Dscam diversity reduced to ~11,000 isoforms by deleting various subsets of Ig2 variants, does not affect mushroom body development (Wang et al., 2004). Intriguingly, single isoforms of Dscam, differing only in their ectodomains or transmembrane domains, substantially rescue loss-of-function defects in Dscam null mushroom bodies (Neves et al., 2004; Zhan et al., 2004). These results imply functional redundancy between Dscam isoforms and hence no requisite for isoform diversity. However, expression of a single Dscam
isoform in a bundle of co-projecting neurons causes defasciculation in larval and adult mushroom body. Conversely, expression of a Dscam isoform in a single mutant neuron in an otherwise wild type environment rescues axonal sister branch segregation (Hattori et al., 2007; Zhan et al., 2004) (Figure 12). Dscam diversity is thus essential to provide neighboring neurons with distinct isoforms for normal mushroom body morphogenesis, but a single Dscam isoform is sufficient for sister branch segregation in individual neurons. Moreover, no specific isoform or a subset of isoforms is required for mushroom body development (Hattori et al., 2009; Wang et al., 2004).

Dscam diversity also ensures dendritic field organization in dendritic arborization (da) neurons. Here, da neurons of different classes, sharing the same dendritic territory, need to express different Dscam isoforms. Expression of one Dscam isoform in single mutant da neurons significantly rescues dendritic self-avoidance but over-expressing the same isoform in da neurons of different classes forces spatial segregation of the once overlapping dendritic fields. Dscam, however, does not play a role in heteroneuronal tiling between da neurons of the same class (Hughes et al., 2007; Matthews et al., 2007; Soba et al., 2007) (Figure 13).

In contrast, the role of Dscam diversity is seemingly different in mechanosensory neurons. Expression of randomly chosen single isoforms can restore initial axonal extension and early branching but cannot rescue connectivity. These neurons show impaired secondary and tertiary axonal branching and fail to cross the midline in the central nervous system to reach their target cells. Also, two independent deletions of five Ig2 variants show allele-specific connectivity defects, which suggest functional differences between different isoforms. Thus, in contrast to mushroom body neurons
Figure 12: Role of Dscam diversity in mushroom body development. In Dscam wild-type brains, each cell in a group of co-projecting mushroom body neurons expresses a unique repertoire of Dscam isoforms supporting intra-neuronal axonal branch bifurcation (arrows). The expression of the same isoform in a population of mushroom body neurons prevents axonal branching due to inter-neuronal repulsion (arrows). Axonal branch bifurcation is not affected if only one of the neurons expresses the Dscam single allele. Adapted from (Hummel, 2007).
Figure 13: Role of Dscam diversity in dendritic field organization of dendritic arborization (da) neurons. (A) Sister dendrites of a single da neuron expressing the same Dscam isoform/repertoire do not cross each other. Heteroneuronal tiling between neighbouring da neurons of the same class is however, independent of Dscam. (B) Neighbouring da neurons of different classes expressing completely different sets of Dscam isoforms have overlapping dendritic fields. Adapted from (Schmucker, 2007).
where a larger deletion of Ig2 variants does not affect axonal bifurcation, mechanosensory neurons require a greater Dscam diversity for axonal branching and connectivity with their target cells (Chen et al., 2006) (Figure 14). A model whereby mechanosensory neurons and their corresponding target cells require expression of identical Dscam isoforms to establish neuronal connectivity via homophilic adhesion forces remains plausible.

In summary, axonal sister branch bifurcation and dendritic field organization require the expression of any one Dscam isoform which underlies a strictly cell-intrinsic process of repulsion between isoneuronal neurites. Fasciculation between co-projecting neurons and overlapping heteroneuronal dendritic fields require each cell to express a different set of isoforms to allow co-existence of neighboring neurons arguing for adhesion forces between them (Hughes et al., 2007; Matthews et al., 2007; Soba et al., 2007). An exceptional case of homophilic binding leading to target connectivity via attractive forces is also conceivable.

1.6.1.4. Diversity of Dscam intracellular domain is important for neuronal development

An additional level of Dscam endodomain diversity due to skipping of exon 19 and/or exon 23 has profound impact in neuronal development. RNAi silencing of exon 19 containing transcripts prevents neurites from crossing the midline in Drosophila embryos. Intriguingly, silencing of transcripts lacking exons 19 and 23 disrupts neural wiring in adult flies. Moreover, isoforms without exon 19 are localized in neuritis of mushroom body neurons, contrary to exon 19 containing isoforms that localize in cell bodies as well (Yu et al., 2009). Similarly, exon 17.1 containing isoforms are targeted
Figure 14: A schematic diagram of post-scuteellar (pSc) mechanosensory neuron projections into the adult fly thoracic central nervous system (CNS). The first panel shows Dscam gene and protein organization along with the wildtype projection pattern of a pSc mechanosensory neuron into the thoracic CNS. The remaining panels show projections of pSc mechanosensory neurons in Dscam mutant flies, with varying degrees of reduced exon 4 isoform diversity, indicated by an expanded view of the exons present in each mutant. The light gray boxes indicate deletions. Ectopic or misrouted branches are highlighted in red; branches prevalent in either Dscam deletion mutant (but not both) are highlighted by blue and green arrowheads, and the blue line denotes the CNS midline. Adapted from (Bharadwaj and Kolodkin, 2006).
to dendrites and exon 17.2 containing isoforms are enriched in axons, supporting the evidence that ectopic expression of exon 17.2 containing Dscam rescues axonal branching in mushroom body (Wang et al., 2004). Unlike the ectodomain diversity, no mechanism has been proposed to explain location dependent neuronal function of Dscam isoforms, arising due to alternatively spliced intracellular and transmembrane domains.

1.6.2. Overview of the immune system

The immune system is a collection of defense mechanisms within an organism, which protects it against disease by identifying and combating foreign particles (bacteria, fungi, viruses and parasitic worms) and tumour cells. Importantly, it distinguishes them from the organism's own healthy cells and tissues – 'self' from 'non-self' (Kvell et al., 2007).

The immune system has two components: innate immunity and adaptive immunity. The innate immune system, present in all organisms, is the first line of defense against any invading foreign particle and is believed to be largely invariable without immunological memory (as explained in 1.6.2.1.). In contrast, adaptive immunity, present exclusively in jawed vertebrates, generates variability in the form of antibodies by somatic recombination to cause an intruder-specific response coupled with immunological memory. Phagocytes, responsible for engulfing foreign particles, and B-lymphocytes, producing antibodies, are key players in the innate and adaptive immune responses, respectively (Kvell et al., 2007; Pham et al., 2007).
1.6.2.1. Insect immune system

Insects do not have antibodies, but have a potent innate immune system. Pattern recognition receptors (PRRs) help insects to recognize pathogen-associated molecular patterns (PAMPs) such as lipopolysaccharide (LPS), peptidoglycans (PGNs) and double stranded RNA (Ghosh et al., 2011). Haemocytes are central to cell-mediated immune responses and induce wound healing and blood coagulation to prevent further entry of invading pathogens.

On penetrating this barrier, pathogens are killed by phagocytosis upon recognition, engulfment and intracellular disintegration by haemocytes. Encapsulation wraps multicellular parasites with a thick wall of haemocytes and eradicates them. Nodulation around a microorganism comprises a central core of melanised haemocytes surrounded by a layer of flattened haemocytes, which also effectively clears pathogens (Marmaras and Lampropoulou, 2009; Rowley and Powell, 2007). Antimicrobial peptides (AMPs) produced under the control of two evolutionarily conserved pathways – TOLL and IMD, are key immunomodulators in the insect humoral immune response. The prophenoloxidase (proPO) cascade produces a dark pigment called melanin, which deposits around pathogens and facilitates capsule and nodule formation. Insects also produce cytotoxic reactive oxygen species (ROS) and reactive nitrogen species (RNS), which eliminate host pathogens (Kurata, 2006; Marmaras and Lampropoulou, 2009). Other effectors such as lectins, which bind pathogen associated sugar molecules, and complement-like factors also function in destroying infectious agents (Beck and Habicht, 1996).
1.6.2.2. Evidence for ‘trained immunity’ in invertebrates

Due to the mostly short lifespan of invertebrates, their immune system has been assumed to lack any adaptive capacity. In recent years, however, evidence has accumulated indicating development of pathogen specific responses and immunological memory, which provides invertebrates with a defense system, independent of somatic recombination, called ‘trained’ immunity. One of the earliest works in this direction showed that earthworms were able to recognize and reject grafts from other earthworms. These worms demonstrated a faster graft rejection upon second exposure (Cooper and Roch, 1986).

A later study reported that male American cockroaches (*Periplaneta Americana*) that were first immunized and later challenged with a lethal dose of *Pseudomonas aeruginosa*, survived better (upto 14 days) than saline injected controls. Animals primed with other microbes (*Serratia marcescens, Enterobacter cloacae, Micrococcus lysodeikticus* and *Streptococcus lactis*) failed to show similar protection when challenged with *P. aeruginosa* (Faulhaber and Karp, 1992). Similarly, bumblebees (*Bombus terrestris*) that were initially primed with *Pseudomonas fluorescens, Paenibacillus alvei* or *Paenibacillus larvae* showed higher survival rates on receiving a homologous challenge (same bacteria) as compared to a heterologous challenge (different bacteria) (Sadd and Schmid-Hempel, 2006). These results suggest that insects primed with a particular species of bacteria, show a long-term pathogen-specific immune protection.

Intriguingly, priming *Drosophila* with a sub-lethal dose of *Streptococcus pneumoniae* or *Beauveria bassiana* (fungal pathogen) protected it against a subsequent lethal challenge, but such protection was not observed in flies immunized with *Salmonella*
typhimurium, Listeria monocytogenes or Mycobacterium marinum. These findings suggest that protection by priming is not a general mechanism against all pathogens in Drosophila. Also, for the first time, phagocytes were shown to rapidly clear microbial challenge from the haemolymph and provide such ‘trained’ immunity (Pham et al., 2007).

An interesting phenomenon observed in arthropods involves maternal transfer of pathogen-specific immunity onto the progeny. The water flea Daphnia magna were primed with two strains of pathogenic bacterium Pasteuria ramosa, strain A and strain G. The offsprings of these two groups when exposed to homologous challenge (primed and challenged with strain A or strain G) showed increased survival over those exposed to a heterologous challenge (primed with strain A and challenged with strain G and vice versa) (Little et al., 2003). Such transgenerational immune priming has also been observed in flour moths, bumblebees and mealworms (Moret, 2006; Rahman et al., 2004; Sadd et al., 2005). However, no studies have been carried out to explain the mechanism behind these observations.

The discovery of ‘trained’ immunity in invertebrates serves as a means to assess the importance of cellular versus humoral defenses in providing pathogen-specific long-term protection. The investigations involving P. fluorescens and S. pneumoniae challenge in previously primed bumblebees and fruit flies, respectively, found no role of AMPs in pathogen-specific immune response (Sadd and Schmid-Hempel, 2006). In vivo studies in Drosophila showed that activation of proPO pathway is not required for their survival against infectious challenge (Leclerc et al., 2006). Relish/NF-κB double mutant Drosophila with fewer or no circulating haemocytes, and hence reduced phagocytic activity, but with fully functional AMP synthesis did not survive
opportunistic bacterial or fungal infections (Matova and Anderson, 2006). Moreover, homologous pathogen challenge post-immunization in lobsters (*Homarus americanus*) resulted in enhanced phagocytic activity of haemocytes (Mori and Stewart, 2006; Paterson and Stewart, 1979). These findings provide convincing evidence about the direct involvement of phagocytosis in challenge-specific long-term immunity in invertebrates. It is thus imperative to further explore the mechanism underlying the recent paradigm shift in invertebrate immunity from being completely innate to being ‘trained’.

*Giardia lamblia*, a human intestinal parasite, undergoes variation in antigen expression to escape the host’s immune system. It transcribes many of its variant-specific surface protein (VSP) coding genes (~190 genes) but expresses only one VSP at any given time and can switch to express another VSP. Silencing of the RNAi machinery, results in the expression of multiple VSPs in individual parasites (Prucca et al., 2008). Generally, RNA-dependent mechanisms could represent ways to generate molecular diversity to evade external pathogenic challenges in invertebrates.

1.6.2.3. *Dscam* splicing regulation provides ‘trained immunity’ in invertebrates

Immunoglobulin domain containing proteins comprise the largest group of cell surface receptors, which play a role in antigen recognition, cell adhesion and signaling in vertebrates and invertebrates. In higher vertebrates, Ig containing B and T cell receptors and antibodies exhibit enormous diversity via somatic recombination and clonal selection (Watson et al., 2005). Invertebrates employ alternative splicing
of Dscam, a well-known PRR, to generate a vast repertoire of spliceforms to recognize and protect against a variety of host pathogens (Kurata, 2006). Watson et al., (2005) showed that Dscam is expressed in Drosophila larval haemocytes, fat bodies and brain tissue and is also secreted in the haemolymph. Haemocytes from hypomorphic and transgenic Dscam RNAi larvae; and Drosophila S2 cells show reduced phagocytic uptake of E. coli by ~25-35%. Moreover, Dscam directly binds to live E. coli in vitro via its extracellular domain. This study demonstrates a role of Dscam in phagocytic uptake of bacteria by direct binding and potentially also opsonizing circulating pathogens in the haemolymph via isoforms cleaved from the membrane. However, the specificity of this observation was not examined with different pathogenic exposures (Watson et al., 2005).

In Anopheles gambiae, Sua5B cells (haemocyte-like cell line), on challenge with bacteria (E. coli, Pseudomonas veronii or Staphylococcus aureus) and bacterial cell surface determinants (lipopolysaccharide or peptidoglycan), causes rapid change in AgDscam exon 4 splicing pattern to yield high affinity binding isoforms in a pathogen-specific manner. RNAi against these isoforms decreases bacterial binding, phagocytic uptake by ~55-60% and survival rates on exposure to the same pathogen (Dong et al., 2006). Also, injecting E. coli and S. aureus in crayfish (Pacifastacus leniusculus) induces pathogen-specific, high affinity binding isoforms of PIDscam. Moreover, coating bacteria with these PIDscam recombinant isoforms hides the binding sites required for bacterial clearance by phagocytosis (Watthanasurorot et al., 2011). Feeding mosquitoes with blood meal infected with either one, two or three genotypes of P. falciparum results in increased AgDscam diversity (Smith et al., 2011). The change in splicing pattern and increase in diversity suggests a
mechanism of increasing the probability to recognize, bind and defend against the invading pathogen.

Intriguingly, shrimp (Liptopenaeus vannamei) LvDscam lacks the characteristic transmembrane and intracellular membrane domains associated with Dscam in other arthropods. Haemocytes express a total of 39 individual LvDscam isoforms – 12 isoforms in white spot syndrome virus (WSSV)-free, 11 isoforms in WSSV-persistent and 16 isoforms in WSSV-acute infection conditions. This result suggests a unique mechanism adopted by shrimps to choose LvDscam isoforms exclusively associated with the intensity of viral infection (Chou et al., 2009).

In conclusion, the results from these studies suggest that change in Dscam splicing pattern plays a key role in the putative invertebrate ‘trained’ immunity. Nonetheless, the mechanism, which regulates this splicing switch, is not understood.

1.7. Model for Dscam splicing regulation

Regulation of Dscam diversity by mutually splicing is quintessential for neural circuit formation, organization of dendritic fields and clearance of pathogens by Dscam mediated phagocytosis. For all these functions, expression of a small unique set of Dscam isoforms is required in individual cells of an otherwise identical population.

Single cell analysis in R7 photoreceptor neurons and haemocytes reveals that each cell expresses a distinct set of about 14 to 50 Dscam isoforms (Neves et al., 2004). Hence, every cell in a population of neurons or haemocytes contributes towards an overall unique Dscam identity. Diversification of Dscam could potentially be achieved in the nervous system and the immune system during independent events. In the nervous system, an extensive array of Dscam isoforms present in undifferentiated
neurons is reduced to a cell-specific complement during differentiation, or a limited diversity of *Dscam* isoforms present initially is altered during differentiation to provide a unique neuronal identity. In the immune system, experimental evidence argues that the initial diversity present in a population of haemocytes diversifies into a different complement of isoforms on encountering pathogens.

The mechanism underlying the choice and inclusion of specific exons is however, poorly understood. Particularly intriguing, signals from the cell surface seem to relay to the splicing machinery to regulate *Dscam* diversity. Encountering a pathogen causes high affinity binding *Dscam* isoforms to feedback positive selection and amplification of the same isoform. Contrarily, *Dscam* homophilic repulsion sends negative feedback to exclude selection of identical isoforms between neighbouring neurons (Figure 15).

### 1.8. Aims

This project is integrated in the long-term aim to identify the mechanisms, which regulate the choice of a single or a specific set of *Dscam* isoforms in the *Drosophila* immune system and the nervous system by alternative splicing.

### 1.9. Objectives

In order to mount an immune response, host organisms must first recognize the pathogen after infection. Despite the importance of this process, little is known about the mechanisms of pathogen recognition in invertebrates. Upon pathogen exposure in mosquitoes, the *Dscam* splicing pattern changes to express isoforms that bind the pathogens with higher affinity (Dong et al., 2006). To validate the generality of *Dscam*
Figure 15: Specification of Dscam isoform expression. (A) In the immune system, when Haemocytes encounter a pathogen, the default repertoire of Dscam isoforms is changed towards isoforms that bind invading pathogens with higher affinity. (B) In the nervous system, undifferentiated MB neurons could express either an extensive or a small repertoire of Dscam isoforms. During differentiation, expression of unique set of Dscam isoforms is established in individual cells either by reducing the initial diversity differently in each cell, or by changing the repertoire in each cell. A unique set of Dscam isoforms then allows mushroom body neurons to project in a common path and at the same time use Dscam mediated homophilic repulsion to bifurcate its axonal branches. Adapted from (Hemani and Soller, 2012).
splicing regulation on pathogen exposure, *Drosophila* S2 cells were challenged with different microbes. Potentially, *Dscam* isoforms binding to the pathogens could signal from the membrane to the nucleus through their associated mRNAs to induce the inclusion of the same exons. In support, *Giardia lamblia* employs RNAi to establish antigenic variation on its surface such that at any given time, it expresses only one out of ~190 variant-specific surface proteins (Prucca et al., 2008). Accordingly, a single *Dscam* isoform was expressed in *Drosophila* S2 cells to test if the endogenous Dscam splicing pattern changes to enhance the inclusion of exons present in the exogeneous isoform. To further explore RNA based mechanisms in altering *Dscam* alternative splicing, mutants of *ago1* and *elav*, which have a similar nervous system defect as observed in *Dscam* mutants were tested. In addition, mutants of other RNA processing factors such as *ago2*, involved in siRNA-directed RNAi and members of the exosome and Trl4/Air2/Mtr4p polyadenylation (TRAMP) complexes such as *rrp6* and *trf4* respectively that are involved in removal of cryptic RNA transcripts were also tested. Since ELAV has no obvious binding site in *Dscam* pre-mRNA, an indirect role of ELAV in regulating *Dscam* splicing was assumed. Consequently, mutants of RNA binding proteins, DNA binding proteins and chromatin remodeling factors that are differentially regulated in *elav* mutants were tested for changes in *Dscam* splicing. To further understand the regulation of *Dscam* splicing, transgenic flies containing a construct heterologously expressing an entire variable cluster were made, which will enable us to identify sequence elements that could play a role in selection of specific exons.
CHAPTER 2
MATERIALS AND METHODS

2.1. *Drosophila* husbandry and genetics

2.1.1. Fly food

Flies were reared in glass vials containing 11 ml of standard cornmeal-agar food (8.5% dextrose (w/v), 6% cornmeal (w/v), 2.5% nipagin (v/v), 1% agar (w/v) in tap water). Once solidified, the food was supplemented with live yeast to encourage egg laying.

2.1.2. Fly maintenance

Stocks and crosses were maintained in a 12 hr day/12 hr night photoperiod at 25°C and 70% relative humidity. Crosses were flipped every second day and stocks were flipped every three weeks to allow healthy development of offsprings. The vials were monitored regularly to prevent overcrowding and dehydration.

2.1.3. Embryo collection

14-18 hr old, stage 17 embryos were used for experiments. A cylindrical chamber sealed at the bottom with a wired mesh and covered on top with grape juice plates (3% agar (w/v) and 25% grape juice (v/v) (Ritchies) in tap water with a smear of live yeast) was used to collect embryos. Flies were acclimatized in the chamber for 2-3 d prior to embryo collection. Flies were allowed to lay eggs on a plate in the evening for
four hr. This plate was incubated at 25\degree for 14 hr after which the embryos (14-18 hr old) were collected and dechorionated by immersing them in 50% sodium hypochlorite solution (v/v) for 90 sec. The dechorionated embryos were washed with tap water on a fine mesh using an aspirator (UniEquip) and then observed under a light microscope or a fluorescent microscope (Leica) if green fluorescent protein (GFP) expression was desired. Embryos similar in gut morphology were selected and stored in 10 µl of double distilled water (ddH\textsubscript{2}O) at -80\degree C.

2.1.4. Fly transgenesis with \textit{UAS Dscam 9L}

\textit{UAS Dscam 9L} (containing \textit{Dscam Ex 7-11}, obtained from M. Soller, Figure 16) was prepared by using the Qiagen plasmid purification kit (as explained in 2.4.11). Integration of \textit{UAS Dscam 9L} in the \textit{Drosophila} genome was achieved by \varphi C31 mediated germline transformation system (Bischof et al., 2007). \textit{UAS Dscam 9L} carries a bacterial attachment site (attB), which allows its site-specific incorporation into the genome of flies containing a pre-determined phage attachment landing site (attP) at position 76A2 (\textit{PBac(y[+]attP-9A)}VK00013). The recombination reaction is initiated by a constitutively expressed \varphi C31 integrase encoded on the X chromosome (\textit{y;} \textit{w\textsuperscript{*}M(vasint.Dm)ZH2A}).

The injection mix was prepared by adding 2.5 µg of \textit{UAS Dscam 9L}, 1X injection buffer (5 mM KCl, 0.1 mM Na\textsubscript{3}PO\textsubscript{4}, pH 7.8), 1 mM MgCl\textsubscript{2} and 0.2 µM filtered food dye in a total volume of 10 µl. The mixture was centrifuged at 16,000g for ten min to remove of any interfering particles and placed on ice. The needles used for microinjection were made from fine borosilicate glass capillary tubes (Intracel), with an internal filament (0.22 mm in diameter), by a micropipette puller (Narishige).
Figure 16: Diagrammatic representation of UAS DSCAM 9L
A capillary was placed in a vertical position and melted by a heating filament and subsequently pulled by attached weights under the force of gravity. Needles were filled with DNA from one end, which spread along the whole filament up till the tip, by capillary action.

Approximately 300-500 flies from the \( y^1 \ w^+ M(\text{vasint.Dm})\text{ZH2A} \; PBac(\text{y[+]}\text{-attP-9A})\text{VK00013} \) stock were transferred into an egg laying chamber (as explained in 2.1.3.) and allowed to acclimatize for 2-3 d before injection. Grape juice plates were changed twice a day during this period. The following steps were carried out in an 18°C room. Embryos were harvested after a 20-25 min laying period and washed with tap water using a squeeze bottle. They were dechorionated in 33% sodium hypochlorite solution (v/v) for two min and thoroughly rinsed with tap water.

Dechorionated embryos (~80 embryos) in the pre-blastoderm were aligned with a paintbrush along the edge of a glass coverslip. Care was taken to align the embryos in the same orientation for ease of injection. Embryos were picked up by gently lowering a coverslip coated with heptane glue such that the embryos stuck to the coverslip. The affixed embryos were dehydrated for ~15 min in a desiccation chamber containing silica gel. This time period was adjusted according to relative humidity levels in the desiccation chamber such that a gap developed between the posterior end of the embryo and the vitelline membrane when observed under a phase contrast. The dehydrated embryos were covered in a thin layer of halocarbon 700 oil (Sigma). The coverslip was fixed onto a glass slide and then placed on the stage of an inverted microscope (Nikon) fitted with a micromanipulator (Nikon). The microscope was also attached to a microinjector (Narishige) and a compressor (Jun-Air) via a needle holder. Embryos were focused such that their posterior ends faced
the needle on the same focal plane. The micromanipulator was used to move the embryos against the static needle. Microinjection was achieved by inserting the tip of the needle inside the posterior end of the embryos by expelling a fixed volume of DNA solution under a fixed air pressure. Embryos exhibiting cellularization were destroyed to avoid unnecessary work after eclosion. The injected slide was tilted for 15 min to run out the excess oil. The slide was then placed in a hydration chamber and larvae were collected after 48 hr. These larvae were transferred in special soft food (1% agar (w/v), 5% sucrose (w/v), 4% yeast extract (w/v), 2% inactivated yeast (w/v), 0.5% 9:1 propionic acid: phosphoric acid (v/v)). Eclosed flies from the injected embryos were out-crossed with yw. Apart from the attB site, UAS Dscam 9L has GFP as a positive transformation marker. Progeny from the outcross were screened for GFP expression in the eyes, indicating that transgenesis has occurred. To stabilize the transgenic line, transformants were crossed to a doubly balanced strain (yw; TM3 Sb/TM6b) and a homozygous transgenic stock was established.

2.2. Molecular biology

2.2.1. RNA isolation

Flies/embryos/tissues/cells were homogenized with a pestle in 50 µl of Trizol reagent (Sigma). The pestle was washed off any bound RNA with 450 µl of Trizol reagent. The homogenized mixture was vortexed for 2 min. Then, 200 µl of chloroform:isoamyl alcohol (49:1) was added to the mixture and again vortexed for 2 min. The tube was kept on ice for five min until the phases separated. The mixture was centrifuged at 16,000g for ten min. Then, 200 µl of the aqueous phase (containing RNA) was carefully transferred into a new tube containing 1 µl of
glycogen. Care was taken to avoid contamination of the aqueous phase from the interphase (DNA) or the organic phase (proteins/lipids). Then, 200 µl of absolute isopropanol was added to the aqueous phase and vortexed for a few seconds. The solution was centrifuged at 16,000g for ten min. The supernatant was discarded and the pellet was washed with 750 µl of 70% ethanol (v/v) at 16,000g for ten min. The supernatant was removed and the RNA pellet was left to air dry for ~5 min. The pellet was dissolved in autoclaved diethylpyrocarbonate (DEPC) treated de-ionized water and stored at -20°C. DEPC treated water was made by adding 0.1% DEPC (v/v) to de-ionized water and allowing it to stand for 2 hours at room temperature with intermittent shaking. DEPC was inactivated by autoclaving.

2.2.2. Reverse transcription (RT)

For a total RT reaction volume of 20 µl, a master mix (13 µl) containing the isolated RNA (9.5 µl), 1 µl of 10X RT buffer (200 mM Tris-HCl, 500 mM KCl, pH 8.3), 1 µl of 100 mM dithiothreitol (DTT), 20 units of ribonuclease inhibitor (RNAsin, Promega) and 1 µl of 10 pM oligo(dT) or 1 µl of 20 µM gene-specific probe was kept at 70°C for 15 min in a thermal cycler (Eppendorf). A second master mix (6 µl) containing 1 µl of 10X RT buffer, 1 µl of 100 mM DTT, 1 µl of 100 mM MgCl₂, 1 µl of 10 mM deoxyribonucleotides (dNTPs) and 20 units of RNAsin were added to the reaction mixture just before the temperature reached 50°C. After ten min of incubation at 50°C the reaction mixture was kept at 46°C when 20 units of Superscript II Reverse Transcriptase (Invitrogen) were added and reverse transcription was carried out for one hr. The final extension was carried out at 70°C for 15 min.
2.2.3. Polymerase chain reaction (PCR)

For a total reaction volume of 50 µl, a master mix was prepared containing 1 µl of template (genomic DNA/cDNA/plasmid), 1X Dream Taq buffer (10mM Tris-HCl, 50 mM KCl, 2 mM MgCl₂, pH 8.3, Fermentas), 0.4 µM forward and reverse primers, 0.2 mM dNTPs and 1.25 units of Dream Taq DNA polymerase (Fermentas). This reaction mixture was placed in a thermal cycler (Applied Biosystems) when temperature was >90°C and subjected to an initial denaturation at 94°C for 30 sec, then 30-40 cycles of denaturation at 94°C for 30 sec, annealing at 56°C for 40 sec and synthesis at 72°C for 45 sec and a final extension for one min (annealing temperature and synthesis time varied with primer type and length of PCR product, respectively). PCRs for cloning were carried out using Phusion DNA polymerase (Finnzymes), a proofreading enzyme, in 1X Phusion buffer. Care was taken to prevent mixing of primers before introducing the tubes in the thermal cycler to avoid primer dimer formation.

2.2.3.1. Single fly PCR

One male fly was introduced into a PCR tube and frozen at -20°C for 30 min. Then, 200 µl of isopropanol was added to the tube and incubated at room temperature for one hr. The isopropanol was removed by suction under vacuum for 30 min in a speed vacuum concentrator (Thermo Scientific). A master mix was immediately added to the tube and placed in a thermal cycler to proceed for a PCR reaction (as explained in 4.2.3.).
2.2.4. Agarose gel electrophoresis

To prepare 1% and 2.5% agarose gels (w/v), 1 g and 2.5 g of electrophoresis quality agarose was melted in 100 ml of 1X TAE buffer (40 mM Tris, 1 mM EDTA, 20 mM glacial acetic acid, pH 8.0), respectively. Care was taken to fully dissolve the agarose. Ethidium bromide (5 µl per 100 ml of gel) was added and mixed by shaking. The gel mixture was poured into a gel cast (with combs) and allowed to solidify at room temperature. The PCR products were mixed with 2 µl of 9X loading buffer (43.5% glycerol (w/v), 50 mM EDTA pH 8.0-8.5, 10 mM Tris pH 7.5, 0.05% xylene cyanol (w/v), 0.05% bromophenol blue (w/v)) and loaded into the gels. The samples were run at 150-200 V for 30 min.

2.3. Radioactivity used to label primers

$^{32}$P $\gamma$-ATP (143 µCi/µl, 6000 Ci/mmol, Perkin Elmer) was used for all experiments involving radioactivity. Specific activity of $^{32}$P $\gamma$-ATP was measured using a scintillation counter with a counting efficiency (E) of 0.5. The counts per minute (cpm) reading from the counter was used to calculate disintegrations per minute (dpm) using the equation dpm = cpm/E. The dpm value was converted to disintegrations per second (dps) by using the equation dps = dpm/60. Activity of the radiolabel was calculated in curies (Ci) using the equation Ci = dps/3.7x10$^{10}$ (1 Ci = 3.7x10$^{10}$ dps). Specific activity of the radiolabel was calculated by determining curies per millimole (Ci/mmol) of the radiolabel measured in the scintillation counter (Packard).
2.3.1. Primer radiolabeling

For a total reaction volume of 20 µl, 10 units of T4 polynucleotide kinase (T4-PNK) were used to label 10 µM primers with 143 µCi $^{32}$P $\gamma$-ATP in 1X PNK buffer (70 mM Tris-HCl, 10 mM MgCl$_2$, 5 mM DTT, pH 7.6). The reaction mixture was incubated at 37°C for 30 min. The enzyme was inactivated by incubating the reaction mixture at 80°C for ten min. For a radioactive PCR, either the forward or the reverse primer (as required) was radiolabeled. All of the $^{32}$P $\gamma$-ATP in the reaction mixture was incorporated into the 10 µM of primers used.

2.3.2. Restriction enzyme digests of PCR products from Dscam variable regions

5 µl of PCR product was precipitated using 0.3 M sodium acetate and 2.5 volumes of absolute ethanol at -80°C for ten min. The solution was centrifuged at 16,000g for ten min. The supernatant was discarded and the DNA pellet was washed with 500 µl of 70% ethanol (v/v) at 16,000g for ten min. The pellet was air dried for ~5 min and resuspended in 30 µl of ddH$_2$O. The precipitated PCR product was digested using appropriate enzyme combinations and buffers depending on the amplified Dscam variable exon cluster (as explained below). For a reaction volume of 50 µl, 5 units of each enzyme were used to carry out restriction digests. The digested product was precipitated (as explained above) and dissolved in 20 µl of denaturing gel loading buffer (98% de-ionized formamide (v/v), 10 mM EDTA pH 8.0, 0.025% xylene cyanol (w/v), 0.025% bromophenol blue (w/v)). The samples were denatured at 95°C for 90 sec and immediately placed on ice before loading into the gel.
Exon 4 cluster was digested with *Mbo*I, *Alu*I and *HinPI* in 1X NEBuffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM DTT, pH 7.9) and 1X bovine serum albumin (BSA) at 37°C for one hr in a total reaction volume of 50 μl. Then, *Taq*I was added to the digest mixture and incubated at 65°C for 1 hr.

Exon 6 cluster was digested with *Mbo*I, *Alu*I and *Mspl* in 1X NEBuffer 2 (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.9) and 1X BSA at 37°C for one hr in a total reaction volume of 50 μl. Then, the reaction volume was adjusted to 75 μl with 1X NEBuffer 2 and 1X BSA and digested with *BstUI*, *BstMI* and *Taq*I at 60°C for 1 hr.

Exon 9 cluster with the forward primer radio-labeled was digested with *HpyCH4IV* in 1X NEBuffer 1 (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.0) and 1X BSA at 37°C for one hr in a total reaction volume of 50 μl. Then, the reaction volume was adjusted to 75 μl with 1X NEBuffer 2 and 1X BSA and digested with *HaeIII*, *XmnI* and *MspI* at 37°C for 1 hr. Finally, the reaction volume was adjusted to 100 μl with 1X NEBuffer 2 and 1X BSA and digested with *BstUI* and *BstMI* at 60°C for 1 hr.

Exon 9 cluster with the reverse primer radio-labeled was digested with *HpyCH4IV* in 1X NEBuffer 1 and 1X BSA at 37°C for one hr in a total reaction volume of 50 μl. Then, the reaction volume was adjusted to 75 μl with 1X NEBuffer 2 and 1X BSA and digested with *Mbo*I, *Alu*I and *HinPI* at 37°C for 1 hr. Finally, the reaction volume was adjusted to 100 μl with 1X NEBuffer 2 and 1X BSA and digested with *BstUI* and *Taq*I at 60°C for 1 hr. All restriction enzymes and buffers were bought from New England Biolabs (NEB).
2.3.3. Denaturing polyacrylamide gel electrophoresis

Denaturing polyacrylamide gel electrophoresis was carried out using the SequaGel System (National Diagnostics). For a total volume of 60 ml of 8% gel, SequaGel concentrate (1.07 M acrylamide, 26 mM methylene bisacrylamide (40:1 acrylamide:bisacrylamide), 2.4 M urea), SequaGel diluent (4.35 M urea) and SequaGel buffer (89 mM Tris-borate, 2 mM EDTA, pH 8.3 and 0.75 M urea) were mixed together. Appropriate proportions of the concentrate, diluent and buffer for a range of gel percentages are listed in Table 1 (National Diagnostics). 0.04% TEMED (v/v) and 0.8% freshly prepared ammonium persulfate (v/v) were added to allow polymerization of acrylamide. The un-polymerized gel was gently mixed and cast between two clean glass plates (one of them silanized on one side) separated by plastic spacers. The comb was inserted at one end and the gel was left to polymerize for 1-2 hr. After polymerization, the gel was made to pre-run at 800-1000V for 15-20 min until the temperature reached 55-60°C after which the samples were loaded. 1X TBE (89 mM Tris-borate, 2 mM EDTA, pH 8.3) buffer was used as running buffer. After the completion of the run (at 1300V for ~2.5-3 hr), the glass plates were separated using a spatula. The gel, still stuck on the silanized surface, was fixed with fixation solution (10% methanol (v/v), 10% glacial acetic acid (v/v)) for ten min. The fixed gel was then transferred onto a clean wet filter paper and heat vacuum dried at 85°C for 1 hr. A Kodak imaging screen (Biorad) was exposed to the radioactive gel for 14 hr before scanning it in a molecular imager (Biorad) using Quantity One software (Biorad). The radioactive signal on the imaging screen was erased against a screen eraser (Biorad).
<table>
<thead>
<tr>
<th>DNA Fragment Size (bp)</th>
<th>% Monomer</th>
<th>SequaGel Concentrate</th>
<th>SequaGel Diluent</th>
<th>SequaGel Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;200</td>
<td>4</td>
<td>16ml</td>
<td>74ml</td>
<td>10ml</td>
</tr>
<tr>
<td>80-200</td>
<td>5</td>
<td>20ml</td>
<td>70ml</td>
<td>10ml</td>
</tr>
<tr>
<td>60-150</td>
<td>6</td>
<td>24ml</td>
<td>66ml</td>
<td>10ml</td>
</tr>
<tr>
<td>40-100</td>
<td>8</td>
<td>32ml</td>
<td>58ml</td>
<td>10ml</td>
</tr>
<tr>
<td>0-50</td>
<td>12</td>
<td>48ml</td>
<td>42ml</td>
<td>10ml</td>
</tr>
<tr>
<td>&lt;20</td>
<td>20</td>
<td>80ml</td>
<td>10ml</td>
<td>10ml</td>
</tr>
</tbody>
</table>

**Table 1:** Denaturing polyacrylamide gel composition. The table shows volumes of SequaGel concentrate, diluent and buffer (total gel volume = 100 ml) for commonly used polyacrylamide gel percentages required to resolve corresponding DNA fragments (National Diagnostics).
2.3.4 Quantification of bands

The average intensity of each band was calculated by using the volume rectangle tool in Quantity One software. It measures the total signal intensity within a defined border drawn around the band by adding the intensities of all pixels within the volume boundary multiplied by the pixel area (intensity units x mm²). The background intensity was subtracted from the band intensity using the local background subtraction method. This tool adds the intensities of all pixels in a 1-pixel boundary around the border drawn and divides it by the total number of boundary pixels. This gives a measure of the average background intensity around each volume drawn, which is then subtracted from the intensity of each pixel within the volume. The intensities of resolved exon bands (digested fragments) were divided with the intensity of their input bands (undigested PCR products) to obtain their respective inclusion levels, which were normalized against the input of the control experiment. Regulation of splicing changes was analyzed by determining fold differences between the normalized inclusion levels of a variant exon against corresponding exon bands in the control experiment. The quantified data were represented as best-fit heat-maps, which were generated using the Matrix2png interface.

2.4. Molecular cloning

2.4.1. Media preparation

Luria Bertani (LB) medium was prepared by dissolving peptone (1% (w/v)), yeast extract (0.5% (w/v)) and sodium chloride (0.5% (w/v)) in ddH₂O. The pH was adjusted to 7.0 with 5 M sodium hydroxide (~1.0 ml). The medium was then autoclaved.
2YT medium was prepared by dissolving peptone (1.6% (w/v)), yeast extract (1% (w/v)) and sodium chloride (0.5% (w/v)) in ddH₂O. The pH was adjusted to 7.0 with 5 M sodium hydroxide (~1.0 ml). The medium was then autoclaved. LB plates were made by adding technical quality agarose (1.5% (w/v)) to LB medium and autoclaving. Appropriate antibiotic selection was added to the medium after it cooled to < 50°C. A list of antibiotics used is mentioned in Table 2.

2.4.2. Preparation of competent cells

2.4.2.1. Chemically competent cells

From a single colony of DH5α cells, an overnight culture was set up at 37°C with vigorous shaking. An inoculum of 2.5 ml from the fresh overnight culture was added to 250 ml of LB medium and grown at 37°C with vigorous shaking until OD₆₀₀ reached 0.4-0.5. Care was taken to not overgrow the culture above an OD₆₀₀ of 0.5. The following steps were carried out at 4°C. Cells were harvested by centrifugation at 1,000g for five min. The supernatant was removed and the pellet was carefully resuspended in 100 ml of 50 mM ice-cold CaCl₂. After a recovery time of 20 min, the cells were centrifuged at 1,000g for five min. The supernatant was removed and the pellet was resuspended in 8 ml of 100 mM CaCl₂ to which 2.4 ml of ice-cold 87% glycerol (w/v) was slowly added whilst constant mixing. Aliquots of this suspension were stored at -80°C.

2.4.2.2. Electro-competent cells

From a single colony of DH5α cells, an overnight culture was set up at 37°C with vigorous shaking. An inoculum of 10 ml from the fresh overnight culture was added to
<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock Concentration</th>
<th>Solvent</th>
<th>Working solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>100 mg/ml</td>
<td>Water</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>34 mg/ml</td>
<td>Ethanol</td>
<td>25 µg/ml for plasmids; 12.5 µg/ml for BACs</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>50 mg/ml</td>
<td>Ethanol</td>
<td>12.5 µg/ml; 3 µg/ml for low copy number</td>
</tr>
</tbody>
</table>

**Table 2:** List of antibiotics used during the study. The table shows the stock concentrations (mg/ml), solvents and working solutions (µg/ml) of all antibiotics used during the study.
1 lit of LB medium and grown at 37°C with vigorous shaking until OD$_{600}$ reached 0.6. Care was taken to not overgrow the culture above an OD$_{600}$ of 0.6. The following steps were carried out at ~0°C. Cells were harvested by centrifugation at 1,000g for 15 min. The supernatant was removed and the pellet was carefully resuspended in 1 lit of ice-cold 10% glycerol. The suspension was centrifuged at 1,000g for 15 min. This step was subsequently repeated with 0.5 lit and 250 ml of ice-cold 10% glycerol. The final pellet was resuspended in ice-cold 10% glycerol such that the concentration of cells was 1-3x10$^{10}$ cells/ml. Aliquots of this suspension were stored at -80°C.

2.4.3. Primer phosphorylation

For a total volume of 20 µl, 10 units of T4-PNK were used to phosphorylate 20 µM primers in 1X T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM MgCl$_2$, 10mM DTT, 1 mM ATP, pH 7.5). The reaction mixture was incubated at 37°C for 30 min.

2.4.4. Phenol-chloroform extraction of PCR products

The PCR product was made (as explained in 4.2.3.) upto a final volume of 200 µl with ddH$_2$O. An equal volume of phenol/chloroform/isoamyl alcohol (50:49:1) was added to the diluted PCR product, mixed and centrifuged at 16,000g for ten min. The upper phase was carefully taken and an equal volume of chloroform/isoamyl alcohol (49:1) was added. The solution was mixed and centrifuged at 16,000g for ten min. DNA was precipitated using 0.3 M sodium acetate and 2.5 volumes of absolute ethanol at -80°C for ten min. The solution was then centrifuged at 16,000g for ten min. The supernatant was discarded and the DNA pellet was washed with 500 µl of
70% ethanol (v/v) at 16,000g for ten min. The pellet was air dried for ~5 min and resuspended in ddH₂O to proceed with appropriate restriction enzyme digests.

2.4.5. Restriction enzyme digests of PCR products

All restriction enzyme digests for cloning procedures were performed as 16-fold over-digestion where 1 unit of enzyme digests 1 μg of plasmid DNA/PCR product per hr. All digestion reactions were set in a 50 μl reaction volume. Enzymatic reactions were inactivated by adding 100 μl of phenol/chloroform/isoamyl alcohol (50:49:1) and centrifugation at 16,000g for five min. From the supernatant, 50 μl was run on a 1% agarose gel (w/v) to proceed with gel extraction of the desired band.

2.4.6. Gel extraction of digested DNA fragments

Gel purification of digested DNA fragments was done using the gel extraction kit from Fermentas as follows. After running restriction digests in a 1% agarose gel (w/v), the desired DNA fragment was excised from the gel and weighed. For every 100 mg of agarose gel, 100 μl of binding buffer was added. The gel mixture was incubated at 55°C for five min until the gel piece was completely dissolved. Silica powder suspension was then added to the mixture (5 μl of silica powder suspension for ≤2.5 μg of DNA and an additional 2 μl of silica powder suspension for every μg if DNA ≥2.5 μg). The mixture was incubated at 55°C for five min with intermediate mixing to allow DNA binding to the silica beads. The mixture was then centrifuged at 16,000g for 5 sec to form a silica pellet. The supernatant was removed and the pellet was resuspended in 500 μl of ice cold washing buffer (2.5% concentrated washing buffer (v/v), 50% absolute ethanol (v/v)). This process was repeated three times after which
the silica pellet was air dried until all ethanol evaporated. Finally, the pellet was resuspended in a desired volume of ddH$_2$O or 1X TE (10 mM Tris pH 7.5, 1 mM EDTA pH 8.0), incubated at 55°C for five min and centrifuged at 16,000g for 1 min. The recovered supernatant was placed in a new tube and centrifuged again to remove residual silica before being used for setting up ligation reactions.

2.4.7. Ligation

Gel purified vector backbone and inserts were run on an agarose gel to determine the relative concentrations of each DNA fragment. Appropriate ratios of the backbone and inserts (total DNA amount of 100-120 ng) were ligated in a final volume of 10 μl using 400 cohesive units of T4-DNA ligase in 1X T4-DNA ligase buffer. The ligation mixture was incubated overnight at 16°C.

2.4.8. Bacterial transformation

2.4.8.1. Heat shock transformation

Heat shock transformation was used only to obtain higher yields of plasmids. Chemically competent _E. coli_ (DH5α strain) cells were thawed on ice. Approximately 0.5-1μg of plasmid DNA was mixed with 20 μl of _E. coli_ cells and incubated on ice for 30 min. The mixture was incubated at 42°C for 90 sec and the tube was immediately placed on ice for 1 min. Subsequently, 400 μl of SOC medium (2% Bacto tryptone (w/v), 0.5% Bacto yeast extract (w/v), 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, 10 mM MgSO$_4$, 20 mM glucose) was added and the cells were regenerated at 37°C for 1 hr. The bacteria were then plated on agar plates with appropriate antibiotic selection and incubated overnight at 37°C.
2.4.8.2. Electro-transformation

Electro-transformation was used while cloning fragments to obtain new constructs. Electro-competent *E. coli* (*DH5α* strain) were thawed on ice. The ligation mixture was dialyzed on membranes (0.025 μm, Millipore) against ddH₂O for 30 min. The dialyzed mixture was washed off the membrane using 8 μl of ddH₂O and this step was repeated thrice. The mixture was gently added to 25 μl of bacterial cells and the final volume was made up to 90 μl with ddH₂O. This suspension was carefully transferred into a cold 1 mm electroporation cuvette (0.2cm, Cell Projects) without introducing any air bubbles. The cuvette was placed in a chilled safety slide of a pulser apparatus (Biorad) and pushed into the chamber until the cuvette was seated between the contacts at the base of the chamber. The cells were pulsed once with 2.0 kV, 200 Ω and 25 μF. The capacitance extender and time constant were expected to be 125 μF and 4.5-5 ms, respectively. The pulsed cells were resuspended gently in 500 μl of SOC medium and incubated at 37°C for 1 hr. The bacteria were then plated onto agar plates with appropriate antibiotic selection and incubated overnight at 37°C.

2.4.9. Recombineering

Recombineering is a technique to obtain a fragment from the donor DNA in the host cell by transforming in a linearized vector with homologous regions (200-250 nts) at each end. For homologous recombination to occur the host cell requires the pRed/ET plasmid which encodes the genes of proteins required for recombination, which are induced by L-arabinaose at 30°C. The plasmid is lost at 37°C due to temperature
sensitive origin of replication. In order to detect efficient recombination, the donor vector needs to have a different selection than the receiving vector.

2.4.9.1. Generation of pRed/ET electro-competent cells

From a single colony of pRed/ET plasmid (Gene Bridges) containing cells (agar plates containing 3 µg/ml tetracycline), an overnight culture was set up at 30°C with vigorous shaking. An inoculum of 5 ml from the fresh overnight culture was added to 500 ml of LB medium (3 µg/ml tetracycline) and incubated at 30°C until OD$_{600}$ reached ~0.2. Care was taken to not overgrow the culture above an OD$_{600}$ of 0.2. Immediately, 15 µl of 10% L-arabinose (w/v) was added to the culture and placed in a 37°C water bath and then in a 37°C shaker for one hr until OD$_{600}$ reached 0.35-0.4. Care was taken to not overgrow the culture above an OD$_{600}$ of 0.4. Cells from this culture were washed with 10% glycerol (w/v) to generate electro-competent cells (as explained in 4.4.2.2.). Linearized donor and recipient fragments with homologous flanking sequences (200-400 bp) were co-transformed to allow recombineering of the insert.

2.4.9.2. Generation of Dscam BAC/pRed/ET electro-competent cells

From a single colony of Dscam BAC (Pacman Resources) containing cells (agar plates containing 12.5 µg/ml chloramphenicol) an overnight culture was set up at 37°C with vigorous shaking. An inoculum of 30 µl from the fresh overnight culture was added to 1.4 ml of LB medium (12.5 µg/ml chloramphenicol) and grown at 37°C with vigorous shaking until OD$_{600}$ reached 0.6. Care was taken to not overgrow the culture above an OD$_{600}$ of 0.6. The following steps were carried out at ~0°C. Cells
were harvested by centrifugation at 10,000g for 30 s. The supernatant was removed and the pellet was carefully resuspended in 1 ml of ice-cold 10% glycerol (w/v). This step was repeated once again and the supernatant was tipped out and dripped on a paper towel. The cells were resuspended in the ~30 µl of glycerol left behind. Cells were electro-transformed using 1-5 ng of pRed/ET and grown at 30°C for 70 min before plating (3 µg/ml tetracycline; 12.5 µg/ml chloramphenicol). The cells were grown on plates in the dark at 30°C for 24 hr.

2.4.9.3. Preparation of Dscam BAC/pRed/ET host for DNA retrieval

From a single colony of Dscam BAC/pRed/ET containing cells (3 µg/ml tetracycline; 12.5 µg/ml chloramphenicol) an overnight culture was set up at 30°C with vigorous shaking. An inoculum of 30 µl from the fresh overnight culture was added to 1.4 ml LB medium (3 µg/ml tetracycline; 12.5 µg/ml chloramphenicol) and grown at 30°C with vigorous shaking until OD_{600} reached 0.2. Care was taken to not overgrow the culture above an OD_{600} of 0.2. Then, 15 µl of 10% L-arabinose (w/v) was added to the culture and immediately placed in a 37°C shaker for one hr until OD_{600} reached 0.35-0.4. Care was taken to not overgrow the culture above an OD_{600} of 0.4. The following steps were carried out at ~0°C. Cells were harvested by centrifugation at 10,000g for 30 s. The supernatant was removed and the pellet was carefully resuspended in 1 ml of ice-cold 10% glycerol (w/v). This step was repeated once again and the supernatant was tipped out and dripped on a paper towel. The cells were resuspended in the ~30 µl of glycerol left behind and electro-transformed with linearized fragments with flanking sequences (200-400 bp) homologous to sequences flanking the region of DNA to be retrieved from Dscam BAC clone.
2.4.10. Plasmid DNA mini prep

From 12-48 individual colonies picked randomly from agar plates, overnight cultures were set up in 3 ml of 2YT medium with appropriate antibiotic selection and grown at 35.5°C for 16-18 hr with vigorous shaking. The shaker was set at 35.5°C and not 37°C because in case the temperature goes above 37°C (due to inconsistency of the thermostat), it has been observed that bacteria produce non-digestible DNA and also cause unwanted recombination events with large plasmids. From the overnight culture, 1.5 ml was centrifuged at 800g for five min. The supernatant was discarded and the cells were lysed with STET (8% sucrose (w/v), 0.1% Triton-X 100 (v/v), 50 mM EDTA, 50 mM Tris pH 8-8.5) and lysozyme (0.7 μg/μl, 0.7μM Tris pH 7.5). The solution was vortexed and allowed to stand at room temperature for five min after which it was boiled for 1-3 min and centrifuged at 16,000g for ten min. The protein pellet was removed with a toothpick after which RNaseA (0.04 μg/μl, 0.2 mM Tris pH 7.5, 0.3 mM NaCl, Roche) and 0.3 M sodium acetate (pH 5.2) was added to the supernatant and allowed to stand at room temperature for five min. Thereafter, 500 μl of absolute isopropanol was added and the solution was centrifuged at 16,000g for ten min. The supernatant was discarded and the pellet was washed with 650 μl of 70% ethanol (v/v) at 16,000g for ten min. The pellet was air dried for ~ five min and then dissolved in 50 μl 1X TE out of which 20 μl was used for analytical restriction enzyme digests to screen for the right clone.

2.4.11. Plasmid DNA midi prep

Plasmid DNA midi preps were done using the Qiagen plasmid purification kit as follows. On obtaining the right clone, a starter culture was inoculated from its
corresponding colony in 3 ml of LB medium with appropriate antibiotic selection. After six hr of culturing, it was transferred into 32 ml of LB medium and incubated overnight at 35.5°C for 16-18 hr in a shaker. Subsequently, the cells were harvested by centrifugation at 800g at 4°C for 15 min. The bacterial pellet was resuspended in 4 ml of buffer P1 (50 mM Tris-Cl, 10 mM EDTA, 100 μg/ml RNase, pH 8.0). Then, 4 ml of chilled buffer P2 (200 mM NaOH, 1% SDS (v/v)) was added and the suspension was mixed thoroughly by inverting the tube 4-6 times and incubated at room temperature for five min. Then, 4 ml of buffer P3 (3 M potassium acetate, pH 5.5) was added and immediately mixed by inverting the tube 4-6 times. The lysate was poured into the barrel of the QIAFilter Cartridge and incubated at room temperature for ten min. In the mean time, the QIAGEN-tip 100 was equilibrated with 4 ml of buffer QBT (750 mM NaCl, 50 mM 3-morpholinopropane-1-sulfonic acid (MOPS), 15% isopropanol (v/v), 0.15% Triton X-100 (v/v), pH 7.0) and the column was emptied by gravity flow. A plunger was inserted in the previously equilibrated cartridge and the lysate was filtered into the QIAGIN-tip 100 and allowed to enter the resin by gravity flow. The QIAGIN-tip 100 was washed two times with 10 ml of buffer QC (1 M NaCl, 50 mM MOPS, 15% isopropanol (v/v), pH 7.0) by gravity flow. Finally, the DNA was eluted from the column by 5 ml of buffer QF (50 mM Tris-Cl, 1.25 M NaCl, 15% isopropanol (v/v), pH 8.5). The eluted DNA was precipitated by adding 3.5 ml of room temperature absolute isopropanol and centrifugation at 2,500g for 30 min. The supernatant was discarded and the DNA pellet was washed with 2 ml 70% ethanol (v/v) by centrifugation at 2,500g for ten min. The supernatant was discarded and the DNA pellet was air dried for ~5 min and dissolved in 50-100 μl of 1X TE.
2.4.12. DNA sequencing

All plasmid DNA midi preps were sequenced in the Functional Genomics, Proteomics and Metabolomics Facility in the School of Biosciences, University of Birmingham. For a total reaction volume of 11 µl, 0.5 µM of the primer was mixed with appropriate amounts of DNA as suggested by Beckman Coulter (Table 3).

2.4.13. Cloning of \textit{pAc5.1A Dscam Mut Exons 4-9}

The complete cloning strategy is diagrammatically represented in Figures 17 and 18. For cloning of the \textit{pAc5.1A Dscam Mut Exons 4-9} construct, the 3’UTR of Dscam was amplified from \textit{pcDNA3(-MCS)/F.L. Dscam 7.27.25.2} (kindly provided by Dr. Woj Wojtowicz, University of California) with primers Dscam cDNA end F1 NotI EcoRV (GTGCTGGCGCGCGATATCCGCCAACTGTGCAGAGGACCAATATCG) and Dscam cDNA end R1 (AAATGCCACGCCCACCGCCAGGGCGCGTACAAATTACACTG) and the 3’ end of the endogenous poly A site was amplified from the Dscam BAC clone with primers Dscam 6kb frag F1 SalI/Xho (AAATGTTTTTGTACATCAATTTTGTTGTCTGTGGTCCG) and Dscam 6kb frag R1 (AACATTAATCGGATATCCTACAAATTACACTG) and cloned into \textit{pAc5.1A} (Invitrogen) using NotI and SalI/Xho sites to generate \textit{pAc5.1A Dscam 3’UTR ends}. The first part of Dscam 5’UTR was amplified from the Dscam BAC clone with primers Dscam 5’UTR1 F1 BsrGI Sall BgIII (GCGTCGTACAGTGCAGACAGATCTAGAAACCGGATTTTCAGCGCTAGTCGCGC) and Dscam 5’UTR1 R1
Table 3: Estimation of DNA amounts corresponding to size for DNA sequencing (Beckman Coulter).

<table>
<thead>
<tr>
<th>Size (kbp)</th>
<th>ng for 100 fmol</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>13</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
</tr>
<tr>
<td>0.40</td>
<td>26</td>
</tr>
<tr>
<td>0.50</td>
<td>33</td>
</tr>
<tr>
<td>1</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>130</td>
</tr>
<tr>
<td>3</td>
<td>195</td>
</tr>
<tr>
<td>4</td>
<td>260</td>
</tr>
<tr>
<td>5</td>
<td>325</td>
</tr>
<tr>
<td>6</td>
<td>390</td>
</tr>
<tr>
<td>8</td>
<td>520</td>
</tr>
<tr>
<td>10</td>
<td>650</td>
</tr>
<tr>
<td>12</td>
<td>780</td>
</tr>
<tr>
<td>14</td>
<td>910</td>
</tr>
<tr>
<td>16</td>
<td>1040</td>
</tr>
<tr>
<td>18</td>
<td>1170</td>
</tr>
<tr>
<td>20</td>
<td>1300</td>
</tr>
<tr>
<td>48</td>
<td>1500</td>
</tr>
</tbody>
</table>
Figure 17: Diagrammatic representation of the initial cloning steps to generate pAc5.1A Dscam Mut Exons 4-9 construct. The pAc5.1A Dscam Mut Exons 4-9 construct includes a single Dscam isoform comprising variable exons 4.6, 6.32, 9.25 and 17.2, which was exogenously overexpressed in S2 cells via transfection to test if the endogenous Dscam splicing pattern alters to favour inclusion of the exogenous Dscam isoform.
Figure 18: Diagrammatic representation of the final cloning step to generate \( pAc5.1A \) Dscam Mut Exons 4-9 construct. The \( pAc5.1A \) Dscam Mut Exons 4-9 construct includes a single Dscam isoform comprising variable exons 4.6, 6.32, 9.25 and 17.2, which was exogenously overexpressed in S2 cells via transfection to test if the endogenous Dscam splicing pattern alters to favour inclusion of the exogenous Dscam isoform.
(ATTGTTAACACTCACACACACACTTGTGAGAGTGGTG) and the second part of Dscam 5'UTR was amplified from pcDNA3(-MCS)/F.L. Dscam 7.27.25.2 using primers Dscam 5'UTR2 F2 (TAATCGCATTAAAAACAATTTGGGCAGCAG) and Dscam 5'UTR2 R2 NotI EcoRV (GCGTCGCGGCGCGCGATATCTCGGGCATGGATTGCCGCTGGCCTTG). Both the 5'UTR fragments were cloned into pAc5.1A Dscam 3'UTR ends using Acc65I/BsrGI and NotI sites to generate pAc5.1A Dscam 5'UTRs and 3'UTR ends. The entire 6kb inter-genic fragment from the Dscam BAC clone was recombined into pAc5.1A Dscam 5'UTRs and 3'UTR ends cut with Swal to generate pAc5.1A Dscam 5'UTR and 3'UTR (Construct A). The Dscam cDNA region was recombined from pcDNA3(-MCS)/F.L. Dscam 7.27.25.2 into pAc5.1A Dscam 5'UTR and 3'UTR cut with EcoRV to generate pAc5.1A Dscam cDNA (Construct B). In parallel to the above steps, the pOT Mut Exons 4-9 construct was cloned. To achieve this, a fragment encoding exons 1-4 was amplified from pcDNA3(-MCS)/F.L. Dscam 7.27.25.2 with primers Dscam Exon 1-4 F1 Xho NotI Xba HindIII Ase (GCGTCCTCGGAGACCGGCGCGCAGTCTAGAATAAGCTTGCGATTAATCGGATTTA AAAAAACAATTTGGCCAGCGGCG) and Dscam Exon 1-4 R1 BsrGI (GATATCGGCGGCTAAAATGATTACAACCGGCTCGCAGCATGTCGATCTCCCG). Exons 4-5 were amplified from embryonic Dscam cDNA with primers Dscam Exon 4-5 F1 BsrG1 (CCGGGATGTACATGTGCAGCGGCGTGTGGAATCAGTTTTACGCGGCAGCATATC) and Dscam Exon 4-5 R1 (GTCGTCCTTTTGTGCGAATCGGTTTTTACGCGGCAGCAG). These two fragments were fused together by an overlapping PCR with primers Dscam Exon 1-4 F1 Xho NotI Xba HindIII Ase.
(GCGTCCTCGAGACGCAGCGCCGCAGTCTAGAATAAGCTTGCATTAAATCGCATTTA
AAAAAAATTTTGCCAGCCGCAG) and Dscam Exon 4-5 R1
(GTCGTCCTTTTGTGGGACTTAATCGGGTTTCTCCG). Exons 5-7 were amplified
from embryonic Dscam cDNA with primers Dscam Exon 5-7 F1
(TAGTCATCACAGAGCCCCGTAGCAGTAGTCCGGCCCAAAATCAATG) and Dscam
Exon 5-7 R1 NgoMIV Spe Mlu BglII EcoRI
(GCGTCGAATTTCAGAGATCTGCACCGCTAGACTAGTGCGCGCGCGACGACCTT
GAGGAATACACTTGGTCGGGGTTCCATG). These two PCR products were cloned
into pOT (kindly provided by Dr. Saverio Brogna, University of Birmingham) using
XhoI and EcoRI sites to give pOT Dscam Mut Exons 4-6. Exons 7-8 were amplified
from pcDNA3(-MCS)/F.L. Dscam 7.27.25.2 with primers Dscam Exon 7-8 F1 NgoMIV
(GCGAGTGCTGAGCTGAAGCTCGGAGGCCGTTTCG) and Dscam Exon 7-8 R1
(GAGATCCTCGAGCAGAGTATCCTTCCTTGGC). Similarly, Exons 8-13
were amplified from pcDNA3(-MCS)/F.L. Dscam 7.27.25.2 with primers Dscam Exon
8-13 F1 (TAGAAGTGCAAGTCATGGTTCCACCCAAAATTACGCCCTTCGACTTCG)
and Dscam Exon 8-13 R1 EcoRI
(GTTGGCACTGAAATTTCGACGCCCTTGTACTTCCATG). The above two PCR
products were cloned into pOT Dscam Mut Exons 4-6 using NgoMIV and EcoRI sites
to generate pOT Dscam Mut Exons 4-9. Finally, the mutated Dscam cDNA from pOT
Mu Exons 4-9 was recombined into pAc 5.1A Dscam cDNA cut with Acc65I and Nhel
and end filled with T4 DNA pol to generate pAc5.1A Dscam Mut Exons 4-9 (Final
construct) (Figure 18). The midi prep of the final construct was confirmed by
sequencing the region of the insert carrying the mutations.
2.5. Cell culture

2.5.1. Pathogen infection of *Drosophila* haemocytes

The type and culture conditions of pathogens used are listed in Table 4. *Drosophila* Schneider 2 (S2) cells were cultured in Insect Express medium supplemented with 10% heat inactivated foetal bovine serum (FBS) and 1% L-glutamine/penicillin/streptomycin (w/v) at 27°C with 0% CO₂. They were sub-cultured in a ratio of 1:4 (cells:medium) to maintain 3-5x10⁶ healthy cells/ml and to avoid growth. They were plated at a density of 3x10⁶ cells/well in six well tissue culture plates (Corning), 24 hr prior to pathogen infection. Pathogens, used in the study, were harvested from an overnight culture by centrifugation at 800g for five min, for an inoculum size of 1x10⁶ cells/ml and washed with sterile 1X PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 1.76 mM KH₂PO₄, pH 7.4). Pathogens were resuspended in 1X PBS and heat inactivated at 70°C for 30 min, 60°C and 57°C each for 2.5 min and five min. S2 cells were replenished with serum free and antibiotic free Insect Express medium immediately before pathogen exposure. S2 cells were then exposed to either live pathogens or inactivated pathogens at an infection ratio 1:10 (S2 cell:pathogen) and incubated at 27°C with 0% CO₂ for 12 hr.

After exposure, the infected S2 cells were harvested and resuspended in 1X PBS. RNA was extracted from infected S2 cells for downstream analysis of alternative splicing. PBS treated cells were used as controls. S2 cells and *S. pombe* were kindly provided by Dr. Saverio Brogna, University of Birmingham and the *Mycobacterium* sp. and *C. glutamicum* was kindly provided by Dr. Apoorva Bhatt, University of Birmingham.
<table>
<thead>
<tr>
<th>Pathogens</th>
<th>Strain</th>
<th>Medium</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em></td>
<td>DH5α</td>
<td>Luria Bertani broth</td>
<td>37°C</td>
</tr>
<tr>
<td><em>Saccharomyces pombe</em></td>
<td>040</td>
<td>Yeast extract with supplements + 3% glucose (w/v)</td>
<td>30°C</td>
</tr>
<tr>
<td><em>Mycobacterium smegmatis</em></td>
<td>MC155</td>
<td>Luria Bertani broth + 0.05% Tween 80 (v/v)</td>
<td>37°C</td>
</tr>
<tr>
<td><em>Mycobacterium marinum</em></td>
<td>Wild type</td>
<td>7H9 broth + 0.05% Tween 80 (v/v)</td>
<td>30°C</td>
</tr>
<tr>
<td><em>Corynebacterium glutamicum</em></td>
<td>13032</td>
<td>Luria Bertani broth</td>
<td>37°C</td>
</tr>
</tbody>
</table>

**Table 4:** List of pathogens used to challenge S2 cells. The table shows pathogen strains with their culture conditions used to challenge S2 cells to test changes in Dscam alternative splicing.
2.5.2. Treatment of *Drosophila* haemocytes with drugs affecting transcription and RNA pol II processivity

S2 cells were cultured as mentioned in 4.5.1. S2 cells were then exposed to working concentrations of 10 µg/ml Actinomycin D (ActD), 100 µg/ml 6-Azauracil, 100 µg/ml DRB and 0.83 mg/ml sodium valproate and incubated at 27°C with 0% CO₂ for 12 hr (Table 5). After exposure, the infected S2 cells were harvested and resuspended in 1X PBS. RNA was extracted from treated S2 cells for downstream analysis of alternative splicing.

2.5.3. Transfection of *pAc5.1A Dscam Mut Exons 4-9* in *Drosophila* haemocytes

Transfection of S2 cells with *pAc5.1A Dscam Mut Exons 4-9* was done by using either didecylidimethylammonium bromide (DDAB) or TransIT-LT1 (non-liposomal formulation of lipid and protein/polyamine mixture with low toxicity, Mirus). One day prior to transfection, 3x10⁶ S2 cells were seeded per well at a concentration of 1.5x10⁶ cells/ml of complete Insect Express medium supplemented with 10% heat inactivated FBS and 1% L-glutamine/penicillin/streptomycin (w/v). On day two, transfection was commenced when S2 cells reached a confluency of 50-80%. For transfection with DDAB, 62.5 µl of serum free medium was mixed with 26 µl of 1X DDAB (0.4 mg/ml) in one tube. In another tube, 62.5 µl of serum free medium was mixed with 3 µg of the *pAc5.1A Dscam Mut Exons 4-9* construct. Both these solution mixtures were incubated separately at room temperature for 30 min and were later mixed together. For transfection with TransIT-LT1, 200 µl serum free medium was
<table>
<thead>
<tr>
<th>Drugs</th>
<th>Stock Concentration</th>
<th>Solvent</th>
<th>Working solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinomycin D</td>
<td>50 mg/ml</td>
<td>DMSO</td>
<td>10 µg/ml</td>
</tr>
<tr>
<td>6-Azauracil</td>
<td>50 mg/ml</td>
<td>NH₄OH</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>DRB</td>
<td>24 mg/ml</td>
<td>DMSO</td>
<td>100 µg/ml</td>
</tr>
<tr>
<td>Sodium valproate</td>
<td>50 mg/ml</td>
<td>H₂O</td>
<td>0.83 mg/ml</td>
</tr>
</tbody>
</table>

*Table 5:* List of drugs affecting transcription and RNA pol II processivity used during the study. The table shows the stock concentrations (mg/ml), solvents and working solutions of all drugs used during the study.
mixed with 3 µl of TransIT-LT1 and 3 µg of the pAc5.1A Dscam Mut Exons 4-9 construct and incubated at room temperature for ten min. To the DDAB/DNA and TransIT-LT1/DNA mixtures, 580 µl and 700 µl of serum free medium was added respectively. These mixtures were then added separately to S2 cells and incubated at 27°C without CO₂ for three hr. After transfection, any floating cells in the wells were removed using a Pasteur pipette attached to an aspirator. The cells were replenished with 2 ml of complete Insect Express medium and incubated at 27°C without CO₂ for 2 days. The cells were then harvested for RNA extraction and downstream analysis of alternative splicing.

2.6. Western Blotting

2.6.1. Sample preparation

In order to check if the transgene UAS Dscam 9L expresses protein, Western blotting was done on larval progeny from a cross between homozygous transgenic UAS Dscam 9L and Elav Gal4 flies. Five UAS Dscam 9L/Elav Gal4 larvae were homogenized in 50 µl of 1X PBS. The homogenized sample was mixed with 50 µl of 2X sample buffer (125 mM Tris-HCl pH 6.8, 4% SDS (v/v), 0.01% bromophenol blue (w/v), 100 mM DTT, 15% glycerol (w/v)) and heat inactivated at 95°C for four min.

2.6.2. Sodiumdodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were resolved by SDS-PAGE (Resolving gel: 8% acrylamide (w/v), 0.1% ammonium persulphate (APS) (w/v) and 0.001% TEMED (v/v) in 1X resolving buffer
containing 375 mM Tris, 0.125% SDS (v/v), pH 8.8; Stacking gel: 3% acrylamide (w/v), 125 mM Tris-HCl pH 6.8, 0.1% SDS (v/v), 0.1% APS (w/v) and 0.001% TEMED (v/v); Running buffer: 25 mM Tris, 191 mM glycine, 0.1% SDS (v/v), pH 8.3), The samples were run at 20 mA in the stacking gel and then 25 mA in the resolving gel.

2.6.3. Transfer

A nitrocellulose membrane and eight strips of 0.53 mm blotting paper (Whatman blotting paper 3MM) cut to the size of the gel were pre-wet in the transfer buffer. A sandwich of these components was assembled in a semi dry blotting apparatus (Biorad) in the following order (from cathode to anode) – four strips of blotting paper, nitrocellulose membrane, SDS gel, four strips of blotting paper. The transfer was carried out at 0.8 mA/cm² for 20-30 min and the membrane was then air dried.

2.6.4. Blocking

After re-wetting the nitrocellulose membrane in 1X TBST, it was washed using 5% dry milk (w/v) dissolved in 1X TBST (0.05% Tween 20 (v/v) in 1X TBS buffer - 25 mM Tris, 137 mM NaCl, 2.68 mM KCl, pH 7.4) at room temperature for 30 min on a shaker.

2.6.5. Blotting

After blocking, the membrane was incubated in 1X TBST/0.5% dry milk solution (w/v), containing anti-HA primary antibody (Roche) in a 1:100 dilution for 1.5 hr on a shaker. After the primary antibody incubation, the membrane was briefly washed
twice followed by four ten min washes in 1X TBST. The membrane was then incubated in 1X TBST/0.5% dry milk solution (w/v), containing a peroxidase-conjugated anti-rat secondary antibody (Amersham) in a 1:10,000 dilution on a shaker for 1.5 hr.

2.6.6. Development

Before developing the blot, the membrane was briefly washed twice followed by four ten min washes in 1X TBST. The blot was exposed to an X-ray film and developed by chemiluminescence (ECL Super Signal West femto, Thermo Scientific).

Note: All commonly used chemicals were bought from Sigma or Fischer Scientific.
CHAPTER 3

RESULTS

3.1. *Dscam* variable exons in clusters 4, 6 and 9 have similar sizes

The *Dscam* gene comprises four variable exon clusters 4, 6, 9 and 17 containing 12, 48, 33 and 2 alternative exons, respectively, that are spliced in a mutually exclusive manner (Figure 5, page 27). RT-PCR (reverse transcription polymerase chain reaction) with primers in constitutive exons flanking each variable cluster resulted in a single band on an agarose gel for exon clusters 4, 6 and 9 (Figure 19). These bands represent a population of variable exons that are indistinguishable due to their identical or very similar lengths. Exon 4, 6 and 9 variants range between 159-171bp, 116-128bp and 279-306bp, respectively. The two exon 17 variants differ in 45bp (exon 17.1=168bp; exon 17.2=213bp) and appeared as two separate bands on an agarose gel.

3.2. Separation of *Dscam* variable exons based on sequence variation

To make regulation of *Dscam* alternative splicing possible, it is required to be able to distinguish inclusion levels of individual exon variants in a *Dscam* variable cluster. Analysis of sequence variation between alternative exons in each cluster revealed that their annotated sequences are divergent enough so that they can be
Figure 19: Amplification of Dscam variable exons by RT-PCR using RNA extracted from wild type embryos (Canton S), with primers in the constitutive exons flanking the variable exon clusters. Gene specific primer (Dscam YH 11RT1) was used in the RT reaction for cDNA (complementary DNA) synthesis. Dscam exon 4 variants were amplified using primers Dscam YH 3F2 and Dscam YH 5R1 (Lane 1), exon 6 variants with primers Dscam YH 5F1 and Dscam YH 7R1 (Lane 2), exon 9 variants with primers Dscam YH 8F1 and Dscam YH 10R1 (Lane 3) and exon 17 variants with primers Dscam YH 16F1 and Dscam YH 18R1 (Lane 4). The amplified products were run on a 2.5% agarose gel.
distinguished after digestion with a combination of restriction enzymes resulting in fragments of distinct lengths. To resolve shorter fragments, the forward primer was radiolabeled with $^{32}$P and the digested exon variants were run on a denaturing polyacrylamide gel. Each band on the gel can be assigned its corresponding variable exon number based on the distance of the restriction site closest to the labeled end. A diagrammatic representation of separating exon 4 cluster variants is shown in Figure 20. Exon 4 variants were identified by digestion with restriction enzymes $Mbo$, $Alu$, $HinP1$ and $Taq^\alpha$. All 12 exon 4 variants were separated by this combination of restriction enzymes according to their annotated sequence as follows (Figure 21; Lane 5 - refer to A1 in appendix). Exons 4.3 and 4.8 were identified by $Mbo$ (Figure 21; Lane 1), exon 4.9 by $Alu$ (Figure 21; Lane 2), exons 4.6 and 4.12 by $HinP1$ (Figure 21; Lane 3) and exons 4.1, 4.4, 4.5, 4.7, 4.10 and 4.11 by $Taq^\alpha$ (Figure 21; Lane 4). Exon 4.2 was identified by its full length due to the absence of a restriction site. An unspecific band, likely due to PCR artifacts was detected and is indicated by a ‘*’ band. Often, two closely spaced bands were observed in the size range expected for exons 4.10 and 4.12. This could likely be due to incomplete denaturation of these fragments thus affecting their mobility on the gel. To avoid ambiguity, both bands were considered for exons 4.10 and 4.12 when analyzing $Dscam$ exon 4 splicing regulation.

Exon 9 variants were identified by digestion with restriction enzymes $HpyCH4IV$, $HaeII$, $XmnI$, $MspI$, $BstUI$ and $BstM$. Theoretically, this combination of enzymes would separate 24 out of the 33 exon 9 variants. Experimentally, however, only 17 variable exons were identified as follows (Figure 22; Lanes 7, 8 and 9 - refer to A2 in appendix). Exons 9.7 and 9.10 were identified by $BstUI$ (Figure 22; Lane 2), exons
Figure 20: Diagrammatic representation showing the separation of Dscam exon 4 variants. The green box represents the population of Dscam exon 4 variants. Red boxes represent the flanking constitutive exons 3 and 5. The solid circle (•) indicates the $^{32}$P radiolabel at the 5' end of the forward primer. Coloured arrows show the combination of enzymes that digest Dscam exon 4 variants at distinct sites resulting in fragments of different lengths that can be separated on a denaturing polyacrylamide gel. Note that exon 4 fragments are not drawn to scale.
Figure 21: Resolution of Dscam exon 4 cluster. Separation of Dscam exon 4 variants after RT-PCR using RNA extracted from wild type embryos (Canton S), with primers Dscam YH 3F2, radiolabeled with $^{32}$P, and Dscam YH 5R1 located in constitutive exons 3 and 5, respectively. The population of Dscam exon 4 variants was digested with MboI, AluI, HinP1I and Taq1I (Lane 5). Each band is assigned its corresponding variable exon number and identifying enzyme in brackets. The population of Dscam exon 4 variants was digested individually with MboI, AluI, HinP1I and Taq1I, respectively (Lanes 1-4). Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 22: Resolution of Dscam exon 9 cluster. Separation of Dscam exon 9 variants after RT-PCR using RNA extracted from wild type embryos (Canton S), with primers Dscam YH 8F1, radiolabeled with $^{32}$P, and Dscam YH 10R1 located in constitutive exons 8 and 10 respectively. The population of Dscam exon 9 variants was digested with BstNI, BstUI,MspI, XmnI, HaellII and HpyCH4IV (Lanes 7-9, digest loaded three times). Each band is assigned its corresponding variable exon number and identifying enzyme in brackets. The population of Dscam exon 9 variants was digested individually with BstNI, BstUI, MspI, XmnI, HaellII and HpyCH4IV, respectively (Lanes 1-6). Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB), phiX174 DNA/HinfI marker (Biotools).
9.4, 9.5, 9.22, 9.23 and 9.26 by MspI (Figure 22; Lane 3), exon 9.25 by XmnI (Figure 22; Lane 4), exons 9.1, 9.2, 9.8, 9.19 and 9.24 by HaeIII (Figure 22; Lane 5) and exons 9.11, 9.20, 9.31 and 9.33 by HpyCH4IV (Figure 22; Lane 6). Exons 9.11 and 9.31 were detected at a very low level and in close proximity to highly expressed exons 9.14, 9.15 and 9.28 and exon 9.8, respectively. To avoid the potential overlap of radioactive signal between strong and weak bands, exons 9.11 and 9.31 were excluded from the analysis of Dscam exon 9 splicing regulation. Exons 9.14, 9.15 and 9.28 as well as exons 9.16, 9.18 and 9.27 appeared as a single band as they were digested by BstNI to produce same sized fragments (Figure 22; Lane 1). Exons 9.3 and 9.32, digested by BstNI, and exon 9.9 digested by BstUI also appeared together as a single band (not shown in figure 22). The remaining seven exon variants, namely 9.6, 9.12, 9.13, 9.17, 9.21, 9.29 and 9.30 were not detected at the predicted size likely because of low level of inclusion. A few unspecific bands, likely due to PCR artifacts, were also detected and are shown as ‘*’ bands (Figure 22).

Although the focus of this study is on exon 4 and 9 clusters, a preliminary analysis of exon 6 cluster was also done. Exon 6 variants were identified by digestion with restriction enzymes Mbol, AluI, MspI, BstUI, BstNI and TaqI. Theoretically, this combination of enzymes would separate 28 out of the 48 exon 6 variants. Experimentally, however, only 23 variable exons were identified based on sequence annotation as follows (Figure 23 - refer to A3 in appendix). Exons 6.8, 6.17, 6.18 and 6.23 were identified by Mbol, exons 6.9, 6.13, 6.19, 6.21, 6.22, 6.32, 6.36, 6.38 and 6.39 by AluI, exons 6.16 and 6.28 by MspI, exons 6.10, 6.15, 6.35 and 6.41 by BstUI, exons 6.1, 6.6 and 6.24 by BstNI and exon 6.44 by TaqI. The unresolved exon variants appeared as single bands at distinct positions in groups of two or more.
Figure 23: Resolution of Dscam exon 6 cluster. Separation of Dscam exon 6 variants after RT-PCR using RNA extracted from wild type embryos (Canton S), with primers Dscam YH 5F1, radiolabeled with $^{32}$P, and Dscam YH 7R1 located in constitutive exons 5 and 7, respectively. The population of Dscam exon 6 variants was digested with MboI, AluI, MspI, BstUI, BstNI and TaqI (Lane 2). Each band is assigned its corresponding variable exon number. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
exons because they were digested by enzymes to produce the same sized fragments. The exon variants that could not be separated include exons 6.2 and 6.26 digested by **Alu** I, exons 6.37 and 6.46 digested by **BstMI**, exons 6.3 and 6.5 digested by **BstUI**, exons 6.7, 6.25 and 6.27 digested by **MboI**, exons 6.33, 6.40, 6.42 and 6.48 digested by **Alu** I and exons 6.30 and 6.45 digested by **TaqαI** and **Alu** I, respectively. Exon 6.12 digested by **Alu** I and exons 6.14 and 6.34 digested by **TaqαI**; and exons 6.4 and 6.31 digested by **MspI** also appeared together as single bands, respectively (not shown in the figure 23).

The remaining five exon variants, namely 6.11, 6.20, 6.29, 6.43 and 6.47 were not detected at the predicted size likely because of low level of inclusion. A few unspecific bands, likely due to PCR artifacts were detected and are shown as ‘*’ bands (Figure 23).

### 3.3. **Dscam** splicing pattern changes on exposure to pathogens in S2 cells

The splicing pattern of **AgDscam** exon 4 cluster in Sua5B cells has been shown to change on exposure to pathogens and pathogenic determinants to express challenge-specific high affinity binding isoforms. Silencing these isoforms by RNAi reduces binding of AgDscam to the inducing pathogen and compromises their phagocytic uptake. Also, RNAi-mediated depletion of AgDscam in mosquitoes reduces their viability after bacterial infection (Dong et al., 2006).

To test if the splicing pattern of **Dscam** exon 4 and 9 clusters changes upon pathogen exposure, S2 cells were challenged with *E. coli, S. pombe, C. glutamicum,*
M. smegmatis and M. marinum that were heat inactivated at 70°C for 30 min (Figures 24A and B). Naïve cells expressed all exon 4 cluster variables, but exons 4.3 and 4.6 were expressed at a very low level (Figure 24A; Lane 1). Challenges with all pathogens increased exon 4.7 inclusion, maximally with S. pombe and M. marinum by an 11.2-fold and 9.2-fold upregulation, respectively (Figure 24A; Lanes 3 and 6). E. coli and C. glutamicum challenge resulted in a 5.5-fold and 6.1-fold increase in exon 4.7 splicing, respectively (Figure 24A; Lanes 2 and 4). Splicing of exon 4.3 was downregulated by 5.8-fold in S2 cells exposed to C. glutamicum (Figure 24A; Lane 4). S2 cells infected with M. smegmatis showed the least overall change in Dscam exon 4 splicing (Figure 24A; Lane 5). To attribute the above changes in splicing pattern purely to the pathogenic exposure, Dscam exon 4 splicing was examined over 4 hr, 8 hr and 12 hr in naïve S2 cells to observe if Dscam splicing changes over time. This experiment revealed that Dscam splicing pattern is largely constant over a time period of 12 hours suggesting that changes in exon 4.7 splicing were purely due to the pathogen exposures (Figures 24C and D). Contrary to exon 4 cluster, S2 cells expressed a very limited repertoire of exon 9 cluster variants with exons 9.24, 9.20 and 9.8 being the predominant exons. A quantification of only these strongly expressed exons revealed no significant differences between different pathogen challenges (Figures 24E and F).
Figure 24 (A,B): *Dscam* exon 4 splicing pattern changes on exposure to heat inactivated pathogens in S2 cells. (A) Analysis of *Dscam* exon 4 splicing pattern using RNA extracted from S2 cells exposed to *E. coli* (Lane 2), *S. pombe* (Lane 3), *C. glutamicum* (Lane 4), *M. smegmatis* (Lane 5) and *M. Marinum* (Lane 6). PBS added to S2 cells served as controls (Lane 1). (B) Heat map representation of *Dscam* splicing changes observed in A. *Dscam* exon variants were separated as explained in figure legend 21. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
Figure 24 (C,D): Dscam exon 4 splicing pattern does not change in naïve S2 cells over a period of 12 hr. (C) Analysis of Dscam exon 4 splicing pattern over a time course of 0 hr (Lane 1), 4 hr (Lane 2), 8 hr (Lane 3) and 12 hr (Lane 4) in naïve S2 cells. (D) Heat map representation of Dscam splicing changes observed in C. Dscam exon variants were separated as explained in figure legend 21. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 24 (E,F): Dscam exon 9 splicing pattern changes on exposure to heat inactivated pathogens in S2 cells. (E) Analysis of Dscam exon 9 splicing pattern using RNA extracted from S2 cells exposed to E. coli (Lane 2), S. pombe (Lane 3), C. glutamicum (Lane 4), M. smegmatis (Lane 5) and M. Marinum (Lane 6). PBS added to S2 cells served as controls (Lane 1). (F) Heat map representation of Dscam splicing changes observed in E. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
3.4. Variation in *Dscam* splicing between different developmental stages, strains and sexes

To elucidate the mechanisms regulating *Dscam* splicing in adaptation to pathogen exposure or during neuronal development, it was thought to test mutants in candidate genes for splicing regulators. Since *Dscam* splicing in mushroom bodies and dendritic arborization neurons changes such that individual cells acquire a unique set of *Dscam* isoforms, it is possible that *Dscam* splicing is generally variable. To exclude that such variability is a key feature of *Dscam*, the pattern of *Dscam* splicing at various developmental stages and in individual flies was analysed. Consequently, *Dscam* exon 4 and 9 splicing pattern was analyzed between nine independent pools of ten *Canton S* embryos, ten individual *Canton S* males and eight individual *yw* females (Figures 25 and 26). These results of this analysis revealed that the choice of exon variants in *Dscam* exon 4 and 9 splicing followed a different trend between different developmental stages, wild type strains and sexes. For *Dscam* exon 4 splicing, exon 4.9 accounted for 10.3% of all spliced exons in *Canton S* embryos, where as in *Canton S* males and *yw* females it accounted for only 3.9%. In *Canton S* males, exon 4.5 represented 12.7% of all splicing events, which was only 6.9% in *yw* females. Similarly, exon 4.8 accounted for 12.5% of all spliced variable exons in *yw* females, which was only 6.9% in *Canton S* males (Figures 25A-G).

To attribute variation observed between *Canton S* males and *yw* females to strain or sex differences, and also to get a better understanding about tissue specific splicing, neuron rich head-thoraces were compared with neuron poor abdomens between *Canton S* and *yw* females. The results revealed that differences in proportions of exon variants included in *Canton S* and *yw* flies were due to strain differences as
Figure 25 (A,B): Variation in Dscam exon 4 splicing pattern between different developmental stages, strains, sexes and tissues. (A) Analysis of Dscam exon 4 splicing pattern using RNA extracted from independent pools of ten 14-18 h old Canton S embryos (Lanes 1-9). (B) Heat map representation of Dscam splicing changes observed in A. p values for each exon variant are mentioned alongside. Dscam exon variants were separated as explained in figure legend 21. Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 25 (C,D): Variation in Dscam exon 4 splicing pattern between different developmental stages, strains, sexes and tissues. (C) Analysis of Dscam exon 4 splicing pattern using RNA extracted from single Canton S males (Lanes 1-10). (D) Heat map representation of Dscam splicing changes observed in C. *p values* for each exon variant are mentioned alongside. Dscam exon variants were separated as explained in figure legend 21. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
Figure 25 (E,F,G): Variation in Dscam exon 4 splicing pattern between different developmental stages, strains, sexes and tissues. (E) Analysis of Dscam exon 4 splicing pattern using RNA extracted from single yw females (Lanes 1-8). (F) Heat map representation of Dscam splicing changes observed in E. p values for each exon variant are mentioned alongside. Dscam exon variants were separated as explained in figure legend 21. Samples were run on an 8% denaturing polyacrylamide gel. (G) Graphical representation of percentage inclusion levels of each exon 4 variant observed in A, C and E. M=phiX174 DNA/HinfI marker (Biotools).
Figure 25 (H,I,J): Variation in Dscam exon 4 splicing pattern between different developmental stages, strains, sexes and tissues. (H) Analysis of Dscam exon 4 splicing pattern using RNA extracted from head-thorax from single Canton S female (Lane 1) and yw female (Lane 2); abdomen from single Canton S female (Lane 3) and yw female (Lane 4). (I and J) Graphical representation of percentage inclusion levels of each exon 4 variant observed in H. Dscam exon variants were separated as explained in figure legend 21. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
observed in exons 4.5, 4.8 and 4.12. The pattern between head-thoraces and abdomens was however, very similar within the same strain (Figures 25H-J). Similar levels of exon variants such as 4.2, 4.9 and 4.10 were also observed between Canton S males and yw females. Only exons 4.6 and 4.11 showed comparable levels of inclusion between all the three sample types (Figure 25G). For Dscam exon 9 splicing, exons 9.7 and 9.8 both accounted for 15.5% of all exon 4 variants in Canton S embryos, which were expressed at levels less than 8.6% and 4.1% in Canton S males and yw females, respectively. Exon 9.24 was maximally included in yw females representing 33.4% of all exons included within the cluster, where as in Canton S embryos and flies it was only 9.2% and 19.32% respectively. Overall, exon 9 cluster showed less variation between the three sample sources as compared to the exon 4 cluster (Figures 26A-G).

The splicing pattern between individual pools of Canton S embryos, Canton S males and yw females was highly reproducible. To investigate reproducibility of Dscam splicing pattern between independent experiments a 2-tailed t-test was performed for each of the resolved exon 4 and 9 variants by comparing their inclusion levels across all experiments. The data set was randomly divided into 3 different pairs of arrays and their mean p value was calculated. The obtained p values for all exon 4 and 9 variants were >0.05, which suggested that the results obtained between independent experiments are very similar. Hence, Dscam splicing pattern is highly reproducible between individual samples of the same source such as embryo pools and individual flies. The p values are shown adjacent to each exon 4 and 9 variants in figures 25 and 26B, D and F, respectively. Between individual Canton S males, exons 4.4, 9.2,
Figure 26 (A,B): Variation in Dscam exon 9 splicing pattern between different developmental stages, strains, sexes. (A) Analysis of Dscam exon 9 splicing pattern using RNA extracted from independent pools of ten 14-18 h old Canton S embryos (Lanes 1-5). (B) Heat map representation of Dscam splicing changes observed in A. p values for each exon variant are mentioned alongside. Dscam exon variants were separated as explained in figure legend 22. UNSpecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
Figure 26 (C,D): Variation in \textit{Dscam} exon 9 splicing pattern between different developmental stages, strains, sexes. (C) Analysis of \textit{Dscam} exon 9 splicing pattern using RNA extracted from single \textit{Canton S} males (Lnes 1-5). (D) Heat map representation of \textit{Dscam} splicing changes observed in \textit{C}. \textit{p values} for each exon variant are mentioned alongside. \textit{Dscam} exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8\% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 26 (E,F,G): Variation in Dscam exon 9 splicing pattern between different developmental stages, strains, sexes. (E) Analysis of Dscam exon 9 splicing pattern using RNA extracted from single yw females (Lanes 1-5). (F) Heat map representation of Dscam splicing changes observed in E. p values for each exon variant are mentioned alongside. (G) Graphical representation of percentage inclusion levels of each exon 9 variant observed in A, C and E. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
9.5 and 9.33 showed the most variation. Females from the *yw* strain showed complete absence of exon 9.33.

### 3.5. Analysis of *Dscam* splicing in mutants of genes involved in small RNA and mRNA processing

*Dscam* mutants exhibit severe nervous system defects manifested in the disruption of connectives and commissures in the ventral nerve cord (Schmucker et al., 2000). Therefore, mutants of genes involved in RNA processing, which have similar nervous system defects to analyze *Dscam* splicing were searched. The genes that were identified with such a phenotype were *elav and ago1* (Kataoka et al., 2001; Simionato et al., 2007) (Figure 27). ELAV is a neuron-specific RNA binding protein that regulates alternative splicing of *erect wing* and *Neuroglian* by binding to their pre-mRNA (Lisbin et al., 2001; Soller and White, 2003). Ago1 is a key component of the RNAi machinery, which functions in the cytoplasm as a regulator of translation and RNA degradation, and in the nucleus as a regulator of chromatin remodeling (Bernstein and Allis, 2005; Hutvagner and Simard, 2008). Given Ago1’s role in chromatin remodeling, it was further reasoned that the inclusion of a specific *Dscam* exon in the variable cluster might be mediated by its RNA sequence through Ago1 by signaling back to the nucleus, potentially involving chromatin remodeling. Thus, Ago1 might play a role in regulating *Dscam* splicing by an unknown RNA based mechanism of signaling from the cell surface to the nucleus to induce selection or repression of a particular *Dscam* isoform (shown later in Figure 35). In addition to Ago1, *Drosophila* has a second Argonaute protein; Ago2, which is primarily involved in siRNA-mediated silencing of exogenous genes in response to, for e.g. viruses or of
Figure 27: ago1 and elav show similar Dscam mutant phenotype. Similar nervous system defects manifested in abnormal connectives and commissures observed in stage 16 mutant embryos of (A) Dscam, (B) ago1 and (C) elav. Adapted from (Kataoka et al., 2001; Schmucker et al., 2000; Simionato et al., 2007).
endogenous genes primarily encoded by transposons (Chung et al., 2008; Okamura et al., 2004; van Mierlo et al., 2012). In this study, mutants in rrp6, the activity determining component of the nuclear exosome and trf4, a component of the TRAMP complex, involved in degradation of aberrant RNAs and spurious transcripts termed, cryptic transcripts were also included (Callahan and Butler, 2010). It was reasoned that such transcripts might be involved in selection of exons in the Dscam variable region. In particular, spurious antisense transcription has been shown to silence genes and such a mechanism could also occur in the Dscam region to suppress inclusion of specific exon variants (Camblong et al., 2007).

Dscam splicing in mutants for the above genes in 14-18h old embryos were analyzed because ago1, rrp6 and elav null mutants are embryonic lethal. To analyze Dscam splicing in ago1 and rrp6 mutants, a genetic trick was applied to select for null mutant embryos because no GFP labeled balancers were available at this time that would allow selection of homozygous mutant embryos. Therefore, the GAL4/UAS system was used to mark the chromosome carrying the mutation with elavGAL4 and a chromosome carrying a deficiency of the locus with UAS GFP (Brand and Perrimon, 1993). If these two strains are crossed together, only the mutant embryos will express GFP. This genetic selection system was used to identify ago1 and rrp6 mutants. For selection of elav null embryos, the C155 enhancer trap GAL4 inserted in the elav gene was used in combination with UAS GFP to label the elav+ progeny from the following cross, elav/C155; UAS GFP/+ x C155/Y; UAS GFP/UAS GFP (refer to A4, A5 and A6 in appendix). Embryos of ago2 and trf4 mutants were directly collected from their stocks because they are homozygous viable.
*Dscam* exon 4, 6 and 9 splicing pattern was analyzed in *ago1, ago2, rrp6, trf4* and *elav* mutant embryos. Analysis of *Dscam* exon 4 splicing pattern revealed that *ago1* mutants produced a severe overall effect in altering *Dscam* splicing. Exons 4.5 and 4.11 were not detected (Figure 28A; Lane 3). Also, splicing of exons 4.3 and 4.10 was reduced by over two-fold. Exon 4.4 splicing was enhanced in *rrp6, ago2* and *elav* mutants by greater than five-fold (Figure 28A; Lanes 2, 5 and 6). Analysis of *Dscam* exon 9 splicing revealed that *elav* mutants produced the most dramatic changes. Exons 9.1, 9.4, 9.19, and 9.23 were all downregulated by over five-fold. The *ago2* mutants also showed downregulation of exons 9.19 and 9.20 by 4.5 and 6.8-fold, respectively (Figure 29A; Lane 3). Thus, exon cluster-specific effects in regulation of *Dscam* alternative splicing were observed with *ago1* affecting exon 4 and *elav* affecting exon 9 splicing. Also, a preliminary study on regulation of *Dscam* exon 6 splicing by the above mutants was carried out. None of the mutants produced any pronounced effects in splicing of exon 6 variants, except *ago1* mutants which downregulated exon 6.6 by 2.3-fold (Figure 30A; Lane 2).

### 3.6. Analysis of *ago1* and *rrp6* maternal mutants and redundancy between Agos and *rrp6* in *Dscam* splicing regulation

Although differences in *Dscam* exon 4 splicing were primarily observed in *ago1* mutants, the full effect of the Argonaute proteins might not have been seen because they might function redundantly in regulating *Dscam* splicing. To test if Ago1 and Ago2 play redundant roles in regulating *Dscam* splicing, double mutants of *ago1* and
Figure 28: Analysis of Dscam exon 4 splicing in mutants of genes involved in small RNA and mRNA processing. (A) Analysis of Dscam exon 4 splicing pattern using RNA extracted from 14-18 h old mutant embryos of rrp6/Df (Lane 2), ago1/Df (Lane 3), elav/Y (Lane 4), ago2/ago2 (Lane 5) and trf4/trf4 (Lane 6). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 21. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
Figure 29: Analysis of Dscam exon 9 splicing in mutants of genes involved in small RNA and mRNA processing. (A) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of ago1/Df (Lane 2), ago2/ago2 (Lane 3), rrp6/Df (Lane 4), trf4/trf4 (Lane 5) and elav/Y (Lane 6). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
Figure 3: Analysis of Dscam exon 6 splicing in mutants of genes involved in small RNA and mRNA processing. (A) Analysis of Dscam exon 6 splicing pattern using RNA extracted from 14-18 h old mutant embryos of ago1/Df (Lane 2), ago2/ago2 (Lane 3), rrp6/Df (Lane 4), trf4/trf4 (Lane 5) and elav/Y (Lane 6). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam most significant splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 23. Samples were run on an 8% denaturing polyacrylamide gel. M= phiX174 DNA/HinfI marker (Biotools).
ago2 were generated. Since ago2 is located on a different chromosome than ago1 and ago2 mutants are viable, the GAL4-UAS system could be used as before to recognize ago1 mutants in an ago2 background (refer to A7 in appendix). Potentially, ago1 and rrp6 could act redundantly in selection of Dscam isoforms and therefore double mutants were also tested. Testing these double mutants was possible because the genes are located on different chromosomes and GFP labeled balancers with no maternal expression of GFP were available (refer to A8 in appendix).

Both, ago1 and rrp6 are expressed during oogenesis. Therefore, both proteins are deposited in eggs and provide functionality during embryogenesis. Hence, zygotic null mutants still contain protein that can last up to the end of embryogenesis. To remove maternal ago1 and rrp6, germ line clones for these mutants were generated (refer to A9 and A10 in appendix). The dominant female sterile technique utilizes the dominant ovoD mutation that does not allow development of oocytes. Mitotic recombination mediated by flp-FRT site specific recombination between homologous chromosomes, containing either an ovoD transgene or a mutation of choice in female germ cells, will result in the loss of ovoD and hence development of functional oocytes (Perrimon, 1998) (Figure 31). To allow the generation of germ line clones, FRT sites were added to mutant chromosome arms by meiotic recombination and validated by PCR (Figures 32A and B).

Analysis of Dscam exon 4 splicing in ago1/Df;ago2/ago2 double mutants showed no significant differences in the splicing pattern (Figure 33A; Lane 3). Also, ago1/Df;rrp6/Df double mutants revealed no obvious differences in the splicing pattern. Exon 4.12 showed a 1.8-fold decrease in inclusion (Figure 33A; Lane 5).
Figure 31: Schematic representation of generating germ line clones using the dominant female sterile (DFS) technique. Mitotic recombination occurs between chromosomes with genotypes DFS FRT/lethal FRT. The FRT is inserted proximally to both DFS and the lethal mutation. On heat shock treatment, the heat shock promoter (on another chromosome) gets activated and provides recombinase activity which catalyses the site-specific chromosomal exchange at the FRT sequences. FLP-recombinase target sequences (FRT). Dominant Female Sterile (DFS). Recessive zygotic lethal mutation (lethal) Adapted from (Perrimon, 1998).
Figure 32: PCR validation showing recombined FRT sites on the same chromosome as the (A) ago1 and (B) rrp6 lethal mutation. Primers pUCHneoF1 and pUCHneoR1 flanking the FRT sites on 42B and primers FRT F1 and Car3'InvR1 flanking the FRT sites on 82B were used to carry out single fly PCRs on FRTago1/CyoDfdGFP and FRTrrp6/TM6DtdGFP flies, respectively. The starting fly lines, elavGAL4ago1/CyO and elavGAL4rrp6/TM3Sb served as negative controls. The FRTL/SM6 and FRTSb/TM6 lines served as positive controls for FRT sites. In both the gels, Canton S males were used as negative controls.
Figure 33: Analysis of ago1 and rrp6 maternal mutants and redundancy between Agos and rrp6 in Dscam splicing regulation (A) Analysis of Dscam exon 4 splicing pattern using RNA extracted from 14-18 h old mutant embryos of ago1 z&m (Lane 2), ago1/Dr; ago2/ago2 (Lane 3), rrp6 z&m (Lane 4) and ago1/Dr; rrp6/Dr (Lane 5). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 21. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. z&m = zygotic and maternal mutants. M=50bp ladder (NEB).
Germline clones for \textit{ago1} and \textit{rrp6}, unexpectedly showed no obvious changes in \textit{Dscam} exon 4 splicing pattern (Figure 33A; Lanes 2 and 4).

### 3.7. Overexpression of a single \textit{Dscam} isoform does not reinforce selection of the same variable exons

The defects that were found in \textit{ago1} and \textit{rrp6} zygotic mutants were not conclusive regarding an RNA mediated mechanism involved in \textit{Dscam} variable exon selection. Since this does not rule out that such a mechanism could apply, it was required to test if overexpression of a single \textit{Dscam} transcript species can reinforce splicing of the same isoform. Therefore, a \textit{Dscam} cDNA construct was generated that includes its entire 5'UTR and also includes the endogenous 3' end-processing site as it was reasoned that cytoplasmic cleavage and/or polyadenylation might be involved in producing a relay of expression information to the nucleus and force selection of specific exons. To be able to distinguish the endogenous \textit{Dscam} transcripts from the overexpressed exogenous isoform, a restriction site was included in front of exon clusters 4, 6 and 9, which is absent in the variable downstream cluster to cut off the label from PCR products that originate from the transfected cDNA (Figure 34).

\textit{pAc5.1A Dscam Mut Exons 4-9}, comprising a single \textit{Dscam} isoform with variable exons 4.6, 6.32, 9.25 and 17.2, was transfected into S2 cells (as mentioned in 4.5.4). The splicing pattern of the endogenous \textit{Dscam} was analyzed for exon clusters 4 and 9 as before, but also including \textit{BsrGI} and \textit{XbaI} to cut the label off from the exogenous transcripts originating from the transfected cDNA construct. This analysis revealed that the splicing pattern of endogenous \textit{Dscam} exon 4 and 9 cluster did not change, in particular did not reinforce the inclusion of the exogenously expressed variable
Figure 34: Diagrammatic representation of distinguishing the endogenous Dscam exon 4 variants from the exogenously overexpressed Dscam single isoform, containing exon variant 4.6. A BsrGI restriction site added before the variable exon cluster was used to cut off the 5' radioactive label. The dark circle (•) represents $^{32}$P used to label the forward primer YH Dscam 3F2.
3.8. Dscam intracellular signaling is not involved in regulating its own splicing pattern

The cytoplasmic domain of Dscam interacts with Dock, an SH3-SH2 adaptor protein, and Pak, a serine threonine kinase, to translate the recognized guidance signals into changes in the actin-based cytoskeleton and promote axon guidance (Schmucker et al., 2000). Individual mushroom body and da sensory neurons from an otherwise identical population acquire a unique splicing pattern resulting in canalization of a broad variety of Dscam isoforms into just a few unique isoforms (Matthews et al., 2007; Zhan et al., 2004).

To test if Dscam intracellular signaling is involved in canalization or in any form of exon selection, a set of transposon inserts that result in secreted Dscam (Figure 3A; mutants A and B), membrane bound Dscam without intracellular domain (Figure 3A; mutants C, D and E) or signaling compromised Dscam (Figure 3A; mutant F) were used. Analysis of Dscam exon 9 splicing showed no pronounced effects in any of the analyzed mutants. Mutants D and E with transposons after the transmembrane domain behaved similarly, by upregulating exons 9.2 and 9.5 and downregulating exon 9.33 by over three-fold (Figure 3B; Lanes 5 and 6). Mutants C and F downregulated exon 9.7 by 7.2-fold. Although minor differences were observed, however, none of the changes observed in splicing pattern were striking (Figure 3B; Lanes 4 and 7). Mutants A and B produced the least effect in changing Dscam exon 9 splicing (Figure 3B; Lanes 2 and 3).
Figure 35: Overexpression of a single Dscam isoform does not reinforce selection of the same variable exons. Analysis of endogenous Dscam variable exon 4.6 (A) and variable exon 9.25 (B) splicing in response to transfection of exogenous Dscam construct pAc5.1A Dscam Mut Exons 4-9 into S2 cells by using DDAB (Lane 2) or TRANS-IT (Lane 3). Untransfected S2 cells served as control (Lane 1). Dscam exon variants were separated in A and B as explained in figure legend 21 and 22, respectively. Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 36: Dscam intracellular signaling is not involved in regulating its own splicing pattern. (A) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of P[GawB]DscamNP3327 (mutant A; Lane 2), P[RS3]DscamCB-0486-3 (mutant B; Lane 3), P[XP]Dscam210584 (mutant C; Lane 4), P[EPgy2]DscamEY08620 (mutant D; Lane 5), PBac[SAstopDsRed]LL01770 (mutant E; Lane 6) and PBac[RB]Dscam604629 (mutant F; Lane 7). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
3.9. Analysis of mRNA methylation for a role in \textit{Dscam} splicing regulation

The mRNAs of higher eukaryotes can be variably methylated at the cap or internally (Motorin). At the cap, the ribose of the first and second nucleotides can be 2'-O-methylated sequentially by methyltransferases MTr1 and MTr2, respectively (Belanger et al., 2010; Kruse et al., 2011; Werner et al., 2011). The first 2'-O-methylation results in the formation of cap1 and the subsequent methylation results in the formation of cap2. Cap methylations are also present in U snRNAs and seem to be essential for splicing regulation (Donmez et al., 2004). Internally, mRNAs can be methylated at the N\textsubscript{6} position of adenosine and this modification is introduced by a multimeric complex containing the catalytic activity of MT-A70 (Bokar et al., 1997). Intriguingly, the \textit{Arabidopsis thaliana} homolog of MT-A70 interacts with the homolog of the \textit{Drosophila} splicing regulator Female-Lethal 2D and m6A modification has been found in the vicinity of splice sites (Zhong et al., 2008). Thus, it was reasoned that nuclear mRNA methylation, either at the cap or internally might affect \textit{Dscam} splicing by fixing a specific splicing pattern or positively marking specific exons for inclusion.

Analysis of \textit{Dscam} exon 9 splicing in mutants for cap 2'-O-methyltransferases, dMTr1 and dMTr2 and double mutant dMTr1/dMTr2 revealed no significant changes in splicing pattern (Figure 37C; Lanes 2, 3 and 4). Analysis of \textit{Dscam} exon 9 splicing in mutants for m6A methylase (\textit{dMT-A70}) and two other putative RNA methylases (CG7818 and CG14906) downregulated the inclusion of exon 9.7 by less than two-fold (Figure 37A; Lanes 2, 3 and 4). No other significant differences were observed.
Figure 37 (A,B): Analysis of mRNA methylation for a role in Dscam splicing regulation. (A) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of internal mRNA methylases dMT-A70 (Lane 2), CG7818 (Lane 3) and CG14906 (Lane 4). Canton S embryos served as control (Lane 1). (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
Figure 37 (C,D): Analysis of mRNA methylation for a role in Dscam splicing regulation. (C) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of internal cap methylases dMTr1 (Lane 2), dMTr2 (Lane 3) and dMTr1/dMTr2 (Lane 4). Canton S embryos served as control (Lane 1). (D) Heat map representation of Dscam splicing changes observed in C. Dscam exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
3.10. Analysis of *Dscam* exon 9 splicing regulation in mutants of genes differentially regulated in *elav* mutants

Since *Dscam* exon 9 splicing is particularly affected in *elav* mutants, the introns of the exon 9 cluster were analyzed for potential ELAV binding sites. Although, some U-rich sequences are present in this exon cluster, these sequences are not *bona fide* ELAV binding sites and also no experimental evidence could be obtained so far which indicates that ELAV directly binds to *Dscam* pre-mRNA (Haussmann and Soller, unpublished). Gene expression analysis in *elav* mutants, however, has shown that numerous genes encoding RNA binding proteins, DNA binding proteins and proteins involved in chromatin remodeling are massively differentially expressed (Figure 38) suggesting that ELAV might regulate *Dscam* splicing indirectly. Therefore, all available mutants were obtained from these categories of genes (56 genes) that are differentially regulated in *elav* mutants to test their contribution in regulating *Dscam* splicing (refer A11 (A-H) in appendix). As a control to this analysis, mutants of metabolic genes were used that were differentially regulated in *elav* mutants as it was anticipated that these genes would not have a role in *Dscam* splicing regulation (Figure 39). Analysis of *Dscam* exon 9 splicing in mutants of genes differentially regulated in *elav* mutants revealed that *Dscam* splicing is affected by many genes, the pattern, however, is not random (refer A11 (A-H) in appendix). When taking exons into consideration that were two-fold or more differentially regulated, it was found that a group of exons tended to be preferentially upregulated in a set of mutants while a different group tended to be preferentially downregulated in a different set of mutants. For upregulated exons, it was found that exons 9.2, 9.7, 9.8, 9.25 and 9.26 (Figure 39; Columns 2, 5, 6, 13 and 14) were mostly upregulated in
Figure 38: Graphical representation of proportions of genes differentially regulated in *elav* mutants. RNA binding proteins, transcription factors and chromatin remodeling genes account for 12%, 11% and 12%, respectively of all genes differentially regulated in *elav* mutants.
Figure 39: Heat map representation of *Dscam* exon 9 splicing analysis in mutants of genes that are differentially regulated in *elav* mutants. CR=chromatin remodeling; T=transcription factors; R=RNA binding proteins; M=metabolic genes; O=others, including zinc finger proteins (CG8108 and CG9293), RNA-dependent RNA polymerase (d-elp1), nuclear receptor coactivator (neos).
mutants of chromatin remodeling factors caf1, His2AV and dek (Figure 39; Rows 8, 18 and 19). Similar upregulation of these exons was also observed in mutants for RNA binding proteins zuc and squ, involved in rasiRNA generation (Figure 39; Rows 38 and 39, (Pane et al., 2007).

For downregulated exons, it was found that exons 9.4, 9.19, 9.20, 9.22, 9.24 and 9.33 (Figure 39; Columns 3, 8, 9, 10, 12 and 15) were mostly downregulated in mutants of chromatin remodeling factors CG9418, CG3995 and mcm7 (Figure 39; Rows 1, 2 and 27). Also, mutants for pnn, a splicing factor and neos, a nuclear receptor coactivator, showed a pronounced decrease in inclusion of most of these exons (Figure 39; Rows 41 and 56).

I also observed that some mutants regulated the splicing pattern of Dscam exon 9 in a similar manner by affecting the same groups of exons. Mutants of chromatin remodeling genes snr1, His2AV, dek, spt4, not and Df-1 (Figure 39; Rows 17-22) showed a very similar Dscam exon 9 splicing pattern by upregulating inclusion of exons 9.7, 9.8, 9.23 and 9.25 (Figure 39; Columns 5, 6, 11 and 13) and downregulating inclusion of exons 9.1, 9.2, 9.19 and 9.20 (Figure 39; Columns 1, 2, 8 and 9). Also mutants of zuc and squ regulated Dscam exon 9 splicing pattern in a very similar way (Figure 39; Rows 38 and 39). Contrary to expectation, however, no mutants were found where only one exon variant was affected, e.g. one exon that was massively upregulated or downregulated. This suggested that exon choice is regulated by combinatorial interactions of RNA binding proteins, DNA binding proteins and chromatin remodeling factors. Furthermore, Dscam exon 9 splicing seems to be generally sensitive to genetic perturbation as changes were found in mutants of metabolic genes (Figure 39; Rows 47-52).
3.11. RNA pol II processivity does not influence splicing of Dscam

The processivity of RNA pol II has been shown to be critical for inclusion of the fibronectin EDI exon such that a slow polymerase favours inclusion while a fast polymerase results in preferential exclusion of the alternative exon (Kornblihtt et al., 2004). Accordingly, it was speculated that polymerase processivity might be important for Dscam variable exon selection. In particular, pausing of RNA pol II at a specific exon variant could relieve its repressed state and result in its inclusion. To test if there is a correlation between the speed of the polymerase and inclusion of Dscam variable exons, the Drosophila C4 mutant of the RNA pol II subunit 215, which has a lower elongation rate was used. In addition, two inhibitors of RNA pol II elongation rate namely, 6-Azauracil and dichlororibofuranosylbenzimidazole (DRB) were used to analyze Dscam exon 9 splicing in S2 cells (Hrzenjak et al., 2006; Iglesias-Gato et al.; te Poele et al., 1999). If indeed a slow polymerase would affect inclusion of variable exons, my expectation was to see preferential inclusion of cluster proximal exons over cluster distal exons resulting in a polar effect. Alternative splicing of exon 23a of Nf1 gene has been shown to be affected by histone deacetylase activity such that inhibiting histone deacetylation results in an increased local RNA pol II elongation rate resulting in skipping of the alternative exon (Zhou et al., 2011). Thus, sodium valproate (Na-valproate), a potent inhibitor of histone deacetylase, was used to test if Dscam splicing is regulated by histone modifications. Analysis of Dscam exon 9 splicing in flies with a C4 mutation (Figure 40A; Lane 2) and S2 cells treated with 6-Azauracil, DRB or sodium valproate (Figure 40A; Lanes 4, 6 and 8) showed no significant difference in the overall splicing pattern. In the C4
Figure 40: RNA pol II processivity does not influence splicing of *Dscam*. (A) Analysis of *Dscam* exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of C4 slow polymerase (Lane 2) and using RNA extracted from S2 cells after 12 h of exposure to 100 µg/ml azauracil (Lane 4), 100 µg/ml DRB (Lane 6) and 0.83 mg/ml Na-valproate (Lane 8). Canton S embryos served as control for the C4 mutant (Lane 1); solvents NH4OH (Lane 3), DMSO (Lane 5) and water (Lane 7) served as controls for azauracil, DRB and Na-valproate, respectively. (B) Heat map representation of *Dscam* splicing changes observed in A. Exon 9 variants not detected are represented in grey. *Dscam* exon variants were separated as explained in figure legend 22. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gel. M=50bp ladder (NEB).
RNA pol II mutant, exon 9.26 was severely downregulated by over ten-fold and exon 9.5 was not detected. However, this change does not support the relation between lower elongation rate and increased inclusion of alternative exons as none of the exons showed significant upregulation.

3.12. *Dscam* diversity is not generated by differential stability of isoforms

Although the sequence contribution of variable exons to the overall length of *Dscam* mRNA is minor (<0.1%) and variable exons have very similar sequences, these sequences could affect stability of *Dscam* mRNAs. To test if sequence variation provided by the variable exons contributed to the stability of *Dscam* isoforms, and this way affected steady state levels of different isoforms, the stability of *Dscam* mRNAs was analyzed in S2 cells by stopping transcription with Actinomycin D (Murph et al., 2007). Analysis of *Dscam* exon 4 splicing pattern over a time course of 2 h, 4 h and 8 h post Actinomycin D treatment showed no significant change in the levels of *Dscam* exon 4 variants, which suggested that Dscam diversity, is not generated by differential stability of its isoforms (Figures 41A and B).

3.13. Development of a *Dscam* exon 9 reporter transgene recapitulating endogenous exon 9 splicing

To analyze the sequence elements directing the inclusion of exon 9 variants, a reporter gene is required. Therefore, the exon 9 variable cluster was inserted with
Figure 41: Dscam diversity is not generated by differential stability of isoforms. (A) Analysis of Dscam exon 4 splicing by using RNA extracted from S2 cells exposed to 10µg/ml Actinomycin D over a time course of 2 h (Lane 4), 4 h (Lane 5) and 8 h (Lane 6). S2 cells treated with DMSO over a time course of 2 h (Lane 1), 4 h (Lane 2) and 8 h (Lane 3) served as controls. (B) Heat map representation of Dscam splicing changes observed in A. Dscam exon variants were separated as explained in figure legend 21. Unspecific bands are indicated by asterisks (*). Samples were run on an 8% denaturing polyacrylamide gels. M=50bp ladder (NEB).
flanking constant exons 7 and 8 at the distal end and exons 10 and 11 at the proximal end into a UAS reporter construct and generated a transgenic line at a defined genomic location using phiC31 recombination (Bischof et al., 2007; Venken et al., 2006). Using the phiC31 transformation system, allows insertion of modified reporter transgenes at exactly the same position to control for position effects. To be able to distinguish endogenous Dscam transcripts originating from the reporter transgene, a reporter sequence for reverse transcription was included. Analysis of a transgene inserted at the landing site at 76A2 on the third chromosome revealed that it recapitulated the splicing pattern observed from endogenous Dscam when expressed with elavGAL4 (Figures 42A; Lane 3 and 43B).
Figure 42: Development of a Dscam exon 9 reporter transgene recapitulating endogenous exon 9 splicing (A) Analysis of endogenous Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old Canton S (Lane 1) and elavGal4/UAS Dscam 9L (Lane 2) embryos; and exogenous Dscam exon 9 splicing pattern in elavGal4/UAS Dscam 9L embryos (Lane 3). Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. (B) Western blot confirming expression of UAS Dscam 9L trasngene driven by elavGal4. M=50bp ladder (NEB).
CHAPTER 4
DISCUSSION

The most striking feature about the Dscam gene is its enormous molecular diversity generated by mutually exclusive splicing of its four variable exon clusters, namely 4, 6 and 9 and 17 comprising 12, 48, 33 and 2 variable exons, respectively. Although, all exons in the variable clusters have splice sites comparably similar to the consensus sequence, only one exon is included at a time and no splicing together of adjacent exons is observed (Schmucker et al., 2000). It is still unclear how these exon clusters are kept in a repressed state and how inclusion of specific exons by release of this repression is regulated. Particularly intriguing, pathogen exposure in mosquitoes results in a change of AgDscam splicing pattern favouring inclusion of isoforms that show higher binding affinity towards pathogen recognition (Dong et al., 2006). These experiments suggest that Dscam splicing is under strict regulation by exogenous and endogenous cues. In the immune system, these cues are pathogens, which are recognized by Dscam and endocytosed by haemocytes and in the nervous system they are the neighbouring neurons that need to acquire a different complement of Dscam isoforms to allow for axon bifurcation in mushroom body neurons and generation of overlapping dendritic fields in dendritic arborization sensory neurons (Dong et al., 2006; Matthews et al., 2007; Zhan et al., 2004).

The work described in the present thesis established a method to analyze Dscam splicing in the variable exon clusters where all exons have nearly the same length, but have sequences divergent enough to allow their separation on denaturing acrylamide gels by prior digestion with restriction enzymes. With this approach, the
exon 4 cluster was completely resolved and more than 50% of the exon 9 variants could be separated. However, the exon 6 cluster resolution was limited due to the large number of variables. The main limiting factor in exon 6 cluster is the short lengths of individual exon variants and their limited sequence variability, thus limiting the choice of restriction enzymes. Given that next generation sequencing techniques, (e.g. Illumina and 454 sequencing) are becoming cost effective, they can be applied to analyze Dscam splicing. One limitation to apply e.g. Illumina sequencing for amplicon analysis is the large capacity these machines have been currently designed for requiring combining multiple samples that are identified by unique bar codes. Next generation sequencing would also aid in recognizing the unexpected bands that were observed during the study, which could either be PCR artifacts or sequence associated mobility shifts. It is highly unlikely that the unexpected bands could be generated due to polymorphisms, resulting in generation or omission of a restriction site, because an almost identical Dscam splicing pattern was observed between numerous different strains.

A key feature of Dscam in mosquitoes is the change in its splicing pattern on pathogen exposure (Dong et al., 2006), and it was demonstrated that pathogen exposure also induced changes in Dscam splicing in Drosophila. Contrary to the expectation, however, the changes observed in Dscam splicing in Drosophila were rather mild with either heat inactivated or untreated E. coli resulting in differential regulation of only a few exons.

It remains to be shown that the isoforms that were highly upregulated show very high binding affinity with the treated E. coli. Adaptation to pathogen exposure by producing high affinity pathogen binding isoforms has also been observed in crayfish.
Moreover, coating pathogens with these high affinity binding isoforms allows the pathogens to escape phagocytosis (Watthanasurorot et al., 2011). Thus, *Dscam* splicing across arthropods appears to have a common feature of being sensitive to pathogenic triggers such that they alter *Dscam* diversity to generate a repertoire that allows better recognition, stronger binding and effective defense to the host against the invading pathogen.

*Dscam* splicing pattern changes upon pathogen exposure or in certain developing neurons to unique signatures of variable exons for immune defense or neuronal wiring, respectively (Dong et al., 2006; Matthews et al., 2007; Neves et al., 2004). However, it was found that the *Dscam* splicing pattern in individual flies is quite robustly maintained and seems not to change between individuals. This suggested that variation in *Dscam* splicing might be specific to certain cell types such haemocytes, mushroom body neurons and da sensory neurons. A different mode of *Dscam* splicing regulation is indeed indicated from the analysis of projection of mechanosensory neurons in the ventral nerve cord. Here, reduced *Dscam* variability results in projection defects, which intriguingly had no effect on axon branching in mushroom body neurons. These results indicate that isoforms, specific to certain neurons, have counterparts in the areas where these neurons project to establish connectivity with their target cells (Chen et al., 2006). In this context, *Dscam* would acquire a cell type specific splicing pattern, which would need to be recapitulated in those cells located in the remote location in the brain. Since many neurons and also their precursors migrate during neuronal development, a common origin could result in the same *Dscam* splicing pattern. For example, segmental identity generated by
hox gene expression might be sufficient to generate a regionalized Dscam spliceform repertoire.

The very reproducible splicing pattern observed in various developmental contexts, makes it possible to interrogate the regulation of Dscam splicing by using mutants in various candidate genes for splicing regulation. One such candidate that was analyzed is ago1, which has similar developmental defects in the nervous system as Dscam (Kataoka et al., 2001).

Argonautes are further attractive candidates as they suggest an RNA mediated mechanism involved in reinforcement of splicing to a specific isoform upon encounter of a trigger via a relay of signal from the cell surface to the nucleus. Further analysis of mutants in genes involved in RNA mediated regulation and over expression of a single isoform did not further support such a hypothesis of an RNA mediated mechanism in canalization of Dscam diversity. Rather, the effect of Argonautes, as well as of other factors identified in this study in the regulation of Dscam splicing are cluster specific suggesting unique regulatory mechanisms directed by cluster dedicated factors. In Giardia lamblia, an RNAi mechanism is involved in generating diversity of variant-specific surface protein (VSP) genes. Here, all VSP genes are transcribed at a similar rate but except for a single transcript, all others are degraded by RNAi. From time to time, expression of VSP protein switches from one to another by changing the transcript complements degraded by RNAi (Prucca et al., 2008).

Such a mechanism could potentially also apply to generate Dscam diversity. If this was the case, it could be expected that different Dscam isoforms would have different half lives resulting in different levels of inclusion of variable exons. When analyzing stability of Dscam isoforms, however no evidence for such a mechanism.
elav mutants show nervous system defects similar to Dscam mutants (Simionato et al., 2007) and indeed it was found that Dscam splicing regulation is particularly affected in exon 9 cluster. Analysis of the sequence in exon 9 cluster together with the lack of experimental evidence so far obtained for the identification of ELAV targets argues that elav regulates Dscam splicing indirectly. Indeed, many of genes differentially regulated in elav mutants encode for regulators of gene expression involving RNA binding proteins, DNA binding proteins and chromatin remodeling factors. Analysis of Dscam exon 9 splicing by mutants of such genes revealed that Dscam splicing is very sensitive to changes in mutants of regulators of gene expression. Intriguingly, many exons were observed to be upregulated or downregulated in a specific set of mutants. However, none of the mutants regulated the splicing of just one exon variant. The screen revealed that Dscam splicing is very sensitive to genetic perturbation.

### 4.1. Future work

Given the biological complexity of higher vertebrates, it is assumed that mammalian genes involved in development of the nervous system are more in number and more complex than invertebrate genes. Dscam is an exceptional example of a highly complex gene in relatively simple arthropods. Despite vast structural differences between the insect Dscam and mammalian DSCAM, they share important roles in neuronal patterning. Understanding the molecular mechanisms of Dscam function in different organisms and species would reveal important clues about the evolution of CAMs into recognition units involved in neural wiring and immune recognition.
The key to understanding the dual role of *Drosophila Dscam* in the immune and nervous systems lies in its unique mechanism of generating extraordinary diversity. It is still unclear, which external cues trigger a change in alternative splicing of *Dscam*. Also, a unique mechanism whereby a neuronal or pathogenic trigger may result in the selection of a specific exon or set of exons in each variable cluster is unknown. Last but not the least, it would be interesting to determine if the *Dscam* sequence itself has a role to play in exon selection.

It is crucial to find an external cue which would trigger a signal for change in *Drosophila Dscam* alternative splicing. S2 cells could be exposed to pure microbial cell surface determinants such as lipopolysaccharides or peptidoglycans and analyzed for changes in *Dscam* alternative splicing. Isoforms undergoing a change in inclusion can be silenced in S2 cells to test their ability to phagocytose pathogens expressing the specific cue. Contrarily, pathogens that are mutant for the specific cue can be used to challenge S2 cells to observe, possibly, no splicing changes in the affected alternative exons. An alternative approach of investigating the role of *Dscam* in the *Drosophila* immune system could be to expose flies, mutant for genes that play a role in the immune system (e.g. *imd, relish, dorsal* and *Myd88*) with live or attenuated pathogens and subsequently test for *Dscam* splicing changes in their haemocytes.

In this study, an exon 9 reporter transgene (*UAS Dscam 9L*), expressed heterologously, has been shown to recapitulate endogenous exon 9 splicing pattern, which suggests that the information regulating exon 9 alternative splicing lies within the cluster sequence itself. To further speculate which sequence elements within the cluster are crucial for selection of particular exon variants, different sequence altering
strategies may be adopted. Firstly, deletion of combinations of consecutive introns and exons such as, deletion of an intron and its preceeding or succeeding exon or an exon and its flanking introns, resulting in fusion of two alternate exons, may affect splicing of neighbouring exons. Secondly, reversing the sequence of an intron may influence the splicing of its adjoining exons. Thirdly, mutating the sequence of a particular intron could disrupt its RNA secondary structure which may alter the splicing of its flanking exons. Finally, swapping selective intronic sequences of one variable cluster into another may regulate its default splicing pattern such that an intron, which causes enhanced inclusion of a particular exon 4 variant, when inserted in the exon 9 cluster before or after an exon variant showing low expression, may allow enhanced splicing of that exon 9 variant. Such an approach may help in understanding whether intronic or exonic sequences are important for selection of a particular Dscam exon variant.

Dscam is an extrordinary model to understand how a single gene plays vital roles in the nervous and immune systems. Unravelling the enigma around the splicing mechanism of this unique pattern recognition receptor may throw light on unexplored mechanisms of immune defence in vertebrates.
Appendices

A1. Separation of Dscam 4 variants. The table shows all Dscam exon 4 variants, their annotated lengths, identifying enzymes and expected lengths on a denaturing polyacrylamide gel after restriction digest.

<table>
<thead>
<tr>
<th>Exon</th>
<th>Variable no.</th>
<th>Variable length</th>
<th>Enzyme</th>
<th>Fragment size</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
<td>162</td>
<td>-</td>
<td>238</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>162</td>
<td>Mbol</td>
<td>173</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>162</td>
<td>TaqI</td>
<td>157</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>162</td>
<td>TaqI</td>
<td>152</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>162</td>
<td>TaqI</td>
<td>151</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>171</td>
<td>Mbol</td>
<td>124</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>171</td>
<td>TaqI</td>
<td>123</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>171</td>
<td>HinP1I</td>
<td>106</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>171</td>
<td>TaqI</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>162</td>
<td>HinP1I</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>159</td>
<td>TaqI</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>168</td>
<td>AluI</td>
<td>61</td>
</tr>
</tbody>
</table>
### A2. Separation of *Dscam* 9 variants.

The table shows all *Dscam* exon 9 variants, their annotated lengths, identifying enzymes and expected lengths on a denaturing polyacrylamide gel after restriction digest.

<table>
<thead>
<tr>
<th>Exon</th>
<th>Variable no.</th>
<th>Variable length</th>
<th>Enzyme</th>
<th>Fragment size</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>24</td>
<td>288</td>
<td>HaeIII</td>
<td>261</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>299</td>
<td>MspI</td>
<td>253</td>
</tr>
<tr>
<td>9</td>
<td>22</td>
<td>288</td>
<td>MspI</td>
<td>241</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
<td>288</td>
<td>HpyCH4IV</td>
<td>233</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>288</td>
<td>HaeIII</td>
<td>231</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>288</td>
<td>HpyCH4IV</td>
<td>197</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>291</td>
<td>Xmnl</td>
<td>174</td>
</tr>
<tr>
<td>9</td>
<td>23</td>
<td>288</td>
<td>MspI</td>
<td>169</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>288</td>
<td>Xmnl</td>
<td>158</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>288</td>
<td>HaeIII</td>
<td>151</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>288</td>
<td>BstNI</td>
<td>134</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>288</td>
<td>BstNI</td>
<td>134</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>288</td>
<td>BstNI</td>
<td>134</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>291</td>
<td>HpyCH4IV</td>
<td>132</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>288</td>
<td>MspI</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>288</td>
<td>BstNI</td>
<td>108</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>288</td>
<td>MspI</td>
<td>107</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>291</td>
<td>BstNI</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>288</td>
<td>BstNI</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>27</td>
<td>288</td>
<td>BstNI</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>306</td>
<td>HaeIII</td>
<td>90</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>291</td>
<td>BstUI</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>291</td>
<td>MspI</td>
<td>85</td>
</tr>
<tr>
<td>9</td>
<td>33</td>
<td>294</td>
<td>HpyCH4IV</td>
<td>85</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>291</td>
<td>BstUI</td>
<td>73</td>
</tr>
<tr>
<td>9</td>
<td>31</td>
<td>285</td>
<td>HpyCH4IV</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>291</td>
<td>HaeIII</td>
<td>66</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td>285</td>
<td>HaeIII</td>
<td>63</td>
</tr>
<tr>
<td>9</td>
<td>29</td>
<td>294</td>
<td>MspI</td>
<td>52</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>288</td>
<td>BstNI</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>291</td>
<td>BstUI</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>32</td>
<td>279</td>
<td>BstNI</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>288</td>
<td>BstUI</td>
<td>32</td>
</tr>
</tbody>
</table>
A3. Separation of Dscam 6 variants. The table shows all Dscam exon 6 variants, their annotated lengths, identifying enzymes and expected lengths on a denaturing polyacrylamide gel after restriction digest.

<table>
<thead>
<tr>
<th>Exon</th>
<th>Variable no.</th>
<th>Variable length</th>
<th>Enzyme</th>
<th>Fragment size</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>20</td>
<td>124</td>
<td>-</td>
<td>152</td>
</tr>
<tr>
<td>6</td>
<td>29</td>
<td>124</td>
<td>Alul</td>
<td>145</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>121</td>
<td>Alul</td>
<td>136</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>124</td>
<td>BstNI</td>
<td>127</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>124</td>
<td>Alul</td>
<td>124</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>124</td>
<td>Alul</td>
<td>124</td>
</tr>
<tr>
<td>6</td>
<td>37</td>
<td>124</td>
<td>BstNI</td>
<td>121</td>
</tr>
<tr>
<td>6</td>
<td>46</td>
<td>124</td>
<td>BstNI</td>
<td>121</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>127</td>
<td>Alul</td>
<td>104</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>124</td>
<td>Alul</td>
<td>101</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>124</td>
<td>BstNI</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>124</td>
<td>Alul</td>
<td>95</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>124</td>
<td>BstNI</td>
<td>94</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>124</td>
<td>Alul</td>
<td>93</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>124</td>
<td>Alul</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>44</td>
<td>124</td>
<td>Taql</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>124</td>
<td>BstUI</td>
<td>89</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>118</td>
<td>Alul</td>
<td>87</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>124</td>
<td>Mbol</td>
<td>86</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>124</td>
<td>Mspl</td>
<td>84</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>124</td>
<td>BstUI</td>
<td>83</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>124</td>
<td>Mbol</td>
<td>81</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>124</td>
<td>BstUI</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>124</td>
<td>BstUI</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>124</td>
<td>Taql</td>
<td>74</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>124</td>
<td>Alul</td>
<td>74</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>124</td>
<td>Alul</td>
<td>72</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
<td>124</td>
<td>Alul</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>124</td>
<td>Mbol</td>
<td>66</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>124</td>
<td>BstUI</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>124</td>
<td>Mspl</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>116</td>
<td>Mspl</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>124</td>
<td>BstUI</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>124</td>
<td>Mbol</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>124</td>
<td>Mbol</td>
<td>53</td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>124</td>
<td>Mbol</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>124</td>
<td>Mbol</td>
<td>53</td>
</tr>
<tr>
<td>6</td>
<td>38</td>
<td>124</td>
<td>Alul</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>127</td>
<td>Alul</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>128</td>
<td>Alul</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>124</td>
<td>Alul</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>48</td>
<td>121</td>
<td>Alul</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>43</td>
<td>124</td>
<td>BstNI</td>
<td>42</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>124</td>
<td>Alul</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>127</td>
<td>Mbol</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>121</td>
<td>Taql</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>133</td>
<td>Mspl</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>121</td>
<td>Mspl</td>
<td>31</td>
</tr>
</tbody>
</table>
A4. Crossing scheme showing generation of zygotic ago1 mutants

\[ \begin{align*}
\text{elav Gal4 ago1} & \quad \text{Df ago1 UAS-GFP} \\
\text{CyO} & \quad \text{CyO} \\
\text{elav Gal4 ago1} & \quad \text{Df ago1 UAS-GFP} \\
\end{align*} \]
A5. Crossing scheme showing generation of zygotic \textit{rrp6} mutants

\[ \text{UAS GFP Df}\textit{rrp6} \quad \text{Elav Gal4}\textit{rrp6} \quad X \quad \text{Tm3Sb} \quad \text{Tm3Sb} \]

\[ \text{Elav Gal4}\textit{rrp6} \quad \text{UAS GFP Df}\textit{rrp6} \]

\begin{tabular}{|c|c|}
\hline
\text{UAS GFP Df}\textit{rrp6} & \text{Tm3Sb} \\
\hline
\text{Elav Gal4}\textit{rrp6} & \text{No GFP Expression} \\
\text{UAS GFP Df}\textit{rrp6} & \text{Tm3Sb} \\
\hline
\text{Elav Gal4}\textit{rrp6} & \text{No GFP Expression} \\
\text{UAS GFP Df}\textit{rrp6} & \text{Tm3Sb} \\
\hline
\text{Tm3Sb} & \text{Lethal} \\
\text{Tm3Sb} & \text{Tm3Sb} \\
\hline
\end{tabular}
A6. Crossing scheme showing generation of zygotic elav mutants
A7. Crossing scheme showing generation of \textit{ago1/Df; ago2/ago2} double mutants
A8. Crossing scheme showing generation of *ago1/Df; rrp6/Df* double mutants
A9. Crossing scheme showing generation of ago1 zygotic and maternal mutants

Crossing scheme showing generation of ago1 zygotic and maternal mutants

Gene    | Cytology | Genetic Location
--------|----------|-------------------
ago1    | 50C9-50C17 | 2-70.0
FRT     | 42B      | 2-55.3
L       | 51A4     | 2-72.0
elav Gal4 unknown × Morgans relative to ago1

Single Males

Df ago1 UAS-GFP elav Gal4 FRT ago1/FRT ago1/elav Gal4 ago1/elav Gal4 FRT/elav Gal4 FRT L/FRT L
CyO DfdGFP ×   CyO   CyO CyO CyO CyO

A B C D Discard
(Lobe phenotype)

A elav Gal4 FRT ago1 CyO DfdGFP Stock

B FRT ago1 CyO DfdGFP Stock

C elav Gal4 ago1 CyO DfdGFP

D elav Gal4 FRT CyO DfdGFP

Neural GFP Expression
Lethal
FRT PCR Positive

Neural GFP Expression
Lethal
FRT PCR Positive

Neural GFP Expression
Lethal
FRT PCR Negative
Discard

Neural GFP Expression
Viable; Straight winged flies
FRT PCR Positive
Discard
continued...

```
hsflp  FRT  RBP9  e5  elavOH  yw  Sco  \\
♀       ♂      ♂              ♂          ♂          ♂  \\
CyO_GFP

hsflp  FRT  RBP9  e5  elavOH  yw  Sco  \\
♀       ♂      ♂              ♂          ♂          ♂  \\
CyO

hsflp  +  FRT  RBP9  e5  elavOH  yw  Sco  \\
♀       ♂      ♂              ♂          ♂          ♂  \\
CyO_GFP

hsflp  +  FRT  RBP9  e5  elavOH  yw  Sco  \\
♀       ♂      ♂              ♂          ♂          ♂  \\
CyO

hsflp  ovoD FRT  FRT  elav Gal4 ago1  \\
♀       ♂      ♂        ♂                ♂  \\
CyO_GFP

hsflp  ovoD FRT  FRT  elav Gal4 ago1  \\
♀       ♂      ♂        ♂                ♂  \\
CyO

(ovaries)  +  ovoD FRT  \\
♀       ♂

Heat Shock Treatment

FRT  elav Gal4 ago1  Df ago1 UAS-GFP  \\
♀       ♂

FRT  elav Gal4 ago1  \\
♀       ♂

FRT  elav Gal4 ago1  Df ago1 UAS-GFP  \\
♀       ♂

Neural GFP expression  No neural GFP expression
```
A10. Crossing scheme showing generation of *rrp6* zygotic and maternal mutants

![Diagram showing the generation of *rrp6* mutants through crossing schemes.](image-url)
continued…

![Diagram of genetic crosses and crosses with specified genotypes](Image)
A11. Analysis of Dscam exon 9 splicing in mutants of genes that are differentially regulated in elav mutants. (A) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of CG9418 (Lane 3), CG3995 (Lane 4), CG8149 (Lane 5), sirt7 (Lane 7) and set (Lane 9). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
(B) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of zuc (Lane 3), snr1 (Lane 4), squ (Lane 5), his2AV (Lane 6), dek (Lane 7), spt4 (Lane 8), not (Lane 9) and df-31 (Lane 10). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
(C) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of neos (Lane 3), zf30c (Lane 4), geminin (Lane 5), CG3523 (Lane 6), CG8105 (Lane 7) and msl-1 (Lane 8). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
(D) Analysis of *Dscam* exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of *hel25E* (Lane 3), *d-elp1* (Lane 4), *CG15514* (Lane 5), *CG14965* (Lane 6), *CG9293* (Lane 7), *CG8460* (Lane 8), *CG4266* (Lane 9) and *CG9512* (Lane 10). *Canton S* embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. *Dscam* exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
(E) Analysis of *Dscam* exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of *pnn* (Lane 3), *CG6650* (Lane 4), *top2* (Lane 5), *CG3797* (Lane 6), *sfmbt* (Lane 7), *mcm7* (Lane 8) and *nufip* (Lane 9). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. *Dscam* exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
(F) Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of xl6 (Lane 2), mcm10 (Lane 3), CG14710 (Lane 5) and sup-07 (Lane 8). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
Analysis of *Dscam* exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of *CG3238* (Lane 6), *swa* (Lane 8), *cathD* (Lane 9), *CG12877* (Lane 10) and *CG8378* (Lane 11). *Canton S* embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. *Dscam* exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/Hinfl marker (Biotools).
Analysis of Dscam exon 9 splicing pattern using RNA extracted from 14-18 h old mutant embryos of tip60 (Lane 3), kdm4a (Lane 4), caf-1 (Lane 5), CG2051 (Lane 6), CG9883 (Lane 7), hp5 (Lane 8), chrac-14 (Lane 9), beaf-32 (Lane 10), rox-8 (Lane 11), ball (Lane 13), jigr-1 (Lane 16), rpd3 (Lane 18) and iswi (Lane 19). Canton S embryos served as control (Lane 1). Lanes with red crosses were not included in the analysis. Dscam exon variants were separated as explained in figure legend 22. Samples were run on an 8% denaturing polyacrylamide gel. M=phiX174 DNA/HinfI marker (Biotools).
### A12. List of primers

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primers used to amplify Dscam variable exon clusters</strong></td>
<td></td>
</tr>
<tr>
<td>Dscam YH 3F2</td>
<td>GCAACCAGTTCCGGAAACCATTATCCTCC</td>
</tr>
<tr>
<td>Dscam YH 5R1</td>
<td>CGAGAGGGCAATTTGATGGTACTTCC</td>
</tr>
<tr>
<td>Dscam YH 5F1</td>
<td>GCCCAAAAGGGAGGATTTGCTATCAG</td>
</tr>
<tr>
<td>Dscam YH 7R1</td>
<td>CGGTTTCTTCTTACTGAATTTGCGATCC</td>
</tr>
<tr>
<td>Dscam YH 8F1</td>
<td>CAGCTCTGGAAATCTAATGATCTGG</td>
</tr>
<tr>
<td>Dscam YH 10R1</td>
<td>GCCTTTATCGGTTGGAAGCAGACTCCTGGGAGGTA</td>
</tr>
<tr>
<td>Dscam YH 16F1</td>
<td>GAATACGACTTTGCCACCTTTACGGTAC</td>
</tr>
<tr>
<td>Dscam YH 18R1</td>
<td>CCAATGTCTTACATTTAACTACATCG</td>
</tr>
<tr>
<td>Dscam YH 11RT1</td>
<td>CGAGACCTTCTTACATTTAACCTACATCG</td>
</tr>
<tr>
<td><strong>Primers used to clone pAc5.1A Dscam Mut Exons 4-9</strong></td>
<td></td>
</tr>
<tr>
<td>Dscam cDNA end F1</td>
<td>GTGCTCGCGCGCGATGATCCGCAAGAGG</td>
</tr>
<tr>
<td>Dscam cDNA end R1</td>
<td>ACCAATATCG</td>
</tr>
<tr>
<td>Dscam 6kb frag F1</td>
<td>AAATGTCGTTTACATCAATTTTCGTGTCTGTGGTCCG</td>
</tr>
<tr>
<td>Dscam 6kb frag R1 Xho Spe</td>
<td>GCGTCGCGGCCGCGATATCTCAGGTCAGAGTGCAATTTACGGATGAGTGT</td>
</tr>
<tr>
<td>Dscam 5’UTR F1</td>
<td>CGGTTCTGTACAGTGCAATATTCCATCTGGGAGGTA</td>
</tr>
<tr>
<td>Dscam 5’UTR1 R1</td>
<td>AGTGGTAAACACTCACACACACTTGTAGTGATGGTG</td>
</tr>
<tr>
<td>Dscam 5’UTR2 F2</td>
<td>TAATGCGATTAAAAAATAATTTGGCCAGCCGAC</td>
</tr>
<tr>
<td>Dscam 5’UTR2 R2 Not1 EcoRV</td>
<td>GCGTCGCGGCCGCGATGATCCGCAAGAGG</td>
</tr>
<tr>
<td>Dscam Exon 1-4 F1</td>
<td>GCCGCTTCTGAGACGCGGCCGACGCTGAAATAGGCTTGA</td>
</tr>
<tr>
<td>Dscam Exon 1-4 R1 BsrGI</td>
<td>GATATCGGCGCCGTAAAACTGATTCACAACGGCTCGCACATG</td>
</tr>
<tr>
<td>Dscam Exon 4-5 F1</td>
<td>CCGGGATGTACATGTGCGAGCCGTTGTGAATCAGTTTTACG</td>
</tr>
<tr>
<td>Dscam Exon 4-5 R1 EcoRI</td>
<td>GTCGTCCGCTTTTGTGGCACTTAATCGGGTTTCTCC</td>
</tr>
<tr>
<td>Dscam Exon 4-5 R1 Xho Not1 Xba HindIII Ase</td>
<td>GCCGCTTCTGAGACGCGGCCGACGCTGAAATAGGCTTGA</td>
</tr>
<tr>
<td>Dscam Exon 5-7 F1</td>
<td>TATGCTCATACAGGCGCTGTTAAGATCTGACGGGGAAAATACAAATG</td>
</tr>
<tr>
<td>Dscam Exon 5-7 R1 NgoMIV Spe Mlu BglII EcoRI</td>
<td>GCGTCGAGATGCGCTGCTGTGAAGCCTGTGCAAGATGGCTG</td>
</tr>
<tr>
<td>Dscam Exon 7-8 F1 NgoMIV</td>
<td>GAGATCGTAGATATGCTGCTGCTGAGATGCG</td>
</tr>
<tr>
<td>Dscam Exon 7-8 R1</td>
<td>GAGATCGTAGATATGCTGCTGCTGAGATGCG</td>
</tr>
<tr>
<td>Dscam Exon 8-13 F1</td>
<td>TAAAGATGCAAGTCACTGTTCCACCAAAATAGCCTGCCCTTC</td>
</tr>
<tr>
<td>Dscam Exon 8-13 R1 EcoRI</td>
<td>GGTGGCGACTTGATTTGCAAGGTTTCTTGAATCTTCCATG</td>
</tr>
<tr>
<td><strong>Primers used to validate presence of FRT sites</strong></td>
<td></td>
</tr>
<tr>
<td>pUCHsneoF1</td>
<td>GCCGCTTTTTCGGATTCACTGACGTGT</td>
</tr>
<tr>
<td>pUCHsneoR1</td>
<td>CAGATCTCGGACAGCAGTGGTTTCTTCCATAT</td>
</tr>
<tr>
<td>FRT F1</td>
<td>CAAAGCGTTCCGGAAAACGAGCGTTC</td>
</tr>
<tr>
<td>Car3invR1</td>
<td>GCTATTGCGGTGGGTTGAATATC</td>
</tr>
</tbody>
</table>

172
**Genotype of fly lines**

<table>
<thead>
<tr>
<th>Genotypes differentially regulated in elav mutants</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>y¹ w¹/+; P[Supor-P]CG9184 [KGO13780]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=Epgy2]CG3995[EY03827]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[y¹+t7.7, w¹+mC]=CyO]CG8149 DG19311</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[y¹+m8]=Mae-UAS.6.11]Top60[GG01739]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>[II]118; Sirt/5.Scer/SceI.RS</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=Epgy2]Set[EY09821]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=Epgy2]CG2051[EY21697]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[w¹+mC]=EP]CG9883[G17999]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=Epgy2]HP5[EY10901]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=BR.E.BR]=Supor-P]mus201[KGO1051] Chrac-14[KGO1051]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=BR.E.BR]=Supor-P]BEAF-32[KGO6904]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=GSV1]His2Av[GS3052]TM3, Sb¹ Ser¹</td>
<td>Kyoto</td>
</tr>
<tr>
<td>P[ry¹+t7.2]=PZ]Rpd3[04556] ry¹[506]/TM3, ry¹[RK] Sb¹ Ser¹</td>
<td>Bloomington</td>
</tr>
<tr>
<td>PBac[RB]Rpd3[e01851]</td>
<td>Harvard</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC] y¹+mDint2]=BR.E.BR]=Supor-P]Iswi[KGO3354]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>P[ry¹+t7.2]=PZ]Snr1[01319] ry¹[506]/TM3, ry¹[RK] Sb¹ Ser¹</td>
<td>Bloomington</td>
</tr>
<tr>
<td>w¹*; His2Av[810]TM3, Sb¹</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]Dek[k09907]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]spit4[k05316]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]notDC24306TM3, Sb¹ Ser¹</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[GSV6]GS16860/SIM1</td>
<td>Kyoto</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]geminin[14019]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[y¹+m8]=Mae-UAS.6.11]Top2[LA00892]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[w¹+mC]=EP]Smbt[G2280]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>w¹/I(wg)67c23, P[w¹+mC]=WH]Mcm7[03462]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[y¹+mDint2]=BR.E.BR]=Supor-P]Mcm10[KGO0233]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ sc¹[1]v¹[1]; P[y¹+t7.7, v¹[+1.8]=TRIP.HMS00608]attP2/TM3, Sb¹ Ser¹</td>
<td>Bloomington</td>
</tr>
<tr>
<td>PBac[RB]jigr1[e03251]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]wg130[Ck02506]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[y¹+mDint2]=BR.E.BR]=Supor-P]CG14965[KGO5766] ry¹[506]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>w¹/I(wg)67c23, P[w¹+mC]=WH]CG14710[006000]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>w¹*; P[GawB]Rox8NP0528 / TM3, Ser¹</td>
<td>Kyoto</td>
</tr>
<tr>
<td>w¹/I(wg)67c3; PBac[w¹+mC]=WH]CG18777[f01800]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[lacW]Heh25K11511/CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[Epgy2]EY06979</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[w¹+mC]=WH]CG4266[G7664]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=lacW]x16[k00230]CyO</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[Epgy2]CG32388[EY10295]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>v¹/swa¹/FM3</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹/I(wg)67c23, P[w¹+mC]=WH]CG8655[C151]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[w¹+mC]=EP]CG3797[G8880]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>y¹ w¹*; P[w¹+mC]=EP]CG88460[G18835]</td>
<td>Bloomington</td>
</tr>
<tr>
<td>C[1;Y]11; y¹[1]P[y¹+mDint2]=BR.E.BR]=Supor-P]Flo-2[KGO0162]C[1;DX, y¹[1][1]; ry¹[506]</td>
<td>Bloomington</td>
</tr>
</tbody>
</table>
### Genes affecting cell signaling

- **y**<sup>1</sup> w<sup>67c23</sup>; P[w+mC]<sup>y</sup><sup>+</sup>int2=EPgy2)/CG8108[EFY14316]/TM3, Sb<sup>1</sup> Ser<sup>1</sup>  
  Bloomington
- y<sup>1</sup> w<sup>67c23</sup>; P[EPgy2]/CG9393/EY13364  
  Bloomington

### Genes involved in mRNA methylation

- y<sup>1</sup> w<sup>67c23</sup>; P{EPgy2}CG9293/EY13364  
  Bloomington
- PBac(PB)CG10535c00296  
  Harvard

### RNA pol II processivity mutant

- y<sup>1</sup> w<sup>67c23</sup>; P{lacW}AGO1<sup>AGO1<sup>kb08121</sup></sup>/CyO  
  Bloomington

### Lines used in generation of germline clones

- y<sup>1</sup> w<sup>67c23</sup>; P{lacW}AGO1<sup>AGO1<sup>kb08121</sup></sup>/CyO  
  Bloomington
- w<sup>1118</sup>; PBac(w+[mC]=RB)/CG7818[ef00875]  
  Bloomington

### Genes involved in small RNA processing

- y<sup>1</sup> w<sup>67c23</sup>; P{lacW}AGO1<sup>AGO1<sup>kb08121</sup></sup>/CyO  
  Bloomington

### Lines expressing embryonic GFP marker

- w baz<sup>2</sup> P(FRT(w<sup>4</sup>))9-2/FM7a, P[Dfd-GMR-nvYFP]<sup>1</sup>  
  Bloomington
- w; noe<sup>CD</sup>/CyO, P[Dfd-GMR-nvYFP]<sup>2</sup>  
  Bloomington
- w; ry<sup>2</sup> <sup>2</sup>D<sup>3</sup>/TM3, P[Dfd-GMR-nvYFP]<sup>3</sup>, Sb<sup>1</sup>  
  Bloomington
- w; ry<sup>2</sup> <sup>2</sup>D<sup>3</sup>/TM6B, P[Dfd-GMR-nvYFP]<sup>4</sup>, Sb<sup>1</sup> T<sup>b</sup>b<sup>1</sup> ca<sup>1</sup>  
  Bloomington


• Harpaz, Y., and Chothia, C. (1994). Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. J Mol Biol 238, 528-539.


Lisbin, M.J., Qiu, J., and White, K. (2001). The neuron-specific RNA-binding protein ELAV regulates neuroglian alternative splicing in neurons and binds directly to its pre-mRNA. Genes Dev 15, 2546-2561.


• McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev 11, 3306-3318.


• Soller, M., and White, K. (2003). ELAV inhibits 3'-end processing to promote neural splicing of ewg pre-mRNA. Genes Dev 17, 2526-2538.


