Hippo pathway target genes and transcription factors in organ growth control

By Katrina Arissa Mitchell

ORCID: 0000-0003-0122-8595

COMPREHENSIVE CANCER DOCTOR OF PHILOSOPHY

AUGUST 2020

Peter MacCallum Cancer Centre

and

The Sir Peter MacCallum Department of Oncology, The University of Melbourne

Submitted in total fulfillment of the requirements of the degree of

Comprehensive Cancer Doctor of Philosophy

Abstract

The Hippo pathway, first discovered in the fruit fly *Drosophila melanogaster*, is an evolutionarily conserved regulator of organ growth. The Hippo pathway controls organ growth by regulating the nuclear activity of the transcriptional coactivator Yorkie (Yki). Yki binds to sequence-specific transcription factors to regulate target gene expression, including the TEAD/TEF family transcription factor Scalloped (Sd). It was recently discovered that the Hippo pathway controls the growth of organs in *Drosophila* by balancing transcriptional activation mediated by Yki and Sd, and transcriptional repression mediated by Tondu-domain-containing Growth Inhibitor (Tgi) and Sd. If this fine balance is disrupted and Yki is hyperactivated, organs can grow beyond their normal size. Importantly, deregulation of Hippo signalling drives tumorigenesis in a variety of human cancers. The high degree of conservation of the Hippo pathway throughout the animal kingdom makes the fruit fly a powerful model organism to understand how this pathway regulates tissue growth during development and cancer.

The Hippo pathway regulates the expression of genes involved in cell proliferation, survival, stemness, differentiation, and migration. However, we lack a clear understanding of the full transcriptional program that the Hippo pathway regulates in growing tissues beyond a small number of well-studied target genes. In this thesis, I aimed to deepen our understanding of the transcriptional program that the Hippo pathway controls in the growing *Drosophila* eye. To do this, I addressed three key questions: 1) What are the target genes of Yki, Sd and Tgi? 2) What gene expression changes occur as a result of hyperactive Yki? 3) What additional transcriptional regulators control Hippo pathway target gene expression and eye growth?

1) Using targeted DamID-seq, I identified target genes of Yki, Sd and Tgi in growing *Drosophila* eye discs. I found a very high degree of overlapping target genes between Yki, Sd, and Tgi, indicating that Sd is the key transcription factor that mediates binding of both Yki and Tgi to their target gene loci. Additionally, I found a strong enrichment of the AP-1 binding motif in Yki, Sd, and Tgi target genes, suggesting that these transcription factors co-regulate many target genes.

2) I performed comprehensive analyses of Yki, Sd, and Tgi target genes, and the changes in gene expression and chromatin accessibility caused by Yki hyperactivation. This highlighted key biological processes and signalling pathways that the Hippo pathway could regulate during eye growth, including: MAPK pathway, glutathione S-transferases (GSTs), cuticle development, extracellular matrix (ECM), and retinal determination and differentiation.

3) Using targeted DamID-seq, I found that the *Drosophila* AP-1 transcription factor Jra/Jun shares 71% of its target genes with Yki and/or Sd. However, using genetic studies, I showed that AP-1 transcription factors do not regulate physiological *Drosophila* eye growth, and are not essential regulators of Hippo pathway target gene expression.

The results presented in this thesis provide new insights into the target genes and transcriptional program that is regulated by the Hippo pathway during *Drosophila* eye growth. They have identified previously unexplored biological processes that the Hippo pathway might regulate in the eye (e.g. cuticle development), and reinforced known roles for the Hippo pathway (e.g. cell fate specification). My findings also extend our understanding of the functional links between AP-1 and the Hippo pathway in growing organs. Importantly, the data generated in this thesis will provide a rich resource for researchers interested in studying the direct target genes of both Hippo signalling and AP-1 transcription factors.

Declaration

This is to certify that

i. *the thesis comprises only my original work towards the PhD except where indicated in the Preface,*
ii. *due acknowledgement has been made in the text to all other material used,*
iii. *the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.*

Katrina Arissa Mitchell, August 2020

Preface

I would like to acknowledge and thank the following people, who contributed to the work presented in this thesis. Thank you to Professor Kieran Harvey, who generated a number of *Drosophila melanogaster* stocks used in this thesis, helped with many dissections, and also developed ideas for the research conducted in this thesis. Thank you to Dr. Joep Vissers, who generated a number of Dam-fusion transgene stocks, and also helped with numerous dissections and procedures for the targeted DamID experiments carried out in this thesis. Thank you to Abdul Saiful, who generated a number of stocks used in experiments described Chapter 5. Thank you to Dr. Jonathan Pojer who aided in the bioinformatic analysis of the targeted DamID data. Finally, thank you to Dr. Jan Schröder who created the bioinformatic pipeline used in this thesis.

Acknowledgements

I am very grateful to a large number of people who made my PhD thesis possible. Firstly, I want to thank Kieran Harvey who welcomed me into his lab over 8 years ago as a research assistant. I'm so thankful to have had this opportunity to join Kieran's research lab group and work on such fascinating and diverse work. Over the many years, Kieran has been a wonderful boss, a kind supervisor, and a motivating and inspiring scientist. I have gained a wealth of skills since joining his lab, and I am very grateful for Kieran's enthusiasm, endless help, and for pushing me to think independently.

Thank you to my co-supervisor Dr. Joep Vissers, who has always been a big driving force behind my work and success. Joep has been an excellent supervisor who has taught me a lot about molecular biology and transcription over the last few years, which has been integral to my studies. Beyond that, Joep has been a great friend, who is always happy to help me, have a chat, and give guidance.

Thank you to my fellow PhD student, Jonathan Pojer. I feel really lucky to have gone through the PhD process with him. I'm very thankful for all of his help with bioinformatics, his scientific input, and his friendship over the last few years. I also would like to thank all of the other members of the Harvey, Richardson, and Cheng lab, both past and present. I feel very grateful to have been surrounded by such hard working, intelligent, innovative, and kind people, many of whom have become close friends. I'd like to thank Dr. Carole Poon and Dr. Xiaomeng Zhang for teaching me and guiding me during my earlier days in the lab, without your help and support I wouldn't have got this far.

I also wish to thank my mentor, Dr. Liz Christie, who has gone above and beyond in mentoring me over the last few years. Liz has been an encouraging, understanding, and inspiring mentor, who always has time for a coffee catch-up.

I am grateful to the many people from the core facilities at the Peter MacCallum Cancer Centre for all of your help and training. Thanks to Gisela Mir Arnau and Tim Semple from the Molecular Genomics core, who helped me with sequencing and library preparation. Thanks to Sarah Ellis and her team at the Centre for Advanced Microscopy and Histology for help and training on the Confocal microscopes. I'd also like to thank all the Media Kitchen staff at PeterMacCallum Cancer Centre.

I am grateful to the Australian Government for providing me with an Australian Postgraduate Award to enable me to complete my PhD, and to Kieran Harvey for support in the final stages of my thesis. I am also grateful to the Peter MacCallum Cancer Centre for providing me with funds to travel and present my work at both local and international meetings, as well as EMBL for awarding me a travel scholarship to attend the 2018 EMBL Chromatin and Transcription meeting in Heidelberg, Germany.

And last, but certainly not least, thank you to the most important people in my life, my friends and family. Thank you for always supporting and believing in me. Thanks to Sara, Silvia, Clelia, Ozlem, Jonathan, AJ, Carole, Kellie, Edel, Ash, Lin, Talhuh, Anthony, Luis, Rosie, Tavi, Fay, Maria, Liz, Kate, Steph, Bridget, Hungky, Marta, Kelly, Nat, Belinda and Kristin for their friendship and support. Most importantly, thank you to my Mum, Dad, Tash, Sarah, Valeria, Andrew, Abeula, Auntie Marta, and Uncle Eddie for their love and endless encouragement.

Table of Contents

Chapter 1: Introduction

The Hippo pathway is an evolutionarily conserved signalling pathway which regulates organ size control by restricting proliferation and promoting apoptosis (Irvine and Harvey, 2015; Pan, 2010). Our understanding of the Hippo pathway and how it controls organ growth has immensely benefited from studies in the simple fruitfly, *Drosophila melanogaster*. These studies have expanded the processes that the Hippo pathway regulates to include cell fate determination, stemness, and regeneration (Misra and Irvine, 2018; Zhao et al., 2011). It has also become clear that the Hippo pathway plays a key role in cancer growth in humans (Harvey et al., 2013). Understanding the mechanisms by which the Hippo pathway regulates tissue growth can aid in the development of novel therapeutic strategies for cancer and regenerative medicine. This introduction will summarise current knowledge of the Hippo pathway, with a focus on the downstream components and mechanisms that control the transcriptional output of the pathway, and Hippo pathway target genes.

1.1 Regulation of organ growth

How an organ is capable of growing to the correct size and shape has been a fundamental and puzzling question in developmental biology. Research into this began in the early 20th century with fascinating experiments in salamanders, in which developing legs were transplanted from a large salamander species and grafted to embryos of a smaller salamander species (Twitty and Schwind, 1931). Interestingly, the grafted limbs grew to the large size of the donor organism despite being grafted to a smaller host species, indicating that organs possess "intrinsic" size control mechanisms (Twitty and Schwind, 1931). Such intrinsic organ size control mechanisms are conserved across animals, and have been found to operate in organs such as the mouse fetal thymus glands (Metcalf, 1963) and the imaginal discs of *Drosophila melanogaster* (Bryant and Levinson, 1985). Since these pivotal studies, a complex picture of growth control has been painted in which the final size of an organ is determined by a diverse interaction between both organ intrinsic factors (e.g. signaling pathways, including the Hippo pathway) and organ extrinsic factors (e.g. nutrition and hormones) (Andersen et al., 2013; Irvine and Harvey, 2015).

1.2 *Drosophila melanogaster* as a model organism

The model organism *Drosophila melanogaster* has played a pivotal role in understanding organ growth control and diseases. In 1918, Mary Stark first identified tumours in *Drosophila* larvae, which she described as resembling the tumours of vertebrates (Stark, 1918). *Drosophila* research has since contributed immensely to our understanding of organ growth and cancer through genetic screens identifying developmental and growth-related genes (Villegas, 2019). Early genetic screens in *Drosophila* include screens carried out by Nüsslein-Volhard and Wieschaus which identified genes involved in embryonic patterning, and led to the discovery of many well-known signalling pathways, including the Hedgehog pathway (Niisslein-Volhard and Wieschaus, 1980). The development of mosaic analysis tools allowed scientists to study homozygous mutant tissue adjacent to populations of cells that are wild type (Golic and Lindquist, 1989). This ground-breaking tool overcame the difficulties of studying genes that are lethal early in development in homozygous flies and allowed tissue specific experiments (Golic and Lindquist, 1989). Genetic mosaic mutagenesis screens led to the identification of mutations that conferred a growth advantage compared to wild type cells, and led to the discovery of the Hippo pathway (Harvey et al., 2003; Jia et al., 2003; Pantalacci et al., 2003; Tapon et al., 2002; Udan et al., 2003). There are several advantages that make *Drosophila* such an attractive model organism to study growth and cancer, notably the high conservation of genes between *Drosophila* and humans (Mirzoyan et al., 2019) . Indeed, approximately 75% of genes involved in human diseases have homologs in flies (Mirzoyan et al., 2019). Additionally, the quick life-cycle, low maintenance cost, and the powerful genetic tools available are all benefits of *Drosophila*.

1.2.1 The *Drosophila* eye

The *Drosophila* eye has served as a powerful model to study diverse biological processes, as well as the genes, pathways and networks controlling them. For example, the eye has been used to understand organ growth, programmed cell death, pattern formation, and differentiation (Kumar, 2018). An epithelial tissue in the larva called the eye-antennal imaginal disc gives rise to most of the structures of the adult head, including the visual system (compound eyes and ocelli), olfactory system (antennae and maxillary palps), and head epidermis (Figure 1.1A-C) (Kumar, 2018). The two major layers of the eye-antennal disc are the disc proper (DP), which

is a columnar epithelium, and the peripodial epithelium (PE), which is a squamous epithelium that covers the DP (Figure 1.1D) (Weasner et al., 2020). The PE contributes to the cuticle of the head, and also supports the growth and patterning of the DP, which gives rise to the adult eye and surrounding cuticle (Weasner et al., 2020). These structures are specified during embryogenesis, which is followed by larval stages during which proliferation expands the pool of precursor cells (Mishra & Sprecher, 2020). During this period of proliferation, a group of genes forming the retinal determination (RD) network begin to be expressed in the DP and induce the fate of the eye (Mishra & Sprecher, 2020). During the third instar larval stage, a wave of differentiation called the morphogenetic furrow (MF), sweeps across the eye disc from the posterior to the anterior resulting in the initiation of retinal differentiation (Mishra & Sprecher, 2020). The differentiating cells give rise to photoreceptor neurons and accessory cells. These cells form ommatidia, which are the hexagonal structures that collectively make up the adult compound eye (Mishra & Sprecher, 2020). Therefore, the development of the eye involves the careful execution of many processes including proliferation, cell fate control, patterning, and the initiation and migration of the MF.

1.3 The Hippo pathway

The Hippo pathway was first identified almost two decades ago in *Drosophila melanogaster* through pioneering genetic screens for modulators of organ growth. These screens identified several genes that when mutated bore the same fascinating phenotype in the adult fly: an overgrown head (Harvey et al., 2003; Jia et al., 2003; Pantalacci et al., 2003; Tapon et al., 2002; Udan et al., 2003). The identification of these genes, named *salvador* (*sav*), *warts* (*wts*) and *hippo* (*hpo*), catalysed scientific investigation into how these genes work together and how they control growth. Since then, a growing body of work has identified over 40 members of the pathway, which integrate diverse signals, such as mechanical cues and cell contact, and drive key developmental processes like growth, proliferation, cell survival, regeneration and cell fate. Importantly, the Hippo pathway is evolutionary conserved and its deregulation occurs in multiple human cancers (Harvey et al., 2013; Moon et al., 2018).

1.4 The core kinase cassette

The heart of the Hippo pathway is formed by a core kinase cassette, consisting of the sterile 20-like serine/threonine kinase Hippo (Hpo) (Harvey et al., 2003; Jia et al., 2003; Pantalacci et al., 2003; Udan et al., 2003; Wu et al., 2003) and the nuclear DBF2-related (NDR) kinase Warts (Wts) (Figure 1.2) (Justice et al., 1995; Xu et al., 1995). Hpo and Wts are activated by phosphorylation and associate with the adaptor scaffolding proteins Salvador (Sav) (Tapon et al., 2002) and Mob as tumour suppressor (Mats) (Lai et al., 2005), respectively. Activated Hpo is able to phosphorylate and activate Wts (Wei et al., 2007; Wu et al., 2003). Following this, Wts phosphorylates and inactivates the downstream effector protein Yorkie (Yki) (Huang et al., 2005). Yki constantly shuttles between the nucleus and cytoplasm; Wts-mediated phosphorylation limits Yki nuclear import (Manning et al., 2018), and therefore its ability to bind to its cognate transcription factors and activate its target genes (Huang et al., 2005; Oh and Irvine, 2008). Deletion of these key members of the Hippo pathway, or overexpression of Yki, result in dramatic overgrowth of *Drosophila* tissues as a result of increased proliferation and inhibition of apoptosis (Harvey et al., 2003; Huang et al., 2005; Tapon et al., 2002). These overgrowth phenotypes are due to elevated transcription of Yki target genes, such as *Cyclin E* (*CycE*) and *Death-associated inhibitor of apoptosis 1* (*Diap1*) (Huang et al., 2005).

The Hippo pathway is highly conserved throughout evolution including in mammals. However, the mammalian Hippo pathway is more complex and there are multiple homologs of most Hippo pathway genes that are only represented by one gene in *Drosophila*. In mammals, the Hpo homologues MST1/2, in cooperation with SAV1, phosphorylate the Wts homologues LATS1/2 (Chan et al., 2005). The activation of LATS1/2 results in the phosphorylation of Yes associated protein (YAP) and transcriptional co-activator with PDZ-binding motif (TAZ), which are the homologs of Yki (Huang et al., 2005; Vassilev et al., 2001; Zhao et al., 2008). As in *Drosophila*, the phosphorylation of YAP/TAZ results in cytoplasmic localisation and exclusion from the nucleus (Zhao et al., 2007). Hyperactive YAP can drive the overgrowth of multiple mammalian organs, most notably the liver (Camargo et al., 2007; Dong et al., 2007; Lee et al., 2010; Song et al., 2010; Zhou et al., 2009), and heart (Heallen et al., 2011; Monroe et al., 2019; Xin et al., 2011; Von Gise et al., 2012). In the liver, inactivation of Hippo signalling or hyperactivation leads to increased liver size, expansion of hepatic progenitor cells and eventually hepatocellular carcinoma (Lee et al., 2010). Additionally, loss of YAP can give rise

to developmental defects in the mouse kidney (Reginensi et al., 2013) and skin epidermis (Schlegelmilch et al., 2011).

1.5 Downstream transcriptional regulators

The Hippo pathway core kinase cassette converges on the co-activator protein Yki, which positively regulates the expression of genes that promote tissue growth (Huang et al., 2005). Yki was first identified as the downstream effector protein of the Hippo pathway through a yeast two-hybrid screen of Wts binding proteins (Huang et al., 2005). Homozygous deletion of *yki* in the germline of animals results in lethality during either embryonic or early larval stages (Huang et al., 2005). Similarly, in the *Drosophila* eye imaginal disc *yki* mutant clones grow poorly and are outcompeted by wild type clones, showing that Yki is essential for normal tissue growth (Huang et al., 2005). Recently, the localisation of Yki has become a focal point for many studies. Notably, live imaging has revealed that Yki dynamically shuttles between the nucleus and the cytoplasm (Manning et al., 2018, 2020). In the nucleus Yki regulates target genes that promote proliferation and prevent apoptosis (Oh and Irvine, 2011; Peng et al., 2009; Wu et al., 2003). In the cytoplasm, Yki plays additional roles that are not dependent on its transcriptional activity, including accumulating at the cell cortex to promote the activation of myosin (Xu et al., 2018). The subcellular localisation of Yki is dictated by its phosphorylation status (Dong et al., 2007; Oh and Irvine, 2008; Zhao et al., 2007). Yki can be phosphorylated by Wts at a number of sites, the most well-known is Ser168 which facilitates the binding of 14-3-3 proteins and sequestration in the cytoplasm (Dong et al., 2007; Oh and Irvine, 2008; Ren et al., 2010). Independently of 14-3-3 proteins, Yki can also be inhibited by Wts through phosphorylation at Ser111 and Ser250 (Ren et al., 2010). The Hippo pathway can also inhibit Yki through phosphorylation-independent mechanisms, for example the FERM-domain protein Expanded (Ex) binds to Yki via its WW domains and relocalises it from the nucleus to the cytoplasm (Badouel et al., 2009; Oh and Irvine, 2008). In this way, the Hippo pathway employs several mechanisms to control the levels of Yki in the nucleus, and therefore its transcriptional activity.

1.6 DNA-binding partners of Yorkie

Yki lacks a DNA binding domain and therefore it interacts with DNA-binding transcription factors (TFs) to bind genes and promote growth (Oh and Irvine, 2011; Pan, 2010). Multiple TFs have been identified in the recruitment of Yki to regulatory loci, including Scalloped (Sd) (Goulev et al., 2008; Wu et al., 2008; Zhang et al., 2008), Homothorax (Hth) (Peng et al., 2009), and Mad (Oh and Irvine, 2011). The most well characterised DNA-binding partner of Yki is the TEAD/TEF transcription factor Sd (Goulev et al., 2008; Wu et al., 2008; Zhang et al., 2008). Similarly in mammals, YAP complexes with TEAD transcription factors, which are required in most contexts for YAP-dependent target gene expression (Vassilev et al., 2001; Wu et al., 2008; Zhang et al., 2008). YAP also interacts with alternative transcription factors, including: p73 (Strano et al., 2005), Tbx5 (Rosenbluh et al., 2012), SMADs (Alarcón et al., 2009; Grannas et al., 2015; Ferrigno et al., 2002), and RUNX1/2 (Zaidi et al., 2004).

1.6.1 The TEAD/TEF transcription factors

Sd is a TEAD/TEF family transcription factor that was first identified as a Yki interacting protein through unbiased protein-protein interaction screens (Giot et al., 2003; Wu et al., 2008). Sd is required for Yki driven overgrowth in both the wing and eye disc, as studies have shown that *sd* mutations completely suppresses the tissue overgrowth phenotypes caused by excessive Yki activity (Goulev et al., 2008; Wu et al., 2008). Sd is also required for the elevated target gene expression resulting from Yki hyperactivity and has been shown to specifically bind to a minimal Hippo responsive element (HRE) in the Yki target gene *diap1* (Wu et al., 2008). Additionally, loss of *sd* completely rescues the undergrowth phenotype due to *yki* loss of function phenotypes (Goulev et al., 2008; Wu et al., 2008). However, *sd*, in contrast to *yki*, is dispensable for the growth of most *Drosophila* tissues, including the eye disc and the notal and hinge regions of the wing disc (Goulev et al., 2008; Wu et al., 2008). The one tissue where *sd* is essential is the wing pouch, which gives rise to the adult wing blade (Campbell et al., 1992). In the wing pouch, Sd forms a complex with Vestigial (Vg) and controls wing development by regulating the expression of target genes including *distalless* (Halder et al., 1998; Simmonds et al., 1998). Interestingly, in the wing pouch, Sd and Yki also complex but regulate a distinct set of genes independently of the Sd-Vg complex (Wu et al., 2008).

1.6.2 Alternative DNA-binding partners

The fact that *sd* is dispensable for the growth of many tissues led to a model whereby Yki must engage with alternative transcription factors to regulate growth (Goulev et al., 2008). In the *Drosophila* eye, Yki interacts with Hth (a TALE family homeodomain protein) and Teashirt (Tsh) (a Zn finger protein) to up-regulate the microRNA (miRNA) *bantam* (*ban*), which is a pro-survival gene (Peng et al., 2009). This process occurs specifically in uncommitted progenitor cells in the anterior compartment of the eye to ensure these cells continue to proliferate and remain undifferentiated (Peng et al., 2009). Interestingly, Hth and Tsh are not required for the expression of other Yki targets, such as *expanded* (*ex*) or *Diap1*, suggesting that Yki is able to bind different TFs to regulate different sets of target genes (Peng et al., 2009). Furthermore, although Hth promotes cell proliferation, it is not essential for this process (Peng et al., 2009). *ban* is also regulated by an interaction between Yki and the TF Mad, which acts downstream of the Decapentaplegic (Dpp) signalling pathway (Oh and Irvine, 2011). In this way, the Hippo pathway and the Dpp pathway can interact to control growth and development in *Drosophila* tissues. Interestingly, Yki and Mad bind to a specific enhancer region of *ban*, which is different to that of Hth-Yki binding, indicating that these complexes can independently regulate genes (Oh and Irvine, 2011).

1.6.3 Default gene repression

Although Yki can interact with multiple DNA- binding partners, none of these proteins alone can genetically account for the *yki* mutant phenotype, i.e. the severe undergrowth of tissues (Goulev et al., 2008; Koontz et al., 2013; Wu et al., 2008). For example, the loss of *sd* results in normal tissue growth in most *Drosophila* tissues. The puzzle of the apparent dispensability of *sd*, but not *yki*, was solved by Koontz et al (2013) in experiments that revealed Sd as a default repressor. Sd binds directly to a co-repressor protein called Tondu-domain containing Growth Inhibitor (Tgi) which results in default gene repression of target genes (Guo et al., 2013; Koontz et al., 2013) (Figure 1.3). Tgi mediated repression may also require two co-repressors named Ctbp and Pits (Vissers et al., 2020). Yki relieves Sd-mediated default repression by competing with Tgi, resulting in the activation of target genes (Guo et al., 2013; Koontz et al., 2013). In this way, the severe undergrowth of *yki* mutant clones is due to the active repression of target

gene expression by Sd and Tgi. On the other hand, the loss of *sd* results in the derepression of Hippo pathway target genes, which is reflected by the normal expression of Yki target genes *Diap1* and *ex* (Guo et al., 2013; Koontz et al., 2013). The dispensability of *sd* for development of tissues such as the eye also highlights the possibility that additional proteins compensate for the loss of *sd* and regulate Sd target genes. Significantly, the default repression model is conserved in mammals, as the mammalian orthologue of Tgi (Vestigial like 4; VGLL4), interacts with TEAD and also suppresses YAP activity in vivo (Koontz et al., 2013). Since these studies, VGLL4 has been identified as tumour suppressor via interacting with TEADs, and the lower expression of VGLL4 in cancer cells (e.g. lung cancer) compared to normal cells is associated with poor survival (Deng and Fang, 2018).

1.6.4 Yki-Sd binding partners

In recent years the list of proteins that can co-operate with both Yki and Sd is expanding, revealing how the Yki-Sd complex acts as a hub for many inputs. Intriguingly, the interaction with different proteins may provide a way for Yki and Sd to diversify the transcriptional output of the Hippo pathway (Nicolay et al., 2011). In the *Drosophila* eye disc, the Yki-Sd complex cooperates with and requires the transcription factor E2F1 at target gene promoters to drive the expression of a distinct set of genes involved in preventing cell cycle exit (Nicolay et al., 2011). Retinoblastoma tumour suppressor protein (Rb) negatively regulates E2F1, and therefore limits the expression of these genes, including *dDP* and *cdc2C* (Nicolay et al., 2011). In contrast, in the wing disc, Rb regulates a balance of Yki/Sd activation complexes and E2F1/Sd repressor complexes (Zhang et al., 2017). The interaction of E2F1 and Sd results in the suppression of Yki target genes, such as *Diap1*, *ex* and *ban*, and promotes apoptosis (Zhang et al., 2017). Therefore, in different tissues E2F1 appears to play different roles: in the eye disc E2F1 is required for Yki-Sd mediated activation of target genes, whereas in the wing disc E2F1 and Sd represses Yki target genes. Mask (Multiple ankyrin repeats single KH domain) has also been identified as a co-factor for Yki and Sd in both the eye and wing disc (Sansores-Garcia et al., 2013; Sidor et al., 2013). Mask can form a complex with Yki and Sd on target gene promoters, and is required for Yki to regulate growth and activate target genes, such as *Diap1* and *ex* (Sansores-Garcia et al., 2013; Sidor et al., 2013). Additionally, Yki interacts with Taiman, which is a receptor co-activator downstream of the steroid-responsive ecdysone pathway

(Zhang et al., 2015; Wang et al., 2016). In hyperactive Yki imaginal discs, the Yki-Tai interaction is required for the expression of a distinct transcription program including the germline stem cell genes *nanos* and *piwi* (Zhang et al., 2015). These genes are normally not expressed in imaginal discs, and their expression contributes to the overgrowth of hyperactive Yki tissues (Zhang et al., 2015).

Increasing evidence also highlights that the AP-1 transcription factor complex interacts with Yki/Sd and YAP/TEAD. AP-1 proteins are the downstream transcriptional effectors of the c-Jun amino-terminal kinase (JNK) pathway, and consist of dimers of Jun and Fos (Kockel et al., 2001). In human cancer cell lines, YAP, TEAD and AP-1 form a complex at target genes involved in S-phase entry and mitosis (Zanconato et al., 2015). Together they synergistically activate target genes through chromatin looping events that link distal enhancers and target gene promoters (Zanconato et al., 2015). Interestingly, AP-1 is essential for the oncogenic growth of breast cancer cells induced by YAP (Zanconato et al., 2015). Additionally, YAP/TEAD and AP-1 also drive the transcription of target genes involved in cancer cell migration and invasion, such as the *Dock-Rac*/*CDC42* genes (Liu et al., 2016). This interaction involves the SRC1-3 co-activator proteins and does not occur downstream of JNK signaling (Liu et al., 2016). Recent evidence also highlights that the interaction between AP-1 and YAP/TEAD is conserved in *Drosophila*. In the alary muscles of the *Drosophila* heart, Yki, Sd and AP-1 synergistically activate genes, such as *Myc* and *piwi*, to initiate naturally occurring alary muscle dedifferentiation (Schaub et al., 2019). In this case, this interaction requires active JNK signaling and inactive Hippo signalling, which activates both Yki and Jun, leading to the formation of Yki/Sd and AP-1 effector complexes (Schaub et al., 2019).

1.7 Mechanisms of gene regulation

Yki and YAP employ complex molecular mechanisms to control target gene expression. For example, Yki interacts directly with chromatin modifying proteins including GAGA Factor (GAF), Mediator complex, the Brahma complex (Oh et al., 2013), and Ncoa6 (Qing et al., 2014). These proteins play significant roles in regulating the transcriptional output of Hippo signalling.

Several studies have functionally linked Yki to the *Drosophila* SWI/SNF-related complex, known as the Brahma-associated proteins (BAP) complex, which is an ATP-dependent chromatin remodelling complex (Hillmer and Link, 2019). Interestingly, depending on the context, the BAP complex can either promote or repress Yki-induced tissue overgrowth. For example, Yki physically associates with the two BAP proteins Moira and Brahma (Oh et al., 2013), and this interaction plays an important role in driving midgut intestinal stem cell (ISC) proliferation (Jin et al., 2013), and promotes Yki-dependent transcription and tissue growth (Zhu et al., 2015). However, studies also show that the BAP complex can suppress the wing overgrowth induced by Yki hyperactivity (Song et al., 2017). In this context, loss of the BAP complex results in upregulation of the Wingless (Wg) and Decapentaplegic (Dpp) morphogens which cooperate with Yki to promote tumour formation (Song et al., 2017). These findings suggest that the function of the BAP complex interaction with Yki is context dependent, and may dependent on the transcription factors that Yki interacts with in different cell types (Song et al., 2017). This complex interaction between SWI/SNF and Hippo signalling has also been found in mammals, where it can also either inhibit or activate YAP/TAZ/TEAD target gene transcription. For example, in human mammary epithelial cells, ectopic TAZ expression switches luminal cells into a basal cell type, which is dependent on SWI/SNF (Skibinski et al., 2014). On the other hand, the SWI/SNF complex can inhibit YAP/TAZ transcriptional activity under conditions of low mechanical stress (Chang et al., 2018). When cells experience high mechanical stress, nuclear F-actin prevents the formation of the SWI/SNF-YAP/TAZ complex allowing YAP/TAZ and TEAD to associate and drive transcription (Chang et al., 2018).

Chromatin regions that are bound by Yki are also highly enriched for GAGA Factor (GAF) DNA recognition sites (Oh et al., 2013). GAGA Factor (GAF) is a transcription factor that plays several roles in transcription and chromatin remodelling. It can recruit RNA polymerase II to activate gene expression, recruit chromatin remodelling complexes to maintain open chromatin, and interact with Polycomb group proteins to transcriptionally silence genes (Hillmer and Link, 2019). GAF can bind to Yki (Oh et al., 2013), and is needed for the full activation of E2F1 and Yki/Sd target genes and proliferation (Bayarmagnai et al., 2012). Yki also recruits the Mediator complex, which links transcriptional activators to RNA polymerase II to facilitate gene transcription (Oh et al., 2013). GAF and Mediator are all required for Yki target gene activation, including common Hippo target genes *expanded* and *Diap1* (Oh et al., 2013). Similarly, YAP is able to recruit the Mediator complex at gene

enhancers in mammalian cells (Galli et al., 2015). This facilitates the recruitment of CDK9 elongating kinase and mediates transcriptional pause release, allowing YAP to drive high transcriptional activity (Galli et al., 2015). The interaction between YAP, Mediator and CDK9 is essential for YAP-induced cancer growth in the mouse liver (Galli et al., 2015).

The posttranslational modification of histones plays an important role in regulating gene transcription. In all eukaryotes, histone H3 lysine 4 methylation (H3K4me) is associated with transcriptionally active regions of the genome (Hillmer and Link, 2019; Mohan et al., 2011). In *Drosophila*, histone H3 methylation is carried out by the COMPASS methyltransferase family complexes, which include dSet1, Trithorax (Trx), and Trithorax-related (Trr) complexes (Hillmer and Link, 2019; Mohan et al., 2011). Interestingly, Yki directly binds to a subunit of Trr called Nuclear receptor coactivator 6 (Ncoa6) (Oh et al., 2013; Qing et al., 2014). Trr and Ncoa6 are required for Yki target gene expression and Hippo-mediated growth of imaginal discs (Oh et al., 2013; Qing et al., 2014). Importantly, this is also conserved in mammals, where the mammalian Ncoa6 subunit interacts with YAP (Oh et al., 2013). Therefore, Yki and YAP are capable of recruiting histone methyltransferases to alter chromatin and activate target genes (Oh et al., 2013; Qing et al., 2014).

1.8 Target genes of Hippo signalling

Only a handful of direct target genes of Yki and Sd have been identified. Broadly these genes can be grouped into the following categories: cell cycle and survival, Hippo pathway, and non-growth related genes.

1.8.1 Cell proliferation, growth and survival target genes

Direct target genes of Yki and Sd include the genes *ban* (Oh and Irvine, 2011; Peng et al., 2009), *diap1*(Wu et al., 2003), *Cyclin E* (Tapon et al., 2002), and *E2F1* (Goulev et al., 2008). These genes cooperate to control cell survival and proliferation. For example, *ban* is a microRNA that increases cell number due to the post-transcriptional suppression of proteins that normally inhibit cell proliferation and promote apoptosis (Brennecke et al., 2003). The cell-death inhibitor Diap1 protein is elevated in Yki overexpressing cells and is involved in increasing cell survival via the inhibition of pro-apoptotic caspases (Wang et al., 1999).

Additionally, Yki regulates the cell cycle gene *CycE*, which is required for cells for cells to transition between the G1 and S phases and commence genome replication (Knoblich et al., 1994). The increased expression of *Diap1*, *ban*, and *CycE* contribute to the overgrowth phenotype seen in Hippo pathway mutants. Yki has also been shown to regulate the activity of *Drosophila Myc*, which is a gene involved in ribosome biogenesis and cellular growth (Neto-Silva et al., 2010; Ziosi et al., 2010). In this context *Myc* promotes a "supercompetitive" behaviour of Yki overexpressing cells so that they can outcompete normal cells, which facilitates tumour progression (Ziosi et al., 2010). Importantly, YAP also regulates the transcription of *Myc*, which plays a role in carcinogenesis in liver cancer in mice (Xiao et al., 2013). YAP also promotes proliferation through target genes including CTGF (Zhao et al., 2008), CYR61 (Zhang et al., 2011), and AXL (Xu et al., 2011), and is thought to prevent apoptosis through target genes of the BCL2 (Chen et al., 2018) and IAP family (Dong et al., 2007).

Hippo pathway targets also include growth-promoting genes that encode proteins involved in other signalling pathways. In the *Drosophila* wing and leg disc respectively, Yki regulates *wingless* (*wg*) which encodes a ligand involved in the Wnt pathway, and *serrate* (*ser*) which encodes a Notch pathway ligand (Cho et al., 2006; Mao et al., 2006). These ligands contribute to Hippo pathway overgrowth phenotypes (Cho et al., 2006; Mao et al., 2006). Yki positively regulates the heparin sulphate proteoglycan genes *dally* and *dally-like* which encode cell surface proteins that are involved in the diffusion of Decapentaplegic, Hedgehog, and Wingless morphogens (Baena-Lopez et al., 2008). Additionally, in the intestinal stem cells (ISC) of the *Drosophila* midgut, Yki activity increases the production of the Upd family of cytokines, as well as EGFR ligands, which activate JAK-STAT and EGFR signalling (Ren et al., 2010a). Therefore, loss of Hippo signalling increases JAK-STAT and EGFR signalling and causes ISC proliferation (Ren et al., 2010a). In this way, the Hippo pathway can transcriptionally control the activity of various signalling pathways.

1.8.2 Hippo pathway feedback loop

Yki and Sd promote expression of genes that encode Hippo pathway members. Studies first identifying this negative feedback loop observed that cells that are mutant for *hpo*, *sav* or *wts* had elevated transcription of the upstream Hippo pathway proteins Expanded (Ex) and Merlin

(Mer) (Hamaratoglu et al., 2006), as well as Kibra (Genevet et al., 2010). In normal cells, these proteins are able to relay signals to the core of the Hippo pathway to control Yki activity, and form a negative feedback loop (Genevet et al., 2010). This negative feedback loop maintains Hippo signalling in a steady state and hence maintains tissue homeostasis (Genevet et al., 2010). In mammals, YAP also induces a negative feedback mechanism, whereby YAP/TAZ regulates the transcription of genes encoding LAT2, NF2 (Chen et al., 2015; Moroishi et al., 2015), and WW1-3 (Xiao et al., 2011), which are the mammalian orthologues of Wts, Mer, and Kibra, respectively. These proteins in turn negatively regulate YAP and TAZ activity.

1.8.3 Non-growth related genes

Interestingly, not all Yki target genes are related to growth control. For example, Yki directly upregulates the expression of genes involved in mitochondrial fusion, including *opa1-like* (*opa1*) and *mitochondria assembly regulatory factor* (*Marf*) (Nagaraj et al., 2012). Mitochondrial fusion occurs as a result of metabolic or environmental stress, and involves mitochondria fusing together and mixing contents to maintain functional mitochondria (Youle and Van Der Bliek, 2012). In the eye and wing imaginal discs, activation of Yki results in fused mitochondria and a reduction in reactive oxygen species (ROS) levels (Nagaraj et al., 2012). Interestingly, the reduction of mitochondrial fusion genes suppresses Yki-induced tissue overgrowth (Nagaraj et al., 2012). Additional non-growth roles for Yki include the regulation of the alary muscle lineage (Schaub et al., 2019). In *Drosophila* alary muscles of the heart, Yki upregulates *Myc* and *piwi*, which are crucial for the dedifferentiation and fragmentation of these cells into mononucleate myobolasts (Schaub et al., 2019).

1.9 Genome wide target genes of the Hippo pathway

Until recently, studies in *Drosophila* had identified a relatively small number of transcriptional targets of the Hippo pathway. Recent large-scale genomic studies have unveiled additional information about Yki target genes and how Yki regulates transcription. ChIP-seq studies in *Drosophila* tissues, such as wing and eye imaginal discs, reveal that Yki binds to thousands of chromosomal loci (Ikmi et al., 2014; Oh et al., 2013; Slattery et al., 2013), and is associated with actively transcribed genes, as indicated by H3K4me3 modification (Oh et al., 2013). Yki-

bound regions are enriched close to promoter regions, with less binding in intronic, exonic and intergenic regions (Ikmi et al., 2014; Oh et al., 2013; Slattery et al., 2013). Interestingly, Yki and YAP regulate a set of target genes that are conserved in *Drosophila* and mammals, which may represent a YAP-dependent ancient gene expression signature involved in organ size control (Ikmi et al., 2014). Despite the cell type differences and evolutionary distance, this study identified 74 genes that are commonly regulated by Yki and YAP in *Drosophila* and mice, most of which are DNA replication and cell cycle machinery genes (Ikmi et al., 2014).

An additional feature of Yki genome binding is that Yki target genes are remarkably similar between different tissues in *Drosophila* (Slattery et al., 2013). These genes include highly active housekeeping genes and genes involved in cell proliferation, implying that Yki promotes tissue growth in a non-tissue-specific manner in *Drosophila*, at least in epithelial imaginal discs. On the other hand, the transcription factors Hth and Sd displayed both tissue shared and tissue-specific binding events (Slattery et al., 2013). In these tissues, Yki's overlap with Sd was more prevalent than its overlap with Hth, indicating that Yki interacts mostly with Sd to regulate gene expression (Slattery et al., 2013). However, an interesting observation from ChIP-seq studies is that there are many Yki target genes that are not bound by either Sd or Hth (Ikmi et al., 2014; Slattery et al., 2013). Thus, this data imply that Yki interacts with DNA binding proteins in addition to Sd and Hth to regulate transcription.

1.10 The Hippo pathway in cancer

As Hippo signalling plays integral roles in growth and development it is not surprising that the Hippo pathway has been linked to a myriad of human diseases, most notably cancer. YAP and TAZ are oncoproteins that have been implicated in playing a role in multiple cancer types. Loss of function mutations in Hippo pathway genes have been identified in certain cancers, albeit rarely. Germline mutation of *NF2*, the homolog of Merlin, causes the syndrome neurofibromatosis type 2, which is associated with the development of schwannomas, meningiomas, and ependymomas (Asthagiri et al., 2009; Rouleau et al., 1993). Additionally, loss of function *NF2* mutations or *LATS1/2* mutations are frequent in malignant mesothelioma, a cancer that develops from asbestos exposure (Bueno et al., 2016). Elevated YAP/TAZ expression also occurs in many cancers, including those of the liver (Perra et al., 2014), breast

(Maugeri-Saccà et al., 2015), colon (Wang et al., 2013), and skin (Feng et al., 2014). Elevated YAP/TAZ expression can arise because of amplification of their gene loci, which occurs in hepatocellular carcinoma (Zender et al., 2006) and medulloblastoma (Fernandez-L et al., 2009). Chromosomal translocations involving YAP/TAZ have also been reported (Antonescu et al., 2013; Tanas et al., 2011). For example, TAZ-CAMTA1 and YAP-TFE3 gene fusions occur in epithelioid hemangioendotheliomas (EHE), which is a malignant vascular cancer (Antonescu et al., 2013; Tanas et al., 2011). The TAZ-CAMTA1 gene fusion, which occurs in approximately 90% of EHE cases, results in a chimeric protein that can translocate to the nucleus, is no longer responsive to Hippo signalling, and activates a TAZ-like transcriptional program (Tanas et al., 2016). Ultimately, YAP and TAZ activation confers cancer cells the ability to proliferate, evade apoptosis, gain stem cell attributes, metastasise and even overcome chemotherapy (Kulkarni et al., 2020; Zanconato et al., 2016).

1.11 Summary and Aims

The Hippo pathway is fundamental for the growth and development of tissues in many species including mammals and *Drosophila.* It is able to relay diverse signals, such as mechanical forces and cell-cell adhesion, to the core of the pathway which controls the localisation and activity of the transcriptional coactivator Yki. The activity of Yki can then regulate distinct transcription programs to drive growth, proliferation, cell survival, regeneration and cell fate. Yki relies on DNA binding proteins, most notably Sd/TEAD, to regulate target genes and utilises complex molecular mechanisms to control target gene expression. Developmental defects and diseases, such as cancer, can arise if the activity of Yki/YAP goes awry.

In the past two decades we have developed a rich knowledge of Hippo pathway and its different components. However, our understanding of how the Hippo pathway controls organ growth remains incomplete. A major focus of past studies has been to identify upstream regulators of Yki, while the downstream transcriptional targets and mechanisms of transcriptional control during organ growth have remained poorly understood. Firstly, we do not understand the full transcriptional program that the Hippo pathway regulates in growing tissues beyond a small number of bona fide target genes, such as *Diap1* and *ban*. Furthermore, the downstream transcriptional program responsive to hyperactive Yki in the *Drosophila* eye

has only been established in differentiating cells, and has thus failed to capture the full repertoire of genes that the Hippo pathway regulates in the eye disc. Interestingly, the fact that *sd* is dispensable for eye growth highlights the possibility that additional transcription factors control growth and Hippo pathway target genes in the absence of Hippo signalling. In this thesis, I attempt to address these questions with the following three key aims:

1. I aimed to identify the target genes of Yki, Sd, and Tgi in growing *Drosophila* eye discs using Targeted DamID-seq. Specifically, I used Targeted DamID-seq during a larval stage in which the eye is composed of undifferentiated eye progenitor cells that are actively proliferating. The results for this are presented in Chapter 3.

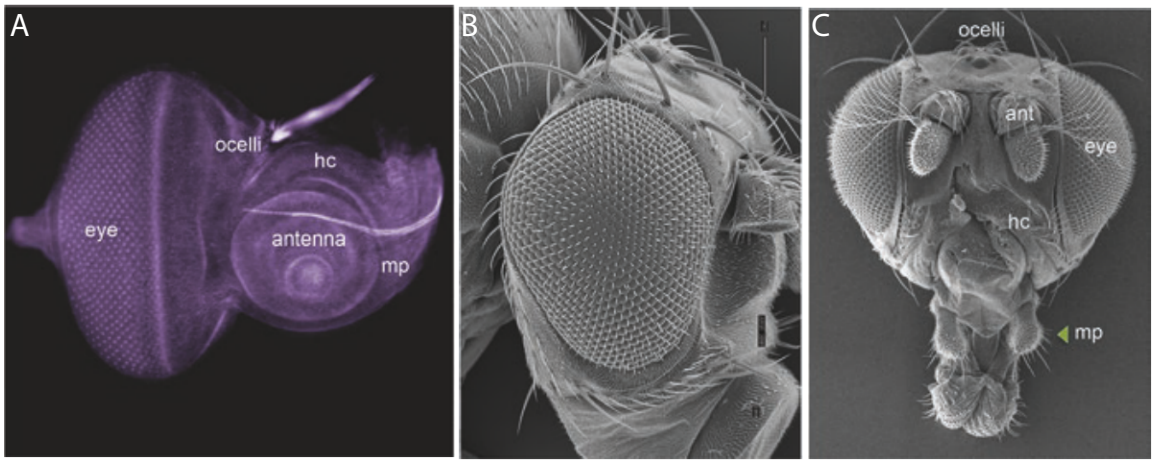
2. I aimed to investigate gene expression changes that occur in eye discs expressing hyperactive Yki, and compared this with Hippo pathway target genes (Chapter 3) and accompanying changes in chromatin accessibility. The results for this are presented in Chapter 4.

3. I aimed to investigate the role that AP-1 plays in regulating growth with the Hippo pathway. Specifically, this included identifying putative AP-1 target genes, and investigating whether Yki, Sd, and AP-1 regulate shared target genes to control the growth of the *Drosophila* eye disc. The results for this are presented in Chapter 5.

Figure 1.1: The structures of the *Drosophila* eye-antennal disc that give rise to the adult head

A. Light microscope image of a *Drosophila* third instar larval eye-antennal disc, which give rise to the adult structures shown in B and C. **B-C.** Scanning electron micrographs of adult heads. hc = head capsule, mp = maxillary palps, and ant = antenna.

D. Schematic of a cross section of the eye-antennal disc showing the cell types and layers. PE is shown in blue, DP is shown in purple, and the optic stalk is shown in grey. Adapted from: Weasner BP, Weasner BM, and Kumar JP, (2020), 'Ghost in the Machine: The Peripodial Epithelium' in Singh A and Kango-Signh M 'Molecular Genetics of Axial Patterning, Growth and Disease in Drosophila Eye', Springer, pp121-141.


A
ocelli
hc
eye
antenna
mp
B
C
ocelli
ant
eye
hc
mp

D

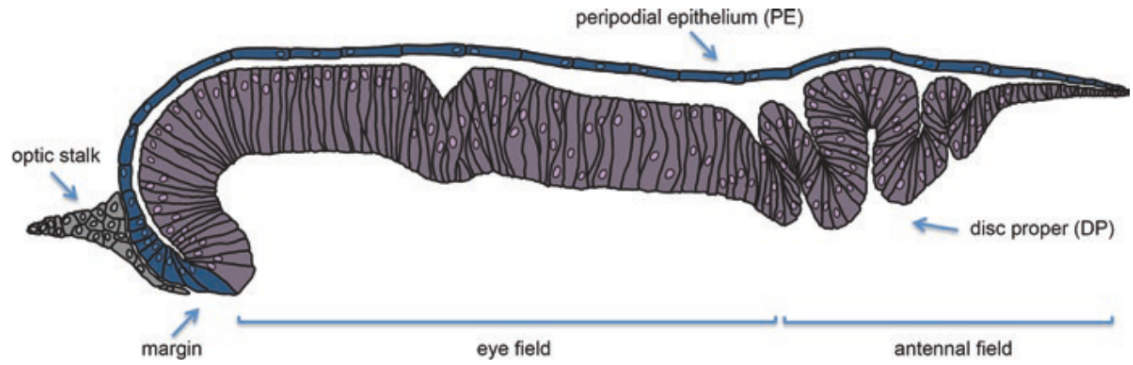

peripodial epithelium (PE)
optic stalk
disc proper (DP)
margin
eye field
antennal field

Figure 1.2: The Hippo signalling pathway in *Drosophila* melanogaster

Schematic diagram of the *Drosophila* Hippo pathway, showing some of the key proteins that are involved in tissue growth. Note that not all identified Hippo pathway proteins and interactions are illustrated in this diagram. Upstream members of the Hippo pathway include proteins that restrict growth (blue) and promote growth (purple). Arrows show interactions that are cooperative, while barred lines show interactions that are inhibitory. Adherens Junction (AJ) and Septate Junction (SJ) complexes are also shown. The core kinase cassette, shown in white, restricts the activity of Yki. Phosphorylated Yki can bind to 14-3-3 proteins in the cytoplasm, and has also been shown to dynamically shuttle between the nucleus and the cytoplasm. When Yki enters the nucleus, it binds to multiple transcription factors, and regulates target gene expression. Diagram created using BioRender, and adapted from Irvine KD, Harvey KF, 2015, 'Control of Organ Growth by Patterning and Hippo Signalling in *Drosophila*, Cold Spring Harbor Perspectives in Biology.

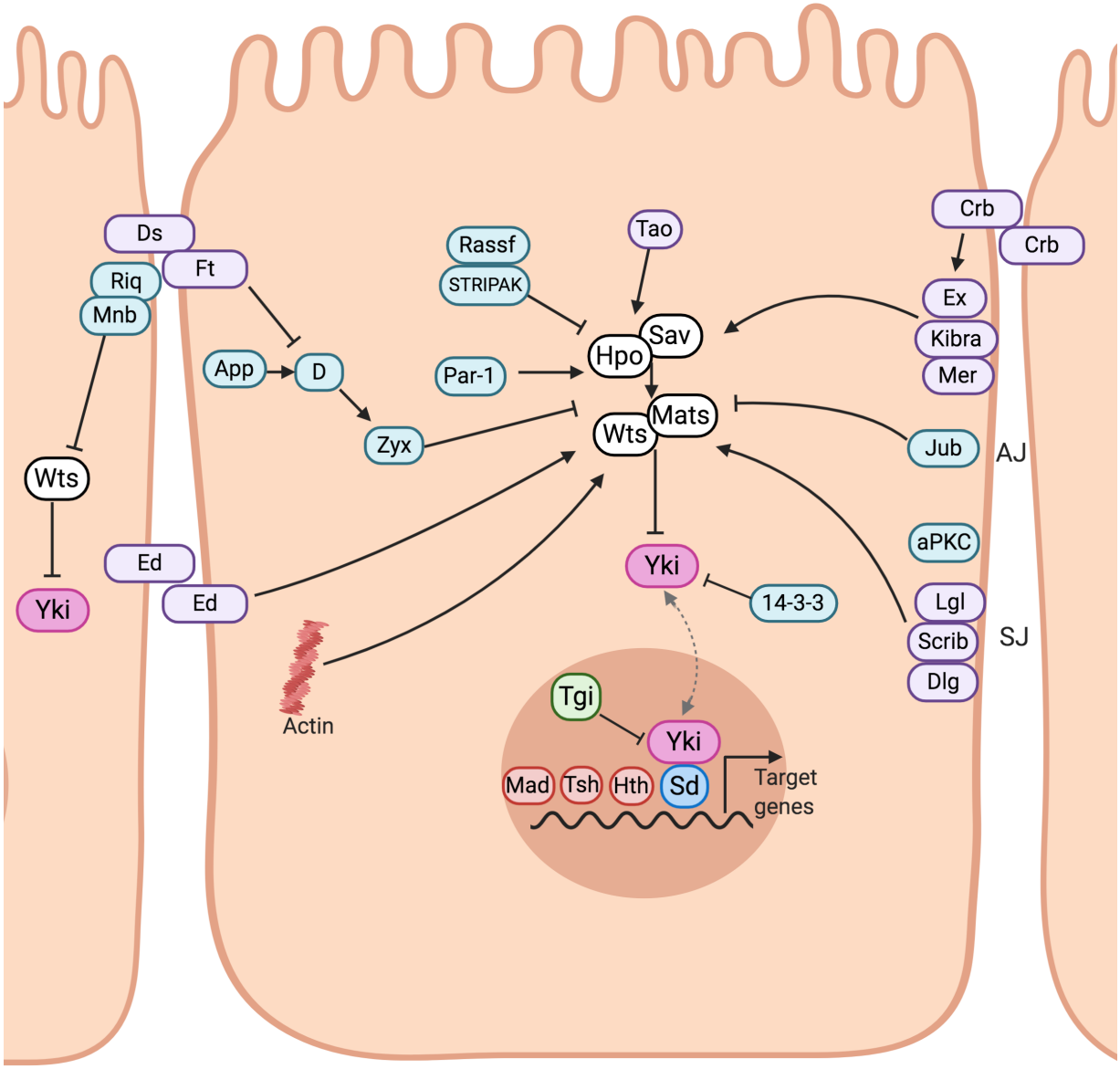
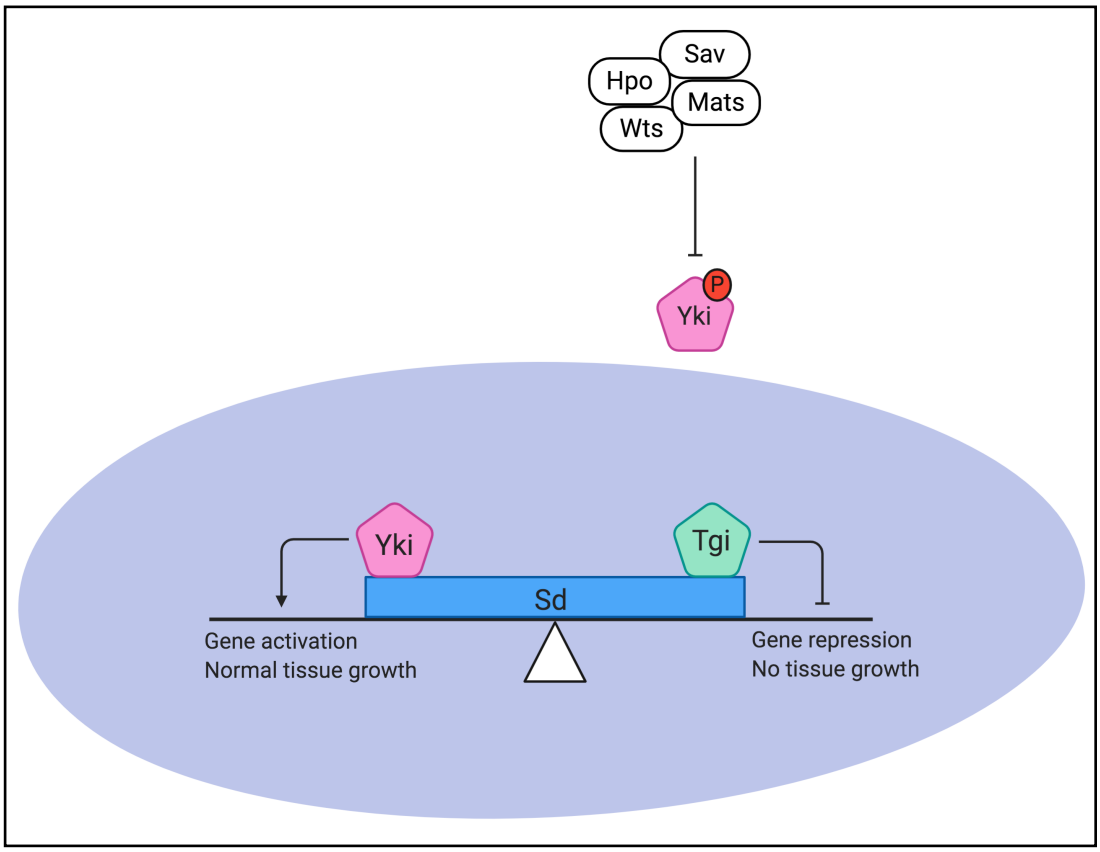

Ds
Ft
Riq
Mnb
Wts
Yki
App
D
Zyx
Rassf
STRIPAK
Tao
Par-1
Hpo
Sav
Wts
Mats
Crb
Crb
Ex
Kibra
Mer
Jub
AJ
aPKC
Lgl
Scrib
Dlg
SJ
Ed
Ed
Actin
Yki
14-3-3
Tgi
Yki
Mad
Tsh
Hth
Sd
Target genes

Figure 1.3: Default transcription repression

Schematic diagram of Hippo pathway default transcription repression model in organ growth. Active Yki binds to Sd in the nucleus and relieves Sd-Tgi default repression. The model shown here includes a scale to represent that the levels of Yki and Tgi direct the outcome of Sd activity. Diagram created using BioRender, and adapted from Koontz LM, Liu-Chittenden Y, Yin F, Yu J, Huang B, Chen Q, Wu S, Pan D, 2013, 'The Hippo Effector Yorkie Controls Normal Tissue Growth by Antagonizing Scalloped-Mediated Default Repression', Dev Cell.


Sav
Hpo
Mats
Wts
P
Yki
Yki
Tgi
Sd
Gene activation
Normal tissue growth
Gene repression
No tissue growth

Chapter 2: Materials and Methods

2.1 Drosophila husbandry and genetics

2.1.1 Drosophila melanogaster husbandry

Drosophila melanogaster stocks were stored at room temperature (22°C) or at 18°C and experimental crosses were carried out at 25°C unless indicated otherwise. Fly stocks and experimental crosses were maintained on Drosophila medium: 4.2% (w/v) yeast, 3.51% (w/v) glucose, 0.34% (w/v) agar, 3.98% (w/v) polenta 1.17% (w/v) Nipagin/Bavistan (10% p-hydroxybenzoic acid methyl ester C8H8O3). Experimental crosses were carried out in vials filled with excess food to guarantee that nutritional availability was not limiting. *Drosophila* strains used in this thesis can be found in Table 2.7.

2.1.2 Gal4-UAS driven transgene expression

The Gal4-UAS system drives the expression of transgenes in specific tissues and at distinct developmental stages (Duffy 2002). Gal4 is a yeast derived activator protein that binds to an upstream activator sequence (UAS) to activate gene transcription (Fischer et al. 1988). This system involves fusing the Gal4 gene to developmentally regulated promoters or enhancers that are expressed in specific cells (Duffy 2002). These Gal4 drivers can then be crossed to a transgene strain with an upstream UAS therefore attaining spatial and temporal expression of the transgene (Duffy 2002).

Gal4 expression can be controlled temporally using a temperature sensitive element called Gal80ts (Elliot & Brand, 2008). Gal80 is a negative regulator of Gal4 and represses transcriptional activation by binding to the activation domain of the Gal4 protein (Elliot & Brand, 2008). Gal80ts is temperature sensitive mutant that is active and represses Gal4 activity at the restrictive temperature of 18°C (Elliot & Brand, 2008). At permissive temperatures above 29°C, the repressive function of Gal80ts is

lost and Gal4 is active. In this way, Gal80ts can be crossed to Gal4 drivers and the activation of transgenes can be controlled by shifting the temperature of the animals.

2.1.3 Generation of homozygous mutant clones using the FLP/FRT system

The FLP/FRT site-specific recombination system is used to generate homozygous mutant clones in *Drosophila* tissues (T. Xu and Rubin 1993). The FLP/FRT system utilises the yeast flipase (FLP) recombinase enzyme which recognises and cleaves FLP recombination target (FRT) sites creating site-specific recombination between them (Golic and Lindquist 1989; Golic 1991). A mutant gene of interest is placed distal to FRT sites and FLP recombinase induced site-specific recombination producing daughter cells that are either homozygous for a wild type gene or homozygous for the mutant gene (T. Xu and Rubin 1993). Mutant clones are marked by the absence of Green Fluorescence Protein (GFP). The FLP recombinase can be expressed in a particular tissue using a tissue specific promoter, which avoids possible lethality associated with the particular mutated gene. In this thesis, FLP was expressed in the eye-antennal imaginal disc under the control of the *eyeless* promoter (Newsome, Åsling, and Dickson 2000). Additionally, experiments were also carried out using a heat-shock promoter that induces recombination at 37°C (T. Xu and Rubin 1993). The heat-shock experiments were carried out 2 days after egg laying (AEL) and involved submerging vials containing larvae in a water bath at 37°C for 15 minutes.

2.2 Sample preparation, Immunofluorescence and Imaging

2.2.1 Staging of *Drosophila melanogaster* larval development

For experiments that required staging of *Drosophila* larvae, crosses were left to lay eggs over a 4-hour period at 25°C, after which adults were then removed from the vials. At 25°C second instar larvae emerge approximately 3 days AEL, early third instar larvae emerged approximately 3.5 days AEL, and late third instar larvae emerged approximately 4 days AEL. The *wtsX1/wtsLacZ* transheterozygous mutant had delayed development, where early third instar larvae emerged approximately 4 days AEL. Early-mid third instar were distinguished morphologically from late instar through the

posterior and anterior spiracles. Larvae were floated from food using 30%w/v glycerol/PBS solution and then subsequently dissected.

2.2.2 Dissection and immunohistochemical staining of larval imaginal discs

Larvae were dissected into PBS on ice and fixed for 20 minutes in 4% (v/v) paraformaldehyde (PFA) at room temperature. Tissues were rinsed three times with 0.3% PBT and subsequently washed three times for 10 minutes in 0.3% PBT. Primary antibodies were prepared at the appropriate dilution in 0.3% PBT with 10% Normal Goat Serum (NGS). The samples were incubated with primary antibodies overnight at 4°C or for 4 hours at room temperature. The samples were then rinsed three times in 0.3% PBT and washed three times for 15 minutes in 0.3% PBT. Secondary antibodies were prepared at 1:500 dilution in 0.3% PBT with 10%NGS and incubated either overnight at 4°C or for 2-4 hours at room temperature. The samples were rinsed three times in 0.3% PBT and washed three times for 15 minutes in 0.3% PBT. DAPI (1:500) or TRITC-Phalloidin (1:200) were added in with the second wash. Tissue samples were then left in mounting medium (90% glycerol in PBS) overnight before being mounted onto a glass slide in either mounting medium or VectaShield. A list of antibodies and stains used in this thesis can be found in Table 2.8.

2.2.3 Confocal microscopy imaging and image analysis

Drosophila tissues were imaged using a Nikon C2 confocal microscope and an inverted Olympus FV3000 confocal microscope. Data analysis was performed using Nikon NIS Elements, Fiji (ImageJ), and Adobe Photoshop CS5.

2.2.4 Imaging of adult *Drosophila melanogaster* eyes

Adult *Drosophila* were frozen at -20°C and heads were then removed. *Drosophila* were carefully positioned to allow visualization of the entire eye. Images were taken using SciTech Infinity 1 camera mounted on an Olympus SZX7 microscope. Images were captured at a 4X magnification using the Infinity Capture 4.6 software, and

subsequently processed using Fiji (ImageJ).

2.3 CRISPR generation of mutant transgenic flies

2.3.1 Guide RNA design

CRISPR technology was used to generate *kay* transgenic mutant stocks, according to the method described by Kondo and Ueda (2013). A 20 bp target sequence was selected for guide design for each gene, with the following two requirements met: 1) The sequence began with a "G", which was needed for transcription initiation from the U6 promoter; and 2) the sequence ended in the PAM sequence "NGG", which was recognised by Cas9. The guides were designed to target as early in the gene sequence as possible, ensuring the target was present in all gene splice variants and also avoiding off targets. Guides oligonucleotides were designed using a combination of several online tools: NIG Fly

(https://shigen.nig.ac.jp/fly/nigfly/cas9/cas9TargetFinder.jsp),

Fly crisprn(http://tools.flycrispr.molbio.wisc.edu/targetFinder/),

and Chop chop (http://chopchop.cbu.uib.no).

2.3.2 Cloning of guide RNAs

The guide oligonucleotides were annealed and phosphorylated. This was carried out by mixing 1µL of 100uM of the forward oligonucleotide, 1µL of 100µM of the reverse oligonucleotide, 1µL of 10X T4 Ligase Buffer, 1µL of T4 Polynucleotide Kinase (PNK), and 6µL of mqH20. This mix was then annealed in a thermocycler at 37°C for 30 minutes, 95°C for 5 minutes and then the temperature was decreased 5°C per minute until it reached 25°C. The double stranded and phosphorylated oligonucleotides were then diluted 1:100. The oligonucleotides were then cloned into the pBFv-U6.2 vector. Firstly, the vector was digested with the restriction enzyme Bbs1 by adding the following: 3µg of vector, 5µl of 10X CutSmart Buffer, 3µl of Bbs1 and mqH20 to a total of 50µl. This mix was digested at 37°C for 3 hours, and in the final hour 1µl of Antarctic Phosphatase (AnP) was added in order to dephosphorylate the vector. The reaction was heat inactivated at 65°C for 20 minutes. The sample was then run on a 1%

agarose gel, the DNA band was excised using a clean scalpel, and gel purified using the QIAGEN gel purification kit according to the manufacturer's instructions. The sample was eluted in 30µl of elution buffer (EB).

The digested vector and annealed oligonucleotides were then ligated in the following mix: 50ng of vector, 2µl of annealed/phosphorylated oligonucleotides (1:100 dilution), 5µl 2X Fast Ligation Buffer, 1µl T4 DNA ligase, and mqH20 to a total of 10µl. The mix was incubated at room temperature for 20 minutes and then transformed into competent DH5alpha *E. coli*. Single colonies were selected and purified using the QIAGEN Miniprep Kit according to manufacturer's instructions. DNA was then sequenced using Sanger sequencing to confirm the oligonucleotides were successfully inserted into the vector. DNA samples were sent to BestGene for injection into flies and integration into the *attP40* landing site on the second chromosome using phiC31 integrase (Markstein et al. 2008). This created transgenic stocks that carried the guide RNA targeted to *kay*.

2.3.3 *Drosophila* genetics

The transgenic guide RNA (gRNA) stocks were crossed to a transgenic stock containing *cas9* driven in the germline using the *Nanos* driver. The progeny from this cross become the "founder" animals that express both the *cas9* and the gRNA, leading to an active *cas9*-gRNA complex. Individual "founder" males that carry the germline *cas9*-gRNA were selected and crossed to female balancer flies. Offspring from this cross were then screened by crossing to multiple deficiencies that covered the gene locus of *kay*. Stocks that were lethal when crossed to the specific deficiency were putative mutants and then Sanger sequenced to confirm the mutation.

2.4 Targeted DamID-seq

2.4.1 Cloning of the Yki DamID contruct

To create a fusion protein with Dam positioned at the C terminus of Yki, the *yki-Dam* fusion sequence was synthesised by Biomatik. This sequence contained the full coding sequence of *yorkie*, followed by a *myc* tag, a short linker and then the *Dam* sequence.

This fusion was sub-cloned into the *pUAST-attB-LT3-NDam* vector using the restriction enzyme Kpn1 so that the Dam sequence was replaced by *yki-Dam*. The full *yki-Dam* sequence was confirmed by Sanger sequencing.

2.4.2 *Drosophila* crosses and staging

Eyeless-Gal4 Gal80ts was crossed to *UAS-LT3-Dam*, *UAS-LT3-Dam-Sd*, *UAS-LT3-Dam-Tgi*, *UAS-LT3-Yki-Dam*, and *UAS-LT3-Dam-Jra*. Crosses were reared at 25°C and left to lay eggs over a 4-hour period. The eggs were then transferred to 18°C (restrictive temperature) for 4 days to repress the activity of the *Gal4*. Following this, second instar larvae were transferred to 29°C (permissive temperature) for 24 hours, which induced the expression of *Gal4* and the Dam fusion proteins. The eye-antennal imaginal discs and mouth-hooks of early third instar larvae were dissected into PBS on ice and then used for DNA extraction. Early third instar larvae were selected based on developmental timing and morphological features that distinguish them from late second instar larvae. Dissected tissues were transferred to 100µL of TENS buffer (100mM Tris (pH 8.0), 5mM EDTA, 200mM NaCl, 0.2% SDS) on ice. The tissue samples were either stored at -80°C until further processed or used immediately. Approximately 300-400 eye-antennal imaginal discs were dissected for each genotype per experiment, and this was carried out in triplicate (n=3).

2.4.3 DNA isolation

DNA isolation on dissected tissues was adapted from Tolhuis et al., 2006. Samples were treated with 2µL Proteinase K (20mg/mL) overnight at 56°C to digest contaminating proteins. The following day, RNA was digested at 37°C for 30 minutes using 0.5µL of RNase A (100 mg/ml). Samples were then transferred to a 1.5mL spin phase lock gel light tube and 100µL of phenol:chloroform:isoamylalchol was added. The samples were gently mixed and centrifuged for 5 minutes at 14000 rpm at room temperature. The upper phase was transferred to a fresh 1.5mL eppendorf tube and the DNA was precipitated using 1µL glycogen, 10µL 3M NaAc, and 300µL EtOH for 30 minutes at -80°C. Each sample was centrifuged for 30 minutes at 14,000 rpm at 4°C, and the supernatant was removed. 500uL of 70% EtOH was added to wash the pellet and followed by a 5 minute centrifuge at 14,000 rpm at room temperature. The ethanol

was carefully removed using a pipette and the DNA pellet was briefly air-dried. The DNA pellet was dissolved using 10-25µL of 10mM Tris ph8 and incubated at room temperature for 2-4 hours. As it is particularly important not to shear the genomic DNA, the DNA pellet was not resuspended by pipetting or vortexing. DNA concentration was measured using Thermo-Fisher NanoDrop 2000 Spectrophotometer and either stored at -80°C or processed immediately.

2.4.4 DamID

The DamID procotol used in my PhD studies was adapted from Marshall et al., 2016, Southall et al., 2013, Vissers et al., 2018, and Vogel et al., 2007. Firstly, isolated genomic DNA from *Drosophila* tissues was digested with the restriction enzyme DpnI, which specifically cuts adenine-methylated GATC sites. 2µg of DNA was digested overnight at 37°C in 0.5µL DpnI (10 units) and 1µL of 10X CutSmart buffer (NEB) in a total reaction of 10µL. A control sample was not incubated with DpnI (-DpnI), which ensured that targeted DamID only amplified methylated DNA and that the DNA has not been sheared. Following overnight digestion, 0.5µL of DpnI was added for an additional 1 hour incubation. The DpnI digestion was heat inactivated at 80°C for 20 minutes. Adaptors for PCR amplification were then ligated to the DpnI digested fragments. A ligation mix was added to each digested DNA sample, which contained 6.2µL H2O, 0.8µL 50uM double stranded adaptor AdR (40pmol), 2µL 10X ligation buffer and 1µL of T4 DNA ligase (5U/µL). A control was included that contained no ligase (-Lig), which ensured that the only the ligated fragments were amplified by PCR. The ligation was carried out for 2 hours at 16°C, and the reaction was inactivated by a 10-minute incubation at 65°C. The samples were then digested with DpnII, which cuts non-methylated GATC sites and prevents amplification of non-methylated DNA. This involved the addition of 1µL DpnII buffer, 0.2µL DpnII (50U), and 28.8µL MQ to each sample. The DpnII digestion was carried out for 4 hours at 37°C. The resulting digested and ligated DNA was then amplified by PCR. The mPCR included adding 20µL 5X MyTaq Reaction Buffer, 2.5µL Primer AdR (50µM), 2µL MyTaq DNA Polymerase, and 50.5µL MqH2O to 25µL of DpnII digested DNA. The DNA was amplified using the following PCR program:

Table 2.1: mePCR program

Temp	Duration	Cycles
68°C	10 min	1
94°C	30 sec	
65°C	5 min	1
68°C	15 min	
94°C	30 sec	
65°C	1 min	3
68°C	10 min	
94°C	30 sec	
65°C	1 min	17
68°C	2 min	
68°C	5 min	1

5µL of mePCR product was loaded on a 1% agarose gel to check the DNA quality. A smear concentrated between 400bp and 2kb indicates a successful amplification of methylated DNA. If a smear was observed in experimental samples, but was not present in the negative control samples (-DpnI and -Lig), the samples were then PCR purified using QIAGEN PCR purification kit according to the manufacturer's instructions. Each sample was eluted twice in 30µL of elution buffer. The concentration of the purified DNA was measured using a Qubit fluorometer (dsDNA High Sensitivity Assay Kit, Thermo Fisher).

2.4.5 Sequencing library preparation

DamID samples were adjusted to contain a total of 350ng of DNA in 55µL of elution buffer. The DNA was sonicated using Covaris S2 instrument to generate an average DNA fragment size of ~300bp. The Covaris S2 was used with the following settings: duty cycle = 10%, intensity = 5 cycles, bust = 200, time = 45 sec, power = ~23W, and bath temperature = ~4°C. Sonicated samples were digested with 1uL of Awl1 at 37°C overnight in order to remove DamID adaptors and the initial GATC sequence. DNA samples were purified using Agencourt AMPure XP beads (1.8X solid phase reversible

immobilisation (SPRI)) and eluted in 30µL of elution buffer. End repair and A-tailing and adaptor ligation was then carried out using the KAPA Hyper Prep Kit (Kapa Biosystems). Unique sequencing adaptors were used for each sample that was sequenced in the same sequencing run. A post ligation DNA purification was carried out using Agencourt AMPure XP Beads (0.8X SPRI) and eluted in 22µL of elution buffer. The libraries were amplified using the KAPA HiFi HotStart ReadyMix (Kapa Biosystems) with the following settings:

***Table 2.2*: sequencing library PCR program**

Step	Temp	Duration	Cycles
Initial denaturation	98°C	45 sec	1
Denaturation	98°C	15 sec	
Annealing	60°C	30 sec	5
Extension	72°C	30 sec	
Final extension	72°C	1 min	1

Libraries were then purified using Agencourt AMPure XP Beads (1X SPRI) and eluted in 20µL of elution buffer. Libraries were quantified using a Tapestation Bioanalyzer 2100 DNA 1000 chip.

2.4.6 NextSeq Sequencing

Sequencing was carried out on Illumina NextSeq instrument, using 75 base paired end (PE75) reads and 20 millions reads per sample.

2.5 RNA-seq

2.5.1 RNA extraction

RNA was extracted using TRIzol™ according to manufacturer's instructions. 100 eye-antennal imaginal discs from mid third instar larvae were dissected into PBS on ice.

Tissues were transferred to 1mL of TRIzol™ and homogenized using a sterile pestle. The samples were incubated at room temperature for 5 minutes to permit the complete dissociation of the nucleoprotein complex. 200uL of chloroform was added and incubated for 2 minutes at room temperature and then centrifuged at 12,000xg for 15 minutes at 4°C. The upper aqueous phase was transferred to a new Eppendorf tube into which 500uL of isopropanol was added and incubated for 10 minutes at room temperature. Each sample was centrifuged for 10 minutes at 12,000xg at 4°C, and the supernatant was removed. 1mL of 75% EtOH was added to wash the RNA pellet and followed by a 5-minute centrifuge at 7,500xg at 4°C. The ethanol was carefully removed using a pipette and the RNA pellet was briefly air-dried. The RNA pellet was dissolved in 20uL of RNase-free water by pipetting up and down. RNA samples were then stored at -80°C until prepared for sequencing libraries.

2.5.2 NextSeq Sequencing

Sequencing was carried out on Illumina NextSeq instrument, using 75 base paired end (PE75) reads and 20 millions reads per sample. RNA-seq samples were sequenced in the same batch and 5 replicates were carried out for each genotype.

2.6 Bioinformatics Analysis

Bioinformatics analysis was carried out for analysing the DamID-seq and RNA-seq sequencing datasets. These analyses were implemented on a macOS operating system using RStudio (Version 3.6.1) and Terminal (Version 2.7.5). All code used for this thesis can be found in Appendix 8.1. Packages and tools used in this thesis are listed in Table 2.5.

2.6.1 DamID-seq analysis

DamID analysis involved the pipelines outlined below, which: 1) aligned the DamID-seq sequencing datasets, 2) applied differential methylation analysis, and 3) called peaks and assign them to genes. Code for this analysis can be found in Appendix 8.1.1.

2.6.1.i Alignment of sequencing data

The following analyses were carried out in Terminal. DamID-seq datasets were processed to remove remaining DamID ligation adaptors and sequencing adaptors using Cutadapt. Alignment to the *Drosophila melanogaster* reference genome (Release 6; dm6) was carried out using the Subread package to generate sequencing reads. FeatureCounts was used to count sequencing reads and assign them to GATC features (subsequently referred to as 'tags').

2.6.1.ii Differential methylation analysis

The count data was imported and analysed in RStudio using the edgeR and limma packages. The EdgeR package in R was used to determine differentially methylated tags of the Dam-fusion compared to the Dam-alone control. edgeR was selected due to its ability to integrate biological replicates and correct for batch effects in the analysis. Firstly, the limma package was used to design a matrix of the DamID-seq data and included group (genotype), batch (biological replicate) and sequencing batch (included when samples were sequenced in separate sequencing runs). A DGEList (differential gene expression list) object was generated to store the count data and various parameters about the data, including the design matrix, and library sizes. The DGEList was filtered so that low counts were excluded from the analysis. Specifically, the data was filtered so that tags that have 0.5 counts per million (cpm) in at least 3 samples were included in the analysis. Normalisation was carried out using the TMM method, which eliminates composition biases between libraries and scales the libraries accordingly. Within group dispersion was also calculated to take into account variability between the replicates. Both the normalisation and the dispersion were added into the DGEList. In order to identify differences in the datasets, a principle components analysis (PCA) was carried out and plotted on a multidimensional scaling (MDS) plot.

A linear model for the data was created using the edgeR package in RStudio. A negative binomial generalised linear model method was implemented, which fits the linear model to the read counts and conducts genewise statistical tests (likelihood ratio and p value) for the DamID datasets. This analysis identified 'significant tags', which

are tags that are significantly different in the Dam-fusion compared to the Dam control ($p<0.05$). Multiple hypothesis testing correction was performed on the linear model to determine differential methylation and calculate the false discovery rate (FDR) value. This analysis generated a list of tags that were differentially methylated higher in the Dam-fusion datasets compared to the Dam-control datasets ('differentially methylated tags').

2.6.1.iii Peak calling

Peak calling analysis was created by Dr Jan Schröder, and the full description of the analysis method will be published elsewhere. Peaks are genomic regions that contain significantly methylated tags that indicate specific binding of the Dam-fusion proteins to DNA. Significantly differentially methylated tags were combined to call peaks using Python (Version 3.7.6) using call_peaks.py in Terminal. In order to be included in a peak, a tag must fulfil the following criteria: have a p value < 0.05, have a log fold change >0, and have a minimum of two significant tags. Additionally, the pipeline also allowed for a tag to be skipped, as some regions had significant tags separated by nonsignificant tags, which lead to the peaks being broken despite strong methylation in surrounding tags. The peaks were then filtered into significant peaks. Firstly, peaks that contained at least one differentially methylated tag were selected. Secondly, peaks that had more than two significant tags were classified as significant. Additionally, a peak was also identified as significant if it contained only two significant tags but all tags were differentially methylated within a 1kb of the centre of the peak. In this way, this takes into account the number of GATC tags within a 1kb window and considers areas of the genome that have fewer GATC tags. Peaks were then associated with the nearest gene within a 5kb distance to the closest transcription start site (TSS) using bedtools closest gene function.

2.6.2 RNA-seq analysis

Alignment of the RNA-seq data to the *Drosophila melanogaster* reference genome (Release 6; dm6) was carried out using the Subread package. FeatureCounts was used to count how many sequencing reads are at each genomic position, which generated a

counts table. RNAseq data analysis was carried out according to the module designed by COMINE-Australia (GitHub: https://github.com/COMBINE-Australia). The count data was imported and analysed in RStudio using the edgeR and limma packages. Lowly expressed genes were filtered and genes that have 0.5 counts per million (cpm) in at least 5 samples were included in the analysis. A DGEList (differential gene expression list) object was generated to store the count data and various parameters about the data, including the library sizes. In order to identify the greatest source of variation in the data, the data was plotted on a multidimensional scaling (MDS) plot. TMM normalisation was implemented in order to eliminate composition biases between libraries. The normalisation factors were added to the DGEList. Following this, a design matrix was created for testing comparisons, which in this case was comparing wild type growth to the hyperplastic samples. Differential expression analysis was carried out using the voom function in limma. Voom was used to transform the read counts in log counts per million (logCPM), taking into account mean-variance in the data. Each gene was fit to a linear model using the lmFit function in limma. lmFit estimates group means and gene-wise variances. The makeContrasts function was utilised to compare between wild type samples and hyperplastic samples, using a null hypothesis that there is no difference between these samples. The contrasts matrix was applied to the contrasts.fit function in limma, which generated statistics. The final step included using the eBayes function in limma to perform empirical Bayes shrinkage on the variances, estimate moderated t-statistics and generate p-values. Code for the analysis can be found in Appendix 8.1.2.

2.6.3 Comparisons between datasets

The intersect function in RStudio was used to identify genes that overlap between datasets. Venn diagrams were created using the eulerr package in RStudio. FlyBase gene identifiers were used when comparing target genes between DamID-seq datasets, between DamID-seq and RNA-seq datasets, and between DamID-seq and ChIP-seq datasets.

bedtools was used to intersect DamID-seq bed peak files in Terminal. bedtools intersect reports the overlaps between two sets of genomic features. In this case, bed peak files were intersected and reported if they share at least 1 bp overlap.

A hypergeometric test was used to estimate the probably of genes overlapping by chance alone (Fury et al. 2006). A hypergeometric test was implemented using the phyper function in RStudio. phyper calculates the p-value for the overlap of two different lists, using the following rule:

phyper(a, b, c, d, lower.tail = FALSE, log.p = FALSE)

Where:

a = number of genes overlapping between lists b and d

b = number of genes in List 1

c = total number of genes in the genome – b

d = number of genes in List 2

Plots comparing log fold changes (LogFC) between DamID-seq datasets were generated in RStudio using ggplot2. Linear regression analysis was carried out using the cor function in RStudio, which measures the strength of association between two variables using the Pearson's correlation coefficient (R). Code for this can be found in Appendix 8.1.3.

2.6.4 Gene Profiles

Gene profiles for DamIDseq datasets were generated using the pipeline published by Marshall and Brand, 2015, and then visualised using the ggplot2 package in RStudio. The full code from Marshall and Brand (2015) can be found at: https://github.com/owenjm/damidseq_pipeline. damidseq_pipeline was performed on each biological replicate bam file in Terminal, using default parameters. The --full_data_files option was used which allowed binning the reads at 75bp instead of binning using GATC sequences. The output from this analysis were bedgraph files that showed the log fold change of the Dam fusion over the Dam-alone control. The

bedgraphs of each biological replicate were averaged using bedtools in Terminal. For visualisation of the gene profiles, the bedgraphs from damidseq_pipeline (Marshall & Brand, 2015) were combined with the peaks generated using the call_peaks.py pipeline described in (Section 2.6.1.iii).

CATaDa profiles were generated using damidseq_pipeline (Marshall & Brand, 2015) with comparing the Dam-alone control datasets from wild type and hyperplastic growth. For CATaDa profiles, the peak files were generated using the find_peaks pipeline from Marshall and Brand (2015) (https://github.com/owenjm/find_peaks).
Gene profiles were created in RStudio, using the ggplot2 package. The code for visualising genes profiles and peaks was created by Dr Jonathan Pojer with edits by Katrina Mitchell. Gene views for each gene profile were created using the annotated *Drosophila melanogaster* genome from FlyBase (release 6.27:
(ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/gtf/). Editions were made in Adobe Illustrator CS6. Code for this can be found in Appendix 8.1.4.

2.6.5 Distribution of genomic features

The distribution of DamID-seq peaks was analysed relative to genomic features. Genomic features were generated from the annotated *Drosophila melanogaster* genome from FlyBase (release 6.27:

(ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/gtf/).

Features that were assigned to peaks included: 3' untranslated region (UTR), 5' untranslated region (UTR), intron, exon, and intergenic. The annotated genome was split into these different genomic features in RStudio. In Terminal, bedtools was used to subtract exon regions from genic regions to obtain intronic regions. Additionally, exonic regions were calculated by subtracting 3' and 5' UTR regions. Then, bedtools was used to overlap DamIDseq peaks with annotated genomic features (using the intersect function). This analysis was also carried out for GATC sites within the genome. The bar plot was generated using the ggplot2 package in RStudio. Code for this can be found in Appendix 8.1.5.

2.6.6 Proximity to the transcriptional start site (TSS)

The proximity of DamIDseq peaks from the transcriptional start site of genes was calculated using the ChIPSeeker package in RSudio. To observe the coverage of the peaks over the Drosophila genome, the covplot function was used. The plotAveProf2 function was used to calculate the peak count frequency and plots this within a 3 kb distance to transcriptional start site of annotated genes. Code for this can be found in Appendix 8.1.6.

2.6.7 KEGG pathway enrichment analysis

KEGG pathway enrichment analysis was performed in RStudio on the DamID-seq datasets and the RNA-seq dataset. Entrez IDs were generated using the AnnotationDbi and org.Dm.eg.db packages. KEGG pathway analysis was carried out on the Entrez IDs using the kegga function from the limma package. The data was then filtered to include significant pathways ($p < 0.05$), and the most significant pathway if pathways were duplicated. The Kegg analysis was plotted in RStudio using the ggplot2 package, and visualized on a log10(p-value) scale. Code for this can be found in Appendix 8.1.7.

2.6.8 GO enrichment analysis

Gene ontology (GO) enrichment analysis was performed in RStudio on the DamID-seq datasets and the RNA-seq dataset. Entrez IDs were generated using the AnnotationDbi and org.Dm.eg.db packages. KEGG pathway analysis was carried out on the Entrez IDs using the goana function from the limma package. The data was then filtered to include: significant terms ($p < 0.05$), and the most significant term if terms were duplicated. The GO analysis was plotted in RStudio using the ggplot2 package, and visualized on a log10(p-value) scale. The code for this can be found in Appendix 8.1.8.

2.6.9 STRING analysis

STRING analysis was performed on the products of the differentially expressed genes in hyperplastic (*wtsX1/wtsLacZ*) eye discs compared to wild type (*FRT82B*) eye discs.

STRING analysis was performed using the STRING webtool: https://string-db.org/. The analysis was carried out using a medium confidence score of 0.4, with the following interactive sources selected: textmining, experiments, databases and co-expression. The results were then exported as a svg file, and edited in Adobe Illustrator CS6.

2.6.10 Motif enrichment analysis

Motif enrichment analysis was carried out using Hypergeometric Optimization of Motif EnRichment (HOMER) in Terminal. The findMotifsGenome.pl tool was used to identify *de novo* motifs enriched in the DamID-seq peak files compared to the background *Drosophila melanogaster* genome (dm6). A false discovery rate (FDR) was calculated from 100 randomisations, and a motif was considered significant according to a p-value < 0.05, and a FDR (q-value) of < 0.05. The results were compiled and visualised in Adobe Illustrator CS6. Code for this can be found in Appendix 8.1.9.

2.6.11 Heatmaps

Heatmaps for the RNAseq dataset were created using the heatmap.2 function from the gplots package in RStudio. The genes of interest and associated logcounts were subset and the matrix created from this was used to create the heatmaps. The RColorBrewer package was used for the colour scheme. Code for this can be found in Appendix 8.1.10.

2.7 *Drosophila* stocks

The following *Drosophila melanogaster* stocks were used in this thesis.

Table 2.3: *Drosophila melanogaster* strains used in this thesis

STOCK	SOURCE
w;; p[pUAS-LT3-Dam] attP2/TM6B	(Southall et al. 2013; Vissers et al. 2018)
w;; p[pUAS-LT3-Dam-Sd] attP2/TM6B	(Southall et al. 2013; Vissers et al. 2018)
w;; p[pUAS-LT3-Yki-Dam] attP2/TM6B	This study

w;; p[pUAS-LT3-Dam-Jra] attP2/TM6B	This study
ey-Gal4	BSC #5534
ey-Gal4, tub-Gal80^{ts}	This study
GMR-Gal4	BSC #9146
Nanos-cas9	BSC #54591
UAS-GFP	BSC#5428
UAS-kay RNAi^{GD}	VDRC #19512
UAS-kay RNAi^{GD}	VDRC #6212
GMR-Gal4,UAS-YFP,UAS-Yki^{S168A}/ SM6a-TM6B	C. Bennett
FRT82B	BSC#2035
ex^{697}	Boedigheimer and Laughon, 1993
wtsX1	BSC#44251
wtsLacZ	Justice et al., 1995
y w eyFlp FRT19A Ubi GFP ; ; FRT82B RFP	NA
hsFLP;; FRT82B/TM6B	NA

2.8 Reagents

The following reagents were used in this thesis.

Table 2.4: Reagents used in this study

REAGENT	SOURCE
Antibodies	
Rat anti-Elav (1:50)	DSHB , Cat# 7E8A10
Mouse anti-Diap1 (1:200)	Bruce Hay
Mouse anti-βgalactosidase (1:100)	Sigma-Aldrich, Cat#G4644
Donkey anti-mouse 647 (1:500)	Life Technologies, Cat#A31571
Chemicals, Peptides, and Recombinant Proteins	
DAPI (1:500)	Sigma-Aldrich, Cat#D9542
Phalloidin-TRITC (1:200)	Sigma-Aldrich, Cat#P1951
Proteinase K (20mg/mL)	Sigma, Cat#E6779

RNase A (100mg/mL)	Qiagen, Cat#19101
Phenol:chloroform:isoamyl alcohol (25:24:1)	Sigma, Cat#P2069
DpnI (10U) T4	NEB, Cat#R0176L
T4 DNA ligase (5 U/µL)	Sigma, Cat#10799009001
DpnII (50U)	NEB, Cat#R0543L
MyTaq HS DNA polymerase	Bioline, Cat# BIO-21112
AlwI	NEB, Cat#R0513L
Agencourt AMPure XP Beads	BeckmanCoulter, Cat#A63881
Bbs1	NEB, Cat# R0539S
T4 Polynucleotide Kinase	NEB, Cat# M0201S
Antarctic Phosphatase	NEB, Cat#M0289S
Kpn1	NEB, Cat# R0142S
DH5alpha *E.coli*	Thermo Fisher Scientific, Cat#18265017
TRIzol	Thermo Fisher Scientific, Cat#15596018
Critical Commercial Assays	
MinElute PCR Purification Kit	Qiagen, Cat#28004
KAPA HyperPrep Kit	KAPA Biosystems, Cat#07962371001
QIAquick Gel purification kit	Qiagen, Cat# 28706
Miniprep Kit	Qiagen, Cat# 27106
Oligonucleotides	
AdRt: CTAATACGACTCACTATAGGGCAGCG TGGTCGCGGCC GAGGA	(Vogel, Peric-Hupkes, and van Steensel 2007)
AdRb: TCCTCGGCCG	(Vogel, Peric-Hupkes, and van Steensel 2007)
AdR_PCR: GGTCGCGGCCGAGGATC	(Vogel, Peric-Hupkes, and van Steensel 2007)
Dam sequencing fwd: TGGACATTATGACGCTGAGG	This study
Dam sequencing rev: TTCATCAGTTCCATAGGTTGG	This study (Dr Joseph Vissers)

Yki-Dam sequencing fwd: GCCCACTTTCCCAGAGAAC	This study
Yki-Dam sequencing rev: TTAGTGTCTTCCCTCCAGATTG	This study
kay crispr gRNA fwd: CTTC GACCACCACGCGCAACATCG	This study
kay crispr gRNA rev: AAAC CGATGTTGCGCGTGGTGGTC	This study
kay crispr PCR screen fwd: GAGGGTTTCGTGCTTGTCAT	This study
kay crispr PCR screen rev: ATTTGAGGTATTCGCGTTGC	This study
kay crispr sequencing fwd: TCATGCACCTTTTCTGCTCTT	This study
Recombinant DNA	
pBFv-U6.2	(Kondo and Ueda 2013)
pBFv-U6.2-*kay* gRNA	This study
pUASt-LT3-Dam	(Southall et al. 2013)
pUASt-LT3-Dam Sd	(Vissers et al. 2018)
pUASt-LT3-Yki Dam	This study
pUASt-LT3-Dam Jra	This study
pUASt-LT3-Dam Kay	This study

2.9 Software, Packages, and toolkits

The following software, packages, and toolkits were used in this thesis. R code was run in RStudio (Version 3.6.1), and bash, perl or python was run in Terminal (Version 2.7.5).

Table 2.5: Software, packages, toolkits and pipelines used in this thesis

PACKAGE/TOOLKIT/PROGRAM	SOURCE
AnnotationDbi (V1.48.0)	(Pagès et al., 2019)
Bedtools (v2.28.0)	(Quinlan and Hall 2010)
BiasedUrn	(Fog, 2015)
call_peaks.py	Dr Jan Schröder - unpublished

ChIPpeakAnno (v3.20.0)	(Lihua J Zhu et al. 2010)
ChIPseeker (v3.10)	(Yu, Wang, and He 2015)
Cutadapt (v1.10)	(Martin 2011)
damidseq_pipeline	(Marshall and Brand 2015)
dplyr (v0.8.3)	(Wickham et al., 2019)
EDASeq (v2.18.0)	(Risso et al., 2011)
edgeR (v3.26.8)	(Robinson et al., 2010)
EnhancedVolcano (v1.2.0)	(Blighe, 2019)
eulerr (v6.0.0)	(Larsson, 2019)
FeatureCounts (v1.6.4)	(Liao et al., 2014)
find_peaks	(Marshall and Brand 2015)
GGally (v1.4.0)	(Schloerke et al., 2018)
ggplot2 (v3.2.1)	(Wickham, 2016)
ggsignif (v0.6.0)	(Ahlmann-Eltze, 2019)
gggenes (v0.4.0)	(Wilkins, 2019)
gplots (v 3.0.1.1)	(Warnes et al., 2019)
Glimma (v1.12.0)	(Su et al., 2017)
GO.db (v3.8.2)	(Carlson, 2019)
gridExtra (v2.3)	(Auguie, 2017)
HOMER (v4.11)	(Heinz et al., 2010)
limma (v3.40.6)	(Richie et al., 2015)
KEGG.db (v3.2.3)	(Carlson, 2016)
org.Dm.eg.db (v3.8.2)	(Carlson, 2019)
purrr (v0.3.2)	(Henry and Wickham, 2019)
RColorBrewer (v1.1.2)	(Neuwirth, 2014)
readr (v1.3.1)	(Wickham et al., 2018)
reshape2 (v1.4.3)	(Wickham, 2007)
Rprojroot (v1.3.2)	(Müller, 2018)
RUVSeq (v1.18.0)	(Risso et al., 2014)
scales (v1.0.0)	(Wickham, 2018)
statmod (v1.4.32)	(Giner and Smyth, 2016)
stringr (v1.4.0)	(Wickham, 2019)

Subread (v2.0.0)	(Liao et al., 2013)
svglite (v1.2.2)	(Wickham et al., 2019)
tibble (v2.1.3)	(Kirill and Wickham, 2019)
tidyr (v1.0.0)	(Wickham and Henry, 2019)
tidyverse (v1.2.1)	(Wickham, 2017)

Chapter 3: Identifying putative target genes of Yorkie, Scalloped and Tgi in the developing *Drosophila* eye disc using Targeted DamID

3.1 Introduction

The Hippo pathway plays a vital function in organ growth control by regulating cell proliferation and survival (Irvine and Harvey 2015; Pan 2010). Over forty components of the Hippo pathway have been identified to date, which relay and integrate a multitude of signals to the core of the pathway (Snigdha et al. 2019). The core members of Hippo signalling converge to regulate the downstream transcriptional coactivator protein Yorkie (Yki) in *Drosophila melanogaster* (Huang et al. 2005), and the orthologues Yes-Associated Protein (YAP) and Transcriptional co-activator with PDX binding motif (TAZ) in mammals (Lei et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008). Yki dynamically shuttles between the cytoplasm and the nucleus where it interacts with DNA-binding transcription factors to regulate the transcription of target genes (Shreberk-Shaked and Oren 2019; Ege et al. 2018; Manning et al. 2018; Elosegui-Artola et al. 2017; Kofler et al. 2018; Franklin et al. 2019). The best characterised binding partner of Yki is the TEF/TEAD family transcription factors, the sole member of which is Scalloped (Sd) in *Drosophila* (Goulev et al. 2008; S. Wu et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008; Zhao et al. 2008). The interaction of Yki with Sd results in the transcriptional activation of target genes which are crucial for organ growth (Goulev et al. 2008; S. Wu et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008; Zhao et al. 2008). Through regulating target genes, Yki promotes the growth of *Drosophila* tissues, including the eye and wing imaginal discs (Huang et al. 2005). Tondu-domain-containing growth inhibitor (Tgi) is a Sd-binding protein that directly competes with Yki for Sd binding and inhibits the transcriptional activity of the Yki-Sd complex (Guo et al. 2013; Koontz et al. 2013). Growth is fine-tuned by a balance of Yki-Sd mediated transcriptional activation and Tgi-Sd mediated transcriptional repression (Guo et al. 2013; Koontz et al. 2013). This fine balance can be disrupted by either the loss, or hyperactivity, of Yki; Yki is crucial for the growth of many organs, whilst hyperactivation of Yki causes organ overgrowth (Zhao et al. 2010). In humans, YAP and its binding partner TEAD are deregulated

in multiple cancers, including breast cancer, gastrointestinal cancer, melanoma and brain cancers (Maugeri-Saccà and De Maria 2018). Identifying the target genes that act downstream of Hippo signalling to control the size of a tissue is vital in understanding both normal and abnormal growth.

The Hippo pathway regulates a range of genes involved in cell proliferation, survival, stemness, differentiation, and migration (Meng, Moroishi, and Guan 2016). Despite this, only a select number of genes have been characterised as direct targets of the Hippo pathway transcriptional regulators Yki and Sd. These genes include the microRNA *bantam* (Oh and Irvine 2011; Peng, Slattery, and Mann 2009), the anti-apoptotic gene *Diap1* (S. Wu et al. 2003), and *expanded* (Hamaratoglu et al. 2006) which encodes an upstream regulator of Hippo signalling. These bona fide target genes have provided sensitive and reliable readouts of Hippo pathway activity in *Drosophila* tissues.

Recently, genome-wide identification of Yki and Sd genome binding sites has uncovered additional insights into the transcriptional regulation and target genes of the Hippo pathway. The results from these studies have been summarised in Figure 3.1. ChIP-based assays in *Drosophila* tissues, such as wing and eye imaginal discs, reported that: 1) Yki and Sd bind to thousands of genomic loci; 2) Yki- and Sd-bound genomic loci are associated with actively transcribed genes; 3) the number of putative Yki and Sd target genes varies between studies; and 4) Yki binds to thousands of additional target genes independently of Sd (Ikmi et al. 2014; Oh et al. 2014). These findings suggest that Yki also interacts with additional transcription factors in addition to Sd (Ikmi et al. 2014; Oh et al. 2014). Indeed, Yki has been reported to bind to and regulate transcription with the TALE-homeodomain protein Homothorax (Hth) (Peng, Slattery, and Mann 2009) and the BMP effector Smad (Oh and Irvine 2011). However, Koontz et al., (2013) have raised doubts regarding these conclusions by showing that the primary function of Yki in normal tissue growth is to relieve Sd mediated transcriptional repression. Furthermore, these ChIP-based studies have several limitations, including that the experiments were carried out on mixed cell populations, they rely on good antibodies and they only provide information on binding occurring at the time of cross-linking as opposed to binding occurring over specific periods of development. Hence, despite genome-wide studies elucidating some novel target genes and biological processes regulated by Hippo

signalling, we still do not know the full repertoire of Hippo pathway target genes that are regulated in order to control the growth of *Drosophila* tissues.

The seemingly contradictory findings between the aforementioned genetic and ChIP studies, the limitations of ChIP-based assays, and the fact that very few target genes have been validated from these studies, highlights the need for a new approach to more robustly identify Hippo pathway target genes. Therefore, I decided to utilise a powerful technique called targeted DamID to identify putative target genes of Yki, Sd, and Tgi in the developing *Drosophila* eye disc. This technique was selected because it does not require antibodies and requires less tissue than ChIP-seq, and enables the *in vivo* identification of target genes in specific cells or tissues (Marshall et al. 2016; Southall et al. 2013). Furthermore, unlike ChIP, targeted DamID can be controlled temporally *in vivo* so that genome binding of transcription factors can be surveyed in precise stages of development. I established targeted DamID in eye cells during the early second instar to early third instar larval stages, when the eye is composed of undifferentiated eye progenitor cells that are actively proliferating (Wolff and Ready 1991). By selecting this time point in eye development, I aimed to identify genes that are specific to a proliferating tissue. Additionally, I also used targeted DamID to identify the genome-wide binding of Tgi, a key player in the transcriptional output of Hippo signalling, which has never been assessed before.

3.2 Results

3.2.1 Establishing targeted DamID in the developing *Drosophila* eye disc

Targeted DamID involves fusing the bacterial DNA adenine methyltransferase (Dam) protein to a chromatin binding protein of interest, which results in methylation near sites in the genome where the protein of interest binds (Figure 3.2) (Marshall et al. 2016; Southall et al. 2013). The Dam-fusion protein can be expressed in specific cell types with temporal control by utilising the Gal4 system (Marshall et al. 2016; Southall et al. 2013). Importantly, a Dam-only control is included to account for non-specific DNA methylation by Dam (Marshall et al. 2016; Southall et al. 2013).

DamID fusion transgenes expressing *Dam*-only control (Vissers et al. 2018), *Dam-Yki*, *Dam-Sd* (Vissers et al. 2018), *Dam-Tgi* were created by Dr Joseph Vissers, and I generated the *Yki-Dam* transgene. These transgenes contained a primary open reading frame (ORF1) followed by two stop codons and a secondary open reading frame (ORF2) encoding the Dam methylase either alone (control), or fused to each protein of interest (Figure 3.2A). The Dam-fusion secondary ORF is translated at low levels due to low rates of ribosomal re-initiation, which allows the Dam-fusion proteins to be expressed at very low levels (Southall et al. 2013). This low expression ensures that the Dam methylase is not toxic to cells. As a consequence the Dam-fusion proteins are unable to be detected by western blotting or immunofluorescence (Southall et al. 2013). As such, expression of mCherry, which is contained in the ORF1, was used as a surrogate to detect the expression of the DamID transgene fusions (Figure 3.2B). I individually crossed the *Dam*-only control, *Dam-Sd*, *Dam-Tgi* and *Yki-Dam* fusion transgenes to the eye-specific Gal4 driver *eyeless* (*ey*-Gal4) and detected the expression of mCherry in the cells posterior to the morphogenetic furrow in late third instar larval eyes (Figure 3.2B-E'). This confirmed that DamID transgene expression can be detected for all four DamID transgene fusions using *ey-Gal4*. Additionally, it appeared as though mCherry formed protein aggregates in the eye, which were visible as bright puncta (Figure 3.2B). This could be a due to the activation of *ey-Gal4* early in the embryo eye primordia and throughout larval development, and could indicate that the Dam-fusion transgene is expressed at high levels (Hauck, Gehring, and Walldorf 1999). As an accumulation of Dam can lead to toxicity, this highlighted a need to restrict the length of time that the Dam-fusion transgenes were expressed.

3.2.2 Characterisation of *eyeless*-Gal4 expression and developing a staging protocol for targeted DamID in the early third instar larvae

I selected the *eyeless*-Gal4 (*ey*-Gal4) driver to assess Yki, Sd, and Tgi gene DNA binding sites in the developing eye disc using targeted DamID, as *eyeless* is active in proliferating cells of the eye-antennal imaginal disc (Hauck, Gehring, and Walldorf 1999). In order to determine how I can use *ey-Gal4* to drive Dam transgene expression in proliferating eye imaginal discs, and avoid contamination from non-eye cells, I investigated the expression pattern of *ey*-Gal4 in 1) multiple tissues, and 2) throughout larval development. Firstly, *ey*-Gal4 was crossed to a

UAS-GFP transgene and expression was examined in tissues within the head region of larvae. I found that *ey*-Gal4 was expressed in the eye (Figure 3.3B-D') and brain, but not in the wing, salivary glands or the mouth hooks. This highlighted a need to remove the brain during dissections of the eye disc and is consistent with research showing a role for *ey* in *Drosophila* brain development (Clements et al. 2009). Secondly, *ey*-Gal4 was crossed to a *UAS-GFP* transgene and the expression over larval instar stages was examined. Larval instar stages were identified by the morphology of the anterior and posterior spiracles which are distinguishable at different stages of larval development (Poças, Domingos, and Mirth 2020). In the eye of second instar larvae and early third instar larvae, *ey*-Gal4 was expressed throughout the eye in uncommitted progenitor cells (Wolff and Ready 1991) (Figure 3.3B-C'). In mid-third instar larvae, *ey*-Gal4 expression shifted towards the posterior region of the eye disc. In late third instar larvae, *ey*-Gal4 expression was restricted to cells posterior to the morphogenetic furrow, which includes cells that are proliferating and cells that are differentiating into the photoreceptor fate (Wolff and Ready 1991) (Figure 3.3D-D'). Based on these findings, early third instar larvae were selected to investigate Yki, Sd, and Tgi genome binding sites, as during this stage the majority of cells in the eye are actively proliferating and minimal differentiation is taking place.

A staging protocol was designed so that early third instar larvae were selected for targeted DamID. Gal4 expression was controlled using a temperature sensitive element Gal80 (Gal80ts) which represses Gal4 at 18°C. *ey*-Gal4 was induced by shifting second instar larval stage to 29°C and early third instar larvae were selected for dissection (Figure 3.3E). In this way, Dam-fusion protein expression was limited to 24 hours coincident with a specific period of eye development when the uncommitted eye progenitor cells are actively proliferating. Adults from this cross were healthy and had normal eye growth, showing that induction of Dam expression for 24 hours had no lasting toxic effect on the cells.

3.2.3 Identification of putative Hippo pathway target genes in the developing *Drosophila* eye disc

3.2.3.i Targeted DamID in the developing *Drosophila* eye disc

Targeted DamID was carried out on early third instar *Drosophila* eye discs in order to identify Yki, Sd, and Tgi DNA binding sites. *Ey*-Gal4 was crossed to *Drosophila* strains carrying transgenes encoding each of the Dam-fusion proteins and a Dam-alone control. The Dam-fusion proteins were expressed for 24 hours, eye discs were dissected from early third instar larvae and DamID was performed. DamID samples were run on an agarose gel to check the quality of the amplified DNA product. A smear was observed between 400bp and 2kb for all experimental samples, revealing successful DNA methylation (Figure 3.4A-B). Negative control samples included: 1) -Dpn1, in which Dpn1 was omitted from the protocol which indicates whether the DNA was sheared; and 2) -Lig, in which T4 Ligase was omitted from the protocol which ensures that only the ligated fragments were amplified by PCR. Negative control samples showed no amplification (Figure 3.4A-B). This indicates that genomic DNA was methylated by the Dam-alone control and Dam-fusion proteins and that this methylation was successfully amplified by PCR.

The targeted DamID vector that was used to create the Dam-fusion transgenes included a multiple cloning site (MCS) upstream of the Dam coding sequence, which permits the insertion of coding sequences encoding a protein of interest. The DamID transgene creates a fusion protein with Dam positioned at the N terminus of the protein separated by a short linker (Figure 3.2A) (Marshall et al. 2016). Initial experiments carried out with the Dam-Yki fusion protein showed successful amplification of methylated DNA (Appendix Figure 8.2.1A). Libraries were then prepared for the Yki-Dam and Dam-alone control samples, and then analysed by next generation sequencing (NGS). Bioinformatic analysis was performed to align reads to the *Drosophila* genome, and identify regions of increased methylation in the Dam-Yki sample compared to the Dam-alone control (differential methylation analysis). NGS results for Dam-Yki revealed a noisy signal, minimal differential DNA binding sites when compared to the Dam-alone control and no enrichment of Yki at known target genes (Appendix Figure 8.2.1B-D). I next investigated whether this was due to a mutation in the Dam sequence, as previous studies have reported spontaneous mutations arising in the Dam sequence after passaging through bacteria (Marshall et al. 2016). However, Sanger sequencing of the Dam coding sequence showed no mutations in the Dam gene, indicating that a mutation in Dam was not reason for the negative DamID sequencing results. I hypothesised that the location of Dam at the N-terminus of the Yki protein could impact Yki function, as the N terminus of Yki is

essential for its nuclear localisation and interaction with Sd (Wang et al., 2016). Therefore, I created a new Yki DamID fusion transgene, with the aim of mitigating problems arising from an N-terminal Dam fusion. The new Yki DamID fusion, called Yki-Dam, involved placing the Dam coding sequence upstream of the Yki coding sequence to create a fusion protein with Dam positioned at the C terminus of Yki. Targeted DamID for Yki-Dam was performed repeating the protocol detailed above. The amplified DNA product was run on an agarose gel, which showed successful amplification of methylated DNA (Figure 3.4A). As above, libraries were prepared for these samples, NGS was carried out, and differential methylation bioinformatics analysis was performed. This analysis revealed clear evidence of Yki-Dam fusion protein binding to known Yki target genes (Figure 3.4).

The Yki-Dam, Dam-Sd, and Dam-Tgi DamID sequencing data was analysed to identify differential occupancy of the Dam-fusions compared to the Dam-alone control. Differential methylation analysis was performed to identify GATC sequences (subsequently referred to as GATC 'tags') that were differentially methylated between the Dam-fusion and Dam-alone control samples. A principal component analysis (PCA) was used to determine the greatest source of variation in the data, in an unsupervised manner. The PCA was visualised using a multidimensional scaling (MDS) plot that plotted the samples based on log2 fold changes. The distance between the samples on the MDS plot approximates the differences in log2 fold changes between the samples. The MDS plot revealed that the samples clustered well according to the specific Dam-fusion protein, with replicates for Dam control samples clustering together, and replicates for each of the Dam-fusion proteins clustered together according to the protein (Figure 3.4C). This indicates that the biological replicates for each sample were reproducible with low variability. The greatest variation was observed between the Dam-alone controls and the Dam-fusion proteins, indicating that the greatest difference exists between these samples.

3.2.3.ii Targeted DamID revealed that Yorkie, Scalloped, and Tgi bind thousands of genomic loci

Peak calling analysis was carried out to identify genomic regions that contain multiple significantly differentially methylated GATC 'tags'. These genomic regions are known as "peaks" and are specific regions of the genome where the Dam-fusion proteins interact with

DNA. Peak calling analysis revealed that Yki significantly bound 4,797 peaks, Sd significantly bound 5,409 peaks and Tgi significantly bound 5,324 peaks. The nearest gene method was used to assign peaks to putative target genes, which indicated that Yki bound to 2,974 genes, Sd bound to 2,960 genes, and Tgi bound to 2,921 genes.

Several genes are commonly used as reporters for Hippo pathway activity, including *Death-associated inhibitor of apoptosis 1* (*Diap1*), *expanded (ex*), *Cyclin E* (*CycE*) and *bantam* (*ban*). These genes are direct targets of Yki and Sd and sensitive to changes in Hippo pathway activity (Hamaratoglu et al. 2006a; Oh and Irvine 2011; Qing et al. 2014; Peng, Slattery, and Mann 2009; S. Wu et al. 2003). Yki, Sd, and Tgi Dam-fusion proteins showed strong methylation and peaks throughout the gene locus of *Diap1*, *expanded*, and *bantam*, and bound to the promoter region of *CycE* (Figure 3.5). This highlights that targeted DamID and the novel analysis pipeline used in this thesis were successful in identifying target genes of Yki, Sd, and Tgi.

3.2.3.iii Yorkie, Scalloped and Tgi bound genomic regions are enriched at promoters of putative target genes

In order to identify regions of the genome that were preferentially bound by Yki, Sd, and Tgi in developing eye discs, I analysed the location of peaks for each Dam-fusion protein within specific annotated regions of the *Drosophila* genome (Release 6; dm6). Peaks for each Dam-fusion protein were analysed for enrichment within the following annotated genomic features: 5'UTR, coding exons, introns, 3'UTR, and intergenic regions. The *Drosophila* genome consists of approximately 5% 3'UTR, 3% 5'UTR, 21% exon, 34% intron, and 37% intergenic regions. The proportion of peaks within these categories for Yki, Sd, and Tgi was compared to the distribution of GATC 'tags' in the genome. The location of GATC 'tags' varies within the *Drosophila* genome, which creates a bias where the Dam-fusion proteins are capable of methylating DNA. GATC 'tags' were enriched within exon regions and underrepresented in intergenic and intron regions of the *Drosophil*a genome (Figure 3.6A). Yki, Sd, and Tgi genome binding peaks were strongly enriched in 5'UTRs and intronic regions, despite these regions containing less GATC 'tags', whilst these peaks were underrepresented within 3'UTRs, exons and intergenic regions (Figure 3.6A).

As Yki, Sd, and Tgi peaks were enriched within the 5'UTR, near gene promoters, I further analysed the location of these peaks within this genomic region. Focusing within a 3kbp window of the transcriptional start site (TSS) of genes revealed that Yki, Sd, and Tgi peaks were preferentially enriched around the TSS (Figure 3.6B). All three Dam-fusion proteins showed a strong peak very close to the TSS. These results indicate that Yki, Sd, and Tgi bind preferentially near the promoter regions of putative target genes. This is consistent with Yki and Sd ChIP-seq data in the *Drosophila* wing disc, which reported Yki and Sd binding predominantly in promoter regions of genes (Oh et al. 2013).

3.2.3.iv Yorkie, Scalloped and Tgi co-occupy thousands of gene loci in the developing *Drosophila* eye disc

To identify candidate genes that are shared targets of Yki, Sd, and Tgi, I examined the overlap of peaks and putative target genes for each Dam-fusion protein. Yki, Sd, and Tgi co-bound a large proportion of genomic peaks, sharing 3,267 peaks between all three proteins (Figure 3.7A). 3,770 peaks overlapped between Yki and Sd, 4,442 peaks overlapped between Sd and Tgi, and 3,426 peaks overlapped between Yki and Tgi (Figure 3.7A). Interestingly, Sd and Tgi had more overlapping peaks compared to either Sd and Yki or Yki and Tgi. Additionally, unique peaks were observed for each protein; Yki bound 868 unique peaks, Sd bound 524 unique peaks, and Tgi bound 1,077 unique peaks (Figure 3.7A). The overlap of putative target genes associated with the peaks for each Dam-fusion protein was also examined. This revealed a similar pattern, whereby 1,811 genes were co-bound by Yki, Sd, and Tgi (Figure 3.7B). Yki and Sd shared 2,070 target genes, Sd and Tgi shared 2,420 target genes and Yki and Tgi shared 1,921 target genes (Figure 3.7B). Additionally, there were target genes unique to each protein; Yki bound 525 unique target genes, Sd bound 280 unique target genes and Tgi bound 391 unique target genes (Figure 3.7B).

In order to understand the similarity between the Dam-fusions, a pairwise correlation analysis was performed. This assesses the strength of association between different pairs of Dam-fusion proteins using the Log fold change (LogFC) of shared peaks. The Pearson score

(represented as an R value) conveys how closely associated two variables are; for example, the correlation of LogFC of a shared peak between two Dam-fusion proteins. Pairwise comparisons between all of the DamID datasets showed a positive correlation (Figure 3.7C-E). The correlation between Yki and Sd peaks (Figure 3.7C), and Yki and Tgi peaks (Figure 3.7D) were moderate; r=0.48 and r=0.41 respectively. Sd and Tgi shared peaks showed the strongest correlation, with a Pearson score of r=0.76 (Figure 3.7E).

3.2.3.v Transcription factor motifs enriched in Yorkie-, Scalloped- and Tgi-bound genomic regions

To identify transcription factor motifs enriched in peaks that were bound by Yki, Sd, and Tgi, Homer analysis was performed. Hypergeometric Optimization of Motif EnRichment (Homer) analysis is a differential motif discovery algorithm (Heinz et al. 2010), and was used to compare the DamID peak genomic regions to randomised background sequences from the *Drosophila* genome. The most significantly enriched consensus motif within Yki peaks was that of Trithorax-like (Trl). This was expected given the fact that Trl is required for Yki target gene activation in the eye and wing disc (Bayarmagnai et al. 2012; Oh et al. 2013). The transcription factor motif identified as the second most enriched motif within the Yki peaks (HSF1) resembles the Sd/TEAD consensus motif, although the Sd/TEAD motif itself was not detected (Figure 3.8A). Homer analysis also revealed the consensus Sd/TEAD motif as a significantly enriched motif within Sd peaks (Figure 3.8B), and Tgi peaks (Figure 3.8C), and peaks that were co-bound by Yki, Sd, and Tgi (Figure 3.8D). The most enriched transcription factor motif in the shared Yki, Sd, and Tgi peaks was the Grainy head consensus motif, which was also identified as a significantly enriched motif in each individual DamID dataset. An additional interesting enriched motif in shared Yki, Sd and Tgi peaks was the AP-1 consensus sequence, which was previously found to be enriched in regions of DNA that neighbour YAP/TEAD DNA binding sites (Zanconato et al. 2015).

3.2.3.vi Enriched signalling pathways among putative Yorkie, Scalloped and Tgi target genes

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis is a commonly used method to uncover signalling pathways that are enriched in large-scale datasets, and was applied to the putative target genes identified from Yki, Sd, and Tgi targeted DamID in the developing eye disc. For all three DamID datasets, there was a significant enrichment of genes involved in several pathways, including MAPK signalling, Hippo signalling, Wnt signalling, apoptosis, ECM-receptor interaction, Hedgehog signalling, and TGF-beta signalling (Figure 3.9). The full list of KEGG analysis results can be found in Appendix Table 8.2.1. Yki and Sd have been shown to regulate the transcription of Hippo pathway genes in a negative feedback loop (Hamaratoglu et al. 2006b; Genevet et al. 2010). As expected, KEGG analysis revealed that Hippo signalling was the second most significantly enriched pathway in all datasets (p value = 4.84^{-10}). Interestingly, MAPK signalling was the most enriched pathway, with a third of the genes in this category (32/93) being shared putative target genes of Yki, Sd, and Tgi (p value = 3.53^{-11}). Several pathways were enriched in only one dataset. For example, Yki target genes were enriched in pathways involved in the biosynthesis of amino acids and glycosaminoglycan biosynthesis (Figure 3.9, Appendix Table 8.2.1B). Sd target genes were uniquely enriched in Toll and Imd signalling, metabolism of alanine, aspartate and glutamate, ABC transporters and circadian rhythm signalling (Figure 3.9, Appendix Table 8.2.1C). Tgi target genes were uniquely enriched in several metabolic pathways (e.g. fatty acid metabolism, nucleotide sugar metabolism, nicotinate metabolism, and sulphur metabolism) (Figure 3.9, Appendix Table 8..2.1D).

Gene Ontology (GO) analysis was also carried out to identify key gene ontology terms that were enriched in the DamID datasets (Figure 3.10). The top 20 significantly enriched GO terms for targeted DamID target genes can be found in Appendix Table 8.2.2. GO analysis revealed a multitude of enriched GO terms in putative Yki, Sd, and Tgi target genes; these included terms related to organismal and tissue development and morphogenesis. These GO terms were expected as the growing *Drosophila* eye disc undergoes a range of developmental and morphogenesis processes.

3.2.3.vii Comparison of putative Yorkie and Scalloped target genes identified by targeted DamID and ChIP

Targeted DamID has been reported to generate comparable results to ChIP. For example, Southall et al., (2013) compared Dam-Pol II binding and ChIP-seq of Pol II in *Drosophila* larval salivary glands, and found that the binding profiles were very similar and there was a significant correlation between DamID and ChIP datasets. Several studies have used ChIP based assays for identifying Yki and Sd target genes in the developing *Drosophila* eye disc (Ikmi et al. 2014; Oh et al. 2014). In these studies, Yki and Sd putative target genes were identified in late third instar larval eye discs (Ikmi et al. 2014; Oh et al. 2014). To identify similarities and differences between Yki and Sd putative target genes identified using targeted DamID versus ChIP-seq in the *Drosophila* eye disc, I compared the putative target genes identified in this thesis to published ChIP datasets. Overall, ChIP for Yki in the eye disc revealed double the number of putative Yki target genes compared to targeted DamID; Ikmi et al (2014) identified 5,732 genes, Slattery et al (2013) identified 5,337 genes, and targeted DamID (this study) identified 2,705 genes (Figure 3.11 A & C). Yki DamID overlapped with 53% of genes identified in Ikmi et al., (2014), and overlapped with 37% of genes identified in Slattery et al., (2013) ChIP data. KEGG pathway analysis of shared ChIP-DamID Yki target genes revealed overlapping genes are enriched in MAPK, Hippo, Wnt, Hedgehog, Notch, and Apoptosis signalling pathways (Appendix Table 8.2.3). Genes unique to Yki DamID included genes involved in the Biosynthesis of amino acids, ABC transporters, Insect hormone biosynthesis, and Purine metabolism (Appendix Table 8.2.3). Genes unique to Yki ChIP datasets included Inositol phosphate metabolism and FoxO signalling (Appendix Table 8.2.3).

Targeted DamID of Sd in the developing eye disc revealed 2,960 putative target genes, whereas ChIP revealed 1,238 targets (Ikmi et al., 2014) and 3,944 targets (Slattery et al., 2013) (Figure 3.11 B & D). Sd DamID overlapped with only 14.7% of genes identified in Ikmi et al., (2014), and overlapped with 35% of genes identified in Slattery et al., (2013) ChIP. KEGG pathway analysis of shared ChIP-DamID Sd target genes revealed that overlapping genes are enriched in MAPK, Hippo, Wnt, Hedgehog, and apoptosis pathways (Appendix Table 8.2.4). Genes unique to Sd DamID included genes from MAPK signalling, Hippo signalling, Purine metabolism, Toll and Imd signaling pathway and additional metabolic pathways (Appendix Table 8.2.4). Target genes unique to the Sd ChIP datasets include genes involved in

Spliceosome, RNA transport, and Proteasome genes (Appendix Table 8.2.4). Differences in target genes identified through ChIP and targeted DamID could reflect one or more of the following: 1) Yki and Sd bind to different regions of the genome at different stages of eye development and/or different cells of the eye, 2) differences in experimental techniques, and/or 3) differences in data analysis methods.

Recently, a publication described the use of targeted DamID to identify Yki target genes in the developing *Drosophila* wing disc (P. Zhang et al. 2017). This publication used a different DamID analysis pipeline than the one used in the current study. Additionally, a different Yki Dam-ID fusion was used in which the Dam was positioned at the N terminus of Yki. I compared the overlap of putative Yki target genes identified in the current study to the target genes identified for Yki DamID in the wing (Figure 3.11 E). Yki bound to a similar number of genes in both the eye and the wing, with an overlap of 39%. KEGG pathway analysis of shared putative Yki DamID target genes between this study and Zhang et al (2017), revealed overlapping genes are enriched in Hippo, MAPK, Wnt, Notch and Apoptosis signalling pathways (Appendix Table 8.2.5), which represent shared genes across different tissues. Putative target genes that were unique to the eye included genes from MAPK signalling, as well as Insect hormone biosynthesis and biosynthesis of amino acids (Appendix Table 8.2.5). Putative target genes that were unique to the wing included those genes involved in Apoptosis and TGF-beta signalling. Putative Yki target genes that did not overlap between eye and wing datasets could represent tissue-specific target genes, or this could also be a result of the different analysis methods applied, or differences in the genome binding of Dam fusion protein (either Dam-Yki vs Yki-Dam). In the future, to help delineate between these possibilities, both datasets should be analysed using the same bioinformatics pipeline. A more thorough analysis of the different datasets could help to reveal interesting insights into shared and tissue-specific target genes of Hippo signalling.

3.3 Discussion

In this chapter, I investigated the genome wide binding of Yki, Sd and Tgi in the early third instar *Drosophila* larval eye using targeted DamID. This revealed that Yki, Sd, and Tgi are associated with thousands of genomic loci in the growing eye, which corresponds to thousands

of putative target genes. Yki, Sd and Tgi binding sites were enriched within promoters and intronic regions. Accordingly, genome-wide studies in *Drosophila* imaginal discs have previously found that regions bound by Yki are enriched near promoters (Oh et al. 2013). Interestingly, not all Yki/Sd/Tgi binding sites occurred near gene promoters, which suggests that Yki/Sd/Tgi also bind distal regulatory regions in the genome. Enhancers are distal regulatory elements that are often located far away from the transcriptional start site of the gene target and interact with promoters of the target gene through long distance interactions (Bulger and Groudine 2010). In human cancer cells, YAP and TEAD have been shown to bind to active enhancer regions via long range chromatin looping (Galli et al. 2015). In the future, it will be important to understand whether long range enhancer-promoter interactions are controlled by Yki/Sd/Tgi in *Drosophila* imaginal discs. This could be assessed by utilising chromosomal conformation capture techniques such as Hi-C, which map interaction between enhancers and promoters (Lieberman-aiden et al. 2009).

Tgi is a transcriptional co-repressor protein that binds directly to Sd and competes with Yki for Sd binding (Koontz et al., 2013). In the absence of Yki, Sd represses target genes through binding with Tgi and this results in impairment in the growth of epithelial tissues (Koontz et al., 2013). In this default repression model proposed by Koontz et al. (2013), Yki relieves the transcriptional repression mediated by Sd and Tgi resulting in the transcriptional activation of target genes (Guo et al., 2013; Koontz et al., 2013). The amount of Yki versus Tgi in a cell has an effect on Yki target genes; for example Tgi overexpression decreases *Diap-1* and *expanded* transcription (Koontz et al., 2013). However, the degree to which this default repression accounts for Yki transcriptional control on a genome-wide scale has not been previously investigated. In this chapter I have described the first analysis of genome-wide Tgi binding and found that Tgi binds to the majority of genomic regions bound by Sd and Yki. This suggests that Sd is the primary transcription factor that mediate the interaction of both Yki and Tgi with the genome.

ChIP studies have suggested that Yki binds to thousands of target genes that are not bound by Sd (Ikmi et al. 2014; Oh et al. 2014). In addition to Sd, Yki can bind to other transcription factors, such as Homothorax (Hth) and Teashirt (Tsh) in the eye (Peng, Slattery, and Mann 2009), suggesting that at some target genes Yki interacts with transcription factors

independently of Sd. However, my targeted DamID experiments revealed that only a small subset of putative Yki target genes are not also bound by Sd, and that these genes are enriched for components of metabolic pathways. This supports the model in which Sd is the main transcription factor that Yki interacts with to regulate target genes, and is consistent with the default repression model proposed by Koontz., et al (2013). The differences in findings obtained using ChIP-based approaches and DamID may be explained by different technical approaches, or due to the different cell types investigated (i.e. proliferating cells vs differentiating cells). In the future, it will be interesting to examine whether Yki and Tgi rely solely on Sd to interact with their cognate target genes, and whether Yki binds to a subset of target genes through alternative transcription factors. One way to address this further would be to perform targeted DamID of Yki and Tgi in *sd* mutant tissue.

My initial experiments using an N-terminal Dam-Yki transgene did not successfully identify putative Yki target genes, whereas a C-terminal Yki-Dam transgene yielded excellent results. This indicated that the position of Dam in the fusion protein is important for Yki protein function in this assay. Indeed, the N terminus of Yki is essential for its nuclear localisation and interaction with Sd and, therefore, it is possible that this interaction was impaired in Dam-Yki (Wang et al., 2016). However, it is interesting to note that a recent publication used a N-terminal Dam-Yki transgene to interrogate the genome occupancy of Yki in the *Drosophila* wing imaginal disc (P. Zhang et al. 2017). This study identified 4,986 binding sites for Dam-Yki compared to the Dam-alone control. It remains unclear why Dam-Yki failed to show Yki occupancy in the genome of the eye disc, but revealed Yki occupancy in the wing disc (P. Zhang et al. 2017), but it could relate to additional differences between the Dam-fusion transgene sequences in the published study versus the current study.

In this Chapter, I identified several transcription factor binding motifs that were enriched within genomic regions bound by Yki, Sd, and Tgi. One of these motifs was the consensus binding motifs of the AP-1 complex (Jun and Fos), which are the downstream effectors of the c-Jun amino-terminal kinase (JNK) pathway and MAPK pathway. AP-1 binding sites are enriched within putative Yki target genes (Pascual et al. 2017), and Yki and AP-1 synergise to induce alary muscle dedifferentiation in *Drosophila* (Schaub et al., 2019). Similarly, in human cancer cell lines, YAP/TEAD and AP-1 synergistically activate target genes important for driving oncogenic growth (Zanconato et al. 2015). Loss of AP-1 blunts

YAP/TAZ oncogenic growth, while gain of AP-1 enhances YAP driven tumour growth (Zanconato et al. 2015). Therefore, it is possible that AP-1 and Yki/Sd also co-regulate target genes involved in the growth of the *Drosophila* eye disc, which requires further investigation. The most enriched consensus binding motif in shared Yki, Sd, and Tgi bound genomic regions was that of the Grainyhead transcription factor. Grainyhead regulates epithelial morphogenesis in *Drosophila* tissues, and is essential for embryonic development and wound healing (Nevil et al. 2017). Grainyhead also acts as a pioneer factor which facilitates chromatin accessibility in larval tissues (Jacobs et al. 2018). It would be interesting to investigate whether Grainyhead also regulates Yki/Sd target genes and the chromatin accessibility at these gene loci.

One striking result from carrying out KEGG analyses on putative Yki, Sd, and Tgi target genes was the observed strong enrichment of MAPK pathway genes. MAPK signalling regulates cell death, proliferation and differentiation in the *Drosophila* eye disc (Freeman 1996; Karim and Rubin 1998). Targeted DamID revealed that Yki, Sd, and Tgi bound to genes encoding many components of the MAPK pathway, including those that encode for the transcriptional factors Pointed (Pnt) and Capicua (Cic), the core MAPK components Ras, Raf, Mitogen-activated protein kinase/Rolled (MAPK/Rl), MEK/Downstream of raf1 (Dsor1), and the upstream components Epidermal growth factor receptor (Egfr), Argos (Aos), Kekkon (Kek) and Sprouty (sty). A recent publication highlighted a functional link between MAPK and Yki/Sd at the transcriptional level, by showing that Yki and Sd regulate the transcription of *cic* and *pnt* (Pascual et al. 2017). Furthermore, the activity of Hippo signalling dictates the outcome of MAPK signalling, where Hippo and the repressor protein Cic prevent hyperproliferation of *Drosophila* epithelial tissues and thereby promote their differentiation (Pascual et al. 2017). Hyperactive Yki in combination with loss of *cic* synergistically activates MAPK pathway target genes to drive tissue overgrowth (Pascual et al. 2017). My findings suggest that Yki, Sd, and Tgi regulate MAPK pathway members in addition to Pnt and Cic, although it will be important to identify whether the expression of these MAPK genes are upregulated or downregulated by the Hippo pathway.

The findings outlined in this Chapter describe a new approach to examine the target genes of Yki, Sd, and Tgi specifically in proliferating epithelial cells of the *Drosophila* eye disc. The putative Yki, Sd and Tgi target genes identified here provide an important starting

point to unravelling the transcriptional program that is regulated by Hippo signalling to control tissue growth. Additionally, the transgenic DamID *Drosophila* strains described here should prove valuable for investigating Hippo pathway target genes in different cell and tissue types. An important question that arises from this data is what the transcriptional outcome of the Yki, Sd, and Tgi genome-wide binding is. This question will be addressed in Chapter 4 by comparing gene expression in normally growing tissues and in hyperplastic tissues that possess Yki hyperactivity.

Figure 3.1: Comparison of putative Yorkie and Scalloped target genes identified in the *Drosophila* eye disc from published ChIP datasets.

A-D. Venn diagrams showing the overlap of putative target genes of Yki and Sd identified in published ChIP datasets. Slattery et al., (2013) utilised ChIP-ChIP in the late third instar eye disc to find Yki and Sd target genes, shown in (A). Ikmi et al., (2014) utilised ChIP-seq in late third instar eye discs to find Yki and Sd target genes, shown in (B). The overlaps of putative target genes from these publications for Yki is shown in (C), and Sd is shown in (D).

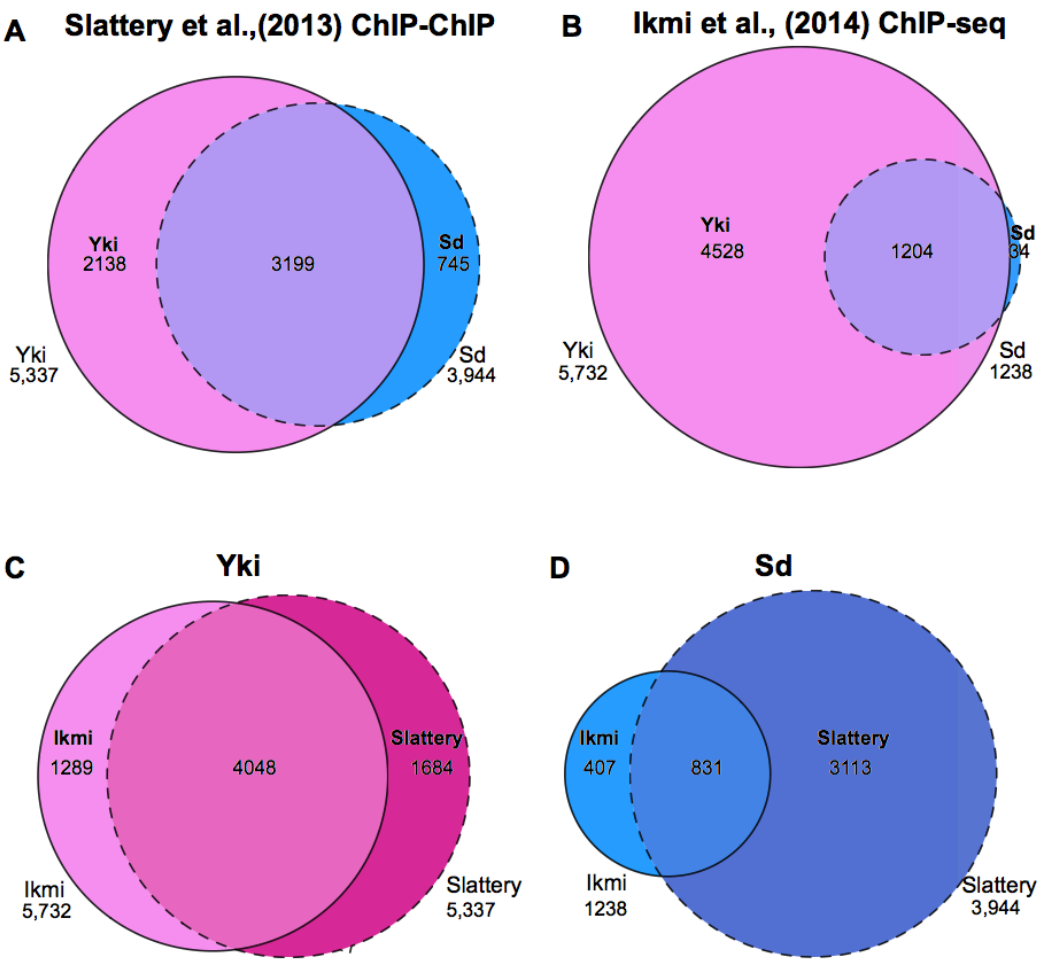

A
Slattery et al.,(2013) ChIP-ChIP
Yki
2138
3199
Sd
745
Yki
5,337
Sd
3,944
B
Ikmi et al., (2014) ChIP-seq
Yki
4528
1204
Sd
34
Yki
5,732
Sd
1238
C
Yki
Ikmi
1289
4048
Slattery
1684
Ikmi
5,732
Slattery
5,337
D
Sd
Ikmi
407
831
Slattery
3113
Ikmi
1238
Slattery
3,944

Figure 3. 2: Targeted DamID overview and expression of Dam-fusion proteins in the developing *Drosophila* eye imaginal disc.

A. Schematic representation of targeted DamID. Transcriptional regulatory proteins of interest were expressed as a fusion to the bacterial DNA adenine methyltransferase (Dam). Dam-fusion proteins were expressed in specific cells using the UAS-Gal4 system. Gal4 is a yeast derived activator protein that binds to an upstream activator sequence (UAS) to activate gene transcription. Dam-fusion proteins were expressed in eye tissue using an eye-specific Gal4 that is under the control of the *eyeless* enhancer. Translation of the Dam-fusion protein is low due to the fact that it follows a primary open reading frame (ORF) that codes for mCherry. As such, only low rates of translation initiation occur at the secondary ORF encoding the Dam-fusion protein. Expression of the Dam-fusion protein *in vivo* leads to methylation of adenines in GATC sequences near sites where the protein interacts with the genome. Following Dam-fusion protein expression, methylated DNA was isolated, amplified by PCR and sequenced, revealing Dam-fusion protein binding in the genome.

B-E'. Confocal microscope images of Dam-fusion protein expression in developing *Drosophila* eye imaginal discs. Images are of late third instar larval eye-antennal discs, anterior is to the left. Dam-fusion proteins were expressed by the *eyeless-Gal4* driver. mCherry (greyscale in B-E, and red in B'-E') was used as a surrogate marker of Dam-fusion protein expression and is visible in the posterior region of the eye disc. DAPI (blue) indicates nuclei (B'-E'). Scale bars = 100μm.

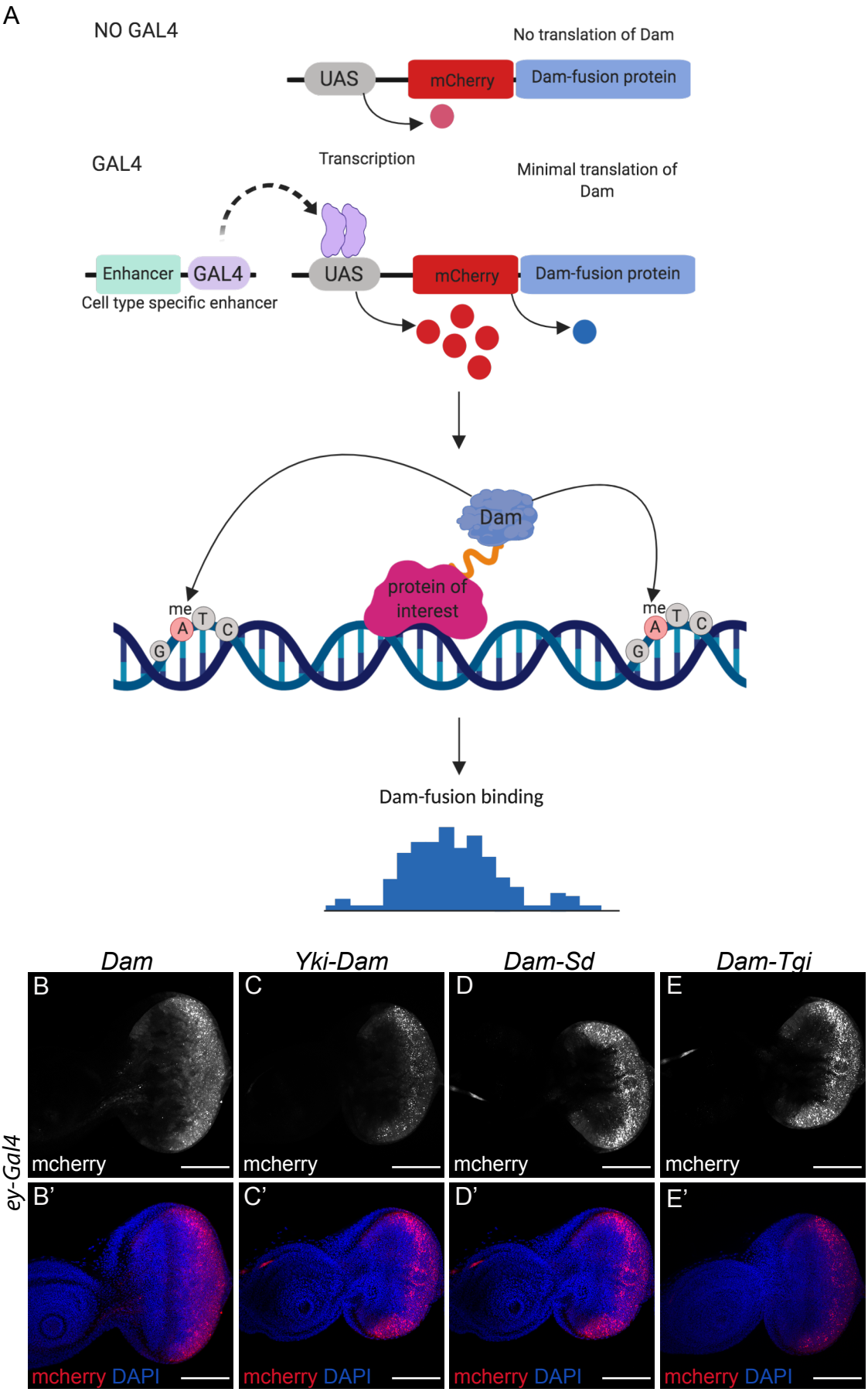

A
NO GAL4
No translation of Dam
UAS
mCherry
Dam-fusion protein
Transcription
GAL4
Minimal translation of Dam
Enhancer
GAL4
Cell type specific enhancer
UAS
mCherry
Dam-fusion protein
Dam
protein of interest
me
A
T
C
G
Dam-fusion binding
Dam
Yki-Dam
Dam-Sd
Dam-Tgi
ey-Gal4
B
C
D
E
mcherry
B'
C'
D'
E'
mcherry DAPI

Figure 3. 3: Characterisation of the *ey-Gal4* driver and development of a targeted DamID staging protocol in the developing *Drosophila* eye imaginal disc.

A. Schematic representation of the development of the *Drosophila* larval eye-antennal disc during second instar larval (L2), early third instar (early L3) and late third instar larval (late L3) stages. The antenna is located in the anterior portion of the disc (left), and the eye is located in the posterior portion (right). During late second instar, the eye-antennal disc is composed of cells that are proliferating asynchronously. During the third instar, the morphogenetic furrow forms and marks the beginning of differentiation of retinal cells. The morphogenetic furrow (MF) moves towards the anterior region of the eye throughout the third instar, creating a wave of differentiation. At the late third instar, the majority of cells that are posterior to the morphogenetic furrow undergo differentiation, while the cells that are anterior to the morphogenetic furrow remain uncommitted and proliferate.

B-D'. *Drosophila* larval eye-antennal discs from the indicated developmental stages, anterior is to the left. UAS-GFP (green in B-D) was driven by *ey*-Gal4, while DAPI reveals the nuclei (blue in B'-D'). In the second instar larvae eye disc (B, B') and the early third instar larvae eye (C, C'), *ey-Gal4* expressed GFP in cells throughout the eye. In the late third instar larvae (D,D'), *ey-Gal4* expressed GFP in cells that are posterior to the morphogenetic furrow. Scale bars = 50μm (B, B'), and 100μm (C-D').

E. Schematic diagram showing the staging protocol established for targeted DamID. The Gal4-UAS system is repressible by Gal80^{ts} at 18°C but not at 29°C. During embryonic and early larval instar stages the Dam-fusion proteins were not expressed due to the inhibition of *ey-Gal4* by Gal80^{ts} at 18°C. During the second instar larval stage, *ey-Gal4* was active by shifting the larvae to 29°C. The Dam-fusion proteins were expressed for 24 hours and the eye-antennal discs were dissected at the early third instar larval stage.

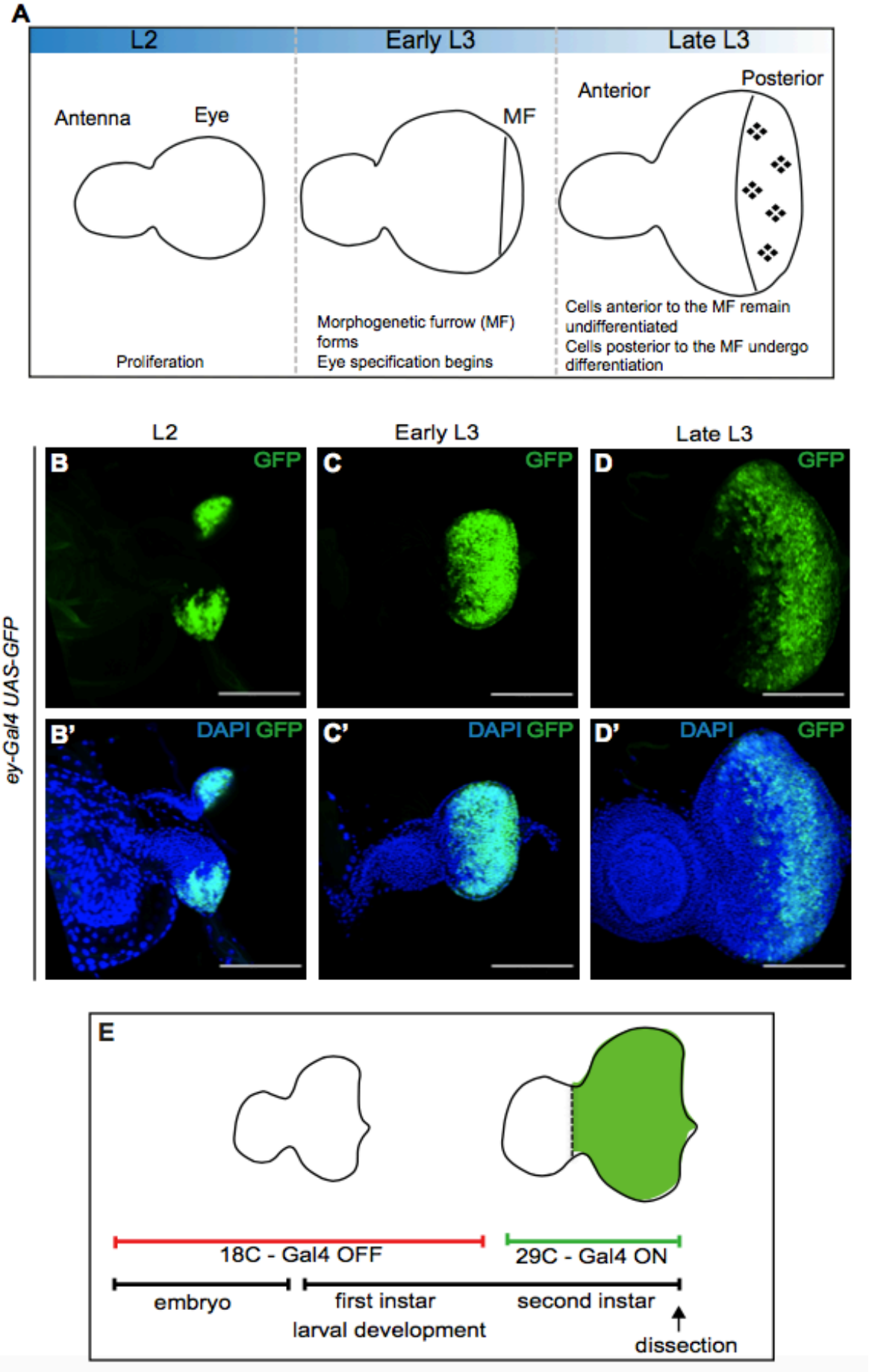

A
L2
Early L3
Late L3
Antenna
Eye
MF
Anterior
Posterior
Proliferation
Morphogenetic furrow (MF) forms
Eye specification begins
Cells anterior to the MF remain undifferentiated
Cells posterior to the MF undergo differentiation
L2
Early L3
Late L3
ey-Gal4 UAS-GFP
B
GFP
C
GFP
D
GFP
B'
DAPI GFP
C'
DAPI GFP
D'
DAPI
GFP
E
18C - Gal4 OFF
29C - Gal4 ON
embryo
first instar
larval development
second instar
dissection

Figure 3. 4: Targeted DamID DNA methylation and Multidimensional Scaling (MDS) plot of DamID datasets.

A-B. Representative PCR-amplified adenine methylated DNA fragments for Yki-Dam, Dam-Sd, Dam-Tgi and Dam-control. An aliquot of 5µl from each DamID PCR sample was separated on a 0.8% agarose gel. The DNA ladder is shown in the 1kb+ lane. A smear concentrated between 200bp and 2kb can be seen in each experimental sample (+), whereas no smears were observed in the negative control lanes (-DpnI and -Lig). The lower band shows primer dimer binding, which was subsequently eliminated at the library preparation stage, prior to sequencing.

C. Multidimensional scaling (MDS) scatterplot representation of Principle Components Analysis (PCA) for DamID-seq samples. DamID-seq samples were plotted in two dimensions so that distances on the plot approximate differences between samples based on log2 fold changes. The x and y axes show the leading LogFC dimension 1 and 2 respectively. The MDS plot displays each DamID-seq sequencing batch (shown as triangles or boxes), which are samples that were sequenced in separate sequencing runs, and the DamID-seq genotype (shown as different colours).

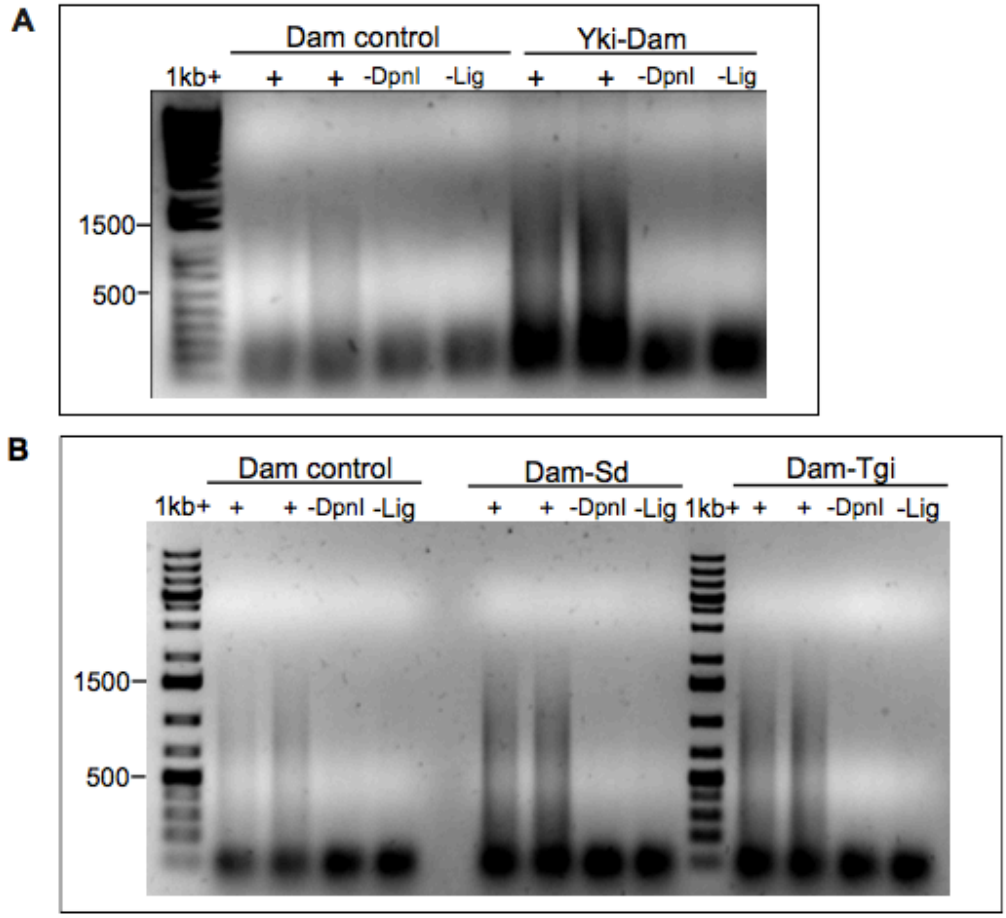

A
Dam control
Yki-Dam
1kb+
+
+
-DpnI
-Lig
+
+
-DpnI
-Lig
1500
500
B
Dam control
Dam-Sd
Dam-Tgi
1kb+
+
+
-DpnI
-Lig
+
+
-DpnI
-Lig
1kb+
+
+
-DpnI
-Lig
1500
500
C
MDS plot for DamID-seq samples
Leading LogFC dimension 2
1000
500
0
-500
-1000
-1500
-4000
-2000
0
2000
4000
Leading LogFC dimension 1
Batch 1
Batch 2
Dam
Yki
Sd
Tgi

Figure 3.5: Targeted DamID gene profiles reveal binding of Yorkie, Scalloped and Tgi at known Hippo pathway target genes.

A-D. Gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusion proteins at the *Diap1* (A), *ex* (B), *CycE* (C) and *ban* (D) genes. Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of *Diap1*, *ex*, *CycE*, and *ban* gene loci and surrounding genes. Scale bars in kb are shown.

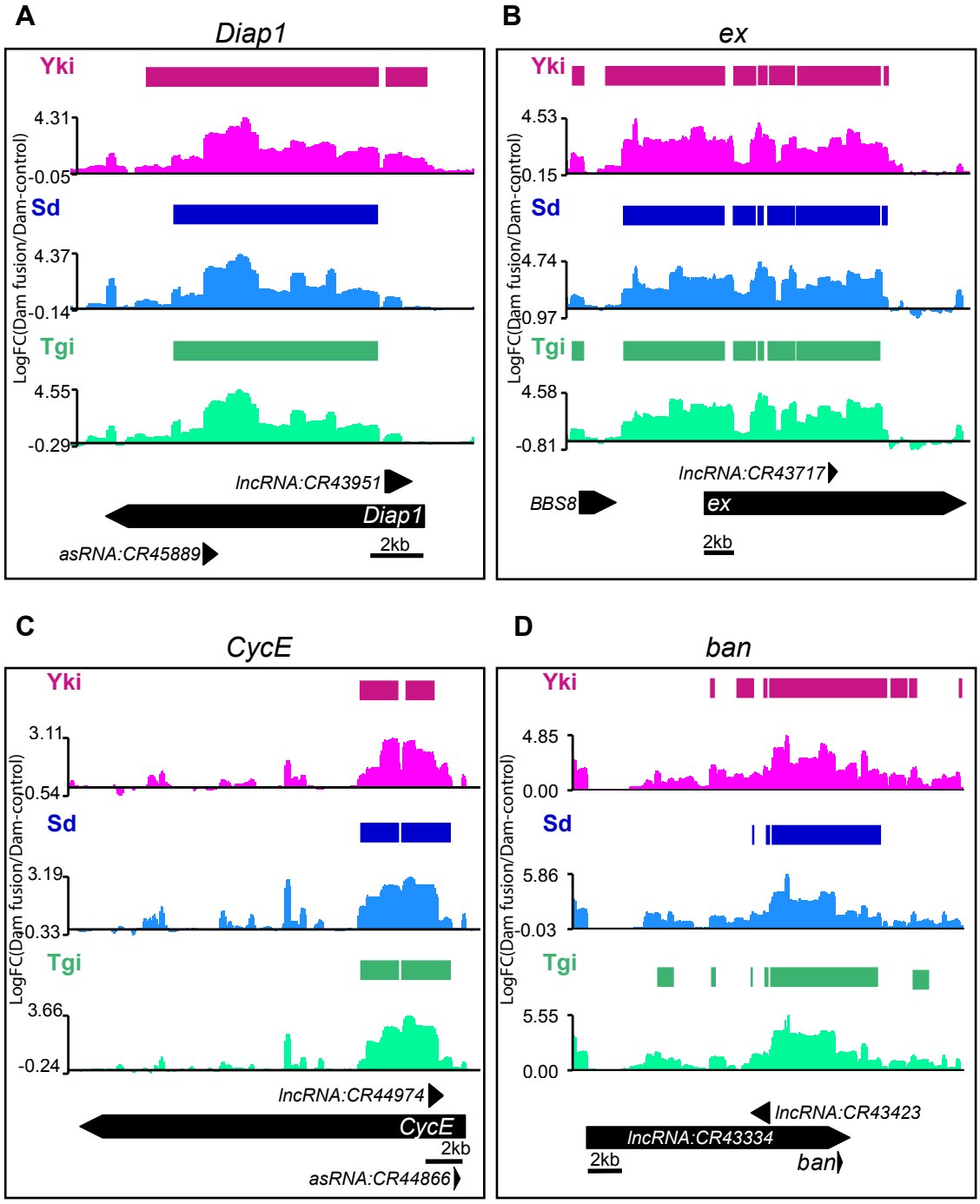

A
Diap1
Yki
4.31
-0.05
Sd
4.37
-0.14
Tgi
4.55
-0.29
LogFC(Dam fusion/Dam-control)
lncRNA:CR43951
Diap1
asRNA:CR45889
2kb
B
ex
Yki
4.53
0.15
Sd
4.74
0.97
Tgi
4.58
-0.81
LogFC(Dam fusion/Dam-control)
lncRNA:CR43717
BBS8
ex
2kb
C
CycE
Yki
3.11
0.54
Sd
3.19
0.33
Tgi
3.66
-0.24
LogFC(Dam fusion/Dam-control)
lncRNA:CR44974
CycE
2kb
asRNA:CR44866
D
ban
Yki
4.85
0.00
Sd
5.86
-0.03
Tgi
5.55
0.00
LogFC(Dam fusion/Dam-control)
lncRNA:CR43423
lncRNA:CR43334
2kb
ban

Figure 3. 6: Yorkie, Scalloped and Tgi DamID peaks are enriched near transcriptional start sites (TSS) of genes.

A. Bar graph comparing the percentage of Dam-fusion protein peaks within the following annotated regions of the *Drosophila* genome (dm6): 3'UTR, 5'UTR, intron, exon, and intergenic. The percentage of the *Drosophila* genome divided into the annotated regions, and the distribution of GATC 'tags' is shown, followed by the percentage of peaks for Yki, Sd, and Tgi within these genomic features.

B. The distance of GATC 'tags' and peaks for Dam-fusion proteins from the nearest transcriptional start site (TSS). The distribution of GATC 'tags' and peaks for Yki, Sd, and Tgi is shown within a 3 kilobase (kb) window upstream and downstream from the TSS's of annotated genes. The y-axis shows the read count frequency, while the x-axis shows the distance from the TSS in kilobases. Confidence intervals are shown as a shaded band.

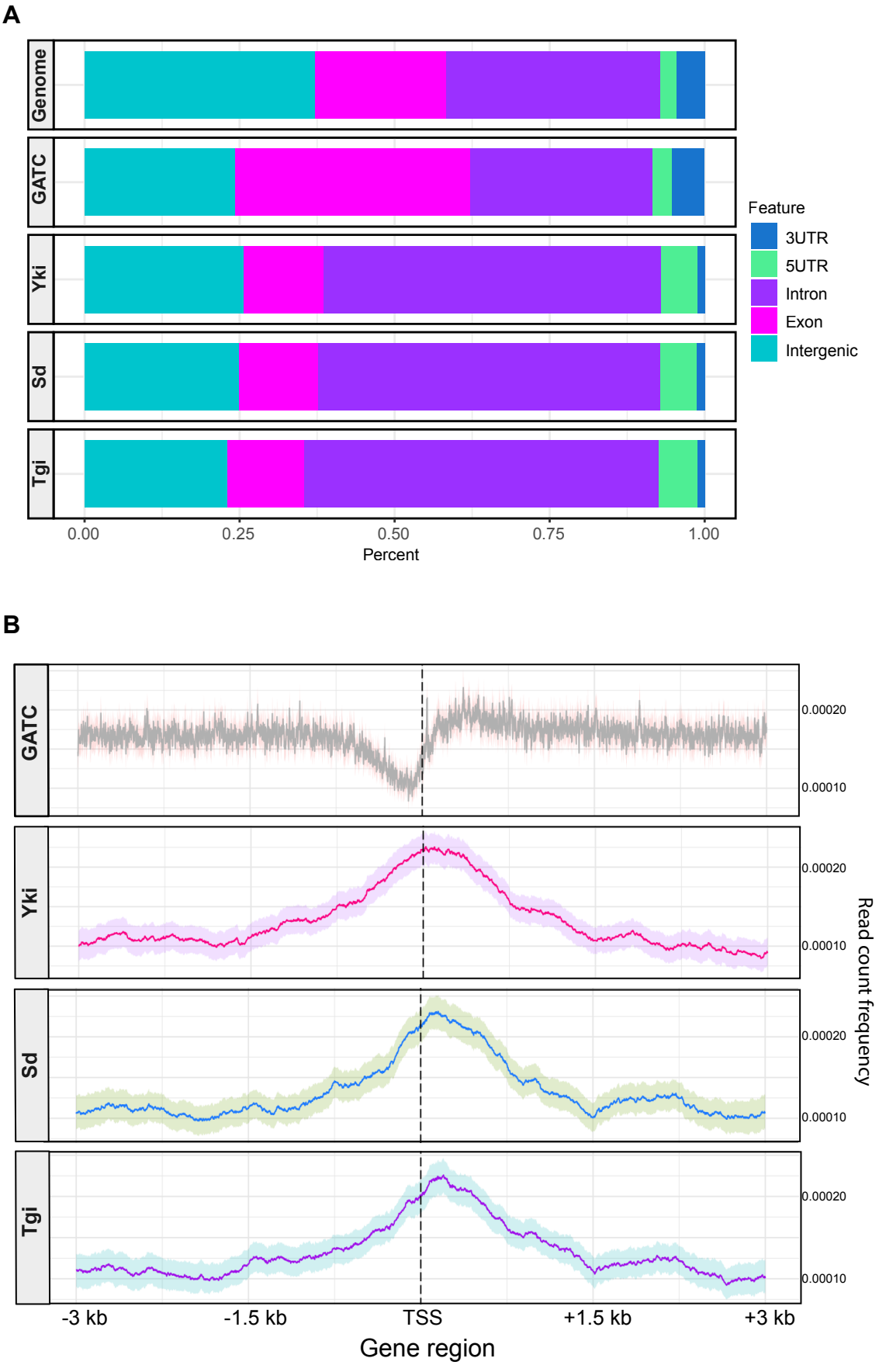

A
Genome
GATC
Yki
Sd
Tgi
0.00
0.25
0.50
0.75
1.00
Percent
Feature
3UTR
5UTR
Intron
Exon
Intergenic
B
GATC
Yki
Sd
Tgi
0.00020
0.00010
Read count frequency
-3 kb
-1.5 kb
TSS
+1.5 kb
+3 kb
Gene region

Figure 3. 7: Yorkie, Scalloped and Tgi co-occupy thousands of gene loci in developing *Drosophila* eye discs.

A-B. Venn diagrams showing the overlap of peaks (A) and the overlap of target genes (B) bound by Yki, Sd, and Tgi. DamID-seq peaks are regions that contain multiple significantly methylated GATC 'tags', and the target genes are genes that are assigned to these peaks based on proximity. The significance of overlapping target genes ($p<0.000$) was assessed by hypergeometric probability analysis.

C-E. Pairwise correlation analyses of the magnitude of peak height shared by Yki, Sd and Tgi, expressed as a Log fold change (LogFC). The LogFC for a peak was calculated by normalising the Dam-fusion protein to the Dam-alone control. Correlations between the LogFC of shared peaks of Yki and Sd are shown in (C), Yki and Tgi in (D), and Sd and Tgi in (E). The Pearson score (shown as an R value) indicates the degree of correlation between samples.

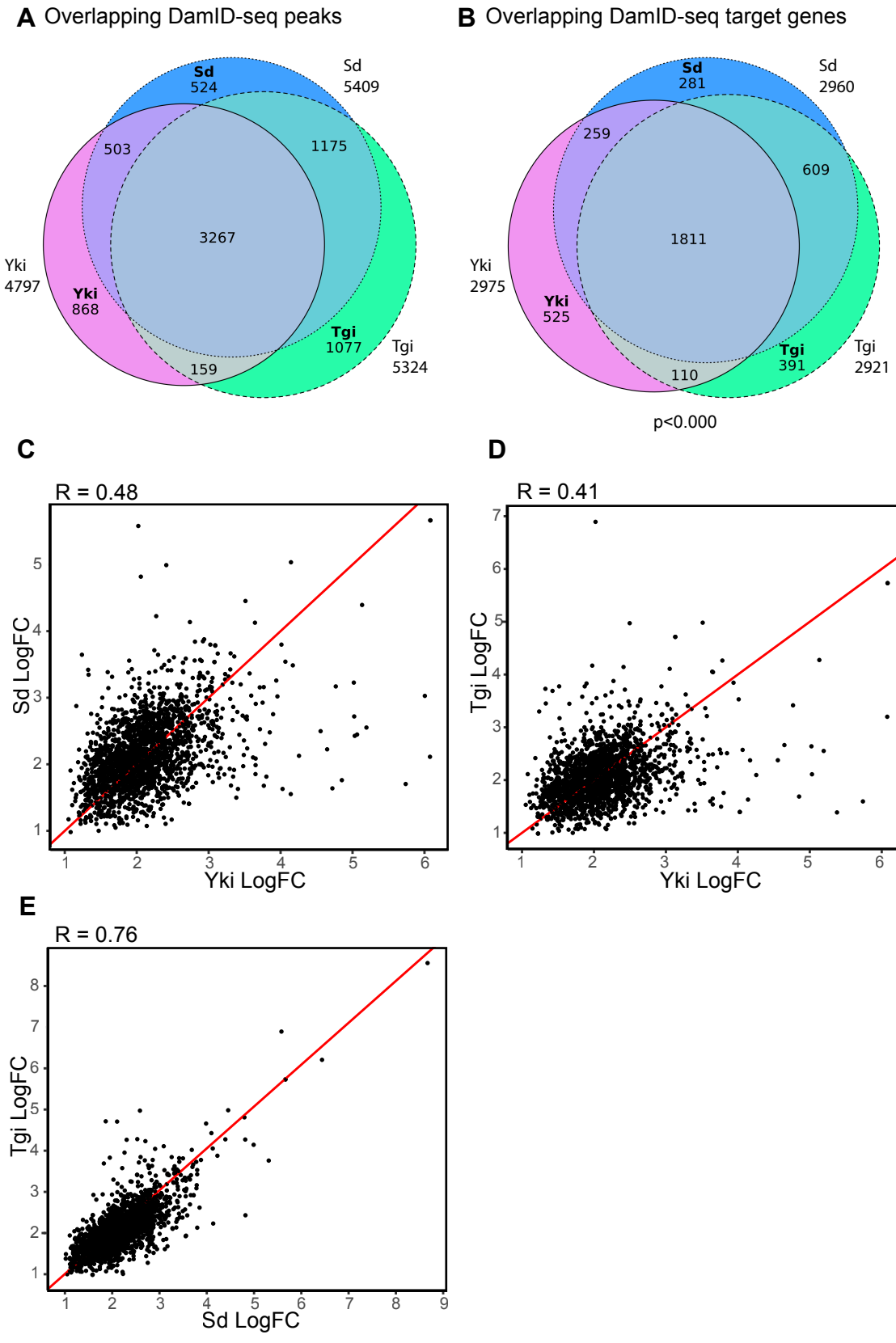

A Overlapping DamID-seq peaks
Sd
524
Sd
5409
503
1175
3267
Yki
4797
Yki
868
Tgi
1077
Tgi
5324
159
B Overlapping DamID-seq target genes
Sd
281
Sd
2960
259
609
1811
Yki
2975
Yki
525
Tgi
391
Tgi
2921
110
p<0.000
C
R = 0.48
Sd LogFC
Yki LogFC
D
R = 0.41
Tgi LogFC
Yki LogFC
E
R = 0.76
Tgi LogFC
Sd LogFC

Figure 3.8: Transcription factor motifs enriched in genomic regions bound by Yorkie, Scalloped and Tgi in the developing *Drosophila* eye disc.

A-D. Tables showing enriched transcription factor motifs in genomic regions bound by Yki, Sd, and Tgi using Homer *de novo* motif discovery analysis. The top 5 enriched transcription factor motifs in Yki peaks (A), Sd peaks (B), and Tgi peaks (C). The top 10 enriched transcription factors motifs in shared Yki, Sd, and Tgi peaks are shown in (D). *De novo* motif enrichment was implemented using HOMER and completed by resampling 100 times using permutation analysis in order to meet a False Discovery Rate of 0.01. Motifs shown in the table were identified using HOMER analysis, and the significance associated with each motif is represented by -log10(p-value).

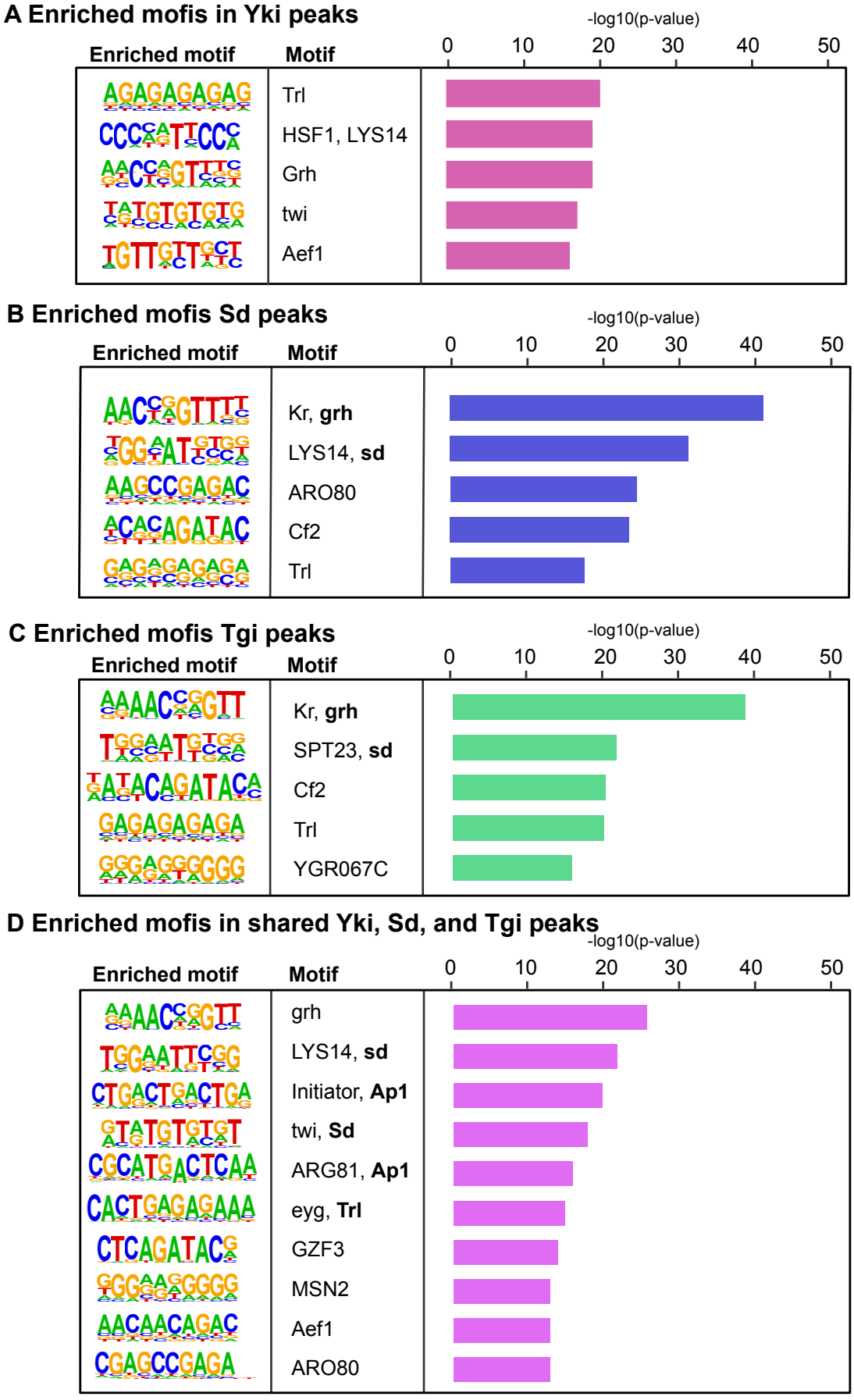

A Enriched mofis in Yki peaks
Enriched motif
Motif
-log10(p-value)
0
10
20
30
40
50
Trl
HSF1, LYS14
Grh
twi
Aef1
B Enriched mofis Sd peaks
Enriched motif
Motif
-log10(p-value)
0
10
20
30
40
50
Kr, grh
LYS14, sd
ARO80
Cf2
Trl
C Enriched mofis Tgi peaks
Enriched motif
Motif
-log10(p-value)
0
10
20
30
40
50
Kr, grh
SPT23, sd
Cf2
Trl
YGR067C
D Enriched mofis in shared Yki, Sd, and Tgi peaks
Enriched motif
Motif
-log10(p-value)
0
10
20
30
40
50
grh
LYS14, sd
Initiator, Ap1
twi, Sd
ARG81, Ap1
eyg, Trl
GZF3
MSN2
Aef1
ARO80

Figure 3.9: Signalling pathways that were enriched among putative Yorkie, Scalloped and Tgi target genes in the developing *Drosophila* eye disc.

A bubble chart showing enrichment of KEGG pathways performed on putative target genes of Yki, Sd, and Tgi that were identified using targeted DamID in the developing *Drosophila* eye disc. The y-axis represents KEGG pathway and the x-axis represents Yki, Sd, and Tgi. The KEGG pathways listed are statistically significant according to a p value of <0.05 to <0.001. The colour of each dot represents the significance of each pathway: higher significance = red, medium significance = purple, lower significance = blue, while the size of the dot represents the gene ratio (the number of enriched genes/the number of genes in the KEGG category).

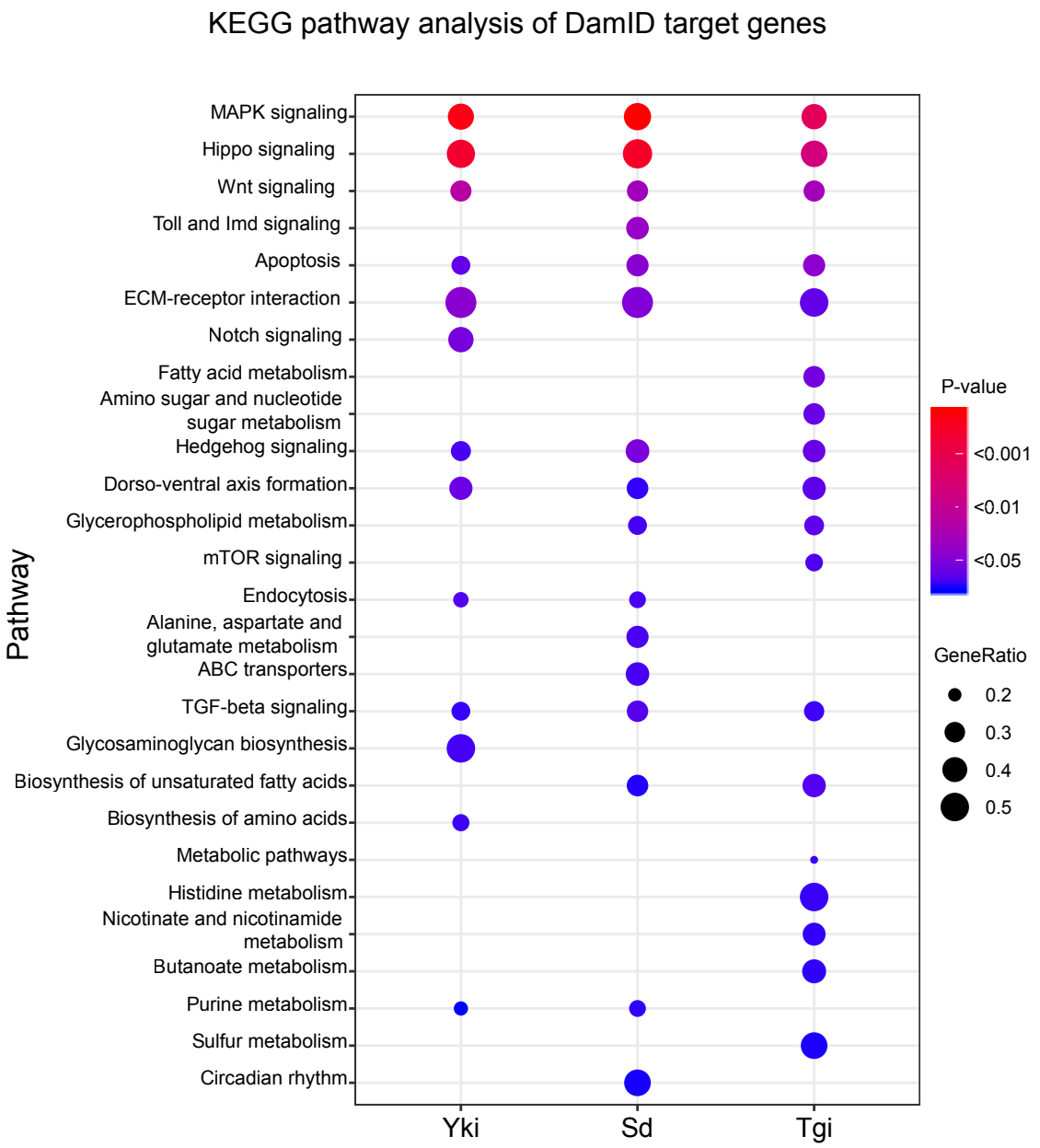

KEGG pathway analysis of DamID target genes
Pathway
MAPK signaling
Hippo signaling
Wnt signaling
Toll and Imd signaling
Apoptosis
ECM-receptor interaction
Notch signaling
Fatty acid metabolism
Amino sugar and nucleotide sugar metabolism
Hedgehog signaling
Dorso-ventral axis formation
Glycerophospholipid metabolism
mTOR signaling
Endocytosis
Alanine, aspartate and glutamate metabolism
ABC transporters
TGF-beta signaling
Glycosaminoglycan biosynthesis
Biosynthesis of unsaturated fatty acids
Biosynthesis of amino acids
Metabolic pathways
Histidine metabolism
Nicotinate and nicotinamide metabolism
Butanoate metabolism
Purine metabolism
Sulfur metabolism
Circadian rhythm
Yki
Sd
Tgi
P-value
<0.001
<0.01
<0.05
GeneRatio
0.2
0.3
0.4
0.5

Figure 3.10: Gene Ontology terms that were enriched among putative Yorkie, Scalloped and Tgi target genes in the developing *Drosophila* eye disc.

A bubble chart showing enrichment of Gene Ontology (GO) terms performed on putative target genes of Yki, Sd, and Tgi that were identified using targeted DamID in the developing *Drosophila* eye disc. The y-axis represents GO term and the x-axis represents Yki, Sd, and Tgi. The top 22 most enriched GO terms are shown, according to a p value of <0.001. The colour of each dot represents the significance of each GO term, and the size of the dot represents the gene ratio (the number of enriched genes/the number of genes in the GO category).

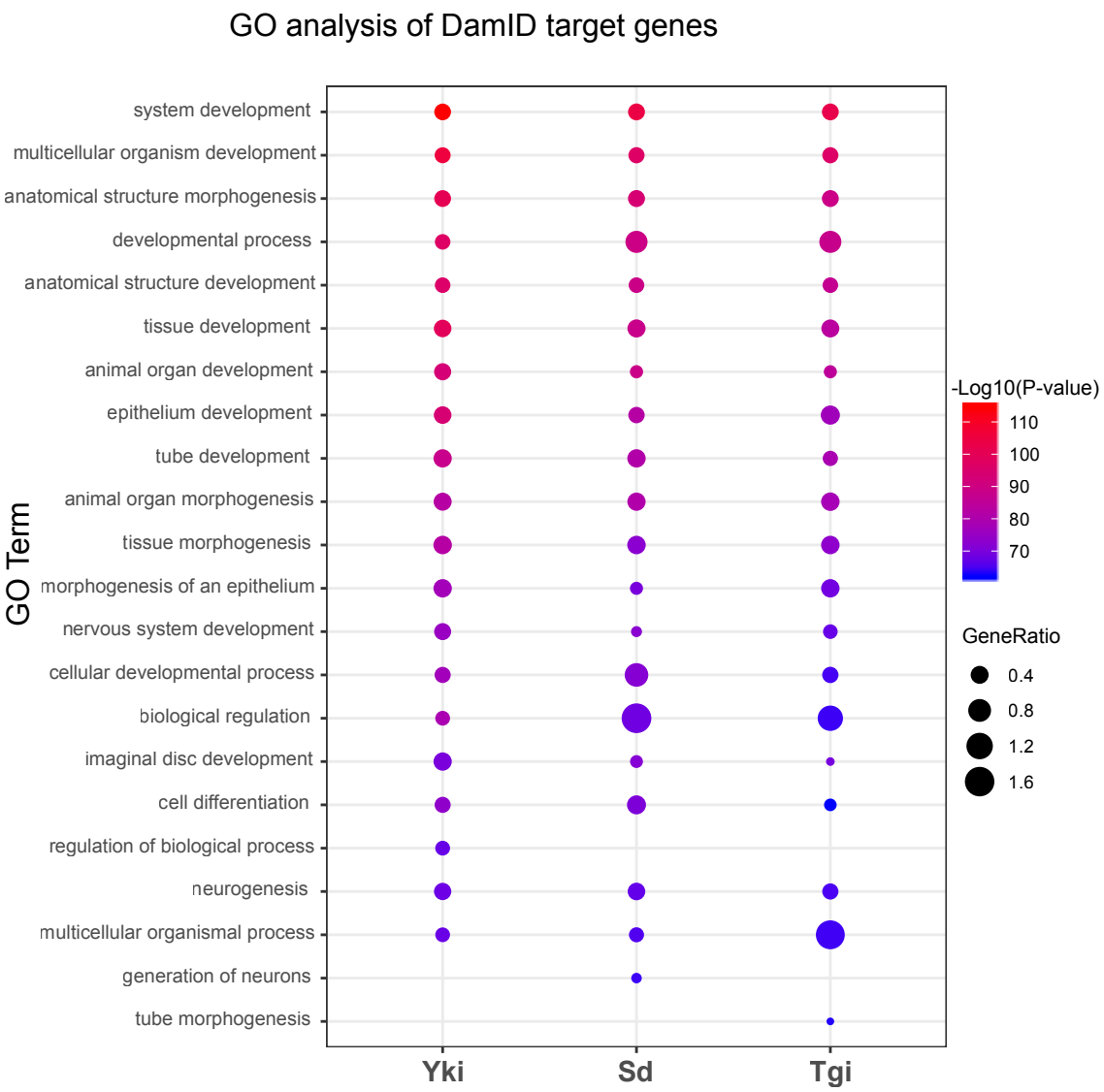
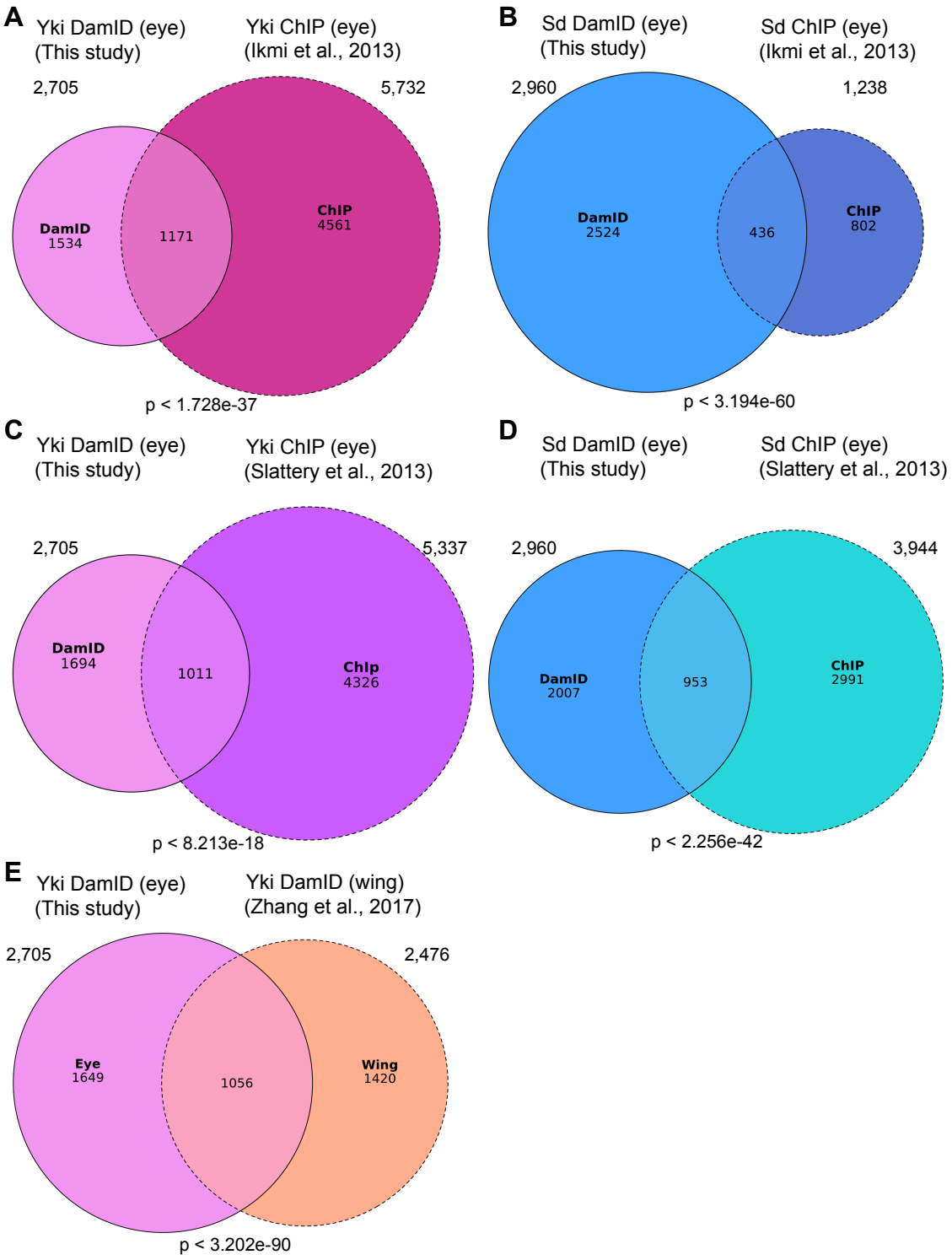

GO analysis of DamID target genes
GO Term
system development
multicellular organism development
anatomical structure morphogenesis
developmental process
anatomical structure development
tissue development
animal organ development
epithelium development
tube development
animal organ morphogenesis
tissue morphogenesis
morphogenesis of an epithelium
nervous system development
cellular developmental process
biological regulation
imaginal disc development
cell differentiation
regulation of biological process
neurogenesis
multicellular organismal process
generation of neurons
tube morphogenesis
Yki
Sd
Tgi
-Log10(P-value)
110
100
90
80
70
GeneRatio
0.4
0.8
1.2
1.6

Figure 3.11: Comparisons of targeted DamID and published ChIP datasets for putative Yorkie and Scalloped target genes.

A-D. Venn diagrams showing the overlap of putative targets genes for Yki and Sd identified from targeted DamID (this study), ChIP-seq (Ikmi et al., 2014), ChIP-ChIP (Slattery et al., 2013), and targeted DamID in the wing (Zhang et al., 2017). Slattery utilised ChIP-ChIP (2013) and Ikmi et al., (2014) utilised ChIP-seq in late third instar eye discs to identify Yki and Sd target genes. Overlapping putative target genes for Yki are shown in (A) and (C), and overlapping putative target genes for Sd are shown in (B) and (D). Zhang et al., utilised targeted DamID to identify genes bound by Yki in the developing wing disc. Overlapping putative target genes for Yki in the eye and wing is shown in (E). The significance of overlapping target genes was assessed by hypergeometric probability analysis.


A
Yki DamID (eye)
(This study)
Yki ChIP (eye)
(Ikmi et al., 2013)
2,705
5,732
DamID
1534
1171
ChIP
4561
p < 1.728e-37
B
Sd DamID (eye)
(This study)
Sd ChIP (eye)
(Ikmi et al., 2013)
2,960
1,238
DamID
2524
436
ChIP
802
p < 3.194e-60
C
Yki DamID (eye)
(This study)
Yki ChIP (eye)
(Slattery et al., 2013)
2,705
5,337
DamID
1694
1011
ChIp
4326
p < 8.213e-18
D
Sd DamID (eye)
(This study)
Sd ChIP (eye)
(Slattery et al., 2013)
2,960
3,944
DamID
2007
953
ChIP
2991
p < 2.256e-42
E
Yki DamID (eye)
(This study)
Yki DamID (wing)
(Zhang et al., 2017)
2,705
2,476
Eye
1649
1056
Wing
1420
p < 3.202e-90

Chapter 4: Identifying gene expression and chromatin accessibility changes in Yorkie driven hyperplastic *Drosophila* eye discs

4.1 Introduction

A crucial role for the Hippo pathway is to keep the growth-promoting activity of Yorkie/YAP in check. There are several mechanisms by which Hippo signalling restricts Yorkie/YAP activity, the best described of which is phosphorylation of Yorkie (Yki) at Serine 168 by the Nuclear Dbf2-related (NDR) kinase Warts/LATS1/2 (Oh and Irvine, 2008). Phosphorylation of Yki at this key site stimulates physical interaction between Yki and 14-3-3 proteins, which increases the cytoplasmic pool of Yki and decreases the nuclear pool (Oh and Irvine, 2008). Extensive research has also illuminated additional methods of Yki regulation, including phosphorylation at multiple additional sites (Oh and Irvine, 2009), phosphorylation-independent mechanisms (Oh et al., 2009) and regulation of Yki protein stability (Sun et al., 2019). This complex regulation of Yki is crucial to limit its downstream transcriptional activity. In *Drosophila*, defects in Hippo signalling and hyperactivation of Yki result in severely overgrown tissues as a result of increased cell proliferation and decreased apoptosis (Kango-Singh and Singh, 2009). Mutations in genes encoding Hippo pathway members and altered expression of the Hippo pathway are also associated with several human cancers (Han, 2019; Calses et al., 2019).

The role of the Hippo pathway in growth control has been well studied using the *Drosophila* eye disc model. In the third instar larval *Drosophila* eye, the morphogenetic furrow travels along the disc, so that cells anterior to the furrow remain undifferentiated and cells posterior to the furrow either differentiate or undergo one final round of synchronous division before differentiating (Heberlein et al., 1995; Wolff and Ready, 1991). The differentiating cells contribute to the ommatidia, and in order to reach the correct cell number excess uncommitted interommatidial cells are normally eliminated through apoptosis (Cagan, 2009). The final adult

eye contains approximately 700 precisely arranged ommatidia comprised of photoreceptor neurons, cone cells and pigment cells (Cagan, 2009). Initially, overexpression of Yki, or loss of different upstream Hippo pathway proteins, in the *Drosophila* eye disc was found to result in an increase in uncommitted interommatidial cells without changing the fate of eye disc cells (Harvey et al., 2003; Jia et al., 2003; Pantalacci et al., 2003; Tapon et al., 2002; Udan et al., 2003). Hence, the Hippo pathway has long since been considered to regulate the proliferation and survival of cells in the larval eye disc, but not to regulate cell differentiation. However, a study by Zhang et al, (2011) implicated Yki in early cell fate specification in the larval eye disc. This study found that Yki and its binding transcription factor Scalloped (Sd) are required for the establishment of the peripodial epithelium layer of the eye disc, which contributes to the cuticle of the fly head (Zhang et al., 2011). Loss of *yki* resulted in the entire disc epithelium transitioning into retina, while constitutive expression of *yki* suppressed eye formation (Zhang et al., 2011). Additionally, several other studies found that loss of upstream Hippo pathway proteins (e.g. Expanded and Fat), or hyperactivity of Yki, cause defects in morphogenetic furrow (MF) progression and photoreceptor specification (Silva et al., 2006; Wittkorn et al., 2015). It was suggested that the timing of loss of gene function during eye development, and/or the contribution of non-cell autonomous signals from wild type cells to surrounding mutant cells in mosaic organs, could account for the apparent contradictory findings between these studies (Zhang et al., 2011).

In order to better understand the role that Hippo signalling plays in developmental growth control and cancer, it is important to identify the transcriptional program that the Hippo pathway regulates. Gene expression profiling studies have revealed that the Hippo pathway regulates many genes, which have been implicated in multiple biological processes including proliferation, apoptosis, stemness, differentiation and migration (Meng et al., 2016). In *Drosophila*, well known target genes that are directly induced by Yki include the anti-apoptotic gene *Death-associated inhibitor of apoptosis* (*Diap1*) (Huang et al., 2005a), the microRNA *bantam* (*ban*) (Oh and Irvine, 2011; Peng et al., 2009), the cell proliferation genes *Cyclin E* (*CycE*) (Huang et al., 2005a) and *E2F1* (Goulev et al., 2008), and the cell growth gene *dMyc*, which is involved in ribosome biogenesis (Neto-Silva et al., 2010; Ziosi et al., 2010). Yki has also been reported to induce the expression of the Hippo pathway genes *expanded* (*ex*), *kibra* (*kib*) and *four-jointed* (*fj*), in negative feedback loops (Cho et al., 2006; Hamaratoglu et al., 2006; Genevet et al., 2010). The mammalian Yki orthologues, YAP and TAZ also activate the

mammalian homolog of *Diap1 (BIRC5)*, *Myc* (Choi et al., 2018), as well as feedback loops and genes from key signalling pathways (e.g. Wnt and Foxo signalling) (Ikmi et al., 2014), which implies at least some degree of conservation between Yki and YAP/TAZ target genes in *Drosophila* and mammals, respectively.

Gene expression is determined not only by the binding of transcription factors (TFs) to target DNA sequences, but also by chromatin accessibility (Klemm et al., 2019). Chromatin accessibility refers to the degree to which proteins, such as TFs, are able to access DNA and is determined by the organisation and modifications of nucleosomes and other chromatin-binding factors (Klemm et al., 2019). Additionally, TFs compete with histones and chromatin-binding proteins to alter nucleosome occupancy facilitating access to target DNA (Klemm et al., 2019). The chromatin landscape can change dynamically in response to external stimuli and developmental cues (Aughey et al., 2018; Klemm et al., 2019). For example, chromatin can undergo reorganisation during cellular differentiation, as well as during cellular transformation into disease states such as cancer (Davie et al., 2015; Klemm et al., 2019). Interestingly, overexpression of YAP increases chromatin accessibility and expression of fetal genes in adult cardiomyocytes (CMs), resulting in the reprogramming of CMs into a fetal-like proliferative state (Monroe et al., 2019). Yki and YAP have also been linked to the recruitment of multiple chromatin-remodelling protein and complexes such as GAGA factor (Bayarmagnai et al., 2012), Ncoa6 (Oh et al., 2014; Qing et al., 2014), the Brahma complex (Jin et al., 2013; Zhu et al., 2015), the Mediator complex (Galli et al., 2015), the SWI/SNF complex (Skibinski et al., 2014), and the NuRD complex (Beyer et al., 2013). The chromatin accessibility changes that occur as a result of increased Yki activity in *Drosophila* tissues currently remain unknown.

The downstream transcriptional program responsive to Yki hyperactivity in the *Drosophila* eye disc has been identified only in the differentiating cells of the eye during late larval development, which includes genes the regulate cell proliferation and survival (Ikmi et al., 2014). To date, no study has identified the gene expression changes that occur in Yki hyperactive eye discs during an earlier stage of larval development, or in entire eye discs. This is important if we are to understand the complex roles that the Hippo pathway plays during the growth and patterning phases of the *Drosophila* eye disc. In this chapter, I identify putative target genes that are responsive to Yki hyperactivity in the developing eye disc and the

accompanying changes in chromatin accessibility. By carrying out these experiments, I aimed to further our understanding of the role of the Hippo pathway in organ growth and cell fate specification via target gene regulation.

4.2 Results

4.2.1 RNAseq in normal and hyperplastic third instar larval eye discs

RNA sequencing was carried out in order to identify changes in gene expression between hyperplastic and normal third instar larval eye discs. RNA was isolated from either homozygous *FRT82B* eye discs ('normal growth') or wts^{X1}/*wtsLacZ* eye discs ('hyperplastic growth'). *warts* (*wts*) mutant tissues overgrow and become hyperplastic (Justice et al., 1995; Xu et al., 1995) as a result of Yki hyperactivity (Huang et al., 2005). A *wts* transheterozygous combination, carrying the above mentioned alleles of the *wts* gene, was selected for these experiments for the following reasons: 1) this combination of *wts* alleles was viable at the third instar larval stage of development, whereas other *wts* mutant combinations were not; and 2) the eye discs were completely mutant for *wts*, which avoided having to purify mutant cells from a mosaic population and also avoided the possibility of non-cell-autonomous effects of wild type cells on the mutant cells. The wts^{X1} allele is a strong loss of function mutant; in which homozygous mutant animals die in late embryonic or early first instar larval stages (Xu et al., 1995). The *wtsLacZ* allele is a hypomorphic mutant (partial loss of function) that was created using a P element insertion (Justice et al., 1995). *FRT82B* was selected as the "normal growth" genotype as the wts^{X1} mutant allele carried the *FRT82B* insertion on the third chromosome and therefore it provided a suitable control to compare to the *wts* transheterozygous mutant. For this experiment, I selected mid third instar larval eye discs as they were feasible to dissect and provided a suitable amount of mRNA for sequencing. At earlier developmental stages this was not feasible and, as such, this experiment used tissues of a slightly older larval stage to the targeted DamID experiments described in Chapter 3.

4.2.2 Differential gene expression analysis in normal and hyperplastic larval eye discs

RNA was isolated from normal and hyperplastic larval eye discs and subsequently run on a Tapestation automated electrophoresis machine to analyse the size, quantity and integrity of the RNA samples (Figure 4.1A). Eukaryotic RNA is commonly separated into two prominent bands at 18S and 28S using electrophoresis, which is used to assess the integrity of extracted RNA. However, upon denaturation the hydrogen bonds of insect RNA are disrupted, which results in two similar sized fragments which migrate closely with 18S rRNA (Winnebeck et al., 2010). As the RNA isolation protocol involved a denaturation step, I expected to see a prominent band of RNA around the size of the 18S rRNA, which runs close to the 2000bp marker. Indeed, I observed a strong band of RNA close to 2000bp in each sample, indicating that the integrity of these samples was high (Figure 4.1A).

RNA was then next generation sequenced and the sequencing reads were aligned to the *Drosophila* genome (Release 6; dm6). The aligned sequencing reads were used to create a counts table and lowly expressed genes were filtered out. A principle component analysis (PCA) was then carried out to determine the greatest source of variation in the data. The PCA was visualised using a multidimensional scaling (MDS) plot which plotted the samples based on log2 fold changes. The distance between the samples on the MDS plot is an approximate measure of the differences in log2 fold changes between the samples, which is calculated as the leading fold change. The MDS plot revealed that the samples clustered well according to genotype, whereby the replicates for normal growth clustered together and the hyperplastic replicates clustered together (Figure 4.1B). The normal growth and hyperplastic groups were separated in leading log fold change dimension 1, showing that the greatest variation was between these two genotype groups. This indicated that the replicates for each group were reproducible.

4.2.2.i Hundreds of genes are differentially expressed in Yorkie driven hyperplastic eye discs

The gene expression data was normalised to account for differences in sequencing depth and composition bias between the samples. Subsequently, differential gene expression analysis was used to identify genes that were differentially expressed between normal and hyperplastic eye discs. A log fold change (logFC) for a particular gene represents the difference in gene expression between hyperplastic eyes discs and normal eye discs using log counts per million. A negative logFC value signifies a gene that is downregulated in hyperplastic eye discs, whereas a positive logFC value signified a gene that is upregulated. An adjusted p value was also calculated for each gene, which takes into account multiple comparison testing. I implemented a fold change threshold of larger than 0.5 and an adjusted p-value threshold less than 0.01. This threshold was applied as it incorporated known target genes that were previously found to be increased in *wts* mutant *Drosophila* eye discs and hence are biologically relevant, including *Diap1* (Huang et al., 2005a) and *CycE* (Huang et al., 2005a). 417 genes were upregulated in hyperplastic eye discs compared with normal eye discs, while 920 genes were downregulated (Figure 4.2A). Interestingly, there were considerably more genes downregulated than upregulated in hyperplastic eye discs, which was unexpected because hyperactive Yki has most often been studied in the context of increased target gene expression.

To better identify transcriptional changes that are a direct result of Yki, Sd and Tgi binding, I next compared putative target genes identified from targeted DamID (Chapter 3) to genes that were found differentially expressed in hyperplastic eye discs. From these comparisons, I found the following: 136 Yki target genes were upregulated, while 219 Yki target genes were downregulated (Figure 4.2B); 143 Sd target genes were upregulated, while 253 Sd target genes were downregulated (Figure 4.2C); 137 Tgi target genes were upregulated, while 242 Tgi target genes were downregulated (Figure 4.2D); and 104 shared (Yki, Sd, and Tgi) target genes were upregulated, while 163 shared target genes were downregulated (Figure 4.2E). A hypergeometric test was carried out on these comparisons to assess whether the overlap between differentially expressed genes and Yki, Sd, and Tgi putative targets genes was significant. This analysis revealed that Yki, Sd and Tgi binding was significantly enriched in both upregulated and downregulated genes, with a stronger enrichment in upregulated genes

(Figure 4.2B-E). Overall, this data reveals that the largest type of gene expression change that occurs in *wts* mutant eye discs is the downregulation of Yki, Sd, and Tgi target genes.

4.2.2.ii Known Hippo pathway target genes are differentially expressed in Yorkie driven hyperplastic eye discs

Several Yki and Sd target genes have been shown to change expression in the *Drosophila* eye disc in response to changes in Hippo pathway activity. This includes the following genes: *Diap1* (Wu et al., 2003), the microRNA *ban* (Peng et al., 2009; Thompson and Cohen, 2006; Nolo et al., 2006), *ex* (Hamaratoglu et al., 2006), *kib* (Genevet et al., 2010), *fat* (*ft*) (Ikmi et al., 2014), *Insulin-like peptide 8* (*Ilp8*) (Boone et al., 2016) and *CycE* (Tapon et al., 2002). I found that all of these genes, with the exception of *ban*, were also upregulated in hyperplastic eye discs compared to normal eye discs, which validates the expression data (Figure 4.3). A heatmap was produced from normalised counts for these genes, which allowed the visualisation of gene expression across replicates and genotypes. The replicates within each genotype (normal or hyperplasic) generally displayed a similar gene expression pattern (Figure 4.3). This indicates that: 1) expected differential gene expression for known Hippo pathway target genes was observed in hyperplastic eye discs and; 2) the replicates for each genotype were reproducible. Unexpectedly, *ban* counts were low (below 0.5 counts per million) and so the expression data for this gene was excluded from the analysis. This could be a result of inefficient mapping of microRNAs to the *Drosophila* genome.

4.2.3 Comprehensive analysis of differentially expressed genes in Yorkie driven hyperplasic eye discs

In order to understand the potential roles of genes that were differentially expressed in hyperplastic eye discs, I carried out a series of analyses, including KEGG, GO and STRING. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis utilises the KEGG database and is a common technique used to uncover biological information in large-scale datasets. Gene Ontology (GO) analysis involves applying statistical tests to identify if genes of interest are associated with specific biological processes (BP), molecular functions (MF), and cellular

components (CC). Lastly, Search Tool for the Retrieval of Interacting Genes (STRING) is an online database that provides a comprehensive network of protein interactions based on experimental and predicted interaction information (Szklarczyk et al., 2011). STRING analysis utilises gene coexpression data, experimental data, and literature searches to build a network of interactions, thus creating clusters of protein networks. These three analyses were performed on the 417 genes that were upregulated and the 920 genes that were downregulated in hyperplastic eye discs compared to normal eye discs. These analyses revealed several signalling pathways and biological processes that were upregulated or downregulated in hyperplastic eye discs. The results are described briefly below, and then in more detail in the following sections.

KEGG and STRING analysis revealed that genes upregulated in hyperplasic eye discs were predominantly enriched in metabolism related pathways, including drug metabolism, metabolism of xenobiotics by cytochrome P450, glutathione metabolism, ascorbate and aldarate metabolism, and retinol metabolism (Figure 4.4A, Figure 4.6, and Appendix Table 8.2.6A). Additionally, the following pathways were also enriched in upregulated genes: non-homologous end-joining, apoptosis, FoxO signalling and biosynthesis of amino acids. KEGG analysis was also carried out on Yki, Sd, and Tgi target genes that were differentially expressed in hyperplastic eye discs (Appendix Table 8.2.6A). This analysis revealed that direct targets of Yki, Sd, and Tgi that are upregulated in hyperplastic eye discs include genes enriched in Hippo signalling, drug metabolism, and ABC transporters. STRING analysis revealed additional pathways and processes that were not found in KEGG analysis, such as early retinal determination genes (RD), SPALT transcription factor related genes, and cuticle related genes (Figure 4.6). GO analysis revealed that 325 GO terms were significantly enriched amongst upregulated genes, including glutathione transferase activity and metabolism, several extracellular matrix related terms, cellular response to gamma radiation, sulphur compound metabolic processes, transferase activity, cuticle structure, and development (Figure 4.5A, Appendix 8.2.7A).

KEGG and STRING analysis revealed that genes downregulated in hyperplasic eye discs were predominantly enriched in different signalling pathways, including MAPK signalling, Dorso-ventral axis formation, and ECM-receptor interaction (Figure 4.4B and

Figure 4.7, Appendix Table 8.2.6B). Several metabolic pathways were also enriched using KEGG pathway analysis, including: glycerophospholipid metabolism, arginine and proline metabolism, and glycine, serine and threonine metabolism (Figure 4.4B). Additionally, direct targets of Yki, Sd and Tgi that were downregulated in hyperplastic eye discs included genes enriched in MAPK signalling, Dorso-ventral axis formation, Pyruvate metabolism, and Hedgehog signalling (Appendix Table 8.2.6B). STRING analysis highlighted that a large proportion of genes that were downregulated in hyperplasic eye discs were involved in retinal determination and photoreceptor differentiation (Figure 4.7). Similar findings were revealed by GO analysis which showed that downregulated genes were significantly enriched in terms related to organ development, eye development, and neurogenesis (Figure 4.4B, Appendix Table 8.2.7B). Interestingly, there were many significant GO terms related to the development of the eye and photoreceptor differentiation. These analyses indicated that genes that are downregulated in hyperplastic eye discs include those that play a role in the development, specification and differentiation of the *Drosophila* eye disc.

4.2.3.i Genes regulated by Spalt transcription factors are upregulated in Yorkie driven hyperplastic eye discs

STRING analysis revealed that genes upregulated in hyperplastic eye discs included those that are regulated by the Spalt transcription factors (SPALT cluster). In *Drosophila*, the Spalt transcription factors play a range of roles in development, including patterning of the wing disc and morphogenesis of the respiratory system (De Celis and Barrio, 2009). In the eye disc, the Spalt transcription factors are important for the late steps of photoreceptor differentiation (Treisman, 2013). In *Drosophila* the Spalt zinc-finger transcription factors are encoded by the genes *spalt major* (*salm*) and *spalt-related* (*salr*) (Organista et al., 2015). Interestingly, many of the genes in the Spalt cluster are directly negatively regulated by Salm/Salr in the *Drosophila* wing disc (Organista et al., 2015). These genes include: *Gadd45, GstE6, GstE8, CG3008, mthl8, ilp8, Irbp, mre11, CG6272, ImpL2, brk, Itl, Arc1, Ady43A, CG15784, hen1, cul2, Cpr11A, CG31875, CG32581, CG32625, CG3448, ago-03,* and *Snap25* (Organista et al., 2015). A heatmap showing the upregulated expression of these genes in hyperplastic eye discs is shown in Figure 4.8A, in which you can see the expression of the Spalt cluster genes across replicates and genotypes. On the other hand, *salm* and *salr* were downregulated in hyperplastic

eye discs. Targeted DamID showed that these Spalt cluster genes, and *salm*/*salr*, were not directly bound by Yki, Sd, or Tgi Dam-fusion proteins (Figure 4.8B&C), suggesting that the Spalt-related genes are not directly regulated by Yki, Sd, and Tgi. I also examined the expression of the gene *brinker* (*brk*), which encodes a transcriptional repressor protein that negatively regulates salm/salr expression (Barrio and De Celis, 2004). Interestingly, I found that *brinker* (*brk*) was shown to be directly bound by Yki, Sd, and Tgi Dam-fusion proteins, and was also upregulated in hyperplastic eye discs (Figure 4.8D). This data could indicate that Yki, Sd and Tgi directly upregulate *brk* which results in transcriptional repression of *salm*/*salr* and thereby an upregulation of Spalt target genes.

4.2.3.ii *Glutathione S-transferase (GST)* genes are putative targets of Yorkie and are upregulated in hyperplastic eye discs

KEGG and STRING analysis highlighted a set of *glutathione S-tansferase* (*GST*) genes that are upregulated in hyperplastic eye discs. The *GST* gene family encodes proteins that play important roles in detoxification, oxidative stress response, regulation of signalling pathways involved in proliferation and apoptosis, and resistance to chemotherapeutic agents (Laborde, 2010; Hayes et al., 2005). *GST* genes that were upregulated in overgrown eye discs included: *GstD3*, *GstD10*, *GstE1*, *GstE3*, *GstE6*, *GstE8*, *GstE10*, *GstT3*, and *GstT4* (Figure 4.9A). Interestingly, using targeted DamID, I found that many these genes were directly bound by Yki, Sd, and Tgi Dam-fusion proteins (Figure 4.9B-E). The *GST* genes are located in clusters in the *Drosophila* genome, for example *GstD1* to *GstD10* are located in a cluster on chromosome 3R, while *GstE1* to *GstE10* are located in a cluster on chromosome 2R. I found that Yki, Sd, and Tgi Dam-fusion proteins were bound throughout the *GstD* gene cluster with significant peaks assigned throughout this genomic region (Figure 4.9B). Additionally, Yki, Sd, and Tgi Dam-fusion protein binding was seen throughout the *GstE* cluster, as well as GstT3, however significant peaks were not assigned at these genes indicating that the binding of the Dam-fusion proteins was below the statistical threshold of our analysis pipeline (Figure 4.9C&D). Yki, Sd, and Tgi Dam-fusion proteins also bound to *GstT4* with significant peaks assigned at this gene region (Figure 4.9E). These results show that Yki, Sd, and Tgi bind directly to multiple *glutathione S-transferase* genes and regulate their transcription in the *Drosophila* eye disc.

4.2.3.iii Cuticle genes are putative targets of Yorkie and upregulated in hyperplastic eye discs

Drosophila eye discs are composed of cells that give rise to the eye, ocelli and cuticle of the fly head (Roignant and Treisman, 2009; Treisman, 2013). STRING analysis revealed that a cluster of genes related to the development of cuticle was upregulated in hyperplastic eye discs (cuticle cluster, Figure 4.6). Insect cuticle is composed primarily of chitin, which is a long-chain polysaccharide (Karouzou et al., 2007; Rebers and Willis, 2001). Genes in the STRING "cuticle" cluster included *TweedleE* (*TwdlE*), *obstructor-E* (*obst-E*), *CG1136,* and many genes from the CPR cuticle protein family including *Lcp65Ag2*, *Lcp65Ag3*, *Cpr78E*, and *Cpr66D* (Figure 4.10A). These genes have been either demonstrated or predicted to be involved in chitin cuticle development, although many of these genes have not been well characterised. Interestingly, some of the upregulated cuticle-related genes were also directly bound by Yki, Sd, and or Tgi Dam-fusion proteins. For example, targeted DamID revealed that Yki, Sd, and Tgi Dam-fusion proteins were strongly bound to *TwdlE* (Figure 4.10B) and *CG1136* (Figure 4.10C). Many of the cuticle protein family genes are located in a gene cluster on the third chromosome and Yki, Sd and Tgi Dam-fusion proteins were strongly bound throughout this gene cluster (Figure 4.10D). This data suggests that the Hippo pathway directly regulates genes that are important for the development of cuticle, and hyperactivity of Yki increases the expression of these cuticle related genes.

4.2.3.iv Early retinal determination genes are putative targets of Yorkie and are upregulated in hyperplastic eye discs

An additional group of genes that STRING analysis highlighted in hyperplastic eyes was early retinal determinant genes (RD) (Figure 4.6). Early retinal determinant genes (RD) genes are important for the development of the entire eye-antennal disc and include *eyeless* (*ey*), *twin of eyeless* (*toy*), *homothorax* (*hth*), *teashirt* (*tsh*), *tiptop* (*tio*), and *twin of eyegone* (*toe*). These genes promote the growth of the early eye disc and also supress the expression of late retinal determination genes, which promote retinal differentiation (Domínguez and Casares, 2005;

Treisman, 2013). Differential gene expression analysis revealed that expression of the early RD genes *tsh*, *tio*, *toy* and *toe* were increased in hyperplastic eye discs compared to normal eye discs (Figure 4.11A). Additionally, *ey*, *hth*, and *wingless* (*wg*) expression was also increased in hyperplastic eye discs, however this upregulation was below the Log fold change threshold of 0.5 used in my study. Targeted DamID revealed that these genes were also directly bound by Yki, Sd, and Tgi Dam-fusion proteins (Figure 4.11B-E). These results suggest that the Yki, Sd, and Tgi likely directly regulate the transcription of early retinal genes in the developing *Drosophila* eye disc, and that hyperactive Yki increases expression of these genes.

4.2.3.v Genes involved in retinal determination and photoreceptor differentiation are putative targets of Yorkie and are downregulated in hyperplastic eye discs

STRING and GO analysis revealed that a large proportion of genes downregulated in hyperplastic eye discs were involved in retinal determination and photoreceptor differentiation (Figure 4.7 & Figure 4.5). During the third instar larval stage, Hedgehog (Hh) and decapentaplegic (Dpp) induce a wave of differentiation that initiates at the posterior end of the eye disc, known as the morphogenetic furrow (MF) (Roignant and Treisman, 2009; Treisman, 2013). Hh and Dpp upregulate the core retinal determination (RD) genes *eyes absent* (*eya*), *sine oculis* (*so*) and *dachshund* (*dac*) (Roignant and Treisman, 2009; Treisman, 2013). Expression of *eya*, *so* and *dac* is essential for the onset of retinal differentiation in cells posterior to the MF (Roignant and Treisman, 2009; Treisman, 2013). Differential expression analysis revealed that the following later retinal determinant (RD) genes were significantly downregulated in hyperplastic eye discs compared to wild type eye discs: *hh*, *dpp*, *eya*, *so*, *ato* and *dac* (Figure 4.12). Additional genes that are important for eye fate specification were also downregulated, including *distal antenna* (dan), *distal antenna related* (*danr*), and *shifted* (*shf*). Targeted DamID data also revealed that Yki, Sd, and Tgi Dam-fusion proteins were directly bound to many of these downregulated RD genes. As shown in Figure 4.13, Yki, Sd, and Tgi Dam-fusion proteins strongly bound to *hh*, *dpp*, *eya*, *so* and *dac*. Additionally, *dan*, *danr* and *shf* were also strongly bound by these Dam fusion proteins. *Optix* was also bound by Yki, Sd, and Tgi Dam-fusion proteins, however this gene was not downregulated and showed a non-significant increase in expression (LogFC = 0.22). These results indicate that eye fate is disrupted upon Yki overexpression, as many RD genes were downregulated. Additionally, my

targeted DamID results suggests that many of these RD genes are direct targets of Yki, Sd, and Tgi.

While RD network genes are necessary for retinal determination, there are additional genes that drive the differentiation of photoreceptor cells, cone and pigment cells during the third instar larval stage. R8 cell fate is specified by the proneural gene *atonal* (*ato*), and then sealed by the zinc finger transcription factor Senseless (Sens) cells (Treisman, 2013). R8 cells express Rhomboid (Rho) and Roughoid (Ru) which cleave and activate Spitz (Spi), which promotes the differentiation of R1-R7 cells, cone and primary pigment cells cells (Treisman, 2013). Many of the genes involved in photoreceptor differentiation were downregulated in hyperplastic eye discs compared to wild type eye discs (Figure 4.14). These included *sens*, *rho*, and *ru*, as well as genes important for R7 differentiation such as *sevenless* (*sev*), *bride of sevenless* (*boss*), *prospero* (*pros*), *phyllopod* (*phyl*), *seven up* (*svp*), *lozenge* (*lz*), and *glass* (*gl*). *Notch* (*N*), *Delta* (*Dl*), and the *Enhancer of split* gene complex were also downregulated, which are required for the differentiation of R7 cells, cone cells and primary pigment cells (Treisman, 2013). Targeted DamID also revealed that Yki, Sd, and Tgi Dam-fusion proteins were bound to many of these genes, including the following genes: *rho*, *ru*, *pros*, *phyl*, *lz*, *gl*, *N*, *Dl*, and *Enhancer of split* gene complex loci. On the other hand, genes that were not bound by Yki, Sd, and Tgi Dam-fusion proteins included *ato*, *sens*, *sev*, *boss* and *svp*. These results indicate that differentiation of retinal cells was disrupted upon Yki hyperactivation and some of these differentiation related genes could be direct target genes of Yki, Sd, and Tgi. On the other hand, some of the downregulated genes do not appear to be direct targets of Yki, Sd, or Tgi, suggesting that the decreased expression of these genes could be an indirect effect of Yki hyperactivity and the disruption of retinal determination that it causes.

4.2.3.vi Extracellular Matrix (ECM) genes are putative targets of Yorkie and are downregulated in hyperplastic eye discs

An additional group of genes that STRING and KEGG analysis highlighted were those that are involved in the Extracellular Matrix (ECM). The ECM is a non-cellular material that is composed of a network of proteins that provides structural scaffolding for cells and also plays important roles in cellular processes such as proliferation, differentiation, adhesion, and

migration (Yue, 2014). Interestingly, differential expression analysis revealed that the following genes were downregulated in hyperplastic eye discs: *Laminin A* (*LanA*) and *Wing Blister* (*wb*) which encode Laminin proteins, the Perlecan *Terribly reduced optic lobes* (*Trol*), the Type IV collagen polypeptides *Cg25C* and *Viking* (*Vkg*), and *BM40/SPARC* (Figure 4.15A) (Broadie et al., 2011). Additional ECM genes that were downregulated in hyperplastic eye discs include *M-spondin* (*mspo*), *Papilin* (*Ppn*), *Multiplexin* (*Mp*), *faulty attraction* (*frac*), *Tiggrin* (*Tig*), and *Peroxidasin* (*Pxn*). *Matrix metalloproteinase 2* (*MMp2*) was also downregulated in hyperplastic eye discs, which plays a role in cleaving the ECM and is highly expressed posterior to the morphogenetic furrow (Llano et al., 2002). Targeted DamID revealed that *wb*, *Mmp2*, *trol*, *LanA*, *Ppn*, and *Mp* were bound by Yki, Sd, and Tgi Dam-fusion proteins and hence likely to be direct targets (Figure 4.15B&C shows *wb* and *Mmp2*). As the Hippo pathway has shown to be responsive to signals from the extracellular matrix (Driscoll et al., 2015; Dupont et al., 2011; Sun et al., 2014; Zhao et al., 2012), it is interesting that Yki, Sd and Tgi also directly bind and downregulate genes involved in ECM.

4.2.3.vii MAPK pathway genes are putative targets of Yorkie and are downregulated in hyperplastic eye discs

STRING analysis on downregulated genes in hyperplastic eye discs also highlighted genes that are involved in the MAPK signalling pathway (Figure 4.7). The MAPK pathway relays signals following activation of receptor tyrosine kinases (RTKs) (for example EGFR and FGFR) and involves a phosphorylation cascade of the MAPK pathway proteins (for example, Ras, Raf, and MEK) (Shilo, 2014). MAPK signalling regulates cell death, proliferation and differentiation in the *Drosophila* eye disc (Freeman, 1996; Karim and Rubin, 1998). Downregulated genes in hyperplastic eye discs included those that are involved in positive regulation of MAPK signalling such as *pointed* (*pnt*), which encodes an ETS-domain transcription factor (Figure 4.16A). Downregulated genes also included those that encode proteins that antagonise MAPK signalling, such as *argos* (*aos*) which encodes a secreted molecule that sequesters the ligand Spitz, *kekkon* (*kek*) which encodes a transmembrane protein that forms inactive complexes with EGFR, and *sprouty* (*sty*) which encodes a cytoplasmic protein that inhibits MAPK signalling (Shilo, 2014) (Figure 4.16A). Importantly, Yki has been shown to negatively regulate the expression of genes encoding the downstream transcription

factors of the MAPK pathway, the activator Pointed and the repressor Capicua (Pascual et al., 2017). Additionally, many of these MAPK related genes play a role in differentiation of photoreceptors downstream of the RTK Sevenless (Raabe, 2000). For example *pnt* is expressed in all photoreceptors and required for their differentiation, *phyllopod* (*phyl*) is important for R1, R6 and R7 fate, and *corkscrew* (*csw*) is important for R7 fate (Raabe, 2000). Targeted DamID showed that many of these genes were directly bound by Yki, Sd, and Tgi Dam-fusion proteins, including *aos*, *kek1*, *kek5*, *pnt*, *phyl*, *RasGAP1*, *sty*, and *pvr* (Figure 4.16B-E). These results suggest that Yki, Sd, and Tgi might directly regulate the transcription of MAPK pathway genes in the developing *Drosophila* eye disc, and hyperactive Yki decreases the expression of these genes. Additionally, these results further indicate that there is a disruption of eye fate and photoreceptor differentiation upon Yki hyperactivity.

Genes involved in fibroblast growth factor (FGF) signalling were also downregulated in hyperplastic eye discs, including genes encoding the ligands Branchless, Pyramus, and Thisbe, as well as the Breathless receptor (Mukherjee et al., 2012) (Figure 4.7). These FGF proteins control morphogenetic movements during larval eye development, such as the proper migration of the retinal basal glial (RBG) cells (Mukherjee et al., 2012). These genes were not directly bound by Yki, Sd, and Tgi Dam-fusion proteins, as indicated by targeted DamID. Downregulation of FGF signalling could be a result of the deregulation of eye development in the hyperplastic eye disc, which is likely to impede the movement of the morphogenetic furrow and therefore repress FGF signalling.

4.2.3 Chromatin accessibility Targeted Damid (CaTaDa) reveals changes in chromatin accessibility in hyperplastic eye discs

Changes in gene expression are accompanied by changes in chromatin organisation (Klemm et al., 2019). Regions of open chromatin often correspond to gene regulatory elements and, as such, are associated with active transcription or transcriptional repression (Klemm et al., 2019). I utilised Chromatin Accessibility profiling using Targeted DamID (CATaDa) to profile chromatin accessibility changes between hyperplastic and normal eye discs. CATaDa involves the use of untethered DNA adenine methylate (Dam) which specifically methylates regions of open chromatin, while it cannot access closed chromatin (regions in which nucleosomes are

stably complexed with DNA) (Aughey et al., 2018) (Figure 4.17A). As with targeted DamID (detailed in Chapter 3.2.1), CATaDa utilises *in vivo* expression of untethered Dam protein in specific cells, with precise temporal control (Aughey et al., 2018). The advantages of CATaDa over more widely used methods to study chromatin accessibility, such as ATAC-seq and FAIRE-seq, include: 1) It avoids isolating cells, fixing and extracting chromatin; and 2) It produces results comparable to ATAC-seq and FAIRE-seq (Aughey et al., 2018).

4.2.3.i Chromatin accessibility profiling using CATaDa in hyperplastic and normal eye discs

ey-Gal4 was crossed to *Drosophila* strains containing a transgene encoding the untethered Dam protein (for information about the Dam transgene, see Chapter 3.2.1). This was carried out in two different genotypes: 1) a "normal growth" control strain; and 2) a "hyperplastic" *wts* transheterozygous strain (wts^{X1}/*wtsLacZ*), which results in severe imaginal disc overgrowth as a result of Yki hyperactivity (Huang et al., 2005). The Dam protein was expressed for 24 hours and DamID was carried out on DNA isolated from early third instar larvae eye discs. Differential methylation analysis was implemented in order to identify differences in chromatin accessibility between normal eye discs and hyperplastic eye discs. As with targeted DamID (Chapter 3), this involved identifying GATC sequences (known as GATC 'tags') that were differentially methylated between untethered Dam in hyperplastic eye discs compared to normal eye discs. In order to determine the greatest source of variation in the data, a principle component analysis (PCA) was implemented. The PCA was visualised using a multidimensional scaling (MDS) plot, which plotted the samples based on log2 fold changes. The distance between the samples on the MDS plot approximates the differences in log2 fold changes between the samples. The MDS plot revealed that the samples clustered according to genotype, where the replicates for normal growth clustered together and the replicates for hyperplastic clustered together (Figure 4.17B). Hence, the greatest variation was between normal eyes and hyperplastic eyes, which indicates that: 1) the replicates for each genotype showed little variation; and 2) Dam accessibility and methylation differ between normal growth and hyperplastic growth.

4.2.3.ii Chromatin accessibility is altered in Yorkie driven hyperplastic eye discs

Peak calling analysis was carried out to identify genomic regions that contain multiple significantly differentially methylated GATC 'tags' in hyperplastic eye discs compared to normal eye discs. These genomic regions are referred to as "peaks" and are specific regions of the genome where the Dam methylase was able to methylate DNA, and thereby indicate accessible chromatin regions. I compared Dam binding between hyperplastic and normal tissues to identify regions of differential methylation. This analysis revealed regions of the genome that were either more 'open' or 'closed' between these two growth settings. Each peak was assigned to genes that were located within a 5kb distance, according to Aughey et al., 2018. This analysis revealed that 1476 peaks, corresponding to 1738 genes, showed increased chromatin accessibility in hyperplastic eye discs compared to normal eye discs (Figure 4.18). Conversely, 932 peaks, corresponding to 1461 genes, showed decreased chromatin accessibility in hyperplastic eye discs compared to normal eye discs (Figure 4.18). Therefore, there was a substantial change in the accessible regulatory landscape between normal eye discs and hyperplastic eye discs, with different sites becoming either more or less accessible in the hyperplastic tissue.

I next compared the gene expression results from RNAseq to the chromatin profiling results from CATaDa, with the expectation that changes in chromatin accessibility would relate to changes in gene expression. The analysis involved comparing genes that were differentially expressed with genes that contained regions that were differentially accessible in hyperplastic growth compared to normal growth (Figure 4.18B&C). The results from this revealed five different groups (previously described by Ma et al., 2019): 1) genes that were upregulated with increased chromatin accessibility in hyperplastic eye discs; 2) genes that were downregulated with decreased chromatin accessibility; 3) genes that were upregulated with decreased chromatin accessibility; 4) genes that were downregulated with increased chromatin accessibility; and 5) genes that were differentially expressed had no change in chromatin accessibility in hyperplastic eye discs. Unexpectedly, the majority of changes in chromatin accessibility between hyperplastic and normal growth were not associated with changes in gene expression. Indeed, only 11.5% of genes associated with regions of 'opening' chromatin were differentially expressed; this overlap was significantly more or less than would be expected by

random chance according to the hypergeometric distribution test ($p < 0.0027$). While only 9.7% of genes associated with regions of 'closing' chromatin were differentially expressed, which was not statistically significant according to the hypergeometric distribution test ($p < 0.41$) (Figure 4.18B&C).

I hypothesised that known Hippo pathway target genes that display altered gene expression in response to Hippo pathway changes would also exhibit chromatin accessibility changes. In order to address this, I visualised differential methylation profiles of Dam in hyperplastic and normal growth at the *Diap1*, *ex*, *kibra*, and *ban* genes (Figure 4.19). In these profiles, positive logFC indicates increased methylation in hyperplastic growth, while negative logFC indicates increased methylation in normal growth. In hyperplastic eye discs, there was increased methylation by Dam throughout *ex*, *kibra*, and *ban* (Figure 4.19 B,C,D). This was particularly evident at the *kibra* locus, which showed very strong differential methylation in hyperplastic eye discs (Figure 4.19C). On the other hand, *Diap1* displayed either increased or decreased methylation at individual genomic positions within the gene loci (Figure 4.19A). Interestingly, the peak showing increased methylation in hyperplastic growth corresponds to a known *Diap1* enhancer (Djiane et al., 2013). These results indicate that changes in expression of select, but not all, known Hippo pathway target genes are associated with measurable changes in chromatin accessibility. An increase in chromatin accessibility within these genes may result from recruitment of chromatin remodelling complexes that are known to bind to Yki, such as the Brahma complex (Jin et al., 2013; Zhu et al., 2015).

The *GSTD* and *GSTE* gene clusters, as well as *brk*, had increased Dam methylation in hyperplastic tissues compared to normal growth (Figure 4.20A-C). This indicates that chromatin was more accessible within these target gene loci in hyperplastic tissues. Increased chromatin accessibility with a corresponding increase in gene expression can be related to transcriptional activators gaining access to their binding sites (Ma et al., 2019). There were also examples of genes that were upregulated in hyperplastic tissue, directly bound by Yki, Sd, and Tgi Dam-fusion proteins, but displayed either no change in accessibility (such as *ey*) or decreased accessibility (such as *hth*) in hyperplastic growth (Figure 4.20D&E). Additionally, the Dam methylation patterns of some RD genes and photoreceptor differentiation genes, such

as *hh*, *eya*, *so*, *ato, Dl,* and *Enhancer of split gene complex* were decreased indicating regions with decreased chromatin accessibility in hyperplastic eye discs (Figure 4.21A-D).

4.2.4 Retinal development is altered in hyperplastic eye discs

The gene expression, targeted DamID, and CaTaDa data presented here indicate that expression of retinal determination and differentiation genes is lowered by hyperactive Yki in hyperplastic eye discs. As such, I investigated eye differentiation during the third instar larval and pupal stages of development in *wts* transheterozygous animals by analysing expression of Elav, which marks differentiating photoreceptors, and F-actin, which marks the morphogenetic furrow (MF). In hyperplastic tissues, there was a clear reduction in the number of cells expressing Elav, compared to wild type tissues (Figure 4.22B' compared to Figure 4.22A', and Figure 4.22D' compared to Figure 4.22C'). Additionally, the morphogenetic furrow was disrupted and could no longer be clearly identified (Figure 4.22B compared to Figure 4.22A, and Figure 4.22D compared to Figure 4.22C). These results confirm a suppression of retinal differentiation in eye tissues with hyperactive Yki. The Elav positive cells that were evident in *wts* mutant eye discs could arise because of residual Wts protein (as this experiment was performed in transheterozygotes that include a hypomorphic *wts* allele) or because some photoreceptor differentiation can occur, even in the presence of Yki hyperactivity. This is consistent with published reports showing that retinal differentiation is strongly impaired in larval eye discs that are largely comprised of wts^{X1} clones (Menut et al., 2007) and Yki overexpressing eyes (Zhang et al., 2011).

wts transheterozygous animals were lethal at the pupal stage of development and, as such, the adult eyes of these animals could not be assessed. Some *wts* mutant pupa survived until the pharate pupal stage, however most died earlier in pupal development. Therefore, I observed the eyes of pharate pupae animals, where flies resemble adults but are yet to eclose (Bainbridge and Bownes, 1981), by removing the pupal case. Consistent with defective differentiation in larval eye discs, pharate *wts* animals displayed a range of gross phenotypes with differing levels of severity; some animals had no discernible eye tissue, while others possessed very small eyes (Figure 4.22F-G compared to Figure 4.22E). The head tissue of these animals was also overgrown and disordered.

4.3 Discussion

In this chapter, I investigated gene expression changes that occurred in *wts* transheterozygous mutant eye discs, and made comparisons with putative Hippo pathway target genes identified using targeted DamID (Chapter 3) and accompanying changes in chromatin accessibility. This is the first time that this combination of analyses has been applied to growing tissues that possess hyperactive Yki. Using this approach, genes that regulate multiple biological processes were found to be differentially expressed in eyes with hyperactive Yki. Yki and Sd have been shown to directly regulate the expression of the genes encoding the downstream MAPK transcription factors Pointed and Capicua (Pascual et al., 2017), and I extended this by showing that Yki regulates the expression of multiple additional members of genes encoding members of the MAPK pathway. I also found a striking alteration in the expression of retinal determination and differentiation genes in Yki hyperplastic eye discs, which extends previous studies linking the Hippo pathway to cell fate specification in the eye disc (Wittkorn et al., 2015; Zhang et al., 2011). Several additional gene groups were also differentially expressed and putative targets of the Hippo pathway, including: glutathione S-transferases (GSTs), Spalt-related, cuticle, and extracellular matrix (ECM). Changes in chromatin accessibility was associated with some, but not all, of these gene expression changes.

4.3.1 Hyperplastic Yorkie eye discs display gene expression changes and eye morphology indicative of a change in cell fate

A striking result from my studies is the number of genes involved in eye cell fate specification that are differentially expressed in hyperactive Yki eye discs. The larval eye-antennal disc gives rise to the eye, head capsule, ocelli (additional light-sensory organ), maxillary palp and antennae (Domínguez and Casares, 2005). During larval eye development an initial proliferative phase establishes the required number of cells, a subset of which then undergo retinal differentiation (Domínguez and Casares, 2005; Treisman, 2013). The early retinal determination genes are important in defining the eye field of the larval eye antennal disc and by third instar larval stage the expression of these genes is restricted to the anterior portion of the eye that will develop into head cuticle (Domínguez and Casares, 2005; Treisman, 2013;

Roignant and Treisman, 2009). I found that putative target genes involved in early retinal determination and cuticle fate are upregulated (Figure 4.11). These upregulated early retinal determination genes include: *ey*, *hth*, *toy* and *wg*. Ey, Hth and Tsh activity promotes the early growth of the eye disc, and *wg* expression is important for the development of head cuticle (Triesman 2013). A previous study by Zhang et al., (2011) has also shown that Yki regulates *hth* expression in the peripodial epithelium (PE) of the eye disc, which is a squamous epithelium layer of the eye disc that contributes to the cuticle of the fly head (Zhang et al., 2011). In this study, Zhang et al., (2011) show that the activation of *hth* by Yki and Sd results in the repression of retinal cell fate in the PE (Zhang et al., 2011). In contrast, Yki levels are lower in the disc proper (DP), which is a columnar cell layer that contributes to the adult eye and surrounding cuticle (Zhang et al., 2011). Yki and Sd also induce wingless signalling, which then blocks the progression of the morphogenetic furrow and photoreceptor differentiation (Wittkorn et al., 2015). My results add to these studies by also showing the upregulation of cuticle genes in *wts* mutant eye discs, such as *TweedleE* (*TwdlE*), *obstructor-E* (*obst-E*), *CG1136,* and the CPR family genes (Figure 4.10). Together, these results indicate that Yki promotes a head cuticle cell fate in the eye disc. Indeed, *wts* mutant pharate pupae have overgrown head tissue with either small or no eyes (Figure 4.22E-G'), suggesting that the extra cells seen in third instar larval eye discs of *wts* mutants (Figure 4.22A-D') contributes to head cuticle instead of eye tissue.

The increase in head and cuticle cell fate in *wts* mutant eye discs is also accompanied by a decrease in the expression of putative Yki and Sd target genes involved in retinal determination, such as *hh*, *dpp*, *eya*, *so*, *ato*, *dac*, *dan* and *danr* (Figure 4.12 & 4.13). The retinal determination genes normally become expressed in the posterior region of the eye disc and progress anteriorly over time in third instar larvae, such that cells in the posterior portion of the disc stop proliferating and begin to differentiate (Domínguez and Casares, 2005; Treisman, 2013; Roignant and Treisman, 2009). Additionally, genes involved in the differentiation of photoreceptors are putative targets of Yki and Sd, and are also downregulated in *wts* mutant eyes (Figure 4.14). Chromatin accessibility at many of these genes is lowered, which might indicate loss of an activator at these gene loci, although further experiments are required to address this possibility (Figure 4.21). Consistent with these findings, retinal differentiation was perturbed in *wts* mutant eye discs and pharate pupae eyes, as determined by eye morphology and antibody staining (Figure 4.22). Interestingly, previous studies have hypothesised that Yki

and Sd block retinal cell fate of the eye disc by activating negative regulators of eye development, such as Hth and Wg, rather than directly downregulating retinal determinant genes (Wittkorn et al., 2015; Zhang et al., 2011). While the results I present in this chapter confirm this block in eye development at the level of gene expression, they also implicate Yki in the direct regulation of numerous eye development, differentiation, and cuticle related genes. Collectively, these findings indicate that the Hippo pathway controls the fate of the eye disc; where Yki promotes cuticle and head cell fate, while repressing retinal cell fate. As Zhang et al., 2011 suggests, this role is separate from its role in controlling proliferation and survival, and may be a result of the amount of active Yki within the PE and DP layers. In the future it will be important to confirm that Yki, Sd, and/or Tgi indeed directly repress candidate retinal determination target genes. Additionally, it would be also interesting to carry out targeted DamID specifically in the PE and DP layers of the eye disc to determine whether Yki/Sd/Tgi target genes differ between these two cell types.

4.3.3 How might Yorkie, Scalloped and Tgi regulate expression of retinal determination and differentiation genes?

An interesting question posed by my studies pertains to the mechanism by which Yki, Sd, and Tgi limit the transcription of retinal determination and differentiation genes. To date, in *Drosophila*, Yki is only known to activate the expression of target genes, while Tgi acts as a co-repressor of Sd (Koontz et al., 2013). On the other hand, several lines of evidence indicate that the mammalian orthologues of Yki and Sd, YAP and TEAD1-4, can repress target gene expression (Beyer et al., 2013; Hillmer and Link, 2019; Kim et al., 2015). For example, YAP and TEAD can interact with the nucleosome-remodeling and deacetylase (NuRD) repressor complex and repress target genes involved in mesendoderm (ME) cell fate, thereby maintaining human embryonic stem cell pluripotency (Beyer et al., 2013). The ability of YAP to repress specific target genes relies on the transcription factors SMAD and OCT4 (Beyer et al., 2013). An additional study also showed that YAP/TEAD recruit the NuRD complex to repress antiproliferative and apoptosis-inducing genes, resulting in the deacetylation of histones and reduced chromatin accessibility at these gene loci (Kim et al., 2015). Furthermore, YAP recruits the polycomb repressive complex 2 to repress the transcription of growth differentiation factor-15 (GD15) in breast cancer cell lines (Wang et al., 2018). Interestingly, chromatin accessibility was decreased at approximately 40% of the retinal determination and differentiation gene loci in *wts* mutant eye discs. Therefore, it is possible that in the developing

Drosophila eye disc, Yki and Sd also recruit chromatin modifiers such as the NuRD complex to reduce the accessibility of these genes and consequently decrease their gene expression. While Yki is well-known to recruit chromatin modifiers to activate target gene expression (Bayarmagnai et al., 2012; Galli et al., 2015; Jin et al., 2013; Skibinski et al., 2014; Zhu et al., 2015), many of these factors can also silence transcription depending on the context. Interestingly, a possible link exists between the brahma-associated protein (BAP) complex and Yki in target gene repression in the wing notum (Song et al., 2017), however the exact mechanism remains unknown. In the future, it will be interesting to address whether Yki/Sd/Tgi recruit chromatin modifiers to repress the expression of retinal determination and differentiation genes. Further experiments could include an investigation of the following 1) which histone modifications are present at downregulated gene loci when Yki is hyperactive; 2) whether the NuRD complex, or other chromatin modifying complexes, are required for the downregulation of these genes; and 3) whether Yki, Sd, and Tgi directly interact with chromatin modifiers such as the NuRD complex. It is also possible that Yki-mediated repression relies on additional transcription factors, as has been shown for YAP and TEAD (Beyer et al., 2013).

4.3.4 GST genes are upregulated and more accessible in Yorkie driven hyperplastic eye discs

Another new finding described here is that, in hyperplastic eye discs, *glutathione S-transferase (GST)* genes are upregulated (Figure 4.9A), display increased in chromatin accessibility (Figure 4.20), and are directly bound by Yki, Sd and Tgi Dam-fusion proteins (Figure 4.9B). Currently, the importance of these genes to Hippo signalling, hyperplastic growth and cell fate is unclear. Glutathione s-transferases (GST) enzymes are encoded by a gene superfamily and catalyse the conjugation of glutathione to various electrophilic compounds (Zhang et al., 2014). Interestingly, high levels of GSTs have been found in some cancers and have also been linked to therapeutic resistance (Zhang et al., 2014). In flies, the GST enzymes have been predominantly studied in the field of insecticide resistance (Tu and Akgül, 2005), and also regulate signalling pathways involved in stress responses, proliferation, differentiation, and apoptosis (Zhang et al., 2014). For example, GSTs can bind to and negatively regulate c-Jun N-terminal kinases (JNKs) and inhibit apoptosis (Laborde, 2010). This binding is dissociated

under conditions of oxidative stress resulting the activation of JNK and cell death (Laborde, 2010). It would be interesting to further investigate GSTs as targets of Hippo signalling given the known links between Hippo signalling and the MAPK pathway, proliferation, apoptosis, as well as therapeutic resistance to anti-cancer drugs (Kim and Myung, 2018). For example, future experiments could include assessing GST-enyzmatic activity in *wts* mutant eye discs, which can be carried out using assays that assess the conjugation of the thiol group of glutathione to the synthetic substrate 1-chloro-2,4-dinitrobenzene (CDNB) (Singh et al., 2001).

4.3.5 Chromatin accessibility is altered in Yorkie driven hyperplastic tissues

Given the published links between Yki/YAP and chromatin modifying complexes, I profiled chromatin accessibility using CATaDa. To date, changes in chromatin accessibility in response to Hippo pathway activity have not been investigated in *Drosophila* tissues. My study revealed thousands of regions with altered accessibility in hyperactive Yki eye discs (Figure 4.18A). Unexpectedly, the majority of changes in chromatin accessibility in hyperplastic growth were not associated with changes in gene expression. Changes in chromatin accessibility that do not correlate with gene expression changes can be found at inactive enhancers that are poised for future activation ('poised enhancers') (Klemm et al., 2019). In the future, it would be interesting to investigate whether regions with altered chromatin accessibility have a poised enhancer signature, which is marked by histone H3 lysine 4 monomethylation (H3K4me1) and low acetylation of histone 3 lysine 27 (H3K27ac) (Koenecke et al., 2017). Interestingly, YAP/TAZ have been linked to poised enhancers via TEAD4 following elimination of the viral *E1A* gene in transformed human embryonic kidney cells (Zemke et al., 2019). In this context, genes bound by TEAD4 and marked by H3K4me1 are poised and then undergo chromatin remodelling and gene activation following YAP binding (Zemke et al., 2019). Additionally, it should also be noted that RNAseq experiments were carried out using mid third instar larval eyes, whereas CATaDa was carried out during the second to early third instar larval stage. As such, correlation between the RNAseq data and CATaDa data may reflect differences in developmental stages; for example, a specific genomic region may become increasingly accessible during mid third instar but not earlier. Therefore, it would be worthwhile carrying out CaTaDa using mid third instar larval eye discs in order to provide a developmentally

matched comparison to the RNAseq data. In the future, it would also be prudent to perform ATAC-seq in order to confirm the chromatin accessibility changes found in this study.

4.3.6 Summary and future directions

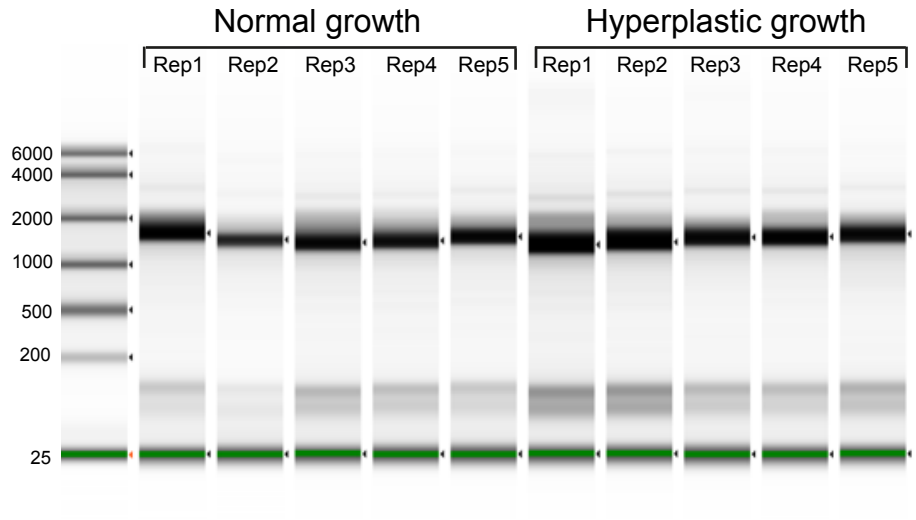
In summary, in this chapter I have presented a comprehensive analysis of gene expression changes and chromatin accessibility in eye tissues with defective Hippo pathway activity and compared these to Yki, Sd, and Tgi target genes. My results highlight a number of biological processes and signalling pathways that are differentially expressed and directly bound by Yki, Sd, and Tgi, providing several promising candidates for further investigation. Interestingly, in Yki driven hyperplastic eye discs there is a dramatic change in gene expression and chromatin accessibility of genes involved in retinal determination and differentiation. In combination with previous research (Zhang et al., 2011; Wittkorn et al., 2015), the data presented here provides evidence for two distinct and fascinating roles of Yki in the developing *Drosophila* eye: 1) it promotes the proliferation of progenitor cells in the eye; and 2) it controls retinal cell fate by directly repressing retinal cell fate genes and promoting head cuticle fate genes. The retinal determination genes studied here are directly bound by Yki, Sd, and Tgi, which suggests a role for these transcriptional effectors in repression of these genes in normal eye development. At present, Yki has only been described as an activator of transcription and hence future studies should be carried out to verify the putative repressor role for Yki identified here. Additionally, future studies should analyse gene expression in tissues that lack Yki, or have increased levels of Tgi. This would help to identify which genes are activated or repressed by these proteins.

Figure 4. 1: RNA quality and Multidimensional Scaling (MDS) plot of RNA-sequencing datasets.

A. Tapestation electrophoresis showing RNA ladder and RNA isolated from normal (*FRT82B*) and hyperplastic (*wts^{X1}/wtsLacZ*) eye discs. 5 replicates were carried out for each genotype. The TapeStation separates the RNA samples based on molecular weight and also analyses the quality (RIN) and quantity of RNA.

B. Multidimensional scaling (MDS) plot of normal (*FRT82B*) and hyperplastic (*wts^{X1}/wtsLacZ*) growth RNAseq samples. Normal growth samples are shown in purple, and hyperplastic growth samples are shown in green. The x and y axes show the leading LogFC dimension 1 and 2 respectively. Distances on the plot approximate differences between samples based on log2 fold changes.

A


Normal growth
Hyperplastic growth
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
6000
4000
2000
1000
500
200
25

B

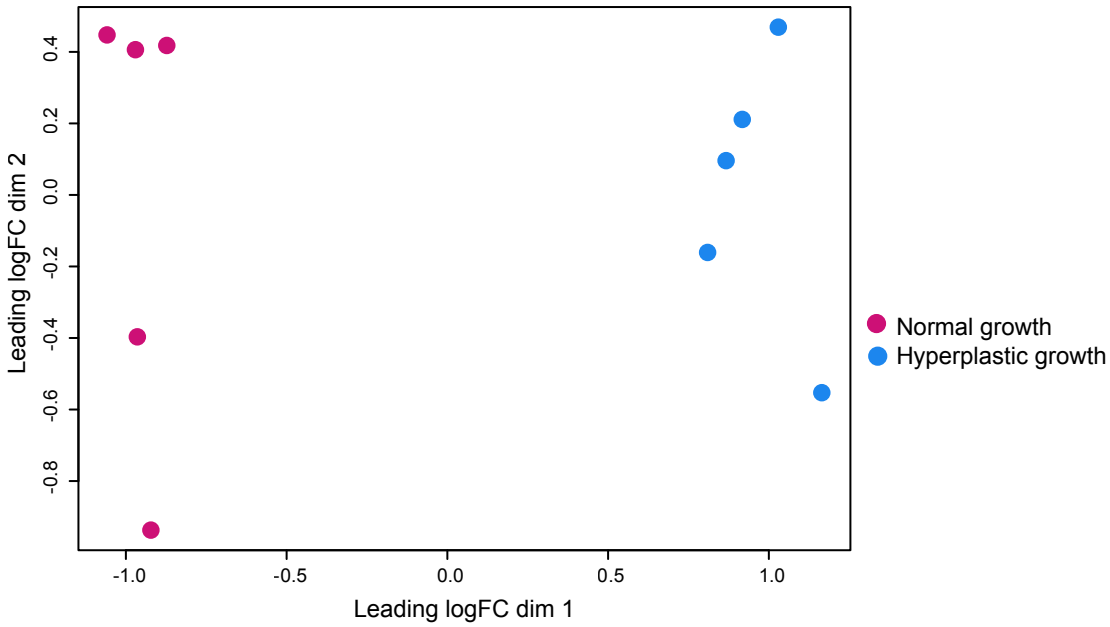

Leading logFC dim 2
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-0.5
0.0
0.5
1.0
Leading logFC dim 1
Normal growth
Hyperplastic growth

Figure 4.2: Yorkie, Scalloped, and Tgi target genes are differentially expressed in Yorkie driven hyperplastic eye discs.

A. Volcano plot of genes whose expression was upregulated (red box) or downregulated (blue box) in hyperplastic (*wts^{X1}/wtsLacZ*) eye discs compared to normal (*FRT82B*) eye discs.

The x and y-axis show the log fold change (logFC), which represents the difference in gene expression between hyperplastic and normal eye discs. Samples shown in each coloured box had a fold change equal to or larger than 0.5 and an adjusted p-value of less than 0.01. Genes that were identified as candidate target genes using targeted DamID are highlighted in pink for Yki, blue for Sd, green for Tgi, and purple for Yki/Sd/Tgi shared genes.

B-E. Targeted DamID candidate genes whose expression was upregulated or downregulated in hyperplastic (*wts^{X1}/wtsLacZ*) eye discs, for Yki-Dam (B), Dam-Sd (C), Dam-Tgi (D), and shared targets (E). The significance (p value) of overlapping target genes was assessed by hypergeometric probability analysis.

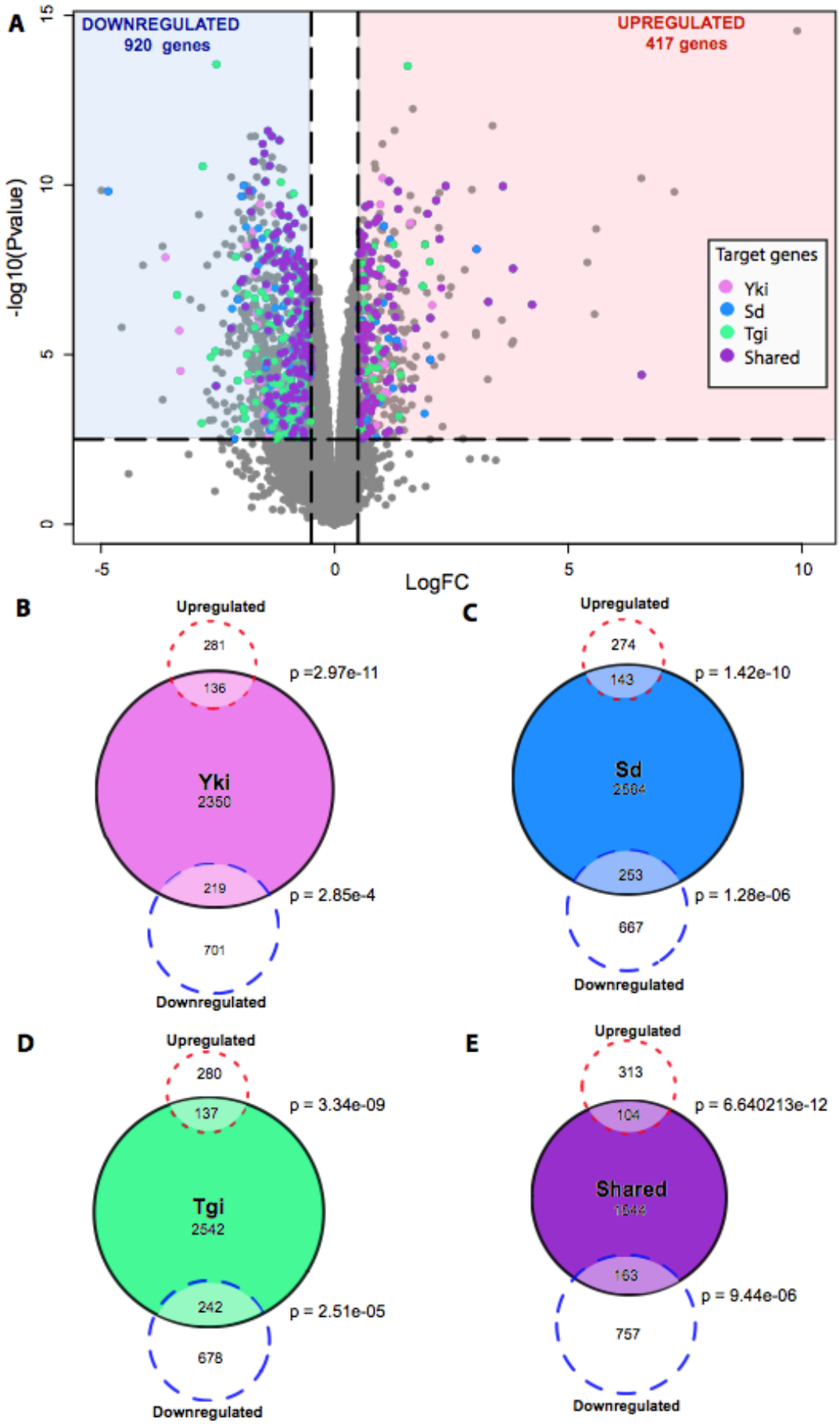
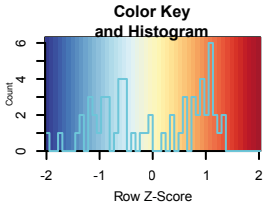

A
DOWNREGULATED
920 genes
UPREGULATED
417 genes
Target genes
Yki
Sd
Tgi
Shared
-log10(Pvalue)
LogFC
B
Upregulated
281
136
p =2.97e-11
Yki
2350
219
p = 2.85e-4
701
Downregulated
C
Upregulated
274
143
p = 1.42e-10
Sd
2564
253
p = 1.28e-06
667
Downregulated
D
Upregulated
280
137
p = 3.34e-09
Tgi
2542
242
p = 2.51e-05
678
Downregulated
E
Upregulated
313
104
p = 6.640213e-12
Shared
1544
163
p = 9.44e-06
757
Downregulated

Figure 4.3: Known Hippo pathway target genes are differentially expressed in Yorkie driven hyperplastic eye discs.

Heatmap showing gene expression levels for the known Hippo pathway target genes *fat* (*ft*), *expanded* (*ex*), *Death-associated inhibitor of apoptosis 1* (*Diap1*), *Cyclin E* (*CycE*), *kibra*, and *Insulin-like peptide 8* (*Ilp8*) in normal (*FRT82B*) (purple) and hyperplastic (wts^{X1}/*wtsLacZ*) eye discs (green). The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic compared to normal growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).


Color Key
and Histogram
Count
0
2
4
6
-2
-1
0
1
2
Row Z-Score

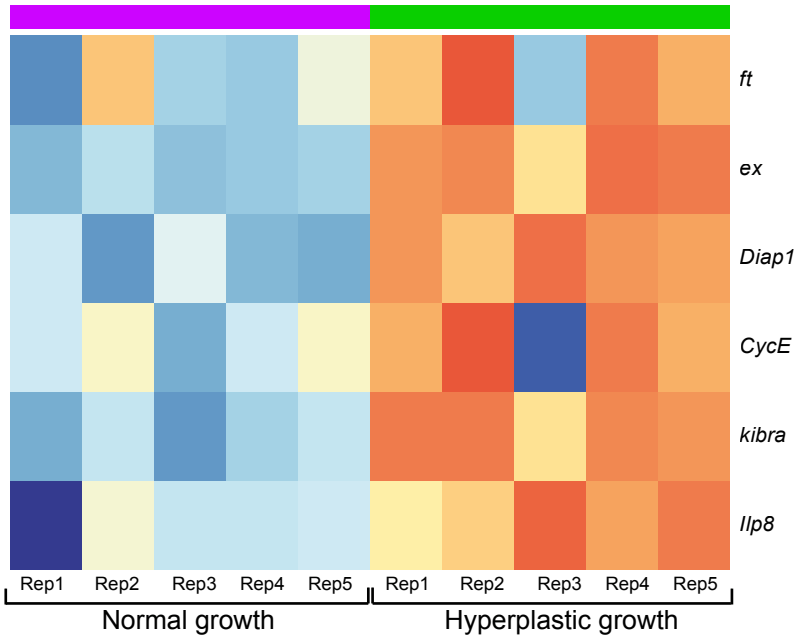
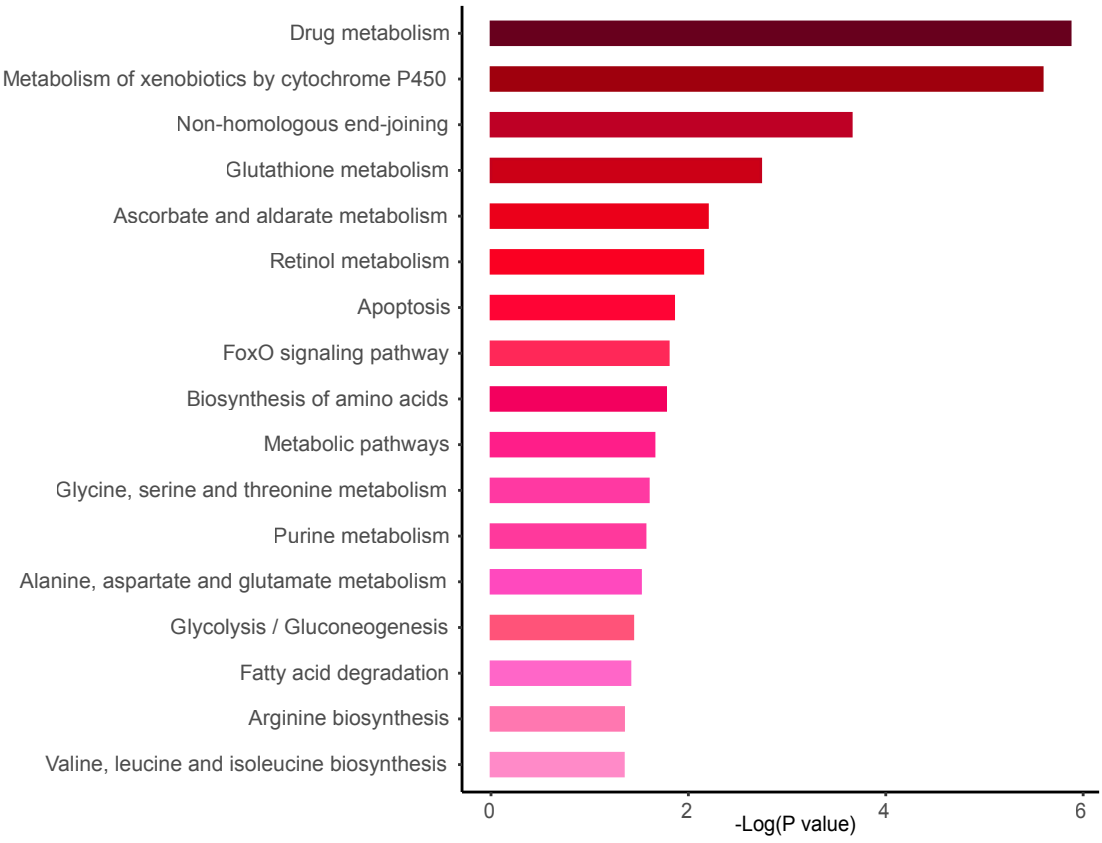

ft
ex
Diap1
CycE
kibra
Ilp8
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth

Figure 4. 4: Signalling pathways that are enriched among upregulated and downregulated genes in Yorkie driven hyperplastic eye discs.

Bar graphs showing enrichment of KEGG pathways performed on differentially expressed genes in hyperplastic (*wts^{X1}/wtsLacZ*) eye discs compared to normal (*FRT82B*) eye discs.

(A) Enriched KEGG pathways from upregulated genes. (B) Enriched KEGG pathways from downregulated genes. The y-axis shows the KEGG pathway and the x-axis shows the negative Log of the p value. The KEGG pathways listed are significantly enriched according to a p value of <0.05.

A


Drug metabolism
Metabolism of xenobiotics by cytochrome P450
Non-homologous end-joining
Glutathione metabolism
Ascorbate and aldarate metabolism
Retinol metabolism
Apoptosis
FoxO signaling pathway
Biosynthesis of amino acids
Metabolic pathways
Glycine, serine and threonine metabolism
Purine metabolism
Alanine, aspartate and glutamate metabolism
Glycolysis / Gluconeogenesis
Fatty acid degradation
Arginine biosynthesis
Valine, leucine and isoleucine biosynthesis
0
2
4
6
-Log(P value)

B

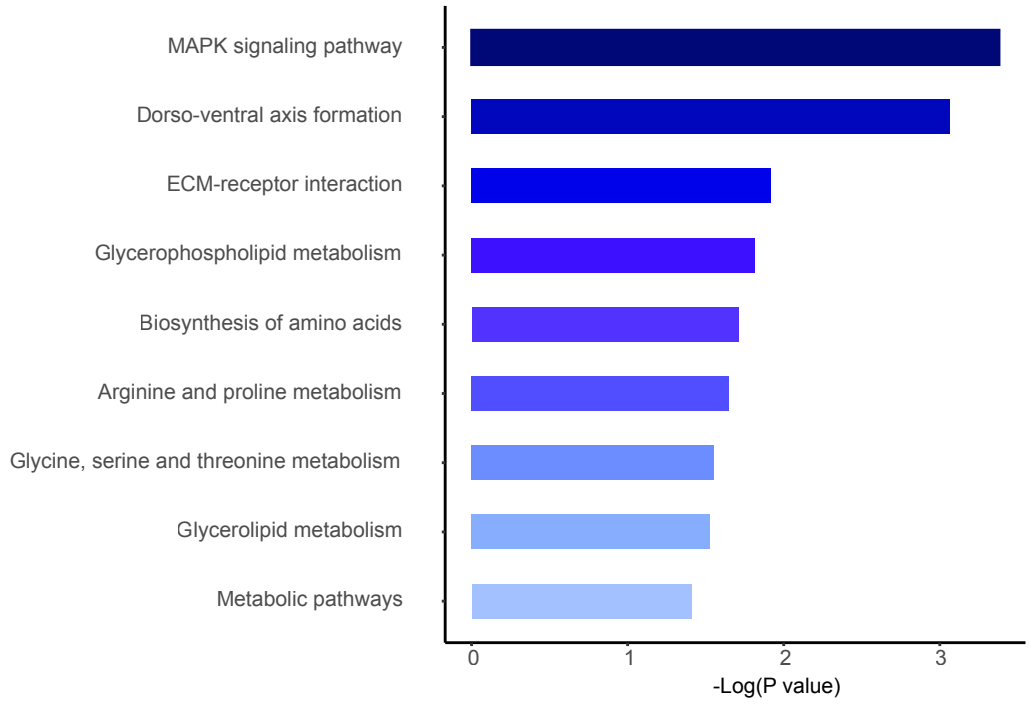
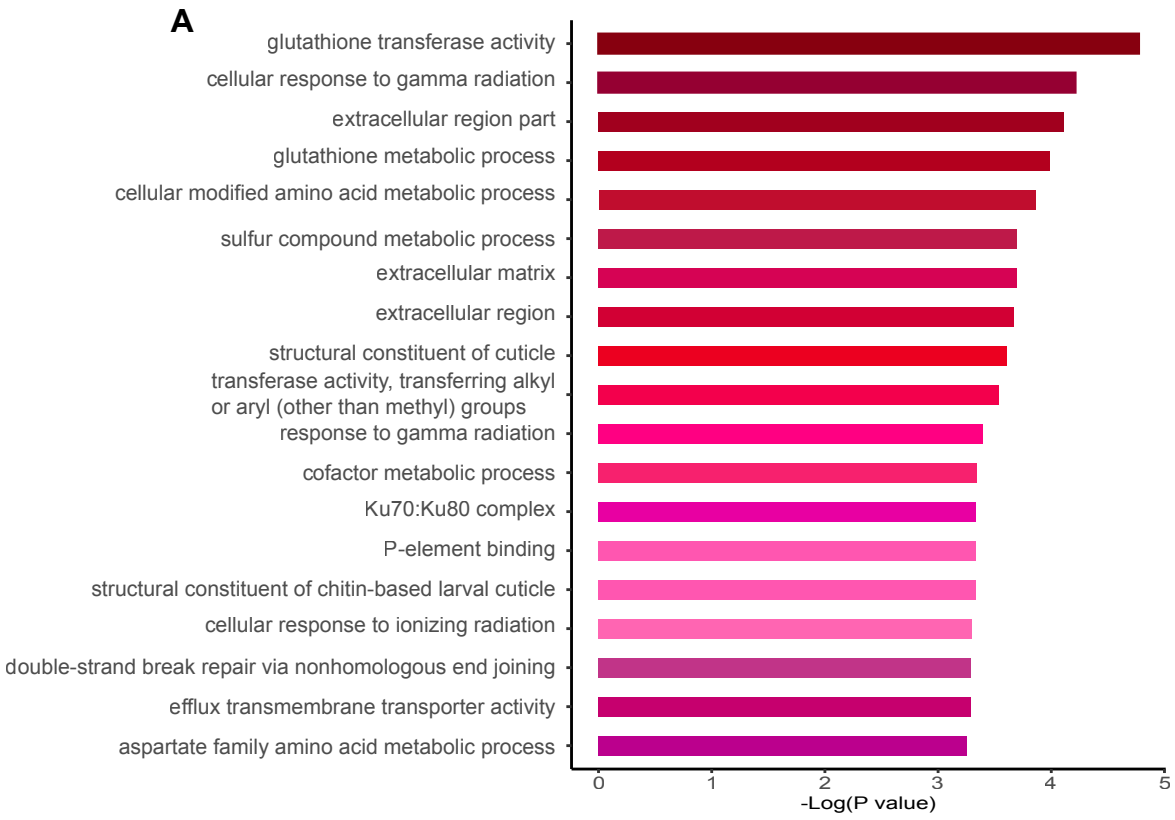

MAPK signaling pathway
Dorso-ventral axis formation
ECM-receptor interaction
Glycerophospholipid metabolism
Biosynthesis of amino acids
Arginine and proline metabolism
Glycine, serine and threonine metabolism
Glycerolipid metabolism
Metabolic pathways
0
1
2
3
-Log(P value)

Figure 4.5: Top Gene Ontology terms that are enriched among upregulated and downregulated genes in Yorkie driven hyperplastic eye discs

Bar graphs showing enrichment of Gene Ontology (G) terms performed on differentially expressed genes in hyperplastic (*wts^{X1}/wtsLacZ*) eye discs compared to normal (*FRT82B*) eye discs. (A) Enriched GO terms from upregulated genes. (B) Enriched GO terms from downregulated genes. The y-axis shows the GO term and the x-axis shows the negative Log of the p value. The top 19 enriched GO terms are shown, according to a p value of <0.001.


A
glutathione transferase activity
cellular response to gamma radiation
extracellular region part
glutathione metabolic process
cellular modified amino acid metabolic process
sulfur compound metabolic process
extracellular matrix
extracellular region
structural constituent of cuticle
transferase activity, transferring alkyl or aryl (other than methyl) groups
response to gamma radiation
cofactor metabolic process
Ku70:Ku80 complex
P-element binding
structural constituent of chitin-based larval cuticle
cellular response to ionizing radiation
double-strand break repair via nonhomologous end joining
efflux transmembrane transporter activity
aspartate family amino acid metabolic process
0
1
2
3
4
5
-Log(P value)

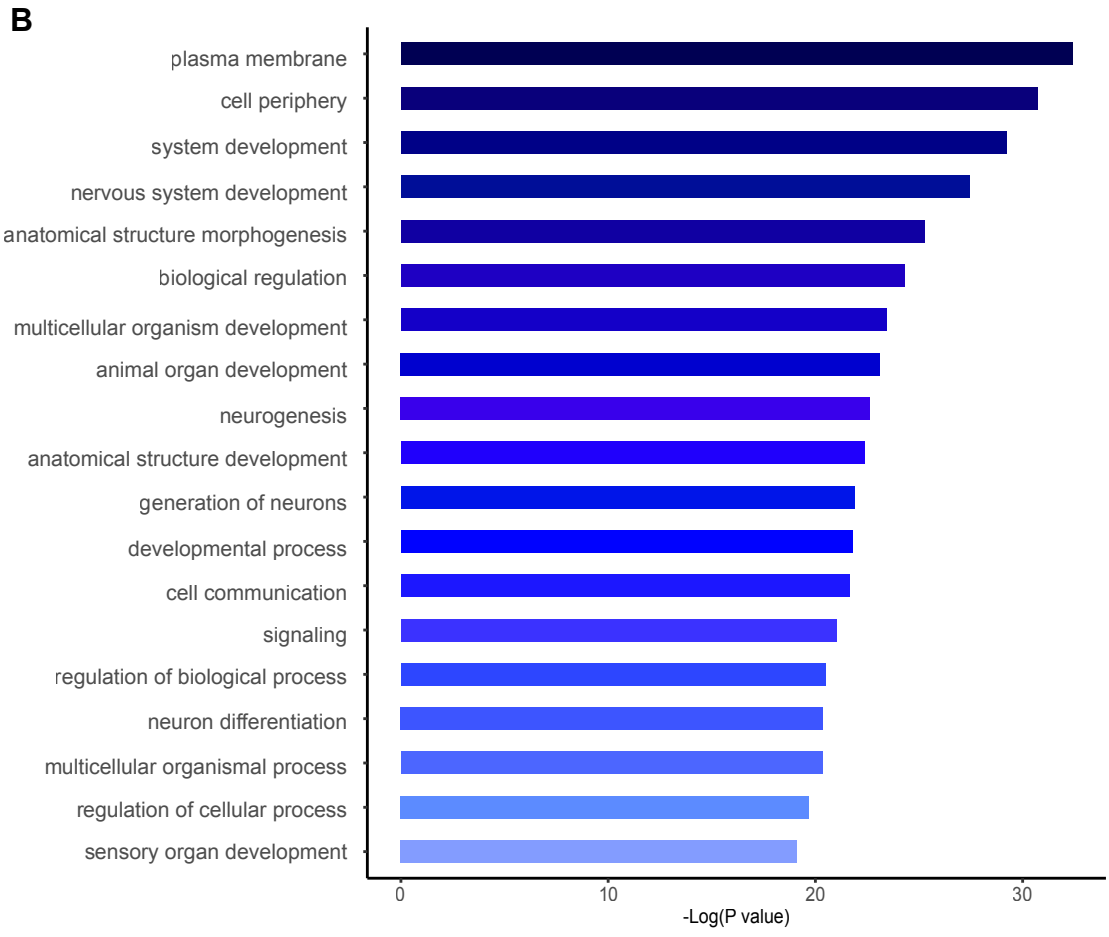

B
plasma membrane
cell periphery
system development
nervous system development
anatomical structure morphogenesis
biological regulation
multicellular organism development
animal organ development
neurogenesis
anatomical structure development
generation of neurons
developmental process
cell communication
signaling
regulation of biological process
neuron differentiation
multicellular organismal process
regulation of cellular process
sensory organ development
0
10
20
30
-Log(P value)

Figure 4.6: STRING analysis highlighting known and predicted interactions of the products of genes upregulated in Yorkie driven hyperplastic eye discs.

STRING analysis network of the protein products of the 417 genes upregulated in hyperplastic (*wts*^{X1}*/wtsLacZ*) eye discs compared to normal (*FRT82B*) eye discs. Analysis was performed using a medium confidence score of 0.4, with the following interactive sources selected: textmining, experiments, databases and co-expression. Clusters are highlighted for early eye determinant factors, Hippo signalling pathway, Transcription by RNA pol II, Gluthione S-transferases (GSTs), SPALT related proteins, and cuticle related proteins. The proteins within each cluster are highlighted, and grey lines indicate interactions between proteins.

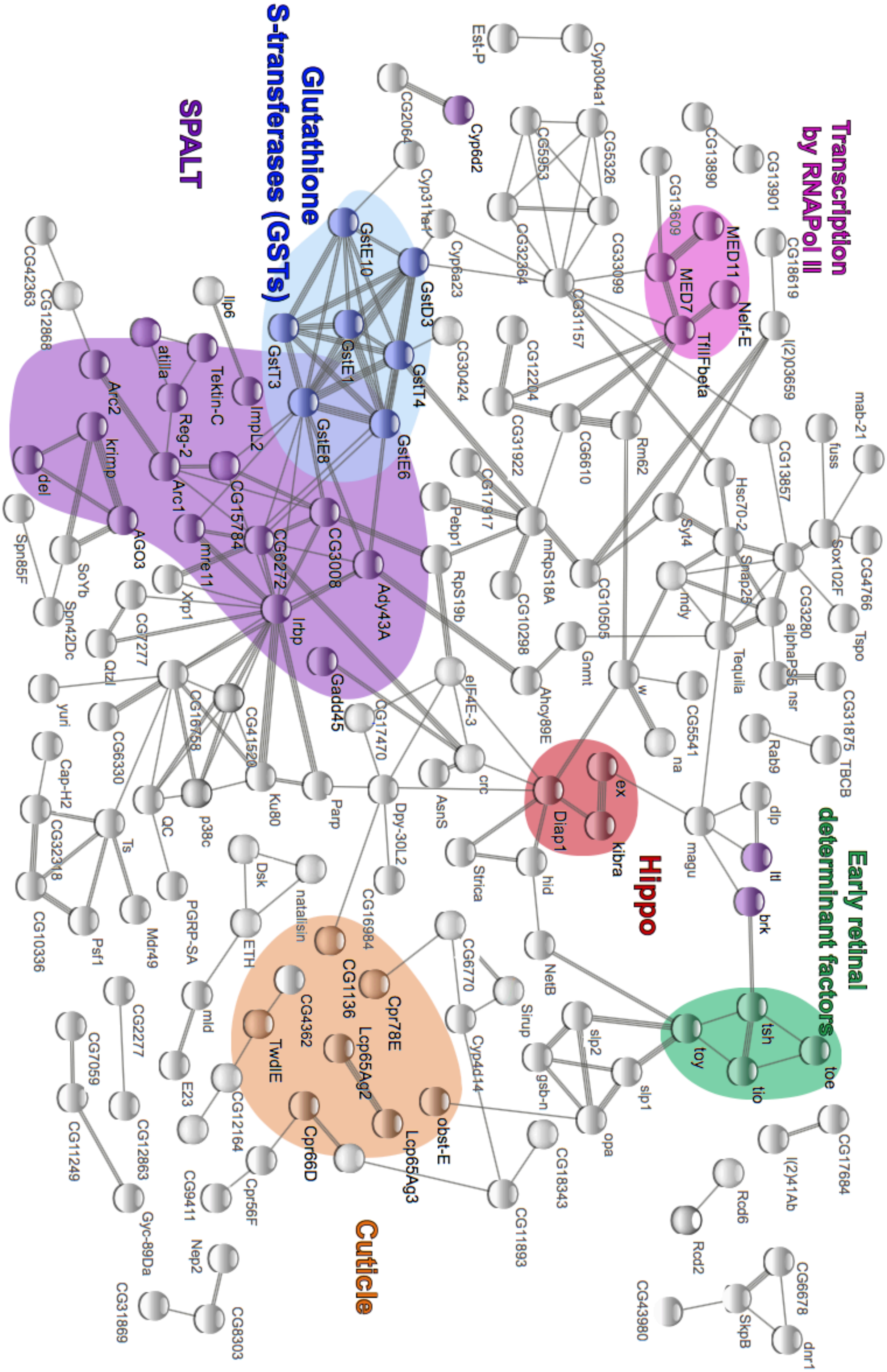

Transcription by RNAPol II
Glutathione S-transferases (GSTs)
SPALT
Early retinal determinant factors
Hippo
Cuticle

Figure 4. 7: STRING analysis highlighting known and predicted interactions of the products of genes downregulated in Yorkie driven hyperplastic eye discs.

STRING analysis network of the protein products of the 920 genes downregulated in hyperplastic (*wts*^{X1}*/wtsLacZ*) eye discs compared to normal (*FRT82B*) eye discs. Analysis was performed using a high confidence score of 0.9, with the following interactive sources selected: textmining, experiments, databases and co-expression. Clusters are highlighted for retinal determination gene and photoreceptor differentiation proteins, MAPK signalling pathway, and ECM related proteins. The proteins within each cluster are highlighted, and grey lines indicate interactions between proteins.

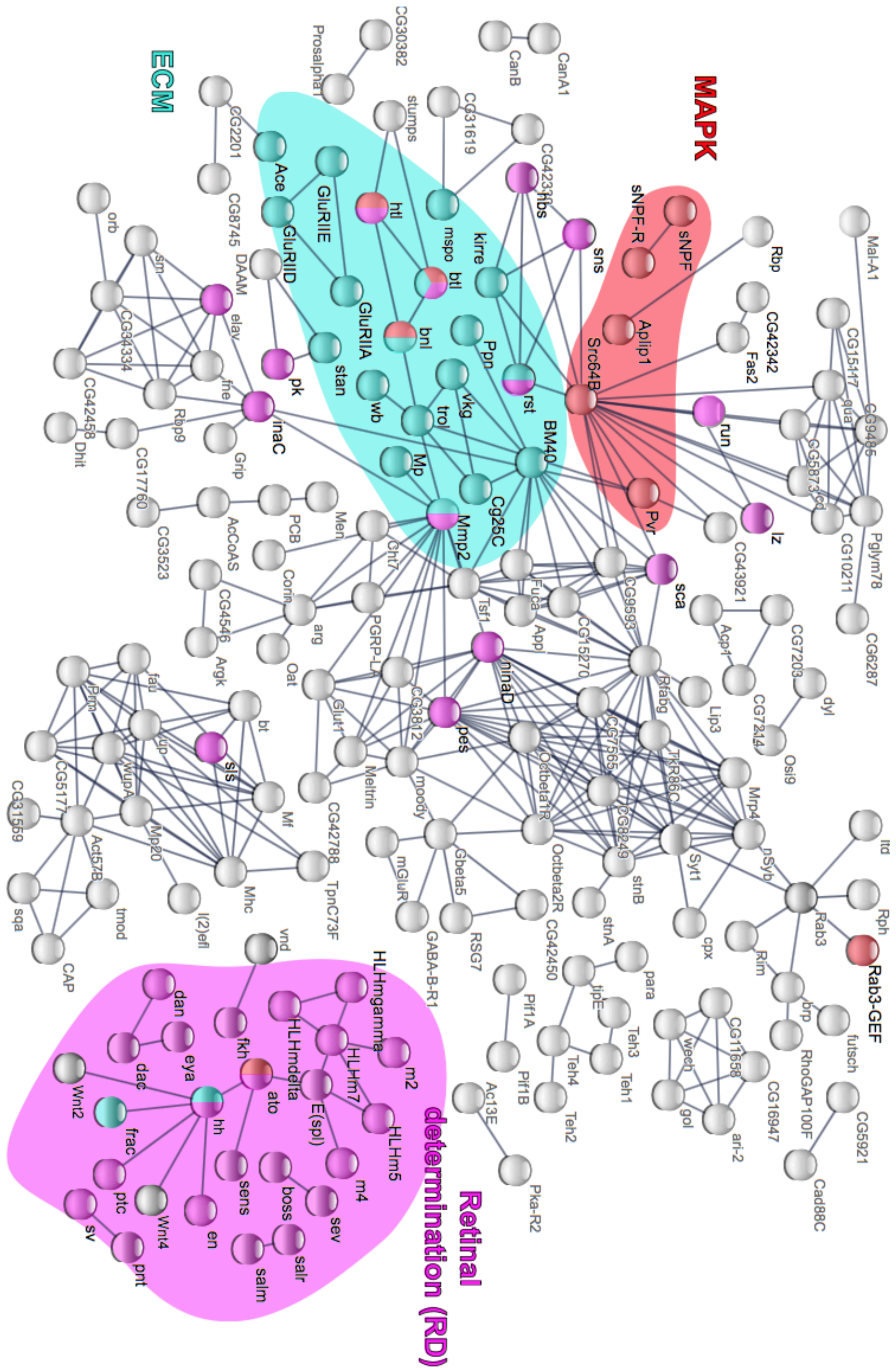

MAPK
ECM
Retinal determination (RD)
Mal-A1
Rbp
sNPF
sNPF-R
CG42342
Fas2
Aplip1
Src64B
CanA1
CanB
CG30382
Prosalpha1
CG31619
stumps
CG42316
hbs
sns
CG2201
CG8745
Ace
GluRIIE
GluRIID
htl
mspo
kirre
btl
GluRIIA
DAAM
bnl
Ppn
rst
stan
pk
wb
trol
vkg
BM40
run
CG15417
qua
CG9485
CG5873
cd
Pglym78
CG10211
lz
Pvr
CG43921
Mp
Cg25C
Mmp2
sca
orb
sn
CG34334
elav
fne
inaC
Grip
Rbp9
CG42458
Dhit
CG17760
CG3523
AcCoAS
PCB
Men
Cht7
Corin
CG4546
arg
Argk
Oat
PGRP-LA
Tsf1
Fuca
Appl
CG9593
CG15270
ninaD
Acp1
CG7203
dyl
CG6287
CG7214
Osi9
Lip3
Rfabg
CG3812
pes
Glut1
Meltrin
moody
CG7565
TkR86C
Mrp4
Octbeta1R
CG8249
Syt1
nSyb
stnB
stnA
Octbeta2R
CG42450
Gbeta5
RSG7
mGluR
GABA-B-R1
CG42788
TpnC73F
ltd
Rph
Rab3
Rab3-GEF
cpx
Rim
brp
futsch
RhoGAP100F
CG5921
Cad88C
CG11658
CG16947
ari-2
gol
wech
para
tipE
Teh3
Teh1
Teh4
Teh2
Pif1A
Pif1B
Ac13E
Pka-R2
fau
bt
Prm
up
sls
Mf
wupA
CG5177
CG31559
Mlp20
Act57B
Mhc
l(2)efl
tmod
sqa
CAP
vnd
HLHmgamma
dan
fkh
HLHmdelta
HLHm7
m2
eya
dac
ato
E(spl)
Wnt2
hh
HLHm5
frac
m4
sens
boss
sev
ptc
Wnt4
en
salr
sv
pnt
salm

Figure 4.8: Genes negatively regulated by the Spalt transcription factors are upregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of Spalt transcription factor regulated genes, the Spalt transcription factors (*salm*/*salr*), and the negative regulator of Spalt, *brinker*/*brk.* Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (wts^{X1}/*wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the *salm* (B), *salr* (C), and *brk* (D) genes. Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of *salm*, *salr*, and *brk* gene loci and surrounding genes. Scale bars in kb are shown.

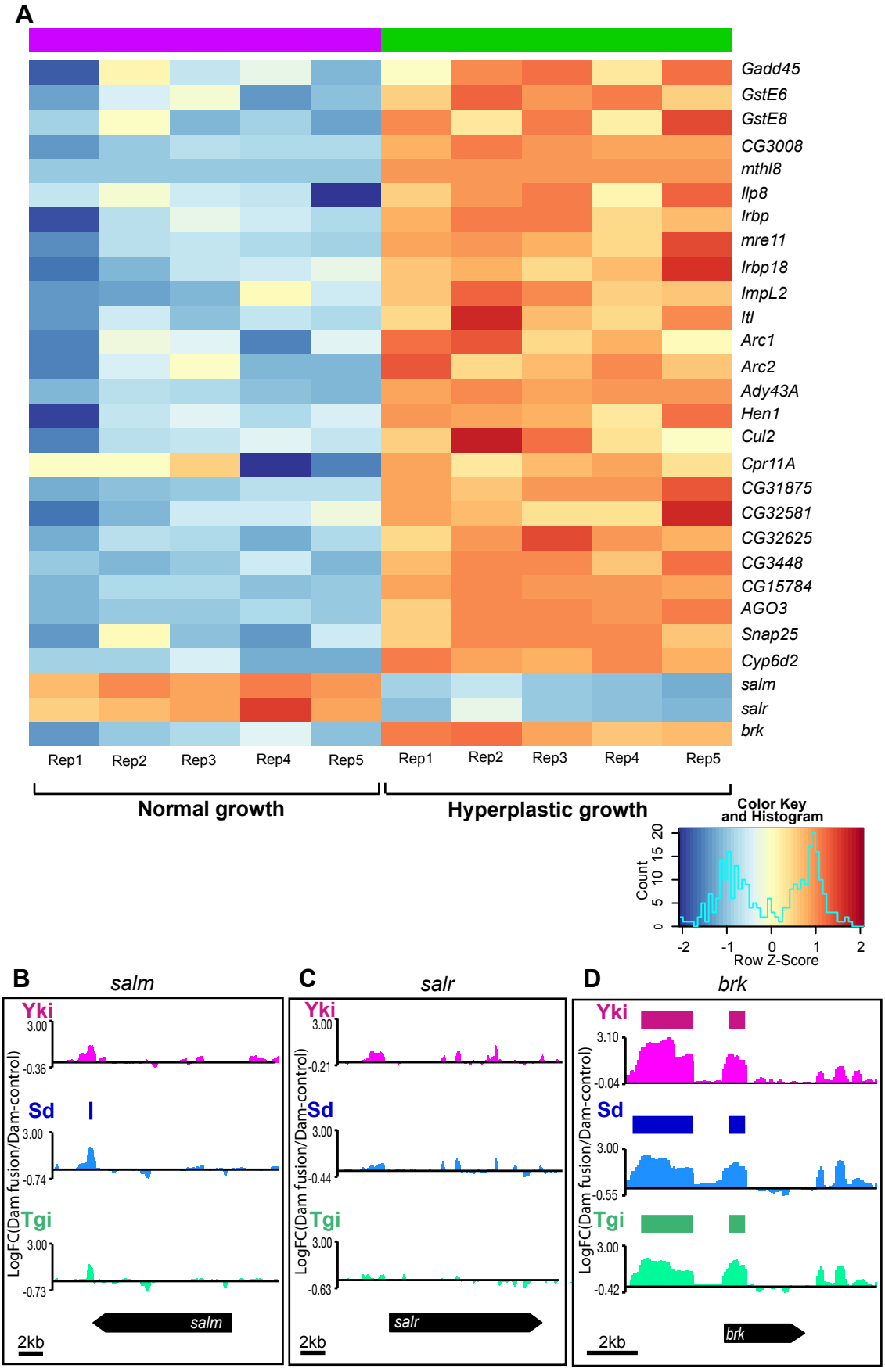

A
Gadd45
GstE6
GstE8
CG3008
mthl8
Ilp8
Irbp
mre11
Irbp18
ImpL2
Itl
Arc1
Arc2
Ady43A
Hen1
Cul2
Cpr11A
CG31875
CG32581
CG32625
CG3448
CG15784
AGO3
Snap25
Cyp6d2
salm
salr
brk
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
Row Z-Score
B
salm
C
salr
D
brk
Yki
Sd
Tgi
LogFC(Dam fusion/Dam-control)
2kb

Figure 4.9: Glutathione S-transferase (GST) genes are upregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of Glutathione S-transferase (GST) genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts^{X1}/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the *GstD* gene cluster (B), *GstE* gene cluster (C), and *GstT3 (D)* and *GstT4* (E) genes. Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

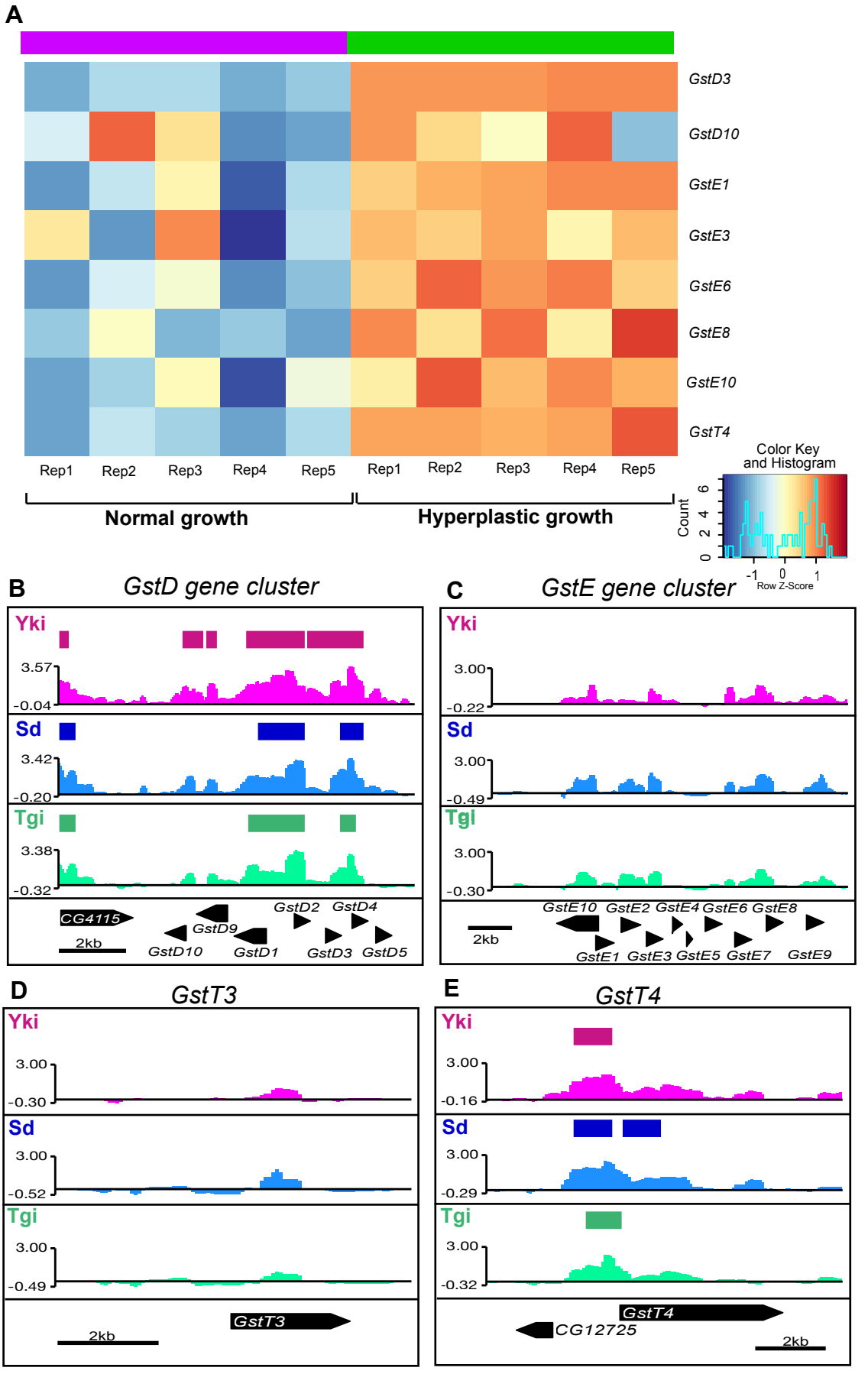

A
GstD3
GstD10
GstE1
GstE3
GstE6
GstE8
GstE10
GstT4
Rep1 Rep2 Rep3 Rep4 Rep5 Rep1 Rep2 Rep3 Rep4 Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
Row Z-Score
B
GstD gene cluster
Yki
Sd
Tgi
CG4115
GstD9
GstD2
GstD4
GstD10
GstD1
GstD3
GstD5
2kb
C
GstE gene cluster
GstE10
GstE2
GstE4
GstE6
GstE8
GstE1
GstE3
GstE5
GstE7
GstE9
D
GstT3
E
GstT4
CG12725

Figure 4.10: Cuticle related genes are upregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of cuticle development related genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts*^{X1}*/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the genes *TwdlE* (B), *CG1136* (C), and the Cuticle Protein gene loci (CPR) (D). Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

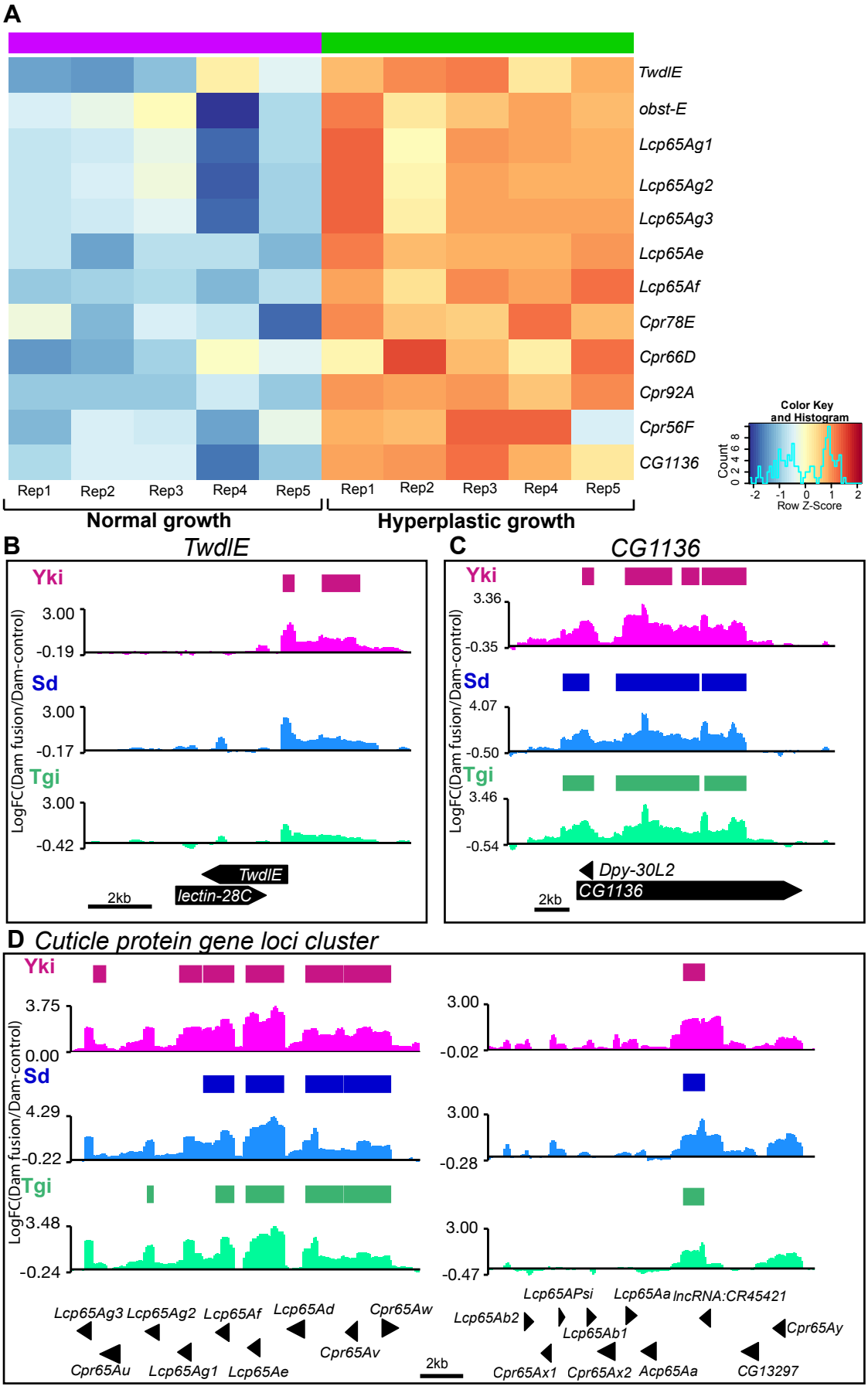

A
TwdlE
obst-E
Lcp65Ag1
Lcp65Ag2
Lcp65Ag3
Lcp65Ae
Lcp65Af
Cpr78E
Cpr66D
Cpr92A
Cpr56F
CG1136
Rep1 Rep2 Rep3 Rep4 Rep5 Rep1 Rep2 Rep3 Rep4 Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
Row Z-Score
B TwdlE
C CG1136
Yki
Sd
Tgi
LogFC(Dam fusion/Dam-control)
3.00
-0.19
3.00
-0.17
3.00
-0.42
TwdlE
lectin-28C
2kb
3.36
-0.35
4.07
-0.50
3.46
-0.54
Dpy-30L2
CG1136
D Cuticle protein gene loci cluster
3.75
0.00
4.29
-0.22
3.48
-0.24
3.00
-0.02
3.00
-0.28
3.00
-0.47
Lcp65Ag3 Lcp65Ag2 Lcp65Af Lcp65Ad Cpr65Aw
Cpr65Au Lcp65Ag1 Lcp65Ae Cpr65Av
Lcp65APsi Lcp65Aa lncRNA:CR45421
Lcp65Ab2 Lcp65Ab1 Cpr65Ay
Cpr65Ax1 Cpr65Ax2 Acp65Aa CG13297
2kb

Figure 4.11: Early retinal determination (RD) genes are upregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of early retinal determination genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts*^{X1}*/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the genes *ey* (B), *toy* (C), *hth* (D), and *wg* (E). Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

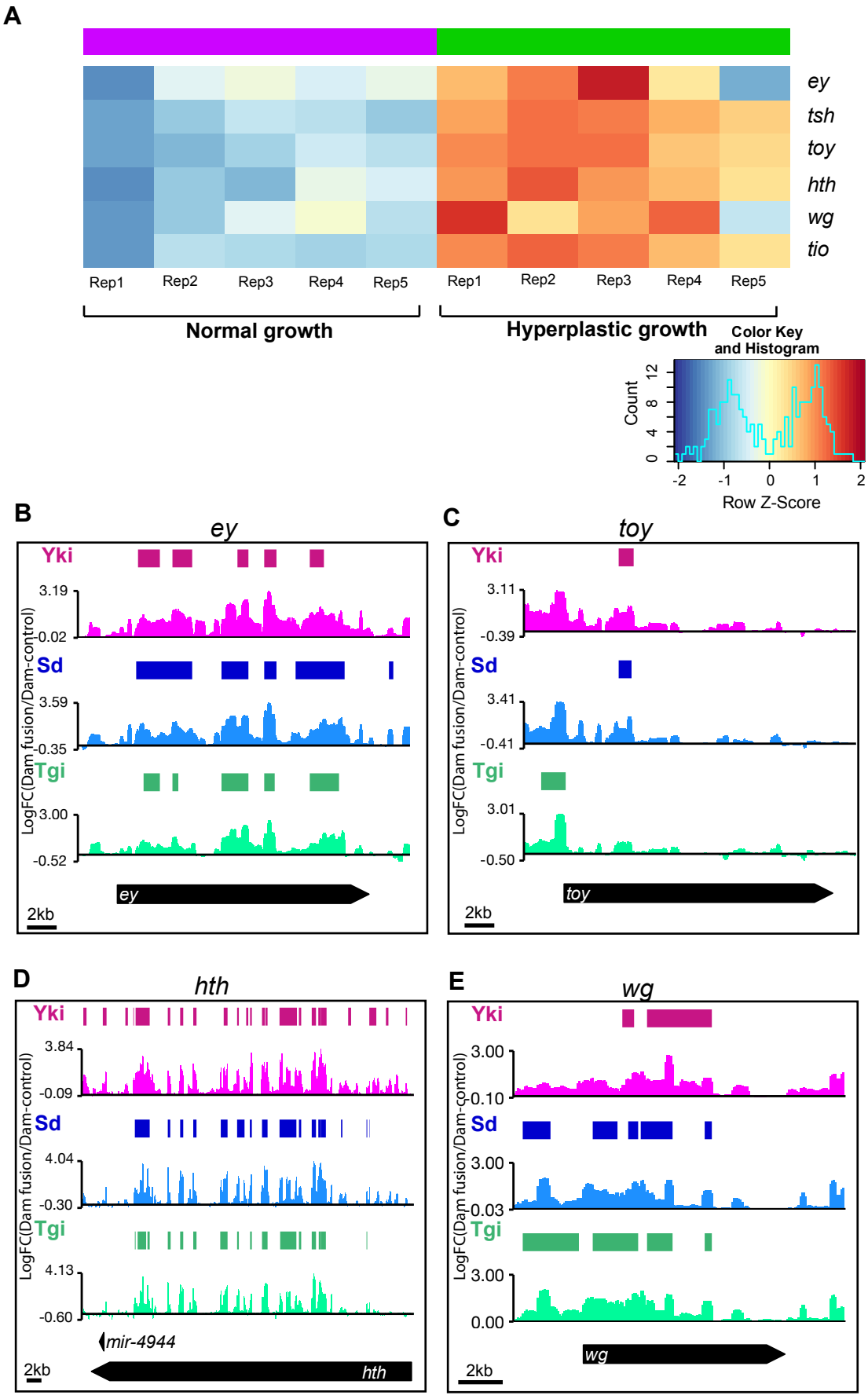
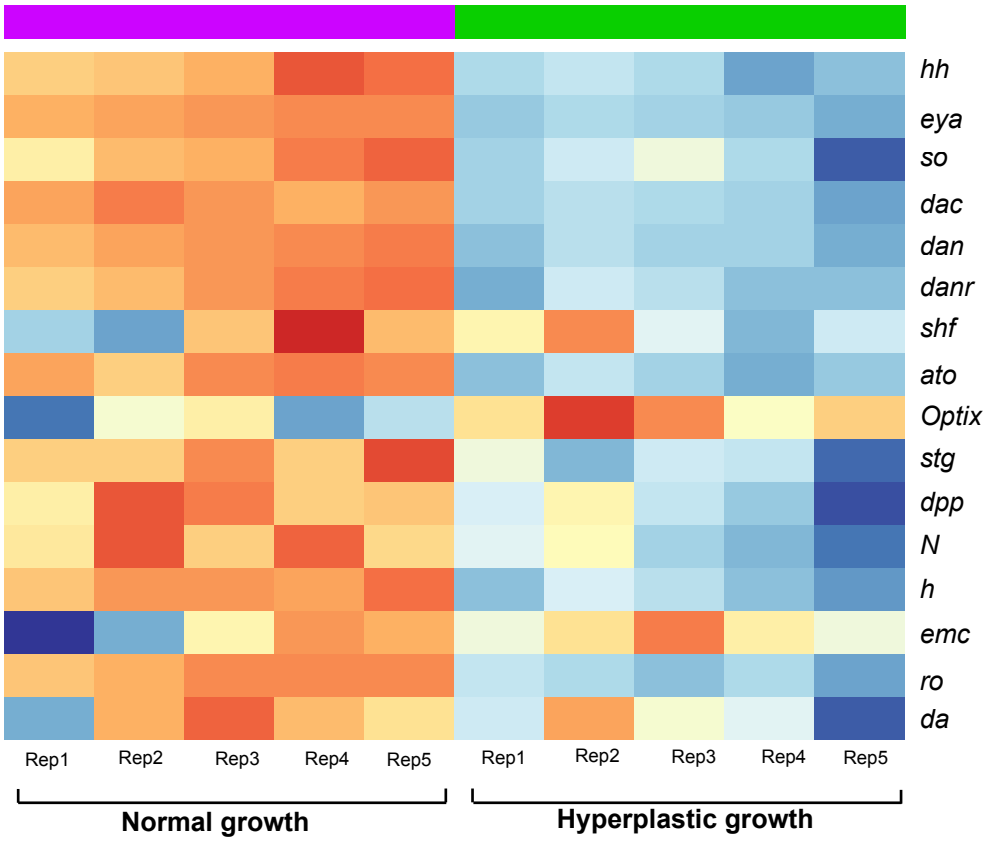

A
ey
tsh
toy
hth
wg
tio
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
0
4
8
12
-2
-1
0
1
2
Row Z-Score
B
ey
Yki
3.19
0.02
Sd
3.59
-0.35
Tgi
3.00
-0.52
LogFC(Dam fusion/Dam-control)
ey
2kb
C
toy
Yki
3.11
-0.39
Sd
3.41
-0.41
Tgi
3.01
-0.50
LogFC(Dam fusion/Dam-control)
toy
2kb
D
hth
Yki
3.84
-0.09
Sd
4.04
-0.30
Tgi
4.13
-0.60
LogFC(Dam fusion/Dam-control)
mir-4944
2kb
hth
E
wg
Yki
3.00
0.10
Sd
3.00
0.03
Tgi
3.00
0.00
LogFC(Dam fusion/Dam-control)
wg
2kb

Figure 4.12: Key retinal determination (RD) genes are downregulated in Yorkie driven hyperplastic eye discs.

Heatmap showing the expression levels of key retinal determination genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts^{X1}/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).


hh
eya
so
dac
dan
danr
shf
ato
Optix
stg
dpp
N
h
emc
ro
da
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth

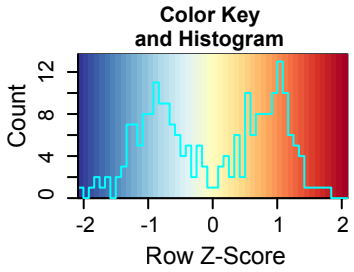

Color Key
and Histogram
Count
0
4
8
12
-2
-1
0
1
2
Row Z-Score

Figure 4.13: Key retinal determination (RD) genes are directly bound by Yorkie, Scalloped, and Tgi.

A-E. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the genes *hh* (A), *dpp* (B), *eya* (C), *so* (D), and *dac* (E). Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

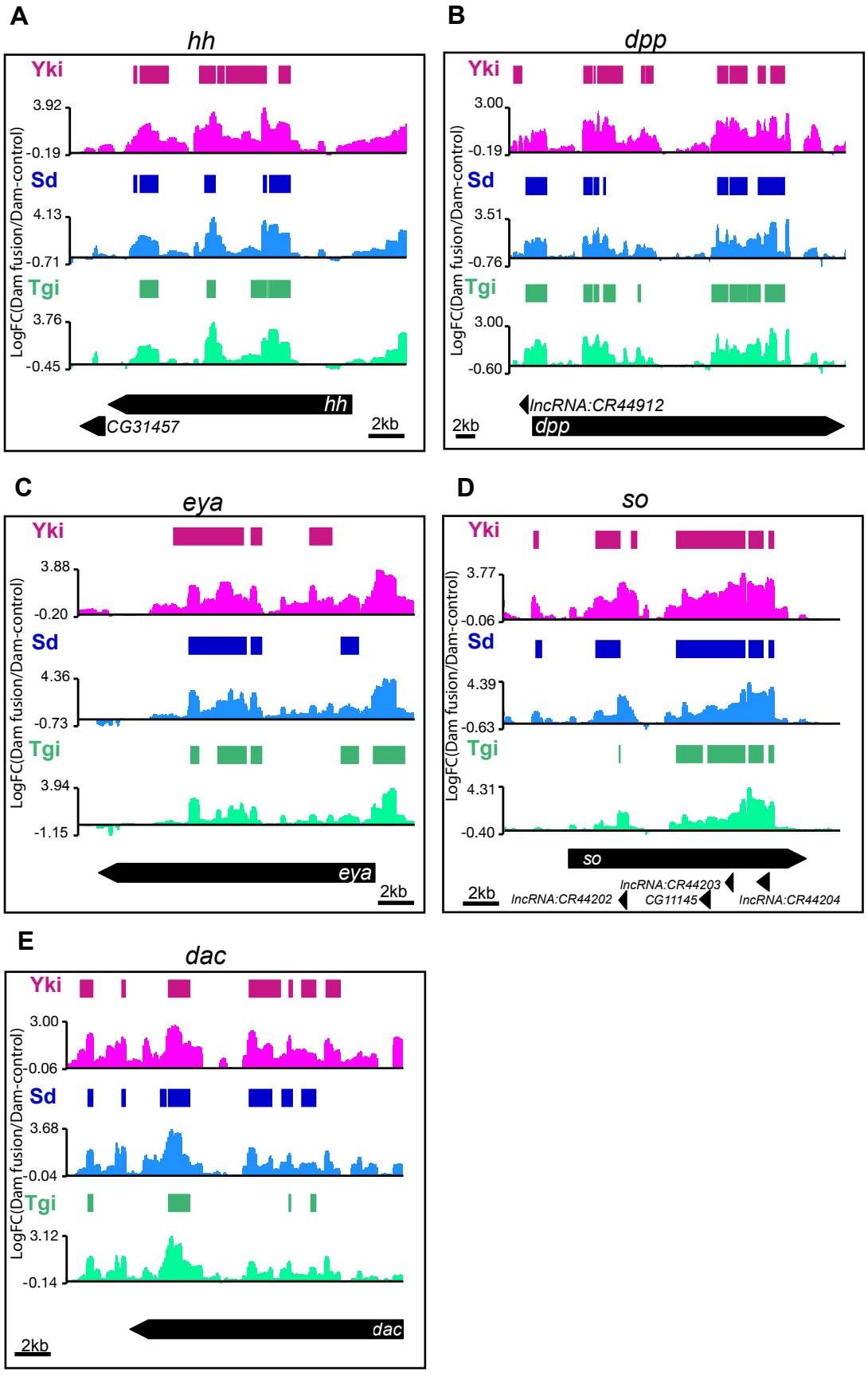

A
hh
Yki
3.92
-0.19
Sd
4.13
-0.71
Tgi
3.76
-0.45
LogFC(Dam fusion/Dam-control)
hh
CG31457
2kb
B
dpp
Yki
3.00
-0.19
Sd
3.51
-0.76
Tgi
3.00
-0.60
LogFC(Dam fusion/Dam-control)
lncRNA:CR44912
dpp
2kb
C
eya
Yki
3.88
-0.20
Sd
4.36
-0.73
Tgi
3.94
-1.15
LogFC(Dam fusion/Dam-control)
eya
2kb
D
so
Yki
3.77
-0.06
Sd
4.39
-0.63
Tgi
4.31
-0.40
LogFC(Dam fusion/Dam-control)
so
2kb
lncRNA:CR44202
lncRNA:CR44203
CG11145
lncRNA:CR44204
E
dac
Yki
3.00
0.06
Sd
3.68
-0.04
Tgi
3.12
-0.14
LogFC(Dam fusion/Dam-control)
dac
2kb

Figure 4.14: Photoreceptor differentiation genes are downregulated in Yorkie driven hyperplastic eye discs.

Heatmap showing the expression levels of photoreceptor differentiation genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts*^{X1}*/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

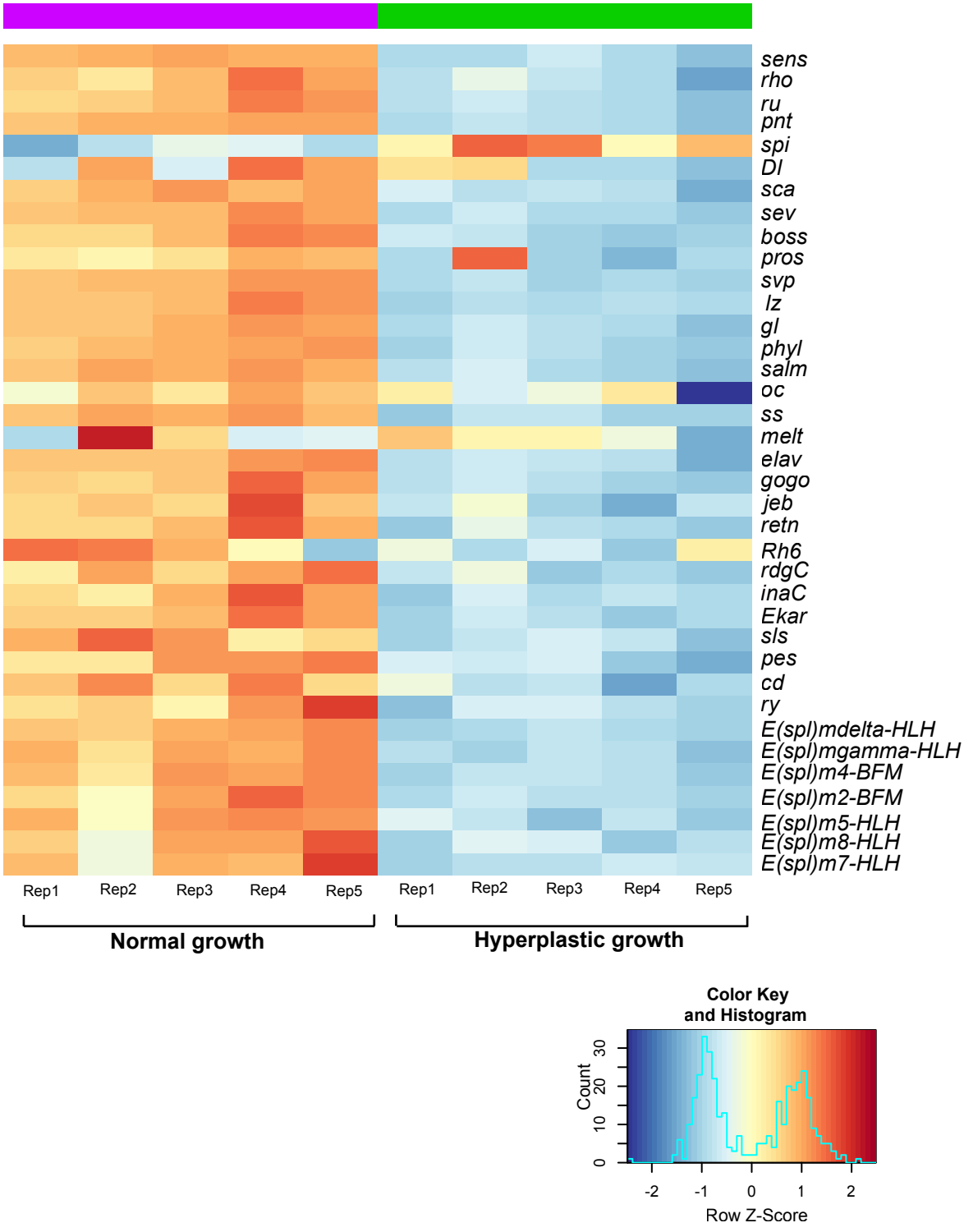

sens
rho
ru
pnt
spi
Dl
sca
sev
boss
pros
svp
lz
gl
phyl
salm
oc
ss
melt
elav
gogo
jeb
retn
Rh6
rdgC
inaC
Ekar
sls
pes
cd
ry
E(spl)mdelta-HLH
E(spl)mgamma-HLH
E(spl)m4-BFM
E(spl)m2-BFM
E(spl)m5-HLH
E(spl)m8-HLH
E(spl)m7-HLH
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth
Color Key
and Histogram
Count
0
10
20
30
-2
-1
0
1
2
Row Z-Score

Figure 4.15: ECM genes are downregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of extracellular matrix (ECM) related genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts^{X1}/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the genes *wb* (B) and *Mmp2* (C). Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

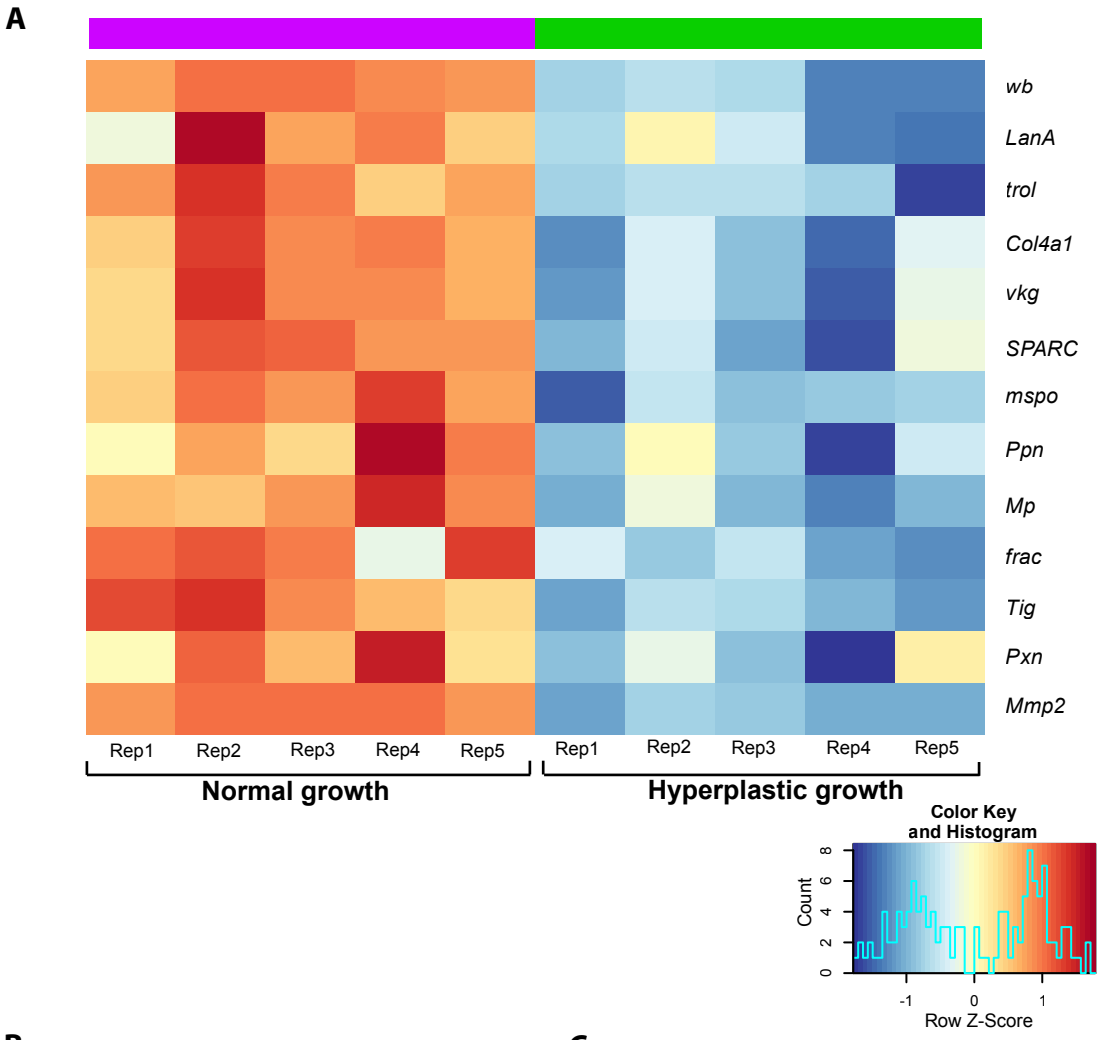

A
wb
LanA
trol
Col4a1
vkg
SPARC
mspo
Ppn
Mp
frac
Tig
Pxn
Mmp2
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
0
2
4
6
8
-1
0
1
Row Z-Score
B
wb
Yki
3.00
-0.53
Sd
3.48
-1.24
Tgi
3.00
-1.16
LogFC(Dam fusion/Dam-control)
wb
lncRNA:CR44729
lncRNA:CR43640
CG32971
lncRNA:CR43639
2kb
C
Mmp2
Yki
3.00
-0.73
Sd
3.22
-0.44
Tgi
3.01
-0.42
LogFC(Dam fusion/Dam-control)
mir-307a
lncRNA:CR44070
Mmp2
2kb
tRNA:iMet-CAT-2-1

Figure 4.16: MAPK genes are downregulated in Yorkie driven hyperplastic eye discs.

A. Heatmap showing the expression levels of MAPK related genes. Replicates from normal growth are shown in purple and replicates from hyperplastic growth are shown in green. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic (*wts^{X1}/wtsLacZ*) compared to normal (*FRT82B*) growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

B-D. Targeted DamID gene binding profiles showing adenine methylation by Yki- (pink), Sd- (blue), and Tgi- (green) Dam-fusions at the genes *pnt* (B), *RasGap1* (C), *aos* (D), and *kek1* (E). Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the Dam-fusion normalised to the Dam-alone control. The gene binding profiles were binned at 75 base pair intervals. The bars in the top of each panel represent peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

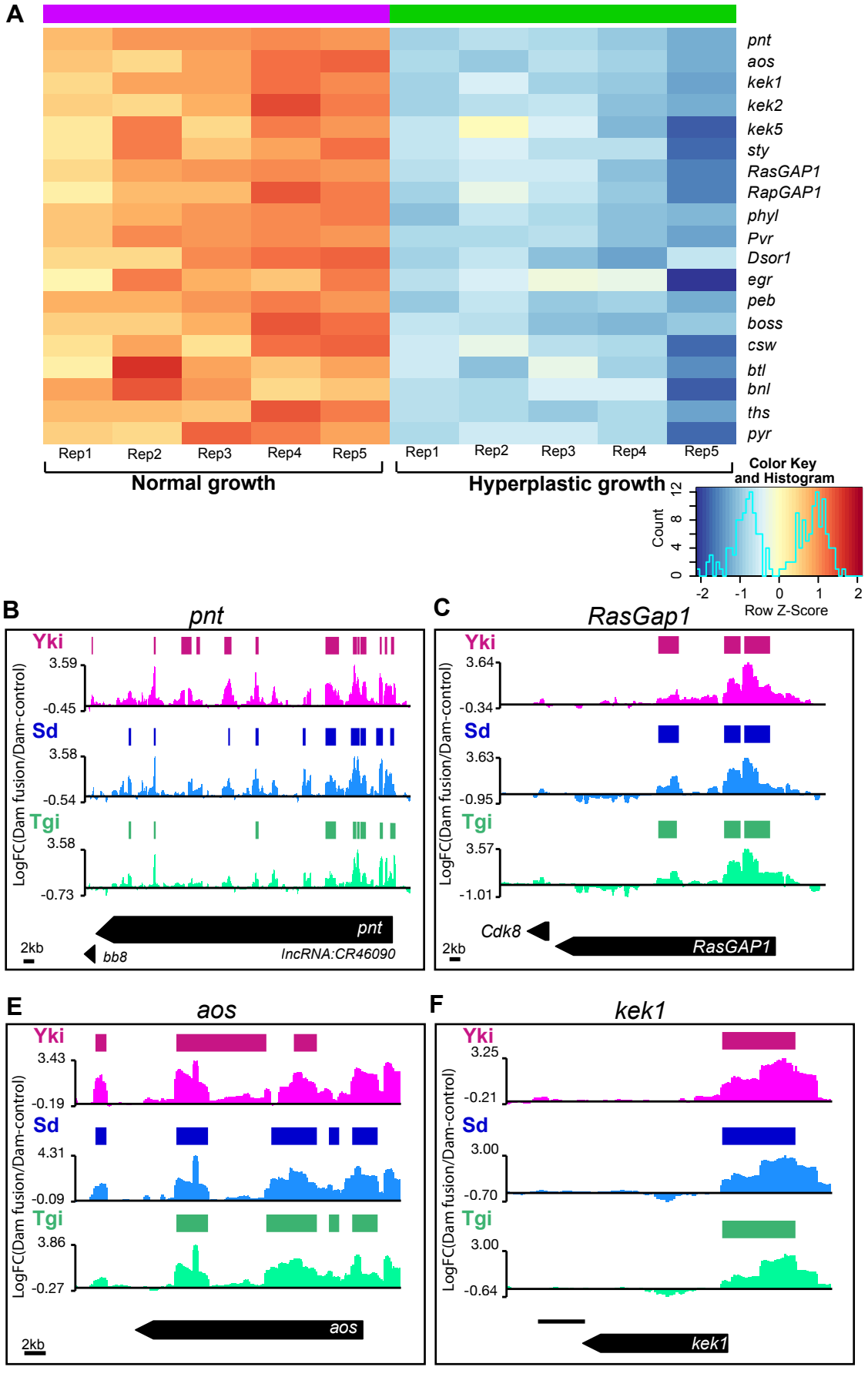
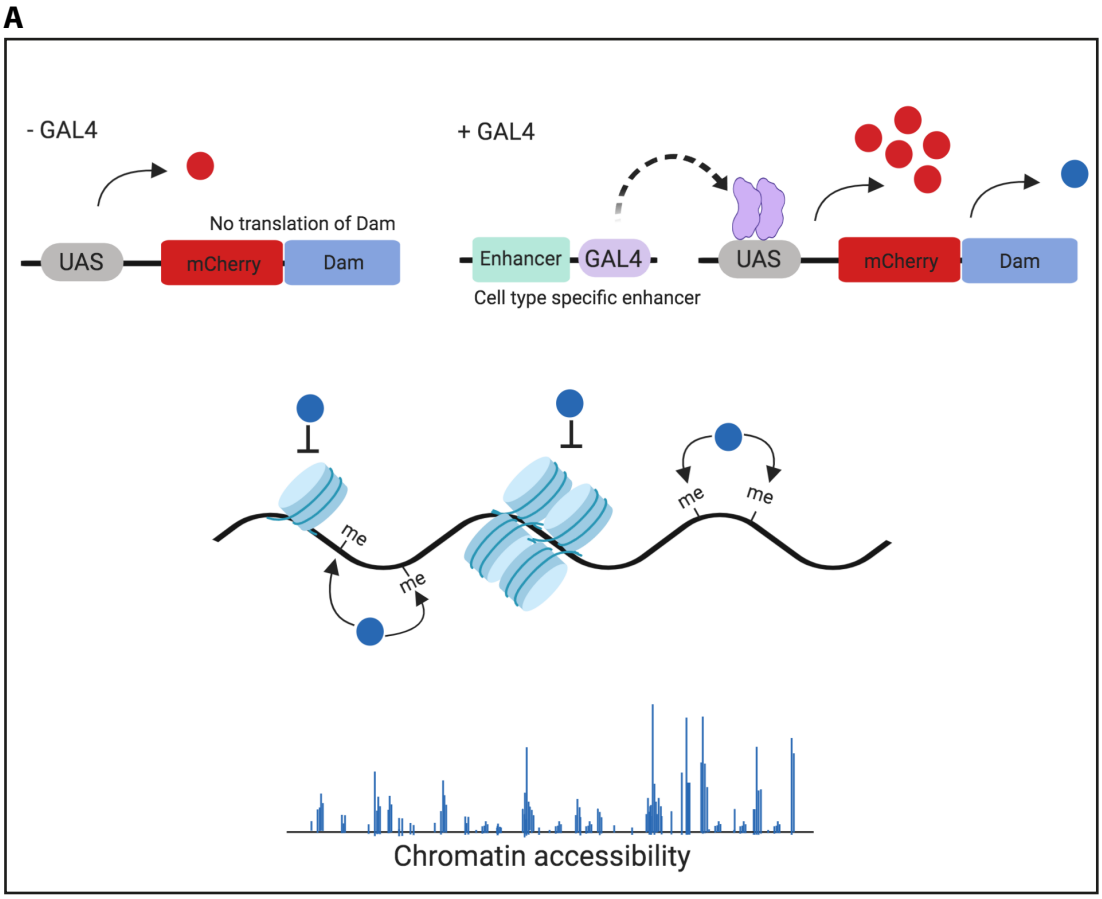

A
pnt
aos
kek1
kek2
kek5
sty
RasGAP1
RapGAP1
phyl
Pvr
Dsor1
egr
peb
boss
csw
btl
bnl
ths
pyr
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth
Color Key and Histogram
Count
Row Z-Score
B
pnt
C
RasGap1
Yki
Sd
Tgi
LogFC(Dam fusion/Dam-control)
2kb
bb8
lncRNA:CR46090
Cdk8
RasGAP1
E
aos
F
kek1

Figure 4.17: Chromatin Accessibility Targeted DamID (CATaDa) reveals chromatin accessibility in normal and hyperplastic eye discs.

A. Schematic representation of Chromatin Accessibility Targeted DamID (CATaDa). Untethered DNA adenine methylase (Dam) is capable of specifically methylating regions of open chromatin while not accessing regions of closed chromatin. Dam was expressed in specific cells using the UAS-Gal4 system. Gal4 is a yeast derived activator protein that binds to an upstream activator sequence (UAS) to activate gene transcription. Untethered Dam was expressed in eye tissue using an eye-specific Gal4 that is under the control of the *eyeless* enhancer. Expression of the Dam protein *in vivo* leads to methylation of adenines in GATC sequences near sites where the protein interacts with the genome. Following Dam protein expression, methylated DNA was isolated, amplified by PCR and sequenced, revealing Dam protein binding in the genome and hence regions of open chromatin.

B. Multidimensional scaling (MDS) plot of normal (*FRT82B*) and hyperplastic (*wts*^{X1}*/wtsLacZ*) growth CATaDa samples. Normal growth samples are shown in pink, and hyperplastic growth samples are shown in blue. The x and y axes show the leading LogFC dimension 1 and 2 respectively. Distances on the plot approximate differences between samples based on log2 fold changes.


A
- GAL4
No translation of Dam
UAS
mCherry
Dam
+ GAL4
Enhancer
GAL4
UAS
mCherry
Dam
Cell type specific enhancer
me
me
me
me
Chromatin accessibility

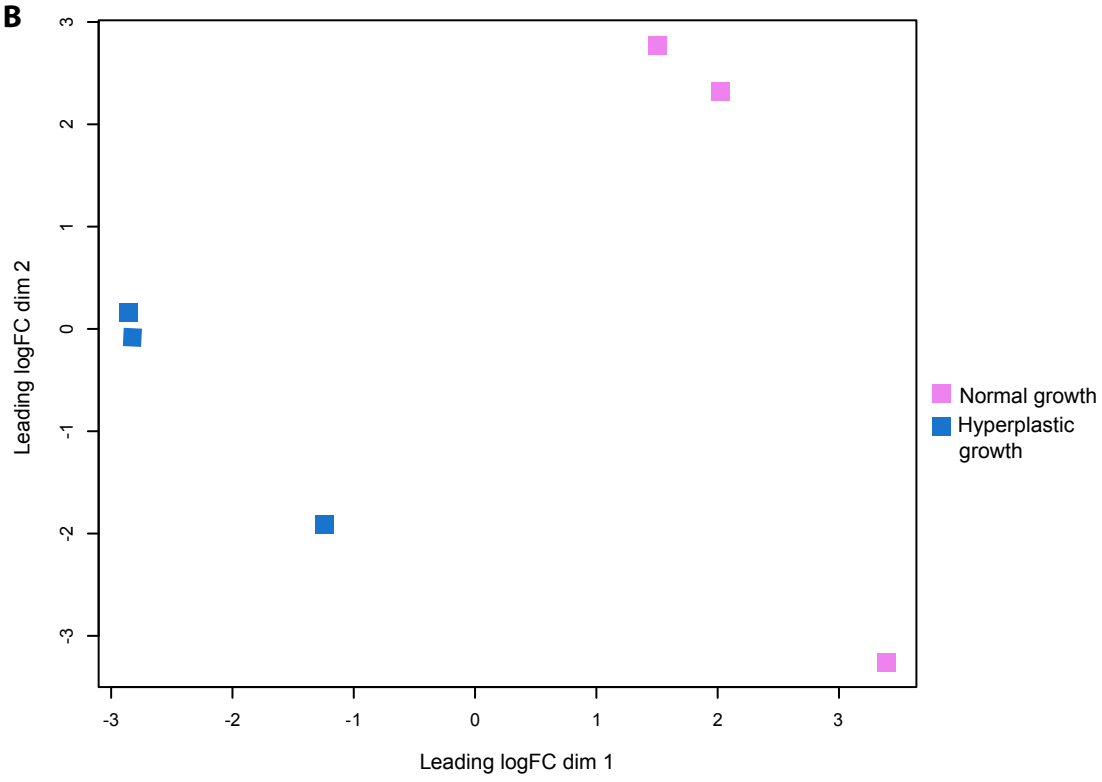

B
Leading logFC dim 2
Leading logFC dim 1
-3
-2
-1
0
1
2
3
Normal growth
Hyperplastic growth

Figure 4.18: Chromatin Accessibility Targeted DamID (CATaDa) reveals that chromatin accessibility is altered in Yorkie driven hyperplastic eye discs.

A. CATaDa peak profile distribution across the *Drosophila* genome. Binding peaks that were differentially bound in hyperplastic eye discs are shown on the top panel (red), and peaks that were differentially bound in normal growth are shown in the bottom panel (blue).

B-C. Venn diagrams comparing CATaDa results and genes either upregulated (B) or downregulated (C) in hyperplastic (*wts*^{X1}*/wtsLacZ*) compared to normal (*FRT82B*) growth. CATaDa peaks were classified as 'opening' if untethered Dam showed increased methylation in hyperplastic compared to normal growth, while CATaDa peaks were classified as 'closing' if untethered Dam showed decreased methylation in hyperplastic compared to normal growth. The significance (p value) of overlapping target genes was assessed by hypergeometric probability analysis.

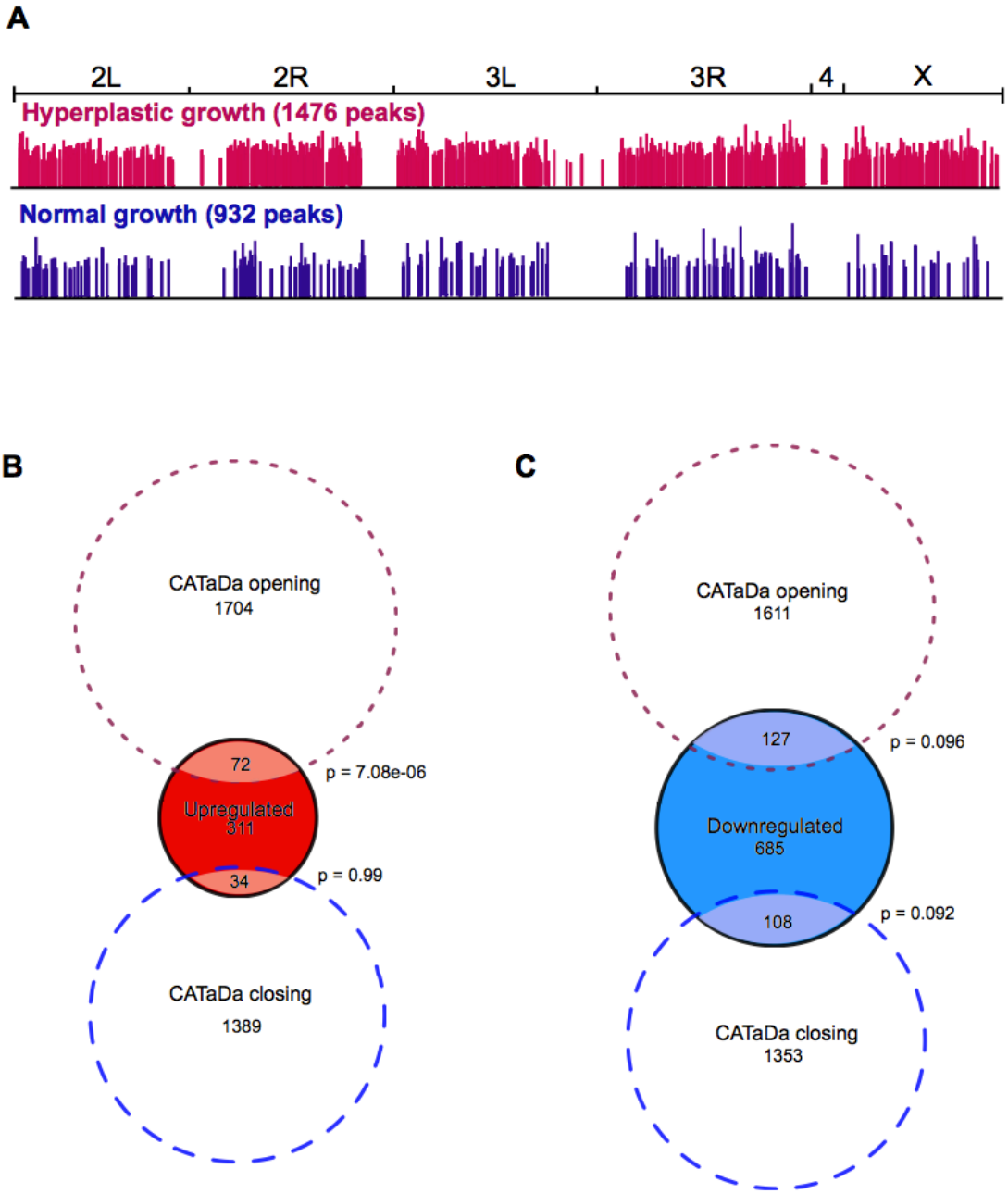

A
2L
2R
3L
3R
4
X
Hyperplastic growth (1476 peaks)
Normal growth (932 peaks)
B
CATaDa opening
1704
72
Upregulated
311
34
p = 7.08e-06
p = 0.99
CATaDa closing
1389
C
CATaDa opening
1611
127
Downregulated
685
108
p = 0.096
p = 0.092
CATaDa closing
1353

Figure 4.19: Chromatin accessibility at known Hippo pathway target genes in Yorkie driven hyperplastic eye discs.

A-D. CATaDa gene binding profiles showing adenine methylation by untethered Dam at the genes *Diap1* (A), *ex* (B), *kibra* (C), and *ban* (D). Binding profiles for untethered Dam represent differential methylation and show the average log (Dam ratio) for all replicates between hyperplastic (*wts*^{X1}*/wtsLacZ*) and normal (*FT82B*) growth. The gene binding profiles were binned at 75 base pair intervals. The bars above and below of each panel represent peaks, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

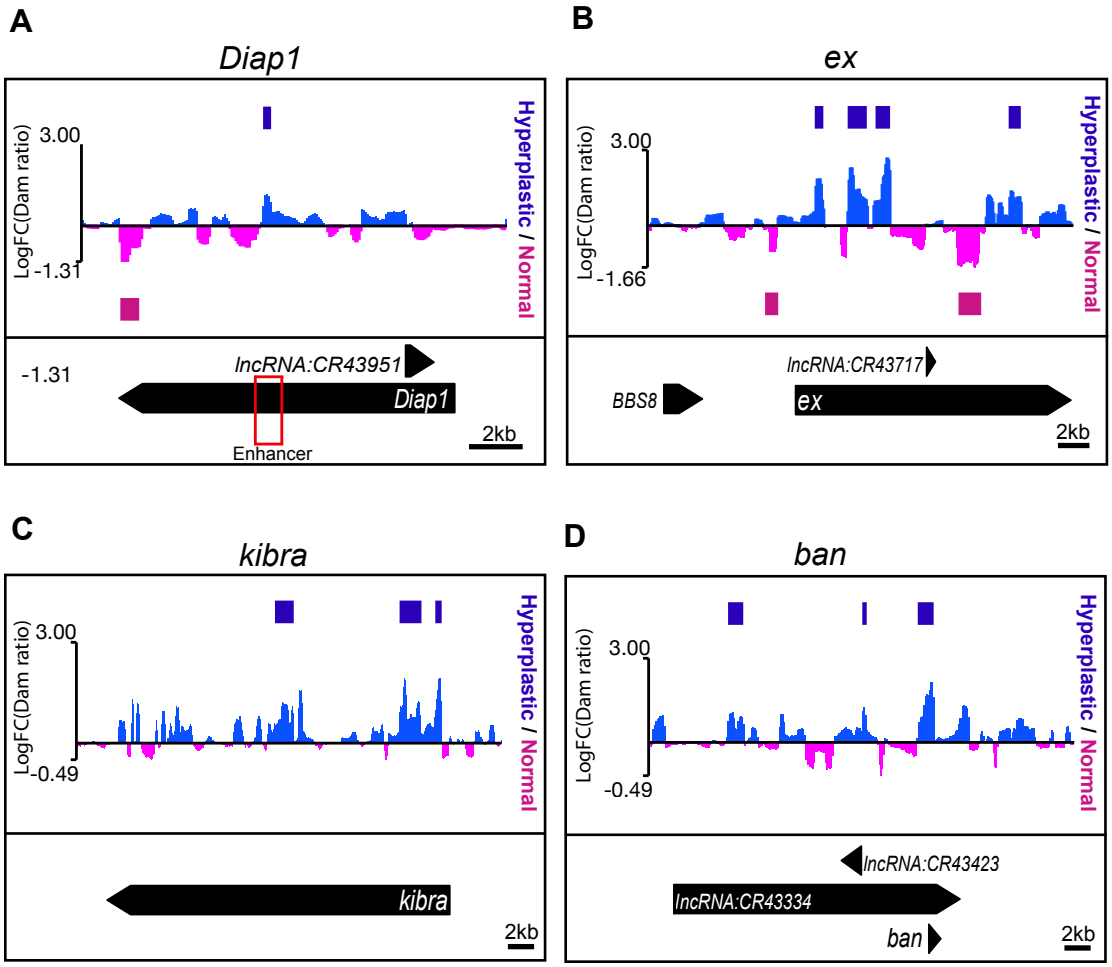

A
Diap1
LogFC(Dam ratio)
3.00
-1.31
Hyperplastic / Normal
-1.31
lncRNA:CR43951
Diap1
Enhancer
2kb
B
ex
LogFC(Dam ratio)
3.00
-1.66
Hyperplastic / Normal
lncRNA:CR43717
BBS8
ex
2kb
C
kibra
LogFC(Dam ratio)
3.00
-0.49
Hyperplastic / Normal
kibra
2kb
D
ban
LogFC(Dam ratio)
3.00
-0.49
Hyperplastic / Normal
lncRNA:CR43423
lncRNA:CR43334
ban
2kb

Figure 4.20: Chromatin accessibility is increased within Glutathione S-transferase (GST) gene loci in Yorkie driven hyperplastic eye discs.

A-D. CATaDa gene binding profiles showing adenine methylation by untethered Dam at the *GSTD* gene cluster (A), *GSTE* gene cluster (B), *brk* (C), *ey* (D), and *hth* (E) gene loci. Binding profiles for untethered Dam represent differential methylation and show the average log (Dam ratio) for all replicates between hyperplastic (wts^{X1}/*wtsLacZ*) and normal (*FT82B*) growth. The gene binding profiles were binned at 75 base pair intervals. The bars above and below of each panel represent peaks, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

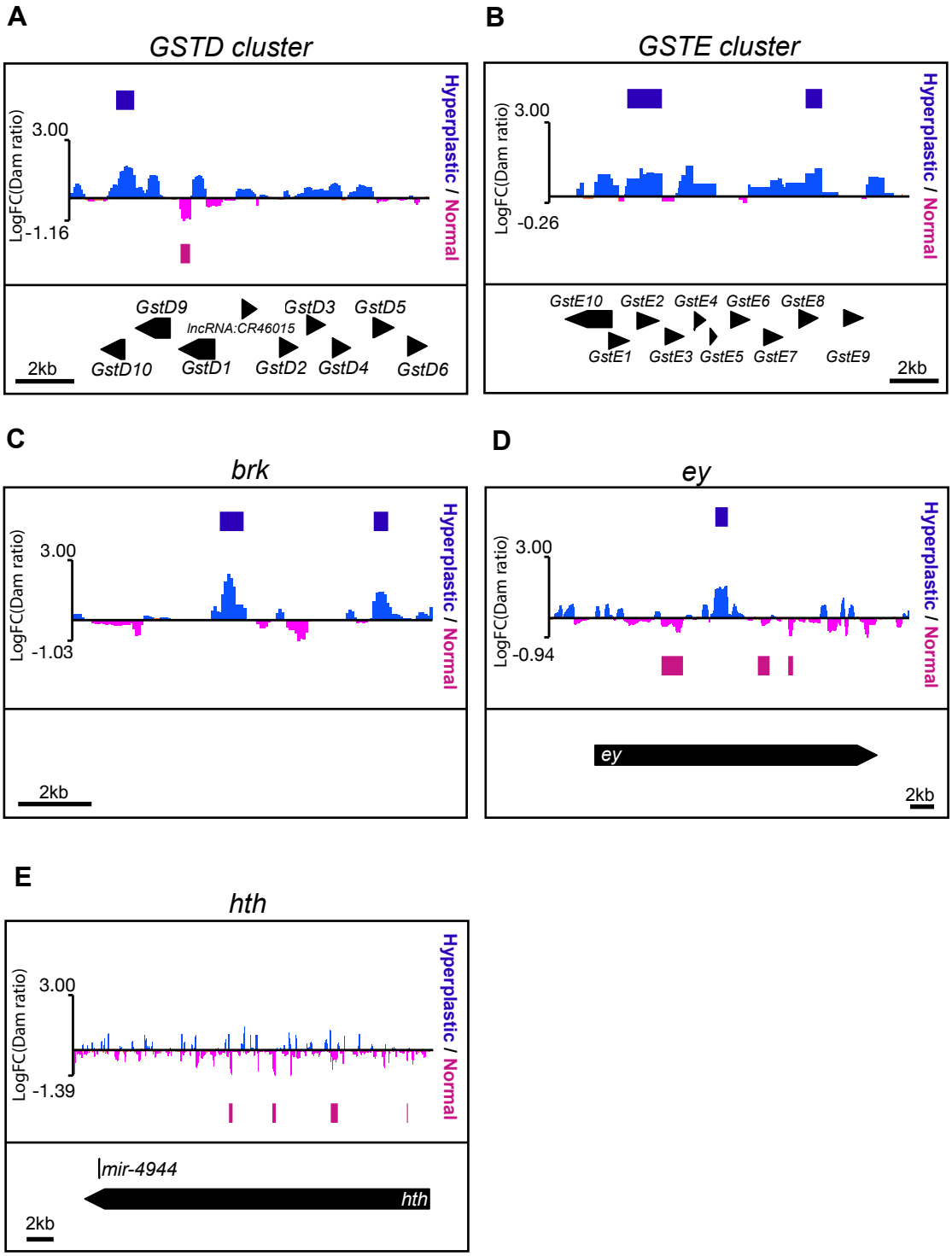

A
GSTD cluster
LogFC(Dam ratio)
3.00
-1.16
Hyperplastic / Normal
GstD9
GstD3
GstD5
lncRNA:CR46015
2kb
GstD10
GstD1
GstD2
GstD4
GstD6
B
GSTE cluster
LogFC(Dam ratio)
3.00
-0.26
Hyperplastic / Normal
GstE10
GstE2
GstE4
GstE6
GstE8
GstE1
GstE3
GstE5
GstE7
GstE9
2kb
C
brk
LogFC(Dam ratio)
3.00
-1.03
Hyperplastic / Normal
2kb
D
ey
LogFC(Dam ratio)
3.00
-0.94
Hyperplastic / Normal
ey
2kb
E
hth
LogFC(Dam ratio)
3.00
-1.39
Hyperplastic / Normal
mir-4944
hth
2kb

Figure 4.21: Chromatin accessibility is decreased in gene loci associated with retinal determination genes and photoreceptor differentiation in Yorkie driven hyperplastic eye discs.

A-F. CATaDa gene binding profiles showing adenine methylation by untethered Dam at the *hh* (A), *eya* (B), *so* (C), *ato* (D), *Enhancer of split* (E), and *Dl* (F) gene loci. Binding profiles for untethered Dam represent differential methylation and show the average log (Dam ratio) for all replicates between hyperplastic (*wts*^{X1}*/wtsLacZ*) and normal (*FT82B*) growth. The gene binding profiles were binned at 75 base pair intervals. The bars above and below of each panel represent peaks, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of the specific gene loci and surrounding genes. Scale bars in kb are shown.

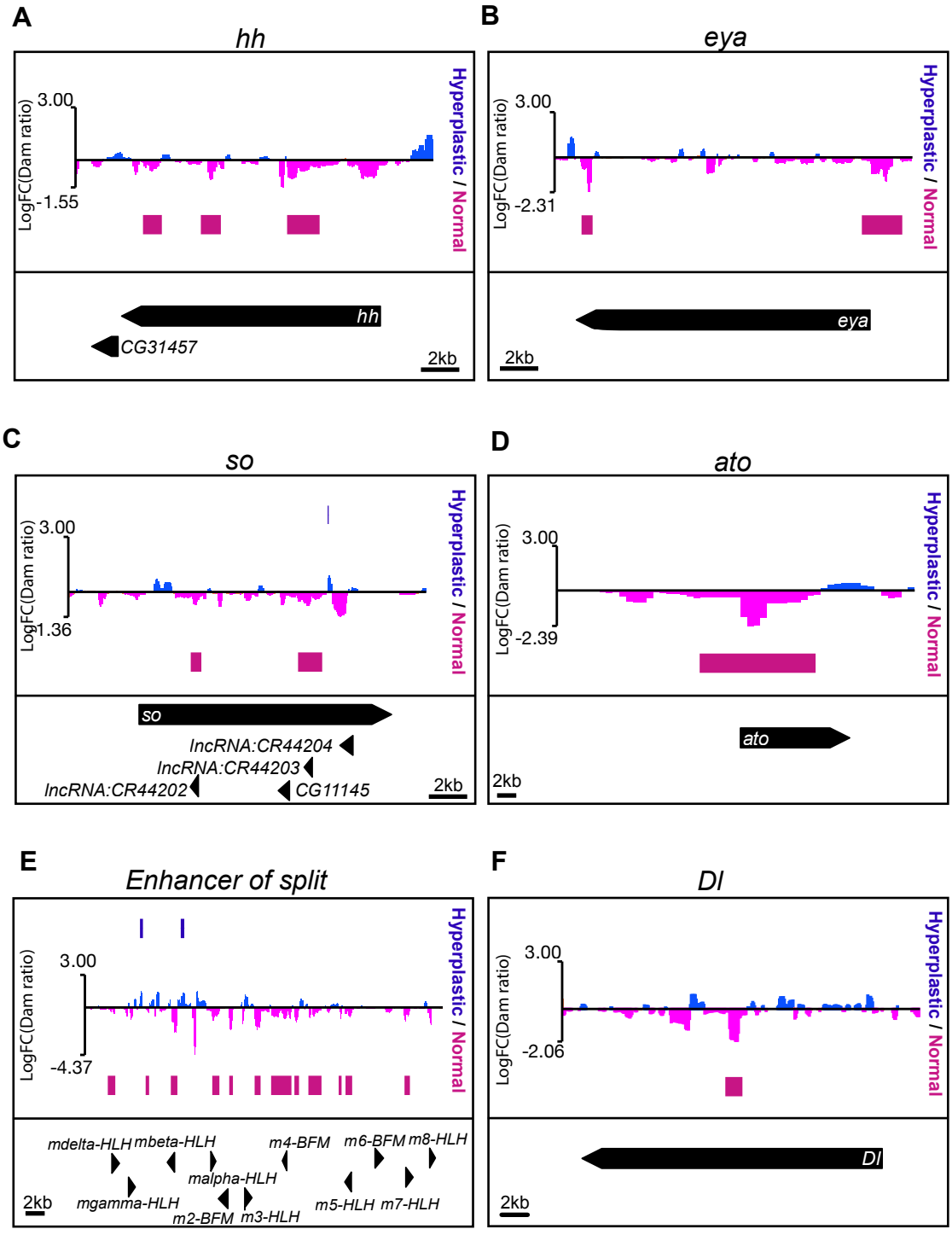
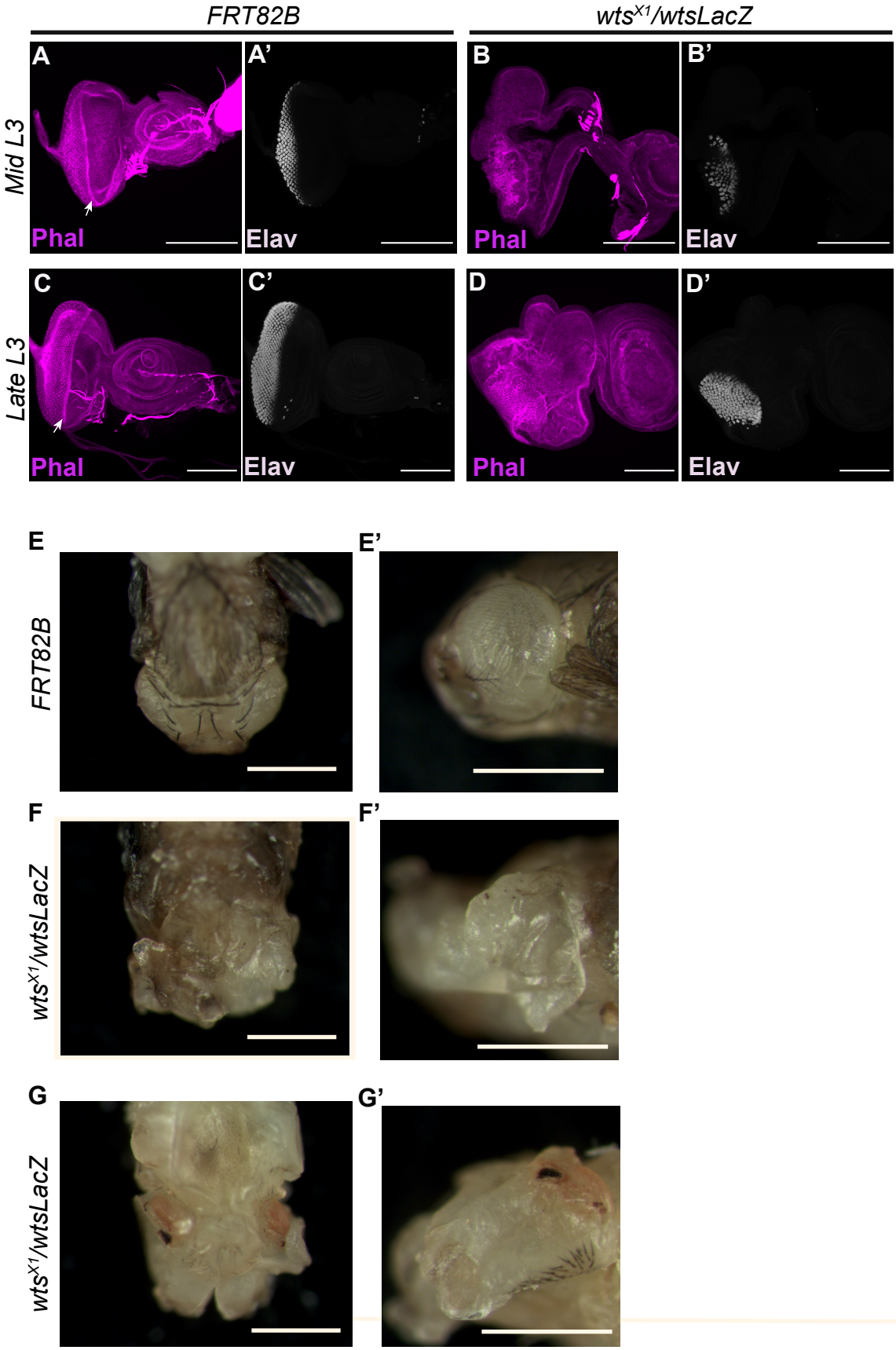

A
hh
LogFC(Dam ratio)
3.00
-1.55
Hyperplastic / Normal
hh
CG31457
2kb
B
eya
LogFC(Dam ratio)
3.00
-2.31
Hyperplastic / Normal
eya
2kb
C
so
LogFC(Dam ratio)
3.00
-1.36
Hyperplastic / Normal
so
lncRNA:CR44204
lncRNA:CR44203
lncRNA:CR44202
CG11145
2kb
D
ato
LogFC(Dam ratio)
3.00
-2.39
Hyperplastic / Normal
ato
2kb
E
Enhancer of split
LogFC(Dam ratio)
3.00
-4.37
Hyperplastic / Normal
mdelta-HLH
mbeta-HLH
m4-BFM
m6-BFM
m8-HLH
malpha-HLH
mgamma-HLH
m2-BFM
m3-HLH
m5-HLH
m7-HLH
2kb
F
Dl
LogFC(Dam ratio)
3.00
-2.06
Hyperplastic / Normal
Dl
2kb

Figure 4.22: Retinal development is altered in Yorkie driven hyperplastic eye discs and pharate pupal eyes.

A-D'. Maximum intensity projections of confocal microscope images of *Drosophila* larval eye-antennal discs from mid third instar (A-B') and late third instar (C-D') larvae. Eye-antennal discs were stained with Rhodamine-phalloidin to detect F-actin (magenta) and Elav antibody (white) to mark photoreceptor cells. Normal growth (*FRT82B*) eye discs are shown in A&C, and hyperplastic (wts^{X1}/*wtsLacZ*) eye discs are shown in B&D. Arrows indicate the morphogenetic furrow. Scale bar is 100μm.

E-G'. Images of female *Drosophila* pharate pupal heads for normal growth (*FRT82B*) (E) and hyperplastic growth (wts^{X1}/*wtsLacZ)* animals (F&G). E,F & G show overhead view, and E',F' and G' show side view of pupal head. Scale bars = 50μm.


FRT82B
wtsX1/wtsLacZ
Mid L3
Late L3
A
A'
B
B'
C
C'
D
D'
Phal
Elav
E
E'
F
F'
G
G'
FRT82B
wtsX1/wtsLacZ
wtsX1/wtsLacZ

Chapter 5: Investigating the AP-1 proteins and Hippo pathway in co-operating to regulate eye growth in *Drosophila*

5.1 Introduction

Organ growth is tightly controlled by a number of extrinsic and intrinsic factors that guarantee that organs grow to the correct size. The Hippo pathway is an example of one such factor that ensures proper organ growth by coordinating regulation of cell proliferation and apoptosis (Irvine and Harvey, 2015; Pan, 2010). The Hippo pathway restricts organ growth by phosphorylating and thereby limiting the nuclear localisation of the transcriptional coactivator Yorkie (Yki) in *Drosophila melanogaster* (Huang et al., 2005), and the orthologues Yes-Associated Protein (YAP) and Transcriptional co-activator with PDX binding motif (TAZ) in mammals (Lei et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008). Yki binds to sequence-specific transcription factors to regulate target gene expression (Slattery, Voutev, Ma, Nègre, White, and Mann 2013). The most well studied binding partner of Yki is the TEAD/TEF family transcription factor Scalloped (Sd), which is orthologous to TEAD1-TEAD4 in mammals (Goulev et al. 2008; S. Wu et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008; Zhao et al. 2008). Yki can also interact with additional transcription factors, such as Homothorax (Hth), Teashirt (Tsh) (Peng, Slattery, and Mann 2009), and Mad (Oh and Irvine 2011).

The interaction between Yki and Sd results in the transcriptional activation of target genes which are crucial for organ growth (Goulev et al. 2008; S. Wu et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008; Zhao et al. 2008). Sd is required for phenotypes resulting from Yki hyperactivity, for example loss of *sd* fully rescues the tissue overgrowth and elevated target gene expression induced by hyperactive Yki (S. Wu et al. 2008). Additionally, the undergrowth and decreased target gene expression in *yki* mutant tissue is rescued by loss of *sd* (Koontz et al., 2013). While *yki* is required for normal growth of the

Drosophila eye, *sd* is genetically dispensable for growth and expression of Hippo pathway target genes in the *Drosophila* eye (Goulev et al., 2008; Huang et al., 2005; Wu et al., 2008; Zhang et al., 2008). These observations are explained by the "default repression" model, whereby Yki regulates target gene expression by relieving transcriptional repression mediated by Sd and its corepressor protein, Tondu-domain-containing Growth Inhibitor (Tgi) (Koontz et al., 2013). In this model, growth is regulated by a careful balance of Yki-Sd mediated transcriptional activation and Tgi-Sd mediated transcriptional repression (Guo et al., 2013; Koontz et al., 2013).

In the larval *Drosophila* eye disc, tissue clones that are double mutant for both *sd* and *yki* grow normally and display normal expression levels of the target genes *Death-associated inhibitor of apoptosis 1* (*Diap1*) and *expanded* (*ex*) (Koontz et al., 2013). Importantly, this indicates that Hippo pathway target genes must be regulated independent of Yki and Sd by additional transcriptional regulatory proteins. Indeed, *Diap1* is also regulated by Stat92E downstream of JAK-STAT signalling following x-irradiation (Betz et al. 2008), while the Hippo pathway target gene *bantam* (*ban*) is also regulated by Notch signalling in neural stem cells (Y. C. Wu et al. 2017). Therefore, it is likely that many Hippo pathway target genes are regulated by transcriptional regulators that act either in cooperation with or in parallel to Yki and Sd.

In studies presented in Chapter 3 of this thesis, I found that the Activator Protein-1 (AP-1) motif is highly enriched in putative Yki and Sd target genes. AP-1 consists of a heterodimeric complex of Fos and Jun family proteins, which are known as Kayak and Jun-related antigen (Jra) respectively in *Drosophila* (Lutz Kockel, Homsy, and Bohmann 2001). AP-1 proteins are the downstream transcriptional effectors of the c-Jun amino-terminal kinase (JNK) pathway and the MAPK pathway (Lutz Kockel et al. 1997). The biological processes that JNK signalling regulates in *Drosophila* tissues are multifaceted and include: dorsal closure of the developing embryo, patterning of imaginal discs, proliferation, apoptosis, wound healing, as well as pro- and anti-tumourigenic activities depending on the context (La Marca and Richardson 2020; Weston and Davis 2007). Interestingly, the dual role of JNK in tumorigenesis is also linked to Hippo signalling: on the one hand, in clones that are mutant for the polarity gene *scribble* (*scrib*), JNK reduces Yki-induced overgrowth; on the other hand, in the presence of

hyperactive Ras, JNK signalling promotes tissue growth and tumorigenesis by inhibiting Hippo signalling and thereby activating Yki (Enomoto et al. 2015). Previous studies have also reported the co-occurrence of Sd and AP-1 DNA consensus binding motifs in gene loci, both in *Drosophila* tissues (Pascual et al. 2017), and human cancer cell lines (Liu et al. 2016; Stein et al. 2015; Verfaillie et al. 2015; Zanconato et al. 2015). Yki, Sd and AP-1 play an important role in lineage reprogramming in the *Drosophila* heart, where they synergistically activate genes that initiate alary muscle reprogramming and heart development (Schaub, Rose, and Frasch 2019a). Additionally, in cancer cells AP-1 forms a physical complex with YAP and TEAD to synergistically activate target genes involved in S-phase entry and mitosis (Zanconato et al. 2015), as well as cancer cell migration and invasion (Liu et al. 2016). YAP/TAZ are also capable of directly promoting *Fos* transcription in a feedforward loop that facilitates further activation of YAP/TAZ/Fos target genes (Koo et al. 2020).

Despite growing evidence for the cooperation between Yki, Sd, and AP-1 complexes, we are still far from having a complete understanding of how these complexes co-operate. Given the many reported links between Yki, Sd and AP-1 in transcription, development and cancer cell biology, it is possible that these transcriptional effectors regulate a set of common target genes to control the growth of *Drosophila* tissues. However, the cooperation between JNK and Hippo signalling has not been studied in the context of normal organ growth, and the global target genes of AP-1 in *Drosophila* are unknown. As such, I investigated whether Yki, Sd and AP-1 regulate shared target genes to control the growth of the *Drosophila* eye disc. In order to address this, I utilised targeted DamID to investigate the genomic binding profile of Jra. Additionally, I carried out genetic experiments to investigate potential co-operative roles for the JNK and Hippo pathways in the control of organ size and gene expression.

5.2 Results

5.2.1 Identification of AP-1 target genes using targeted DamID in the developing *Drosophila* eye disc

In *Drosophila*, the AP-1 complex consists of a heterodimer of Jun-related antigen (Jra) and Kayak (Kay), which have functional and structural similarities with mammalian Jun and Fos

respectively (Alfonso-Gonzalez and Riesgo-Escovar 2018). Although a number of AP-1 target genes have been identified in *Drosophila* tissues (Bunker et al. 2015; Külshammer and Uhlirova 2013; Martin-Blanco et al. 1998; Mcewen and Peifer 2005; Pearson et al. 2009), its genome-wide targets remain unknown. In order to understand the relationship between AP-1 and the Hippo pathway, it is essential to identify the full range of target genes of AP-1. Thus, I utilised targeted DamID to identify genes bound by the AP-1 protein Jra. Targeted DamID (detailed in Chapter 3.2.1 of this thesis) was performed on early third instar *Drosophila* eye discs. Briefly, targeted DamID involves fusing the bacterial DNA adenine methyltransferase (Dam) protein to a protein of interest, which results in methylation of DNA near genomic loci where the protein binds (Figure 3.2 & Figure 5.1A) (Marshall et al. 2016; Southall et al. 2013). The Jra-Dam fusion transgene was created by amplifying the *Jra* coding sequence, which was inserted into the multiple cloning site (MCS) upstream of the Dam sequence. The Dam-only control, *Yki*-Dam, and Dam-*Sd* DamID fusion transgenes are detailed in Chapter 3.2.1. These transgenes contained a primary open reading frame (ORF1) followed by two stop codons and a secondary open reading frame (ORF2) encoding the Dam methylase either alone (control), or fused to each protein of interest (Figure 5.1A). To confirm the expression of the DamID fusion transgenes, the Dam-only control and the Dam-*Jra* fusion transgene were crossed to the eye specific Gal4 driver *eyeless* (*ey*-Gal4). *mCherry* is contained in the ORF1 of the DamID fusion transgenes, and was used as a surrogate to detect the expression of the DamID transgene fusions (Figure 5.1B-C'). mCherry was detected in cells posterior to the morphogenetic furrow in late third instar larval eyes, confirming expression of the DamID transgenes (Figure 5.1B-C').

Targeted DamID was subsequently carried out on early third instar *Drosophila* eye discs, as outlined in Chapter 3. Differential methylation analysis was performed to identify GATC sequences (subsequently referred to as GATC 'tags') that were differentially methylated between the Dam-fusion and Dam-only control samples (Chapter 2.6.1.ii). A principal component analysis (PCA) was implemented to determine the greatest source of variation in the data. As I was interested in comparing the target genes bound by Jra, Yki, and Sd, I carried out the PCA on all of these targeted DamID datasets (Figure 5.1D). The PCA was visualised using a multidimensional scaling (MDS) plot that plotted the samples based on log2 fold changes. The distance between the samples on the MDS plot is an approximation of the differences in log2 fold changes (Dam-fusion compared to the Dam control) between the

samples, which is calculated as the leading fold change (Figure 5.1D). The MDS plot revealed that: 1) replicates for each DamID fusion protein clustered together closely and hence were reproducible; 2) the greatest source of variation was observed between the Dam-only controls and the Dam-fusion proteins; and 3) the second largest source of variation was observed between Jra and Yki/Sd Dam-fusion proteins.

Peak calling analysis was then performed to identify genomic regions that contain multiple significantly differentially methylated GATC 'tags' (Chapter 2.6.1.iii). These genomic regions are known as 'peaks' and are specific regions of the genome where the Dam-fusion proteins preferentially interact with DNA. The peaks were assigned to the nearest single gene within a 5kb distance. Peak calling analysis revealed 6001 peaks and 3461 Jra target genes, 5194 peaks and 2859 Yki target genes, and 6393 peaks and 3581 Sd target genes (Figure 5.2A&B). It should be noted that the number of peaks and target genes assigned for the Yki-Dam and Sd-Dam fusion proteins are different to the number reported in Chapter 3 (Figure 3.7A&B). This is a result of batch correction, which corrects for variation introduced when comparing datasets that were sequenced separately.

In order to gain further insight into the genome binding pattern of Jra, I analysed the location of peaks for Dam-Jra within specific annotated regions of the *Drosophila* genome (Release 6; dm6). The distribution of Dam-Jra peaks were analysed within the following annotated genomic features: 5'UTR, coding exons, introns, 3'UTR, and intergenic regions (Figure 5.3A). The *Drosophila* genome consists of approximately 5% 3'UTR, 3% 5'UTR, 21% exon, 34% intron, and 37% intergenic regions. As a control, I analysed the location of GATC 'tags' across the genome. The location of GATC 'tags' varies within the *Drosophila* genome, which creates a bias in where the Dam-fusion proteins are capable of methylating DNA. GATC 'tags' are enriched within exon regions and underrepresented in intergenic and intron regions of the *Drosophil*a genome (Figure 3.6A). Dam-Jra peaks were strongly enriched in introns and 5'UTRs, while underrepresented in exons, 3'UTRs and intergenic regions (Figure 5.3A). Focusing on gene promoters, specifically a 3 kilobase (kb) window either side of the transcriptional start site (TSS), revealed that Dam-Jra peaks were preferentially enriched around the TSS (Figure 5.3B), with a strong average peak observed immediately downstream of the TSS.

5.2.2 Comparisons between Jra, Yki, and Sd target genes

To identify shared candidate target genes of Jra, Yki and Sd, I examined the overlap of peaks and putative target genes for each Dam-fusion protein. Since AP-1 consensus transcription factor binding motifs have been found within Yki and YAP bound regions of the genome (Liu et al. 2016; Pascual et al. 2017; Zanconato et al. 2015; and Chapter 3.8 of this thesis), I predicted that I would find a significant number of overlapping target genes between Jra, Yki, and Sd. Indeed, 3114 peaks and 1737 putative target genes were shared between all three DamID fusion proteins (Figure 5.2A&B). This corresponds to 72% of putative Yki target genes and 60% of putative Sd target genes. Hence, Jra, Yki, and Sd share a high percentage of putative target genes, which suggests that they might coregulate expression of these genes. Yki and Sd shared more peaks and putative target genes than either protein did with Jra, which is expected given that Yki and Sd are known to form a physical complex and co-regulate target genes downstream of Hippo signalling (Goulev et al. 2008; S. Wu et al. 2008; L. Zhang, Ren, Zhang, Chen, Wang, and Jiang 2008; Zhao et al. 2008). Additionally, there were 1601 peaks and 1010 putative target genes that were uniquely bound by Dam-Jra, indicative of Yki and Sd independent genome binding by Jra (Figure 5.2B).

Pairwise correlation analysis was performed to assess the strength of association between the different pairs of Dam-fusion proteins using the Log fold change (LogFC) of shared peaks. The Pearson score (represented as an R value) conveys how closely associated two variables are; for example, the correlation of LogFC of a shared peak between two Dam-fusion proteins. Pairwise comparisons between all of the DamID datasets showed a positive correlation. The correlation between Jra and Yki shared peaks, and Jra and Sd shared peaks were moderate; $r = 0.31$ and $r = 0.45$ respectively (Figure 5.2C&D). As expected, Yki and Sd showed the strongest correlation, with a Pearson score of $r = 0.51$ (Figure 5.2E).

5.2.3 Transcription factor motifs enriched in Jra bound genomic regions

In *Drosophila*, Jra and Kay form heterodimers and have been shown to recognise the AP-1 consensus binding motif (Perkins et al. 1990). Importantly, AP-1 consensus binding motifs have been found to co-occur within Yki and YAP target genes (Liu et al. 2016; Pascual et al. 2017; Zanconato et al. 2015; and Chapter 3.8 of this thesis). Specifically, in a range of human cancer cell lines AP-1, YAP and TEAD co-occupy regions, at both distal enhancers and promoters, which contain TEAD and AP-1 consensus binding motifs (Liu et al. 2016; Zanconato et al. 2015). In the *Drosophila* wing disc, AP-1 consensus binding motifs are found at putative Yki target genes that are responsive to Yki hyperactivity (Pascual et al. 2017). To identify enriched transcription factor motifs in peaks bound by Jra, I performed motif enrichment analysis using Hypergeometric Optimization of Motif EnRichment (HOMER) (Heinz et al. 2010). As expected, the top most enriched motif within Jra peaks was the AP-1 consensus binding motif, which gives confidence that Jra DamID-seq was successful in identifying DNA loci bound by Jra (Figure 5.4A). The next most enriched motif included the consensus motif for YAP1, which is the Yeast AP-1 protein (Moye-Rowley, Harshman, and Parker 1989). Importantly, the Sd binding motif was also enriched within Dam-Jra binding sites. Additional enriched motifs include those for Grainyhead (Grh), Trithorax-like (Trl) and the basic leucine zipper domain (bZIP). Homer analysis was also carried out on peaks that were co-bound by Jra, Yki and Sd Dam-fusion proteins (Figure 5.4B). This analysis revealed that the top two consensus motifs included the AP-1 motif and the Sd motif, as expected. The Grh and Trl consensus motifs were also highly enriched in shared Jra, Yki and Sd peaks.

5.2.4 Jra target genes are enriched in Hippo and MAPK signalling pathways

KEGG analysis was implemented to identify signalling pathways that were enriched among putative Jra, Yki and Sd target genes. For all Dam-fusion datasets, the most significantly enriched pathway was MAPK signalling (Figure 5.5). In KEGG, the MAPK pathway includes the following three pathways: Extracellular signal-regulated kinase (ERK)/MAPK pathway, c-Jun N-terminal kinase (JNK) pathway and the p38 kinase pathway. The enrichment of MAPK signalling in KEGG analysis is consistent with Jra functioning downstream of both JNK and ERK/MAPK signalling (L. Kockel et al. 1998). Putative target genes of Jra, Yki and Sd within the MAPK pathway included genes involved in JNK feedback loops such as *puckered* (*puc*)

and *kayak* (*kay*), as well as genes that play important roles in JNK mediated tumour growth and invasiveness such as *cheerio* (*cher*), *unpaired1* (*upd1*), and *matrix metalloprotease 1* (*Mmp1*) (La Marca and Richardson 2020). Genes involved in ERK/MAPK signalling were also bound by Jra, Yki and Sd, including *pointed* (*pnt*), *capicua* (*cic*), *spitz* (*spi*), *argos* (*aos*), *kekkon* (*kek*), and *Epidermal growth factor receptor* (*Egfr*). Interestingly, Jra-Dam fusion bound genes were also enriched in Hippo pathway target genes, such as *kibra*, *expanded* (*ex*), and *warts* (*wts),* which encode Hippo pathway proteins. These genes are known transcriptional targets of Yki and Sd and form a negative feedback loop (Hamaratoglu et al. 2006b; Genevet et al. 2010). The Jra Dam-fusion protein also bound other well-known Yki and Sd target genes, including *Death-associated inhibitor of apoptosis 1* (*Diap1*), *Cyclin E* (*CycE)*, *bantam* (*ban*), and *Myc*.

Both JNK/AP-1 and the Hippo pathway regulate apoptosis (Dnasekaran and Premkumar Reddy, 2017; Huang et al., 2005). In KEGG analysis, shared target genes of Jra, Yki and Sd were enriched in apoptosis pathways (Figure 5.5). Shared putative Jra, Yki, and Sd Dam-fusion target genes included the pro-apoptotic genes *hid*, *reaper*, and *Dark*, as well as the anti-apoptotic genes *puc*, *bantam* and *Diap1*. Additionally, there was a significant enrichment of genes involved in the following signalling pathways: Endocytosis, Wnt signalling, ECM-receptor interaction, Hedgehog signalling, and TGF-beta signalling (Figure 5.5). These results indicated that Jra, Yki, and Sd co-bind genes involved in multiple signalling pathways and processes. The full list of KEGG analysis can be found in Appendix Table 8.8.

Several pathways were also uniquely enriched among Jra target genes, but not shared by Yki and Sd, including Pyrimidine metabolism, ubiquitin mediated proteolysis, ether lipid metabolism and autophagy (Figure 5.5). Autophagy, which is the process of lysosomal degradation of cytoplasmic material that involves autophagy-related (Atg) proteins, is activated by JNK in response to oxidative stress (Y. Y. Zhou et al. 2015). Interestingly, previous research has shown that JNK can induce the expression of several Autophagy-relation genes via dFoxo (H. Wu, Wang, and Bohmann 2009). Targeted DamID of Jra revealed that Jra binds to multiple Autophagy-related genes, such as *Atg4a*, *Atg5*, *Atg12*, *Atg13*, and *Atg18*. These results suggest that Jra might also regulate autophagy-related genes, presumably downstream of JNK signalling, in the developing *Drosophila* eye disc.

5.2.5 Jra, Yki, and Sd bind to known AP-1-regulated genes

I next investigated the binding of Jra, Yki and Sd at the known AP-1-regulated genes *puckered* (*puc*), *misshapen* (*msn*), *cheerio* (*cher*), and *unpaired3* (*upd3*). These genes are responsive to AP-1 activity downstream of JNK signalling, however it is not known whether these genes are directly bound by AP-1 (Bunker et al. 2015; Martin-Blanco et al. 1998; Pearson et al. 2009).

puc encodes a phosphatase that inhibits JNK activity in a negative feedback loop during dorsal closure of the developing embryo (Martin-Blanco et al. 1998), and also prevents apoptosis in the eye and wing imaginal discs by restricting JNK signalling (McEwen and Peifer, 2005). Targeted DamID revealed that the Jra Dam-fusion strongly methylated the entire *puc* locus and peaks were assigned throughout the gene indicating binding of Jra to the *puc* locus (Figure 5.6A). Yki and Sd Dam-fusion proteins also methylated *puc* and several peaks were assigned throughout the gene locus, indicating that Yki and Sd also bind the *puc* locus. The pattern of methylation between Jra, Yki and Sd was similar, however, the degree of methylation, represented by the height of the LogFC between the Dam-fusion and Dam control, was much weaker and there were fewer peaks for Yki and Sd compared to Jra.

msn encodes an upstream activating kinase (MAPKKKK) in the JNK pathway, which is activated by TNF in response to cellular stress (Lutz Kockel, Homsy, and Bohmann 2001). *msn* is also directly regulated by AP-1 as part of a negative feedback loop and is activated in epidermal cells surrounding wounds in *Drosophila* embryos and larvae (Pearson et al. 2009). I found that the *msn* gene locus was methylated by the Jra Dam-fusion, with several methylation peaks (Figure 5.6B). Yki and Sd Dam-fusion proteins also methylated *msn* to a similar degree as Jra, and all three fusions showed several overlapping peaks.

cher encodes an F-actin crosslinking protein that is downstream of JNK signalling and required for the growth of *Drosophila* invasive Ras tumours with epithelial cell polarity defects (Pearson et al. 2009). I found that the *cher* gene locus was strongly methylated by the Jra Dam-fusion with several methylation peaks (Figure 5.6C). Yki and Sd Dam-fusion proteins strongly resembled the methylation pattern and peaks in the *cher* gene locus as those observed for Jra.

upd3 encodes a cytokine-like Unpaired ligand that activates JAK/STAT signalling. In *Drosophila*, *upd3* is also known to drive tumour overgrowth downstream of JNK signalling in *scribble* (*scrib*) mutants, which are invasive tumours with epithelial cell polarity defects (Bunker et al. 2015). Interestingly, in the *Drosophila* wing disc, a polarity responsive enhancer within *upd3* (*upd3.3* enhancer) can be simultaneously activated by: 1) Fos downstream of JNK signalling; and 2) Yki downstream of apicobasal cell polarity-regulated aPKC, which is independent of Hippo signalling (Bunker et al. 2015). Indeed, I found that Jra, Yki, and Sd Dam-fusions all strongly methylated the *upd3* gene region (Figure 5.6D), which confirms previous findings (Bunker et al. 2015). Jra showed the strongest methylation of *upd3*, with larger peaks than Yki or Sd. However, Yki and Sd Dam-fusions did not directly methylate the *upd3.3* enhancer, and the Jra Dam-fusion protein exhibited a low amount of methylation of this enhancer (Figure 5.6D). Instead, Jra and Sd Dam-fusion proteins methylated the *upd3* locus directly upstream of the enhancer (Figure 5.6D), indication that Jra and Sd Dam-fusion proteins bind near the enhancer region. Overall, these results show that Jra binds to known JNK/AP-1 target genes in the developing eye disc and further supports the validity of the Jra DamID data. The fact that Yki and Sd bind to the well-known AP-1 target genes, *puc*, *msn*, *cher*, and *upd3*, suggests that the Hippo pathway might regulate at least some AP-1 target genes, although further experiments are needed to confirm this. Additionally, consensus motifs for AP-1 and Sd were broadly distributed throughout the *puc*, *msn*, *cher* and *upd3* gene loci, which are shown as vertical lines below each gene view (Figure 5.6). As shown previously, I also found that the *upd* 3.3 kb enhancer contained multiple AP-1 and Sd consensus binding motifs (Figure 5.6D).

5.2.6 Jra binds to known Hippo pathway target genes

Direct target genes of Yki and Sd include the cell death inhibitor gene *Death-associated inhibitor of apoptosis 1* (*Diap1*), the Hippo pathway gene *expanded (ex*), the cell cycle regulator gene *Cyclin E* (*CycE*) and the microRNA *bantam* (*ban*) (Hamaratoglu et al., 2006; Oh and Irvine, 2011; Qing et al., 2014; Peng et al., 2009; Wu et al., 2003). I visualised the binding of Jra, Yki and Sd Dam-fusion proteins at these genes (Figure 5.7) and found that: 1) there were significant methylation peaks induced by each of the Jra, Yki and Sd Dam-fusion proteins; 2) Yki and Sd Dam-fusion proteins displayed the most similar methylation pattern and peak locations; 3) The Jra Dam-fusion protein closely resembled the methylation pattern

and peak locations of the Yki and Sd Dam-fusion proteins but was slightly different. Specifically, Jra methylation was present throughout the *Diap1*, and *ex* gene loci and closely resembled Yki and Sd methylation patterns (Figure 5.7A-B), and Jra, Yki, and Sd peaks were concentrated close to the CycE promoter (Figure 5.7C). The microRNA *ban* also contained overlapping Jra, Yki and Sd methylation peaks (Figure 5.7D). These results show that Jra binds to several well-defined Hippo pathway target genes and indicates that Jra might regulate these target genes either independently or in cooperation with Yki and Sd. Sd and AP-1 consensus motifs were broadly distributed throughout the gene loci of *Diap1*, *ex*, *CycE*, and *ban* (Figure 5.7A-D). As shown previously, I also found Sd consensus binding motifs within a 1.8kb enhancer of *Diap1*, which is known to be responsive to Hippo signalling (Figure 5.6D).

5.2.7 Differential expression of JNK pathway target genes in eye discs with defective Hippo signalling

As my results revealed that Jra, Yki and Sd bind to many of the same genes, I next investigated if Jra target genes were differentially expressed in *warts* (*wts*) transheterozygous mutant eye discs, which overgrow and become hyperplastic (Justice et al. 1995; Tian Xu et al. 1995) as a result of Yki hyperactivity (Huang et al., 2005). For this, I compared Jra target genes that were identified by targeted DamID with genes that were differentially expressed in *wts* mutant hyperplasic eye discs compared to wild-type control eye discs (described in Chapter 4). Briefly, I used differential gene expression analysis to calculate the log fold change of genes in *wts* mutant versus *wild-type* eye discs. A negative logFC value denotes a gene that was downregulated in *wts* mutant eye discs, whereas a positive logFC value denotes a gene that wass upregulated. I implemented a fold change threshold of larger than 0.5 and an adjusted p-value threshold of less than 0.01, as this threshold incorporated known Hippo pathway target genes that were increased in *wts* mutant *Drosophila* eye discs. From these comparisons, I found that 154 putative Jra target genes were upregulated and 268 putative Jra target genes were downregulated in *wts* mutant eye discs compared to *wild-type* eye discs (Figure 5.8A). A hypergeometric test was carried out on these comparisons to assess whether the overlap between differentially expressed genes and Jra target genes was statistically significant. This analysis revealed that Jra binding was significantly enriched in both upregulated and downregulated genes in *wts* mutant eye discs, and more significantly enriched among upregulated genes (Figure 5.8A). Overall, approximately 12% of genes that were bound by Jra

in DamID studies were differentially regulated in *wts* mutant eye discs. In addition, 60% of the differentially expressed putative Jra target genes were also putative target genes of both Yki and Sd.

The expression of numerous genes has been shown to change expression in response to activation of JNK signalling. To identify whether canonical JNK target genes were differentially expressed in *wts* mutant tissue and therefore responsive to Yki hyperactivation, I analysed their expression. JNK target genes that are important for embryonic dorsal closure include: *chickadee* (*chic*) (Jasper et al. 2001), *puc* (Martin-Blanco et al. 1998), *jra*, *reaper* (*rpr*), *Zasp52*, *scarface* (*scaf*), *decapentaplegic* (*dpp*) and *Rab30* (Rousset et al. 2017). Additionally, multiple genes are upregulated by JNK signalling in invasive *Drosophila* tumours with apico-basal polarity defects (M. Wu, Pastor-Pareja, and Xu 2010; Pagliarini and Xu 2003), including *cher,* which encodes a dimeric F-actin cross-linking protein (Külshammer and Uhlirova 2013), the ETS-domain factor *Ets21c*, the nuclear receptor *Ftz-F1* (Kulshammer et al. 2015), the unpaired genes (*upd1*, *upd2*, *upd3*) which encode cytokines (M. Wu, Pastor-Pareja, and Xu 2010), the matrix metalloprotease gene *Mmp1* (Uhlirova and Bohmann 2006), and the pro-apoptotic gene *head-involution defective* (*hid*) (Luo et al. 2007). These genes were all methylated by Jra, Yki and Sd Dam-fusion proteins in my DamID studies (Figure 5.6). The majority of these genes were not differentially expressed in *wts* mutant eye discs (Figure 5.8B), with *hid* being the sole gene that was upregulated, with a logFC of 0.84. Additionally, select genes were significantly differentially expressed but did not meet the logFC threshold implemented in this thesis, including *Mmp1* (logFC = 0.47), *upd1* (logFC = -0.37) and *upd3* (logFC = -0.37). Collectively, these results indicate that the expression of most well-known JNK/AP-1 target genes are not responsive to Yki hyperactivation despite the fact that Yki and Sd reside at these genes, as determined by DamID.

5.2.8 Investigating a potential role for AP-1 in growth control

The role of JNK signalling in regulation of growth of *Drosophila* tissues is not well understood and insights come from only a select number of publications. JNK has been reported to activate Yki in a stripe along the centre of the developing wing disc which then results in the activation of Yki target genes (Willsey et al. 2016). Loss of JNK signalling along this stripe results in the

reduction in wing size (Willsey et al. 2016). Additionally, a link between AP-1 and Yki at the transcriptional level has been recently reported in growth control, whereby loss of a co-repressor called CtBP in the wing disc leads to Yki and AP-1-mediated activation of the growth-promoting microRNA *ban*, a Yki target gene (Sumabat et al. 2019). Early studies investigating AP-1 in *Drosophila* have shown that *jra* mutant clones in the eye disc display only rare and mild defects in photoreceptor differentiation and ommatidial assembly, suggesting that Jra is not important for eye growth (Lutz Kockel et al. 1997). However, many questions remain unknown, including whether AP-1 cooperates with the Hippo pathway to control the growth of the *Drosophila* eye, and if eye growth is controlled through a set of target genes shared by AP-1, Yki and Sd. The targeted DamID data presented here showed that Jra, Yki and Sd Dam-fusion proteins bound many of the same target genes, suggesting that Hippo pathway target genes could be controlled by AP-1, either independent of, or together with, Yki and Sd. As such, I next sought to investigate a potential role for AP-1 in growth control in the *Drosophila* eye. Here, I specifically aimed to investigate the possibility that AP-1, Yki and Sd co-regulate expression of genes that are critical for organ growth. Although I studied Jra in the context of targeted DamID to identify putative AP-1 target genes, in the following experiments I studied the Jra binding partner Kay. The reasoning for this was that Kay forms stronger homodimers than Jra, although both homodimers are not as efficient in binding AP-1 consensus sequences and activating transcription as Jra/Kay heterodimers (Ciapponi and Bohmann 2002). Additionally, evidence also suggests that Kay functions independently of Jun during the early stages of embryogenesis, either as homodimers or heterodimers with other transcription factors (Riesgo-Escovar and Hafen 1997a). Hence, the function of Jra and Kay heterodimers could be substituted by Kay homodimers, a possibility that can be circumvented by *kay,* but not *jra,* loss of function experiments.

5.2.8i Depletion of *kay* reveals a role for AP-1 in growth control

I investigated whether the depletion of the AP-1 gene *kay* had an effect on the overgrowth caused by Yki hyperactivity in the *Drosophila* eye. To assess this, I utilised a mutant form of Yki, in which a critical Wts phosphorylation site is altered from a Serine to an Alanine (YkiS168A), and is therefore unable to be inhibited by Wts (Dong et al., 2007; Oh and Irvine, 2008). Expressing YkiS168A specifically in the eye using the *Glass Multiple Reporter (GMR)-*

Gal4 driver resulted in severely overgrown adult eyes, consistent with published studies (Figure 5.9D&D') (Dong et al., 2007; Oh and Irvine, 2008). As controls, first I expressed two *kay* RNAi lines in the eye alone to assess if reduction of *kay* affects eye growth. This resulted in relatively normal eye growth, although BL6212 did show a mild reduction in eye size (Figure 5.9C). Depleting *kay* using two independent RNAi lines supressed the overgrowth caused by YkiS168A (Figure 5.9E&F). The *kay* RNAi line BL27722 showed the strongest phenotype, whereby the YkiS168A eye overgrowth was rescued to wild-type size (comparing Figure 5.9E to 5.9A). These results indicate that *kay* is important for Yki-mediated overgrowth in the *Drosophila* eye.

Next, I assessed whether the depletion of *kay* affected the growth of tissues in which *sd* was also depleted. Previous studies have shown that Sd is dispensable for normal eye development due to its role as a default repressor of transcription (Koontz et al., 2013), which indicates that additional transcription factors regulate growth in the absence of Sd. Consistent with previous findings (Koontz et al., 2013), I found that depletion of *sd* using an RNAi line resulted in no or only a very mild reduction in growth (Figure 5.9G-G'). Interestingly, depletion of *kay* resulted in a further reduction in eye size when combined with *sd* RNAi (Figure 5.9H-I). In particular, the *kay* RNAi line BL27722 showed the strongest phenotype in which there was a strong reduction in eye size (Figure 5.9H-H'). Therefore, the combined reduction of Sd and Kay negatively effects eye growth, whereas loss of either alone have no or very minor effects on eye growth. These results suggest that both Sd and Kay are required for normal eye growth, and in the absence of either one, the other protein is able to compensate and maintain the growth of the eye.

5.2.8ii Mutagenesis of *kayak* by CRISPR Cas9 genome editing

RNAi lines can have off-target effects or elicit inefficient gene knockdown, which can give either false-positive or false-negative result. For example, approximately 25% of the VDRC KK RNAi library causes enhancement of expression of Hippo pathway target genes in the eye leading to false-positives in Hippo pathway screens (Vissers et al. 2016). Therefore, to confirm the RNAi-based results above I sought to verify these studies using mutant *kay* alleles. There are two mutant *kay* alleles described in the literature that were created by ethyl

methanesulfonate (EMS) mutagenesis: *kay*^{1} is an amorphic allele that also harbours a mutation in *Diap1* (Riesgo-Escovar and Hafen 1997a), and *kay*^{2} which is a weaker temperature sensitive hypomorphic allele (Zeitlinger et al. 1997). As these were not ideal for studying the role of *kay* in growth control, I utilised CRISPR Cas9 genome editing to create new mutant *kay* alleles. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) utilises the sequence specific Cas9 endonuclease whose target specificity is determined by the sequence of its guide-RNAs (gRNA) (Jinek et al. 2012). The 20 nucleotide gRNA can be designed to match desired DNA sequences to which the Cas9 can be targeted (Bassett and Liu 2014). The target sequence also must contain a photospacer adjacent motif (PAM) that is required for Cas9 to cut the DNA target site and create a double stranded break (DSB) (Bassett and Liu 2014). The DSB is repaired by nonhomologous end joining (NHEJ), which is error prone and therefore can create deletions or insertions (indels) at the target DNA loci (Bassett and Liu 2014).

The *kay* locus encodes for five different *kay* transcripts through alternative splicing (Figure 5.10A) (Alfonso-Gonzalez and Riesgo-Escovar 2018). All isoforms share the last two exons, where the bZIP domain is encoded, which is important for DNA binding and dimerization (Zeitlinger et al. 1997). To create a mutant *kay* allele, I designed a gRNA targeted to exon 6 of the *kay* sequence, as this exon was common to all *kay* transcripts (Figure 5.10A). This guide was cloned into the pBFv-U6.2 vector which was inserted into the *Drosophila* genome to create transgenic stocks (Kondo and Ueda 2013). Germline *kay* mutation stocks were created, screened and sequenced to confirm the mutation. PCR was carried out using the DNA extracted from potential mutant *kay* alleles in order to amplify the *kay* sequence targeted by the CRISPR gRNA. DNA from the PCR reaction showed two different sized bands on an agarose gel, while a wild-type stock showed only a single band (Figure 5.11B). Sequencing revealed that the mutant *kay* allele contained a 28 base pair deletion that resulted in a frame shift and a premature STOP codon (Figure 5.10C&D). For the remainder of this chapter, I will refer to this mutant *kay* allele as *kay*^{KM}. Animals that were homozygous mutant for *kay*^{KM} died early in larval development, which suggests that it is a loss of function allele. The STOP codon occurs before the basic region and leucine zipper (bZIP) domain sequence of *kay* which is needed for DNA binding and dimerization (Zeitlinger et al. 1997). Therefore, I postulated that the *kay*^{KM} allele encodes for a truncated protein that is not able to perform its normal function, including DNA binding and homo or heterodimerization. However, as we did not have access to a *Drosophila* Kay antibody, I was unable to confirm this using a Western blot. In the future,

it will be crucial to fully characterise the *kay*^{KM} allele. Additionally, to confirm that *kay*^{KM} is indeed a loss of function allele I could also assess whether it gives rise to embryonic dorsal closure phenotypes, which have been associated with other *kay* alleles (Alfonso-Gonzalez and Riesgo-Escovar 2018).

5.2.8iii Clones of *kay* mutant eye tissue displayed no overt growth defects

In order to confirm the *kay* RNAi-based results, I carried out experiments using the *kay*^{KM} allele. Firstly, I investigated whether the loss of *kay* results in a synergistic reduction of tissue growth, and Diap1 protein expression, when combined with a *sd* loss of function allele. To do this, I generated *sd kay* double mutant clones in the developing eye disc using a crossing scheme similar to that published by Yu & Pan 2018 (Figure 5.11A). *sd* and *kay* are located on different chromosomes, and, as such, creating double-mutant clones required combining an eye-specific flippase (FLP) with two flippase recognition target (FRT) chromosomes containing *sd* and *kay* mutations in trans to two FRT chromosomes containing GFP and RFP markers (Figure 5.11A). As controls, *sd* and *kay* single-mutant clones were generated using a FRT chromosome containing the mutation and a wild-type FRT chromosome in trans to two FRT chromosomes containing GFP and RFP (Figure 5.11A). This experiment allowed the marking of wild-type clones in GFP and RFP as follows: single mutant *sd* clones lacking GFP; single mutant *kay* clones lacking RFP; and *sd kay* double mutant clones lacking both GFP and RFP. Based on the eye undergrowth phenotype of Kay and Sd double RNAi (Figure 5.8G-H'), I expected to see that loss of *sd* and *kay* alone to have no effect on clone size, but loss of *sd* and *kay* together to impair clone size. Additionally, as Jra, Yki, and Sd all bind to the *Diap1* locus (Figure 5.7A), I expected that the combined loss of *sd* and *kay* would reduce the expression of Diap1 protein (as depicted in Figure 5.11B). Clones that contained a strong mutant allele *sd*^{47M} (Srivastava et al. 2004), grew normally (Figure 5.12A-A'''), consistent with published studies (Koontz et al., 2013; Wu et al., 2008). Additionally, these clones displayed no change in Diap1 protein expression compared to wild-type clones, although Diap1 expression is not uniform throughout the eye disc (Figure 5.12A'''). *kay*^{KM} eye disc clones also grew normally and displayed no change in Diap1 levels (Figure 5.12B'B'''). Unexpectedly, *sd*^{47M} *kay*^{KM} double mutant clones also grew normally and displayed no change in Diap1 levels (Figure 5.12C-C'''). Adult eyes from these crosses were also examined for growth phenotypes (Figure5.13).

In this experiment, the red tissue represented homozygous wild-type clones, the orange tissue represented clones with a single mutant or heterozygous tissue, and the white tissue represented the double-mutant clones. As controls, *sd* and *kay* single mutant clones containing the relevant mutation and a wild-type FRT chromosome were also represented by white tissue. The overall size of the adult eye was similar between the sd^{47M} kay^{KM} double mutant genotype and the sd^{47M} and kay^{KM} single mutants (comparing Figure 5.13C&C' to A-B'). These results indicate that the combined loss of *sd* and *kay* has no obvious effect on the growth of clones in developing eyes.

I next investigated whether the loss of *kay* would rescue the tissue overgrowth caused by a Hippo pathway mutation, and whether this would impact upon Yki activity, as measured by *ex-LacZ* expression. Salvador (Sav) is a scaffold protein that is phosphorylated by Hippo (Hpo), which then facilitates Wts phosphorylation by Hpo (Tapon et al. 2002; S. Wu et al. 2003). In *sav* mutants, the phosphorylation of Wts is blocked, which leads to overgrown tissues and elevated expression of Yki target genes (Tapon et al. 2002; S. Wu et al. 2003), including *ex* (Hamaratoglu et al., 2006). I created a recombinant stock that contained the kay^{KM}, the sav^{3} mutation, and a FRT on the same chromosome arm. Homozygous clones were created in the eye by crossing eyeless-FLP to lines harbouring kay^{KM}, sav^{3}, and kay^{KM} sav^{3} alleles. kay^{KM} mosaic eyes possessed a relatively equal amount of mutant tissue (white clones) and wild-type tissue (red clones) (Figure 5.14A&A'). As shown previously (Tapon et al. 2002), sav^{3} mosaic eyes displayed an excess of mutant compared to wild-type tissue, as well as tissue overgrowth and folding (Figure 5.154B & B'). Eyes that were mosaic for both kay^{KM} and sav^{3} also showed an excess of mutant eye tissue and overgrowth, resembling sav^{3} mosaic eyes (Figure 5.14C&C'). Therefore, the loss of *kay* did not rescue the eye overgrowth phenotype induced by sav^{3}. Additionally, I also investigated if the loss of *kay* would rescue the elevated *ex* levels in sav^{3} mutant tissue (Figure 5.15). *ex* is a well-characterised target gene of Yki, which acts in a Hippo pathway negative feedback loop (Hamaratoglu et al., 2006). I utilised a LacZ enhancer trap inserted into the *ex* locus as a reporter for *ex* transcription (Boedigheimer and Laughon 1993). The heat-shock FLP system was used to generate clones of tissue containing sav^{3} alone or kay^{KM} sav^{3}. As shown previously (Hamaratoglu et al., 2006), *ex* expression was elevated in sav^{3} clones in the larval eye disc. kay^{KM} sav^{3} clones also displayed elevated expression of *ex* (Figure 5.15B & B'), revealing that the loss of *kay* did not rescue the increase in Yki activity in eye sav^{3} tissue. In conclusion, while RNAi experiments suggested a role for Kay in

regulating eye growth together with the Hippo pathway, these results were not verified using loss of function studies with a newly generated kay allele.

5.3 Discussion

In this chapter, I investigated the genome wide binding of *Drosophila* AP-1, as well as a potential role for AP-1 in regulating organ growth. Targeted DamID revealed that the AP-1 protein Jra shares 71% of its target genes with the transcriptional effectors of the Hippo pathway. RNAi-mediated depletion of the AP-1 gene *kay* suppressed Yki-driven overgrowth in the eye, and reduced eye growth when *sd* was also depleted. This suggests that AP-1, Yki, and Sd cooperate to promote growth of the *Drosophila* eye, and also that this cooperation could directly converge at the level of common target genes. Unexpectedly, however, experiments using a mutant *kay* allele indicated no obvious role for AP-1 in regulating growth. Furthermore, *kay* was not required for the expression of the Hippo pathway target genes *Diap1* and *ex*.

5.3.1 Genome-wide binding of AP-1, Yorkie, and Scalloped

Targeted DamID studies revealed that Jra, Yki, and Sd shared a high percentage of putative target genes. Indeed, 71% of putative Jra target genes overlapped with putative Yki and/or Sd target genes, indicating that together these proteins may coregulate several thousands of target genes. Shared putative Yki, Sd, and Jra target genes were also enriched with Sd and AP-1 consensus binding motifs, providing further evidence that Yki, Sd, and Jra share many target genes. This is consistent with previous findings of AP-1 consensus motifs co-occuring within Yki and YAP target genes from the *Drosophila* wing disc (Pascual et al. 2017), and human cancer cell lines respectively (Liu et al. 2016; Stein et al. 2015; Zanconato et al. 2015; Verfaillie et al. 2015). Additionally, KEGG analysis on shared target genes highlighted a strong enrichment of MAPK/JNK and Hippo pathway genes. Taken together, these results suggest that AP-1, Yki, and Sd converge at the level of common target genes. However, while AP-1 complexes commonly consists of heterodimers of Jra and Kay, in some contexts such as embryogenesis, Kay can function independently of Jun (Riesgo-Escovar and Hafen 1997a). Therefore, in order to fully understand the genome wide binding of AP-1 complexes, it will

also be important to establish the genome-binding profile of Kay. For example, in the future, targeted DamID could be used to identify Kay target genes in the eye disc.

I found that Jra, Yki and Sd all reside at the JNK target genes *puc*, *msn*, and *cher*. However, these genes and additional well-known AP-1 target genes were not responsive to hyperactive Yki, indicating that Yki and Sd are unlikely to be key regulators of these genes during eye growth. I also found that *upd3* was a shared target gene of Jra, Yki and Sd. *upd3* encodes a cytokine-like Unpaired (Upd) ligand, which drives neoplastic tumour growth, which are tumours with disrupted epithelial structure, through activating the JAK-STAT pathway (Amoyel, Anderson, and Bach 2014). In my studies, *upd3* was downregulated in hyperactive Yki eye discs, although it did not meet the LogFC threshold implemented in this study. In contrast, in the wing disc, Yki binds to a *upd3* polarity-responsive enhancer which is upregulated by a hyperactive form of Yki (Bunker et al. 2015), and *upd1/2/3* are highly upregulated in *wts* mutant clones (Pascual et al. 2017). These contrasting results could be explained by tissue specific differences in Yki target gene regulation; in *scrib* mutant wing discs, which form neoplastic tumours, *upd3* can be simultaneously activated by JNK-dependent Fos and aPKC-activated Yki (Bunker et al. 2015). In the future, it will be interesting to understand if a similar mechanism of gene regulation by Yki and AP-1 also occurs at shared target genes in the neoplastic tumours of the *Drosophila* eye.

The JNK target gene *Mmp1* was found in my study to be a shared target of Jra, Yki and Sd and also upregulated in response to hyperactive Yki. *Mmp1* encodes a protein that is required for basement membrane degradation and also contributes to the invasiveness of tumour cells (Hua et al. 2011). For example, *Mmp1* is regulated by AP-1 and its expression is crucial for the invasiveness of *Ras scrib* tumours in *Drosophila* (Uhlirova and Bohmann 2006). A previous study has shown that Yki also regulates *Mmp1* in wing discs, indicating that this is a target gene of Yki in both eye and wing discs (X. Ma et al. 2015). Interestingly, in colorectal cancer cell lines, YAP induces the expression of the human orthologue of *Drosophila Mmp1*, called *MMP-7*, in response to substrate stiffness, resulting in enhanced cell proliferation (Nukuda et al. 2015). Additionally, Fos and Jun can also activate the promoter of *MMP-7* (Shi et al. 2010). This suggests that Yki and AP-1 regulation of matrix metalloproteinase genes is conserved between *Drosophila* and mammals. My results highlight that Yki, Sd, and AP-1 may co-regulate *Mmp* genes in the *Drosophila* eye disc, which requires further investigation.

5.3.2 Hippo pathway target gene regulation by AP-1

I found that there was the strong enrichment of Jra at several Hippo pathway target genes, including the well-defined targets *Diap-1*, *ex*, *Cyclin-E*, and *ban*. Predicted Sd and AP-1 binding sites were located throughout these gene loci. This raises the possibility that AP-1 might regulate Hippo pathway target genes, either in cooperation with or independently of Yki and Sd. However, genetic studies showed that the loss of *kay* had no effect on the Hippo pathway target genes *Diap1* and *ex* in the *Drosophila* eye disc. For example, the loss of *kay* did not rescue elevated *ex* expression in *sav* mutant tissue, and the combined loss of *sd* and *kay* did not affect Diap1 levels in the eye disc. In the future, it will be interesting to examine a potential role for Kay in regulating other Hippo pathway target genes, such as the anti-apoptotic microRNA *ban* (Brennecke et al. 2003). Indeed, it was recently shown that a minimal enhancer of *ban* is regulated by Yki and AP-1 in the *Drosophila* wing disc (Sumabat et al. 2019). In normal growth conditions, CtBP restricts *ban* expression by antagonising AP-1 and Sd, and therefore limiting tissue growth, and upon loss of the CtBP transcriptional co-repressor *ban* expression was activated (Brennecke et al. 2003). It will be interesting to investigate whether this same mechanism operates in the *Drosophila* eye disc, and whether this occurs at multiple Hippo pathway target genes.

5.3.3 AP-1 regulation of eye growth

Given that Jra, Yki and Sd share thousands of target genes, which contain both Sd and AP-1 binding sites, I posited that growth is cooperatively regulated by AP-1 and Yki/Sd which converge on a set of common target genes. Depletion of *kay* using RNAi suppressed Yki induced eye overgrowth and also synergistically reduced tissue growth in combination with *sd* depletion. These results pointed to the possibility that both Sd and Kay are required for normal eye growth, and each protein can compensate and maintain eye growth in the absence of the other. This was a fascinating result that could explain the fact that loss of Sd has no effect on growth or Hippo pathway target genes (Koontz et al., 2013). However, studies using a mutant *kay* allele revealed no obvious role for Kay in synergistically regulating target genes and

growth with Yki and Sd. These conflicting results can possibly be explained by the following: 1) the *kay* RNAi lines had off target effects which resulted in false-positive results; or 2) the kay^{KM} allele produced a functional Kay protein; or 3) the loss of *kay* is substituted in this context by Jra homodimers. In the future, it will be important to confirm if the kay^{KM} allele indeed is a loss of function allele. For example, I can determine the size of the Kay protein generated by this allele to confirm that it is truncated. Additionally, I could verify whether the kay^{KM} allele gives rise to known *kay* mutant phenotypes, such as embryonic dorsal closure defects (Ciapponi and Bohmann 2002; Riesgo-Escovar and Hafen 1997b; Zeitlinger et al. 1997). It will also be critical to determine whether kay^{KM} has lost the ability to regulate its target genes, which could be confirmed by observing decreased expression of an in vivo AP-1 sensor construct (e.g. TRE-GFP) (Chatterjee and Bohmann 2012). Additionally, as Kay and Jra complex as heterodimers (Perkins et al. 1990), it will be interesting to investigate if Jra plays a role in synergistically regulating growth with Yki and Sd, which can be assessed using a mutant *jra* allele.

It was recently proposed that different *kay* isoforms can determine whether AP-1 acts as either a gene activator or repressor at the *ban* enhancer (Sumabat et al. 2019). Knocking down all isoforms of *kay* resulted in no effect on *ban* expression, while overexpression of isoform D activated *ban* expression (Sumabat et al. 2019). The authors posited that knocking down all isoforms of *kay* averaged the opposing effects of the different isoforms, however they did not show that certain isoforms are capable of repressing *ban* (Sumabat et al. 2019). An additional example of distinct roles for *kay* isoforms includes the requirement of isoform D in regulating the circadian rhythm gene *clock* (Ling, Dubruille, and Emery 2012). Thus, it is possible that the kay^{KM} allele created in this thesis, which targets all isoforms, results in a balance of activation and repression of target genes. In the future it could be worth investigating the ability of different *kay* isoforms to activate Yki target genes in the *Drosophila* eye disc.

An interesting question that arises from my studies is that if AP-1 does not regulate eye growth with the Hippo pathway, then why do Jra, Yki and Sd bind to thousands of the same target genes? One possibility is that AP-1 and Yki/Sd only regulate their shared target genes in response to specific signals. For example, this regulation could occur in neoplastic tumours in *Drosophila* imaginal discs. In this context, several lines of evidence have linked JNK signalling and the Hippo pathway. Firstly, imaginal discs that are entirely mutant for *scrib* activate aPKC

and JNK activity, which in turn elevates Yki and Kay levels, respectively (Bunker et al. 2015). Yki and Kay converge on the target gene *upd3*, which drives the tumour overgrowth of *scrib* mutant tissues (Bunker et al. 2015). Secondly, in *scrib* mutant clones that are surrounded by wild-type tissue, JNK signalling inhibits Yki activity, resulting in the elimination of these clones (C. L. Chen et al. 2012). Given my data shows that Jra binds to many Yki/Sd target genes, it is possible that in the context of neoplasia Yki target genes are repressed by AP-1. Additionally, Ras hyperactivity prevents the cell death of *scrib* mutant cells, resulting in massive invasive tumours and increased levels of Yki target genes, such as *ex* and *fj* (Doggett et al. 2011; Enomoto et al. 2015). In this context, F-actin is accumulated through the activity of the LIM domain protein Ajuba, resulting in the inhibition of Wts and thereby Yki activation (Enomoto et al. 2015). Interestingly, knockdown of Yki in Ras *scrib* clones does not completely rescue tumour overgrowth, indicating that additional mechanisms are involved in promoting tumorigenesis in this setting (Doggett et al. 2011). Therefore, future studies should address whether AP-1 is involved in regulating Yki target genes and tumour growth in Ras *scrib* clones.

5.3.4 Summary

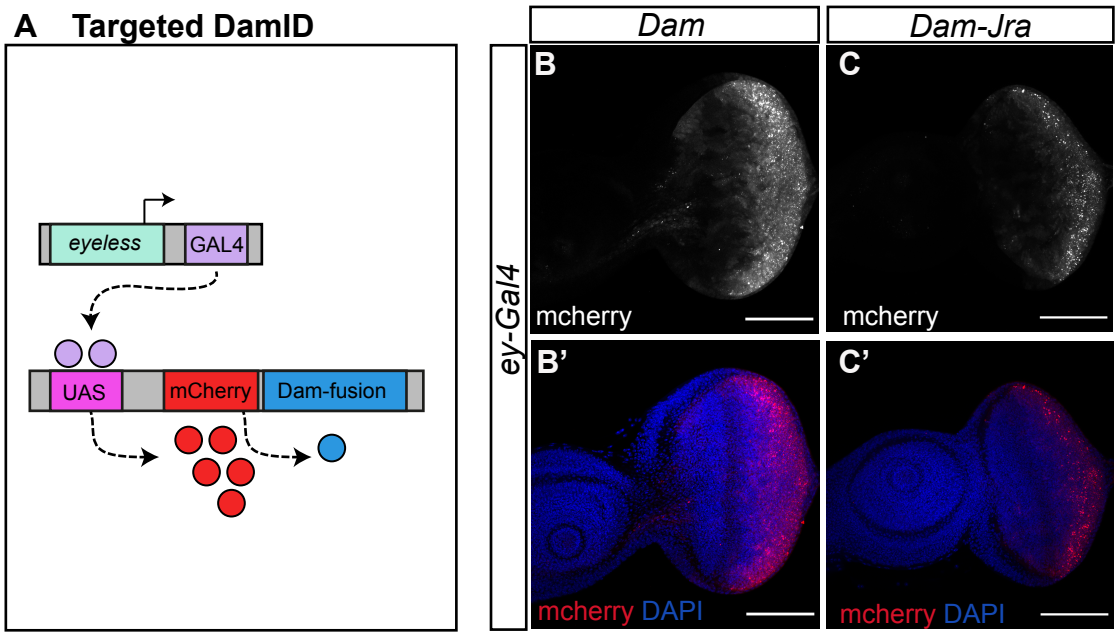
As the loss of Sd has no effect on eye growth and Hippo pathway target gene expression (Koontz et al., 2013), additional proteins must control growth and target gene expression. To understand this, I investigated the role of AP-1 in eye growth and Hippo pathway target gene expression. I found that Jra, Yki and Sd share thousands of putative target genes. However, I found conflicting results for AP-1 in controlling eye growth and found no evidence for AP-1 in regulating Hippo pathway target genes under normal growth conditions. Given these contradictory results, it will be important to carry out additional experiments, including further validation of the mutant *kay* allele generated in this thesis and the investigation of further shared target genes. Furthermore, given the high percentage of shared Jra, Yki and Sd target genes, future efforts should aim to understand the specific cellular contexts under which these genes are regulated by these proteins, for example in neoplastic tumours. Such studies will be important in understanding the role that the Hippo pathway and AP-1 play in regulating organ growth and gene expression. Importantly, these studies could also give further insight into how AP-1 and the Hippo pathway cooperate in tumorigenesis.

Figure 5.1: Targeted DamID overview and the expression of Jra-Dam in the developing Drosophila eye disc.

A. Schematic representation of targeted DamID. Transcriptional regulatory proteins of interest were expressed as a fusion to the bacterial DNA adenine methyltransferase (Dam). Dam-fusion proteins were expressed in specific cells using the UAS-Gal4 system. Gal4 is a yeast derived activator protein that binds to an upstream activator sequence (UAS) to activate gene transcription. Dam-fusion proteins were expressed in eye tissue using an eye-specific Gal4 that is under the control of the *eyeless* enhancer. Translation of the Dam-fusion protein is low due to the fact that it follows a primary open reading frame (ORF) that codes for mCherry. As such, only low rates of translation initiation occur at the secondary ORF encoding the Dam-fusion protein.

B-C'. Confocal microscope images of Dam-fusion protein expression in developing *Drosophila* eye discs. Images are of late third instar larval eye-antennal discs, anterior is to the left. Dam-fusion proteins were expressed using the *eyeless-Gal4* driver. mCherry (greyscale in A & B, and red in A' & B') was used as a surrogate marker of Dam-fusion protein expression and is visible in the posterior region of the eye disc. DAPI (blue) indicates nuclei (A' & B'). Scale bars = 100μm.

D. Multidimensional scaling (MDS) scatterplot representation of Principle Components Analysis (PCA) for DamID-seq samples. DamID-seq samples were plotted in two dimensions so that distances on the plot approximate differences between samples based on log2 fold changes. The x and y axes show the leading LogFC dimension 1 and 2 respectively. The MDS plot displays each DamID-seq sequencing batch (shown as triangles, boxes, or circles), which are samples that were sequenced in separate sequencing runs, and the DamID-seq genotype (shown as different colours).


A Targeted DamID
eyeless
GAL4
UAS
mCherry
Dam-fusion
Dam
Dam-Jra
ey-Gal4
B
C
mcherry
mcherry
B'
C'
mcherry DAPI
mcherry DAPI

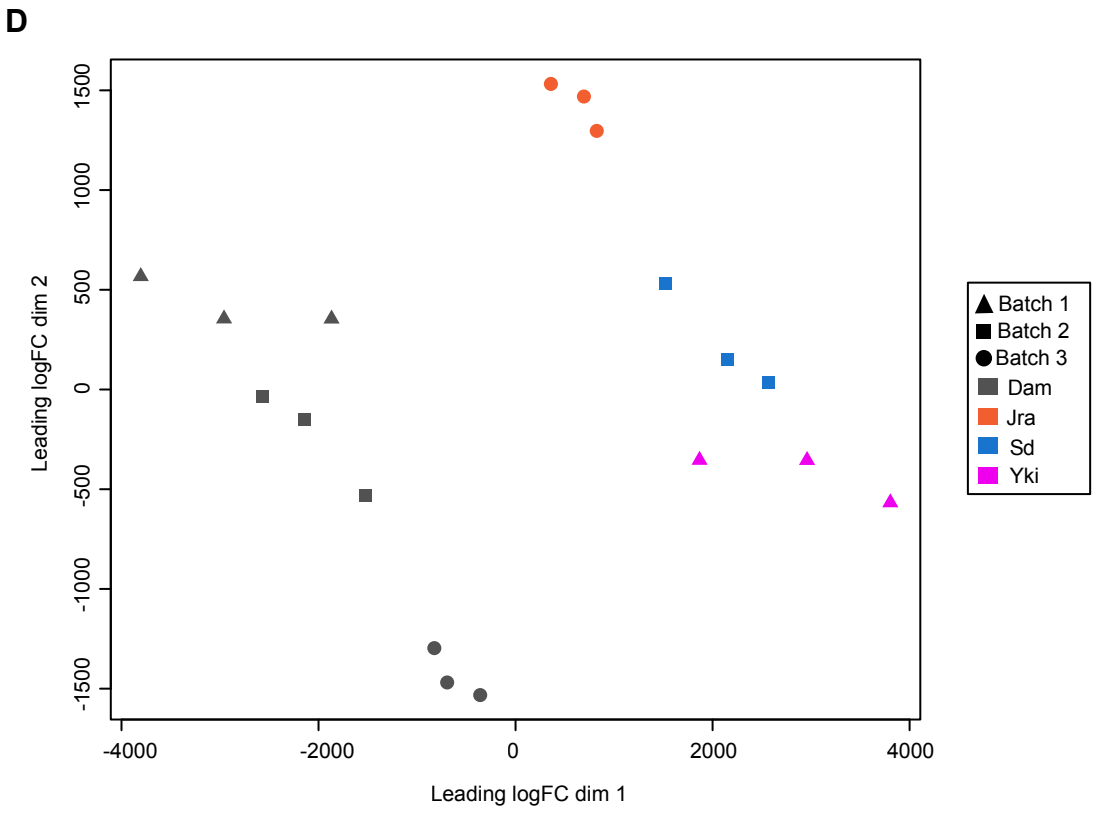

D
Leading logFC dim 2
1500
1000
500
0
-500
-1000
-1500
-4000
-2000
0
2000
4000
Leading logFC dim 1
Batch 1
Batch 2
Batch 3
Dam
Jra
Sd
Yki

Figure 5.2: Jra, Yorkie and Scalloped co-occupy thousands of gene loci in developing *Drosophila* eye discs.

A-B. Venn diagrams showing the overlap of peaks (A) and the overlap of target genes (B) bound by Jra, Yki and Sd. DamID-seq peaks are regions that contain multiple significantly

methylated GATC 'tags', and the putative target genes that are assigned to these peaks based on proximity. The significance of overlapping target genes ($p<0.000$) was assessed by hypergeometric probability analysis.

C-E. Pairwise correlation analyses of the magnitude of peak height shared by Jra, Yki, and Sd, expressed as a Log fold change (LogFC). The LogFC for a peak was calculated by normalising the Dam-fusion protein to the Dam-alone control. Correlations between the LogFC of shared peaks of Jra and Yki are shown in (C), Jra and Sd in (D), and Yki and Sd in (E). The Pearson score (shown

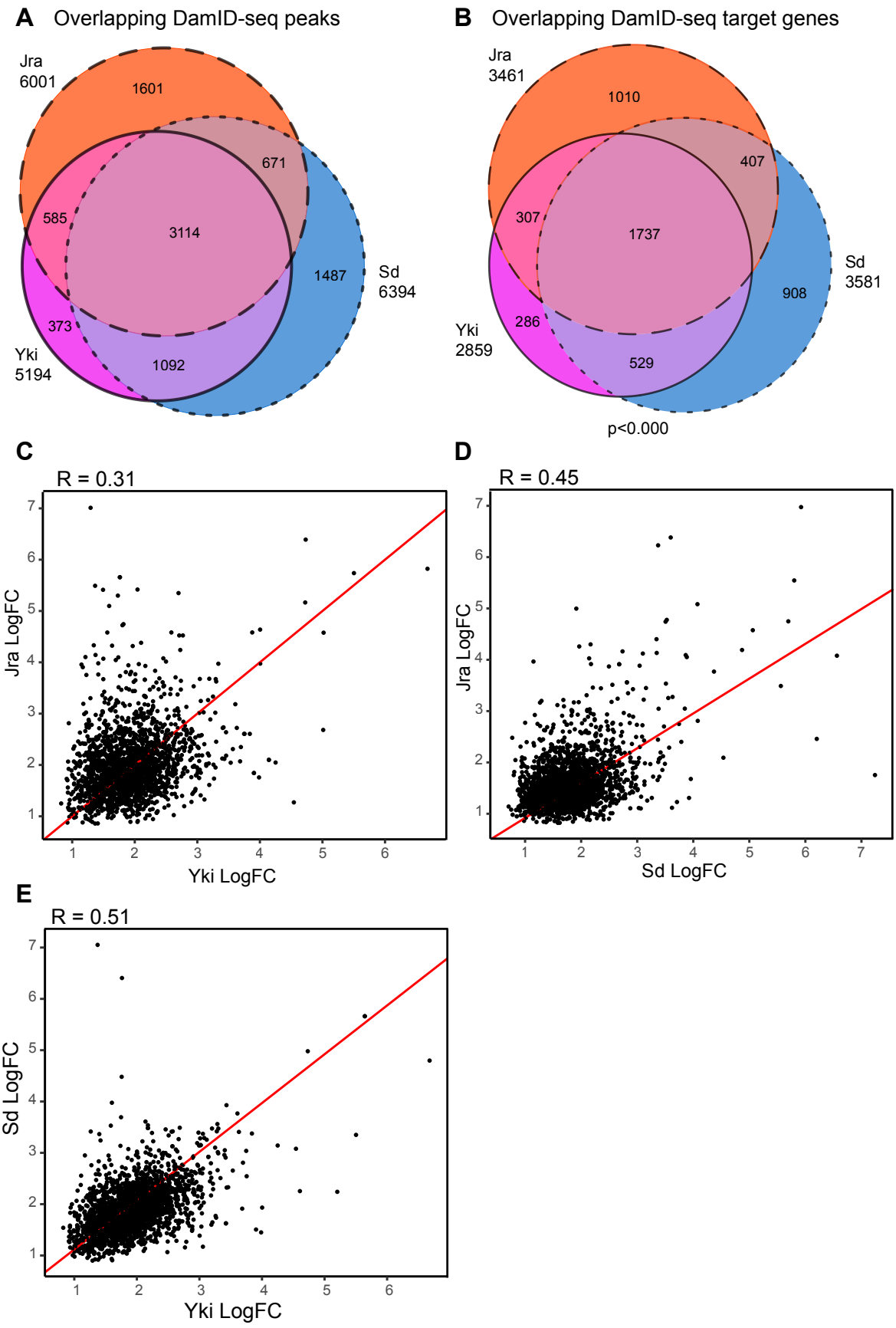

A Overlapping DamID-seq peaks
Jra
6001
1601
671
585
3114
1487
Sd
6394
373
1092
Yki
5194
B Overlapping DamID-seq target genes
Jra
3461
1010
407
307
1737
908
Sd
3581
286
529
Yki
2859
p<0.000
C
R = 0.31
Jra LogFC
Yki LogFC
D
R = 0.45
Jra LogFC
Sd LogFC
E
R = 0.51
Sd LogFC
Yki LogFC

Figure 5.3: Jra DamID peaks are enriched near the transcriptional start site (TSS) of genes.

A. Bar graph comparing the percentage of Dam-fusion protein peaks within the following annotated regions of the *Drosophila* genome (dm6): 3'UTR, 5'UTR, intron, exon, and intergenic. The percentage of the *Drosophila* genome divided into the annotated regions, and the distribution of GATC 'tags' is shown, followed by the percentage of Dam-Jra peaks within these genomic features.

B. The distance of GATC 'tags' and Dam-Jra peaks from the nearest transcriptional start site (TSS). The distribution of GATC 'tags' and Dam-Jra peaks are shown within a 3 kilobase (kb) window upstream and downstream from the TSS's of annotated genes. The y-axis shows the read count frequency, while the x-axis shows the distance from the TSS in kilobases. Confidence intervals are shown as a shaded band.

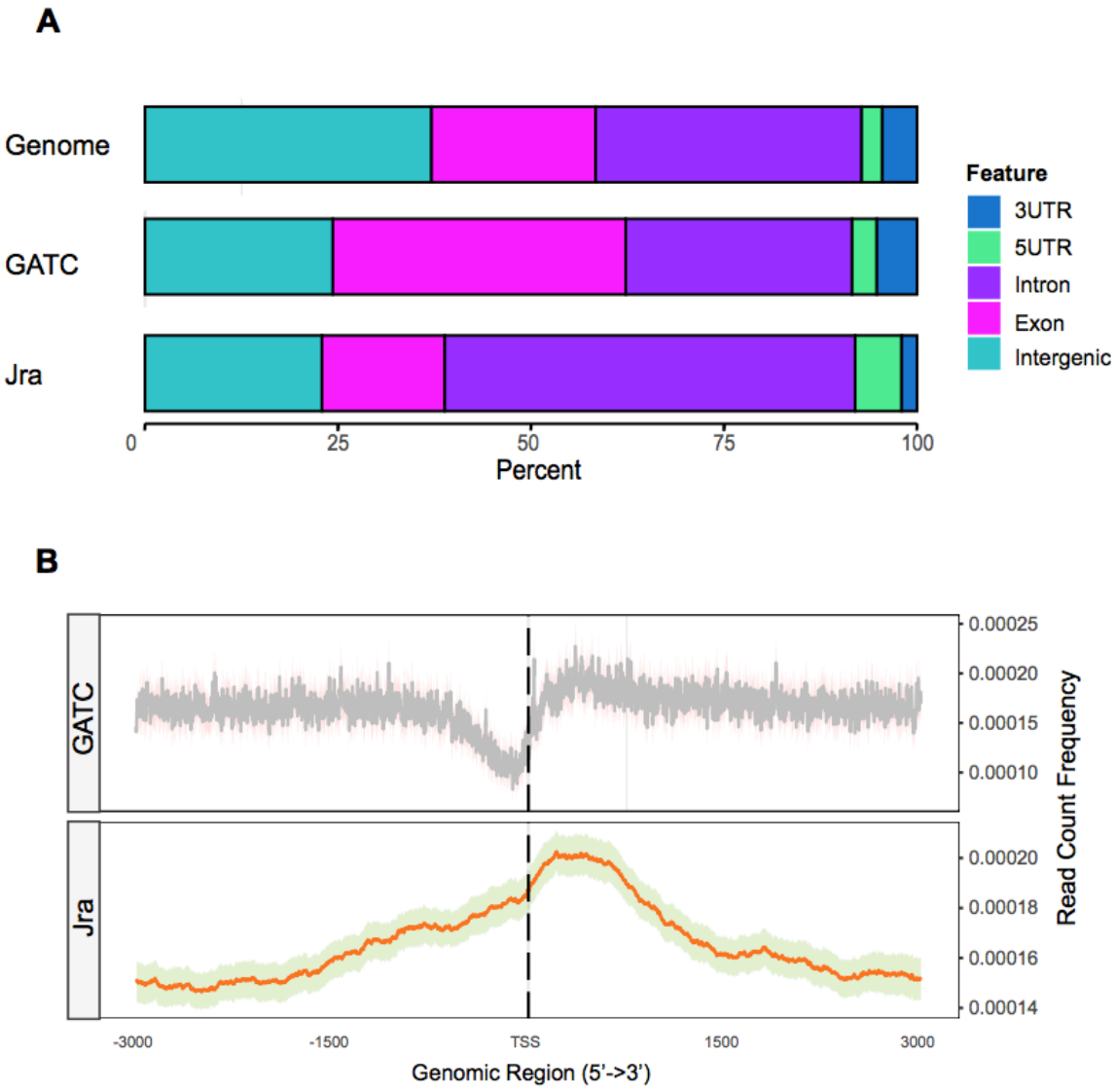
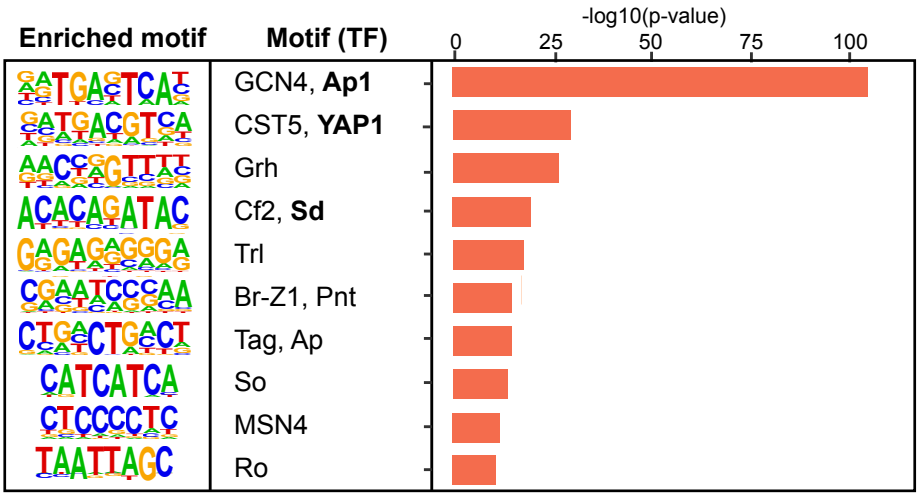

A
Genome
GATC
Jra
0
25
50
75
100
Percent
Feature
3UTR
5UTR
Intron
Exon
Intergenic
B
GATC
Jra
0.00025
0.00020
0.00015
0.00010
0.00020
0.00018
0.00016
0.00014
Read Count Frequency
-3000
-1500
TSS
1500
3000
Genomic Region (5'->3')

Figure 5.4: Scalloped and AP-1 motifs are enriched in Jra, Yorkie, and Scalloped binding sites in the developing *Drosophila* eye disc.

A-B. The top 10 enriched transcription factor motifs in genomic regions bound by Jra (A) or co-bound by Jra, Yki and Sd (B). *De novo* motif enrichment was implemented using HOMER and completed by resampling 100 times using permutation analysis in order to meet a False Discovery Rate of 0.01. Motifs shown in the table were identified using HOMER analysis, and the significance associated with each motif is represented by -log10 (p-value).

A **Enriched mofis in Jra peaks**



B **Enriched mofis in shared Jra, Yki, and Sd peaks**

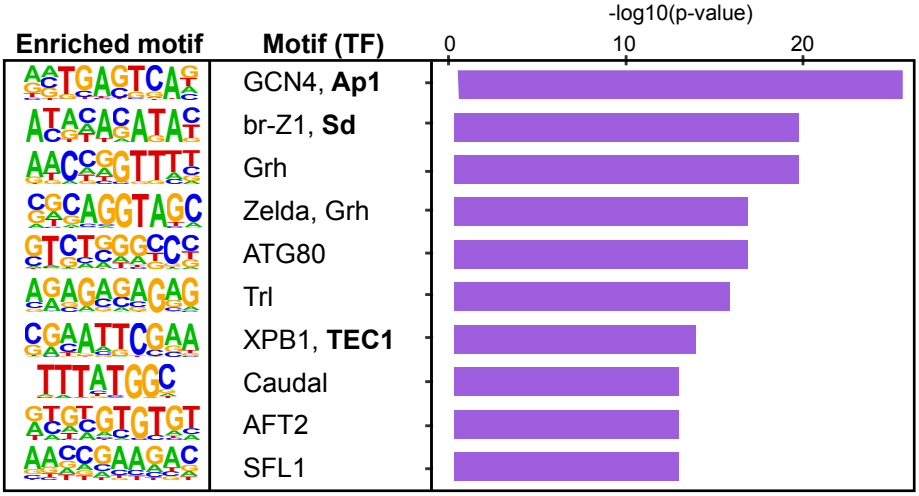


Figure 5.5: Signalling pathways enriched among putative Jra, Yorkie, and Scalloped target genes in the developing *Drosophila* eye disc.

A bubble chart showing enrichment of KEGG pathways performed on putative target genes of Jra, Yki, and Sd that were identified using targeted DamID in the developing *Drosophila* eye disc. The y-axis represents KEGG pathway and the x-axis represents Jra, Yki, and Sd. The KEGG pathways listed are statistically significant according to a p value of <0.05 to <0.001. The colour of each dot represents the significance of each pathway: higher significance = red, medium significance = purple, lower significance = blue, while the size of the dot represents the gene ratio (the number of enriched genes/the number of genes in the KEGG category).

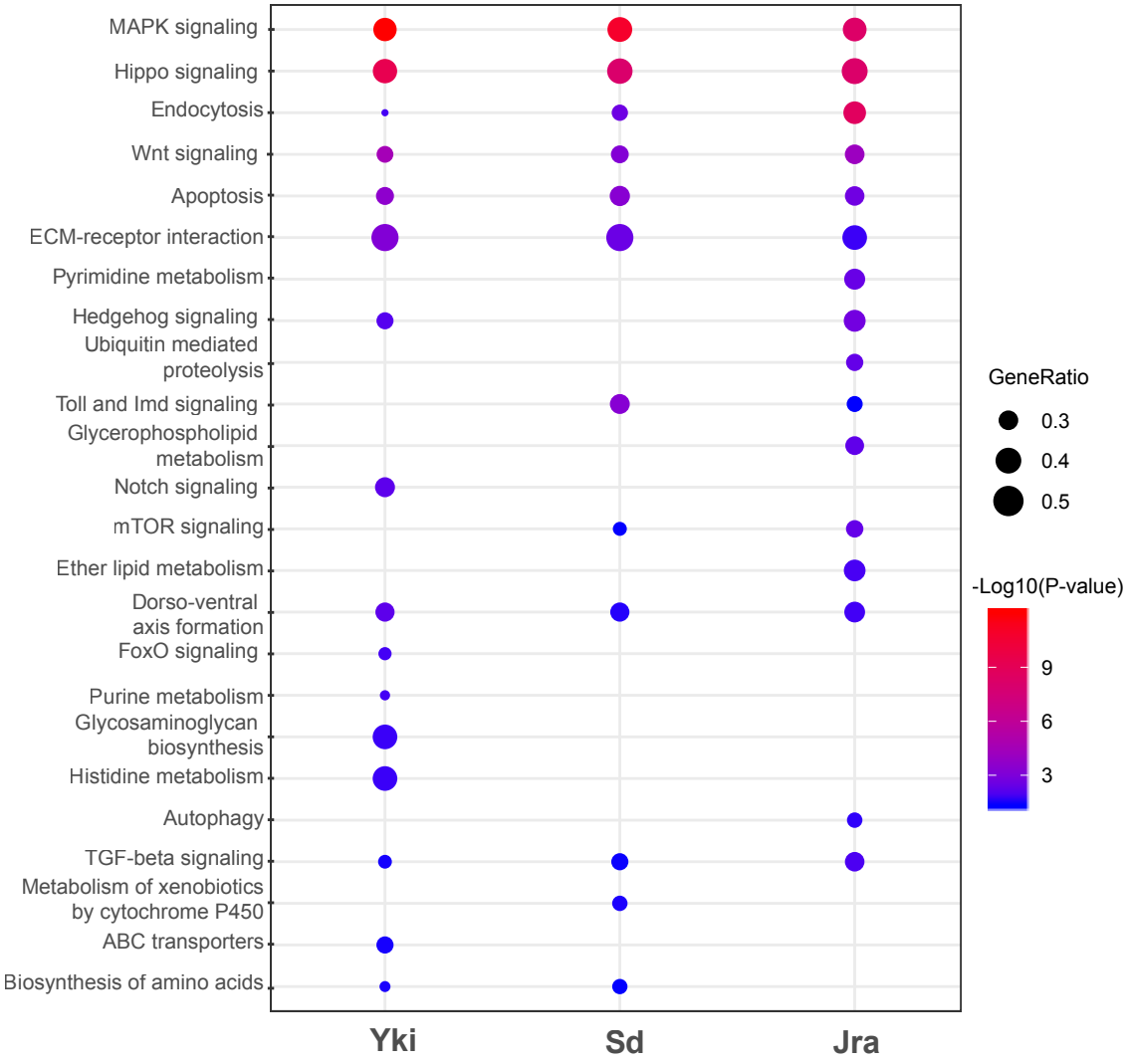

MAPK signaling
Hippo signaling
Endocytosis
Wnt signaling
Apoptosis
ECM-receptor interaction
Pyrimidine metabolism
Hedgehog signaling
Ubiquitin mediated proteolysis
Toll and Imd signaling
Glycerophospholipid metabolism
Notch signaling
mTOR signaling
Ether lipid metabolism
Dorso-ventral axis formation
FoxO signaling
Purine metabolism
Glycosaminoglycan biosynthesis
Histidine metabolism
Autophagy
TGF-beta signaling
Metabolism of xenobiotics by cytochrome P450
ABC transporters
Biosynthesis of amino acids
Yki
Sd
Jra
GeneRatio
0.3
0.4
0.5
-Log10(P-value)
9
6
3

Figure 5.6: Targeted DamID gene profiles reveal binding of Jra, Yorkie, and Scalloped at known JNK target genes.

A-D. Gene binding profiles showing adenine methylaton by Jra- (orange), Yki- (pink), and Sd- (blue) Dam-fusion proteins for *puc* (A), *msn* (B), *cher* (C), and *upd1* (D). The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the methylation for the Dam-fusion normalised to the Dam-alone control. Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The bars in the top of each panel represent significant peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of *puc*, *msn*, *cher*, and *upd1* gene loci and surrounding genes. Sd consensus motifs are shown in blue and AP-1 consensus motifs are shown in red. Location of Sd and AP-1 consensus motifs in the upd3.3 enhancer (D') (Bunker et al. 2015). Scale bars in kb are shown.

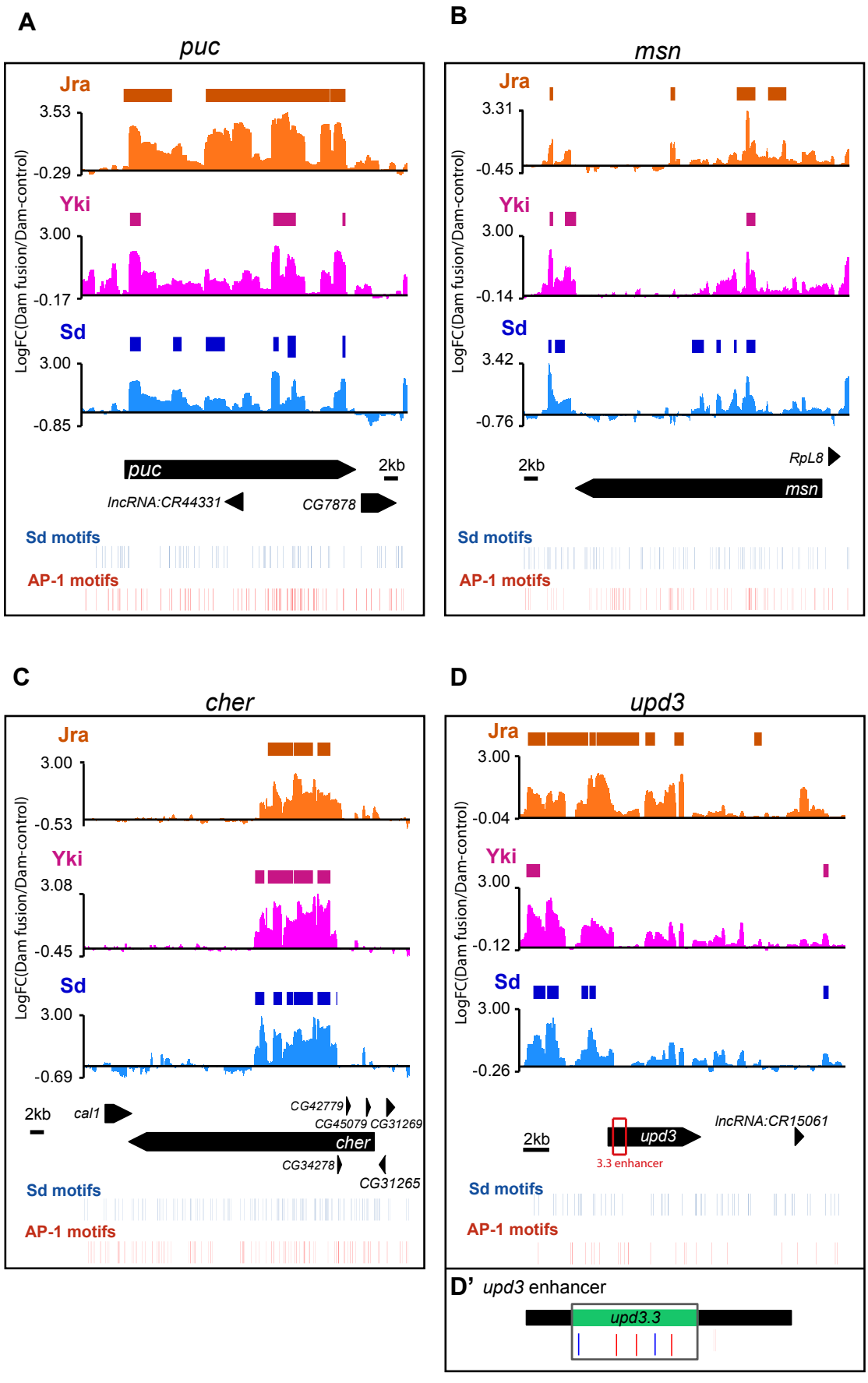

A
puc
Jra
3.53
-0.29
Yki
3.00
-0.17
Sd
3.00
-0.85
LogFC(Dam fusion/Dam-control)
puc
2kb
lncRNA:CR44331
CG7878
Sd motifs
AP-1 motifs
B
msn
Jra
3.31
-0.45
Yki
3.00
-0.14
Sd
3.42
-0.76
LogFC(Dam fusion/Dam-control)
2kb
RpL8
msn
Sd motifs
AP-1 motifs
C
cher
Jra
3.00
-0.53
Yki
3.08
-0.45
Sd
3.00
-0.69
LogFC(Dam fusion/Dam-control)
2kb
cal1
CG42779
CG45079 CG31269
cher
CG34278
CG31265
Sd motifs
AP-1 motifs
D
upd3
Jra
3.00
-0.04
Yki
3.00
-0.12
Sd
3.00
-0.26
LogFC(Dam fusion/Dam-control)
2kb
upd3
lncRNA:CR15061
3.3 enhancer
Sd motifs
AP-1 motifs
D' upd3 enhancer
upd3.3

Figure 5.7: Targeted DamID gene profiles reveal binding of Jra, Yorkie, and Scalloped at known Hippo pathway target genes.

A-D. Gene binding profiles showing adenine methylaton by Jra- (orange), Yki- (pink), and Sd- (blue) Dam-fusion proteins for *Diap1* (A), *ex* (B), *CycE* (C), and *ban* (D). The gene binding profiles for each Dam-fusion protein are shown on the y-axis and represent the Log2 of the methylation for the Dam-fusion normalised to the Dam-alone control. Binding profiles for each Dam-fusion protein represent differential methylation, where GATC sequences were methylated higher by the Dam-fusion protein compared to the Dam-alone control. The bars in the top of each panel represent significant peaks for each Dam-fusion protein, which are genomic regions that contain multiple significantly methylated GATC sites. The x-axis shows the genomic location of the gene binding profiles and peaks. The bottom panels show the gene view of *Diap1*, *ex*, *CycE*, and *ban* gene loci and surrounding genes. Sd consensus motifs are shown in blue and AP-1 consensus motifs are shown in red. Location of AP-1 and Sd consensus motifs in the *Diap1* 1.8kb enhancer (A') (L. Zhang, Ren, Zhang, Chen, Wang, and Jian 2008). Scale bars in kb are shown.

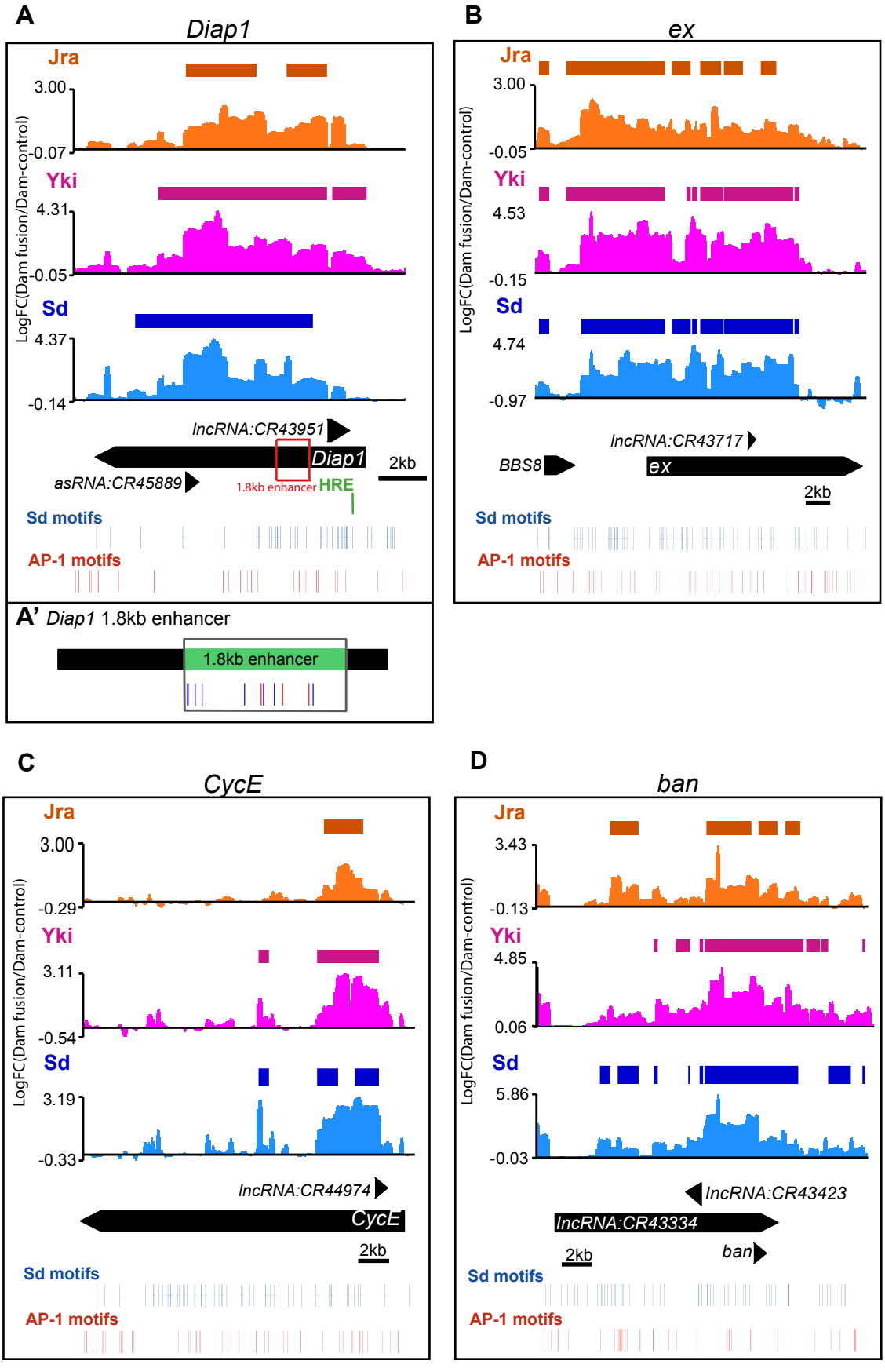
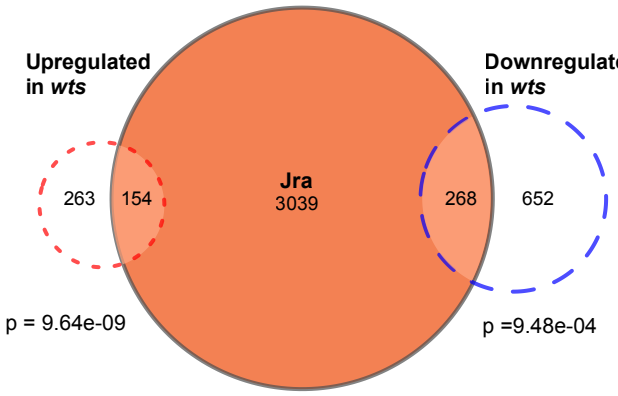

A
Diap1
Jra
3.00
-0.07
Yki
4.31
-0.05
Sd
4.37
-0.14
LogFC(Dam fusion/Dam-control)
lncRNA:CR43951
Diap1
2kb
asRNA:CR45889
1.8kb enhancer
HRE
Sd motifs
AP-1 motifs
A' Diap1 1.8kb enhancer
1.8kb enhancer
B
ex
Jra
3.00
-0.05
Yki
4.53
-0.15
Sd
4.74
-0.97
LogFC(Dam fusion/Dam-control)
lncRNA:CR43717
BBS8
ex
2kb
Sd motifs
AP-1 motifs
C
CycE
Jra
3.00
-0.29
Yki
3.11
-0.54
Sd
3.19
-0.33
LogFC(Dam fusion/Dam-control)
lncRNA:CR44974
CycE
2kb
Sd motifs
AP-1 motifs
D
ban
Jra
3.43
-0.13
Yki
4.85
0.06
Sd
5.86
-0.03
LogFC(Dam fusion/Dam-control)
lncRNA:CR43423
lncRNA:CR43334
2kb
ban
Sd motifs
AP-1 motifs

Figure 5.8: Expression of Jra target genes in *wts* mutant developing *Drosophila* eye discs.

A. A Venn diagram showing Dam-Jra candidate genes whose expression was upregulated or downregulated in hyperplastic (*wts^{X1}/wtsLacZ*) eye discs. The significance (p value) of overlapping target genes was assessed by hypergeometric probability analysis.

B. A heatmap showing gene expression levels for known JNK pathway target genes in normal (*FRT82B*) (purple) and hyperplastic (*wts^{X1}/wtsLacZ*) (green) eye discs. The colour and intensity of each box indicates the level of change in gene expression for each gene in hyperplastic compared to normal growth. Blue indicates genes with reduced expression (downregulated) and red indicates genes with increased expression (upregulated).

A


Upregulated
in wts
Downregulated
in wts
263
154
Jra
3039
268
652
p = 9.64e-09
p =9.48e-04

B

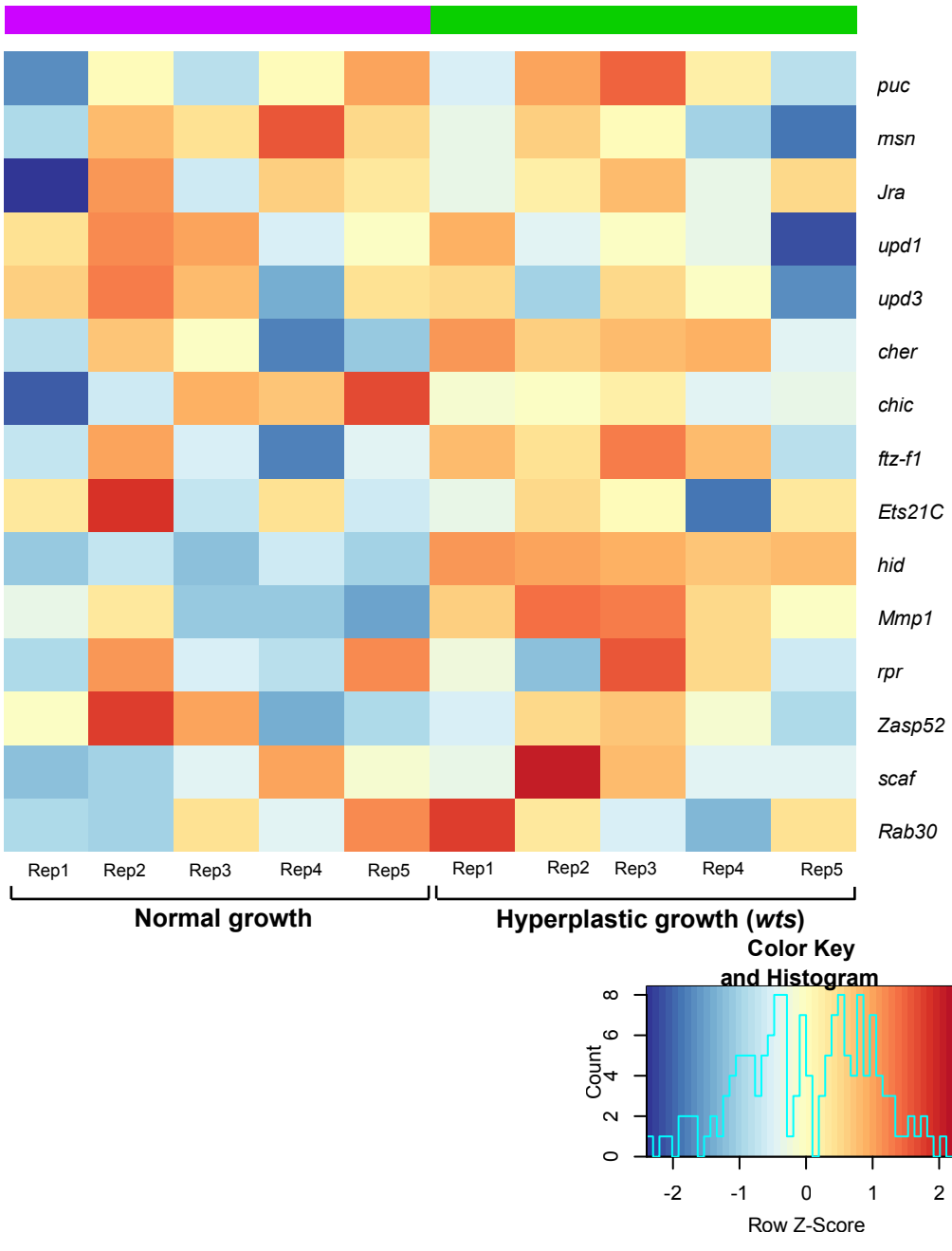

puc
msn
Jra
upd1
upd3
cher
chic
ftz-f1
Ets21C
hid
Mmp1
rpr
Zasp52
scaf
Rab30
Rep1
Rep2
Rep3
Rep4
Rep5
Rep1
Rep2
Rep3
Rep4
Rep5
Normal growth
Hyperplastic growth (wts)
Color Key
and Histogram
Count
0
2
4
6
8
-2
-1
0
1
2
Row Z-Score

Figure 5.9: Kayak depletion suppresses Yorkie-induced eye overgrowth and Kayak and Scalloped depletion synergistically reduces eye growth.

Images of female adult *Drosophila* heads in which *GMR-Gal4* drives the expression of *kayak* UAS-RNAi lines (A-C'), *UAS-yki*^{S168A} and *kayak* UAS-RNAi lines (D-F'), or *UAS-sd RNAi* and *kayak* UAS-RNAi lines in the eye. A-I shows the lateral view of adult heads, while A'-I' shows the anterior view.

A-C'. Control retinas expressing *GMR-Gal4* crossed to *βgalRNAi* (A, A'), *kay RNAi VDRC#27722* (B, B'), and *kay RNAi VDRC#6212* (C, C'). **D-F'**. *GMR-Gal4 UAS-yki*^{S168A} crossed to *βgalRNAi* (D,D'), *kay RNAi VDRC#27722* (E, E'), and *kay RNAi VDRC#6212* (F, F'). **G-I'**. *GMR-Gal4 UAS-sd RNAi* crossed to *βgalRNAi* (G,G'), *kay RNAi VDRC#27722* (H, H'), and *kay RNAi VDRC#6212* (I, I'). Scale bar is 25mm.

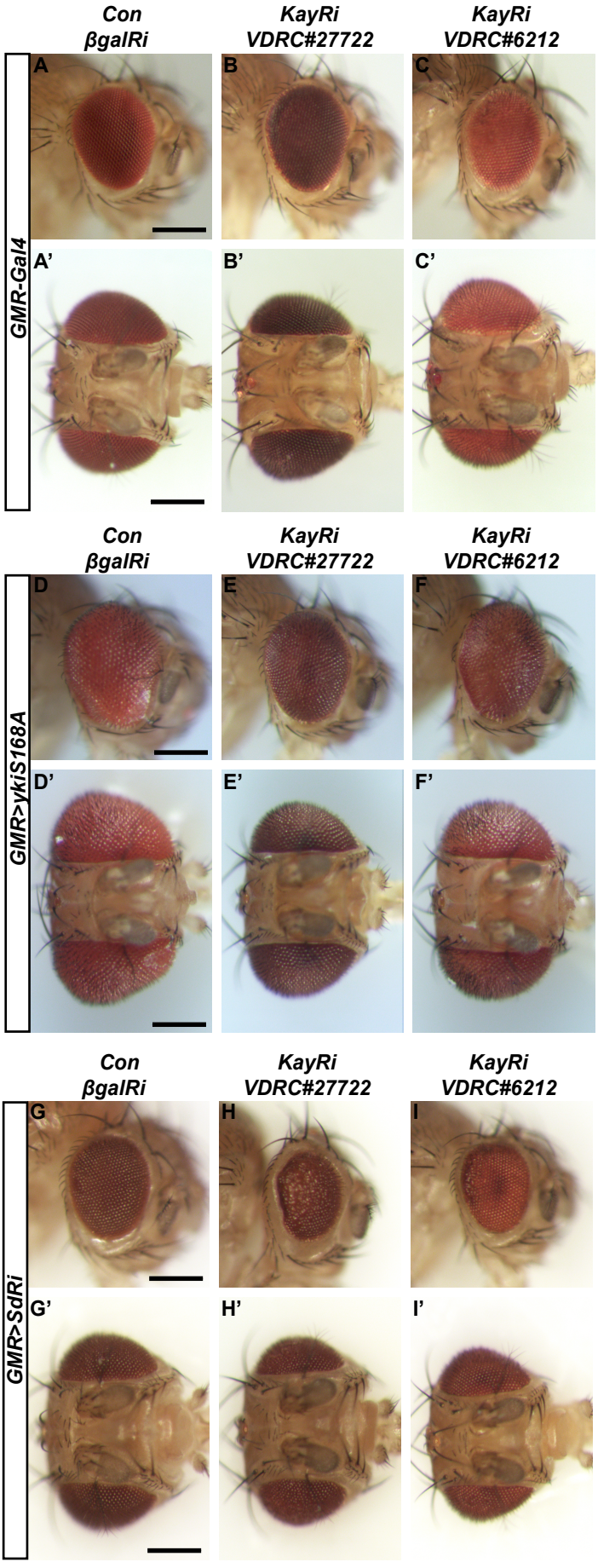

Con
βgalRi
KayRi
VDRC#27722
KayRi
VDRC#6212
GMR-Gal4
A
B
C
A'
B'
C'
Con
βgalRi
KayRi
VDRC#27722
KayRi
VDRC#6212
GMR>ykiS168A
D
E
F
D'
E'
F'
Con
βgalRi
KayRi
VDRC#27722
KayRi
VDRC#6212
GMR>SdRi
G
H
I
G'
H'
I'

Figure 5.10: Mutagenesis of the *kayak* gene by CRISPR Cas9 genome editing.

A. Schematic diagram of the *D. melanogaster kayak* locus. Exons from each *kay* isoform are shown as purple boxes and introns are represented as black lines. The location and sequence of the gRNA target and PAM are shown below.

B. PCR analysis of the *kay* target region separated on a 2% agarose gel. The wild type *kay* locus is shown in the first lane, followed by two mutant stocks that are heterozygous for the *kay* mutation.

C. Nucleotide sequence of the *kay* locus following CRISPR Cas9 genome editing. The wild-type sequence is shown at the top with the gRNA target sequence highlighted in red and the PAM highlighted in blue. The *kay* mutant sequence is shown below. Dashes indicate deleted nucleotides. Sequencing chromatograms of wild-type *kay* and the heterozygous mutant are shown in the two bottom panels. The deleted region is highlighted in yellow.

D. Schematic diagram of wild type (WT) and Kay mutant proteins. The target site of the gRNA is shown in red, and the basic region and leucine zipper (bZIP) domain is shown in blue. The *kay* mutation results in a frameshift that induces a premature STOP codon and a truncated protein.

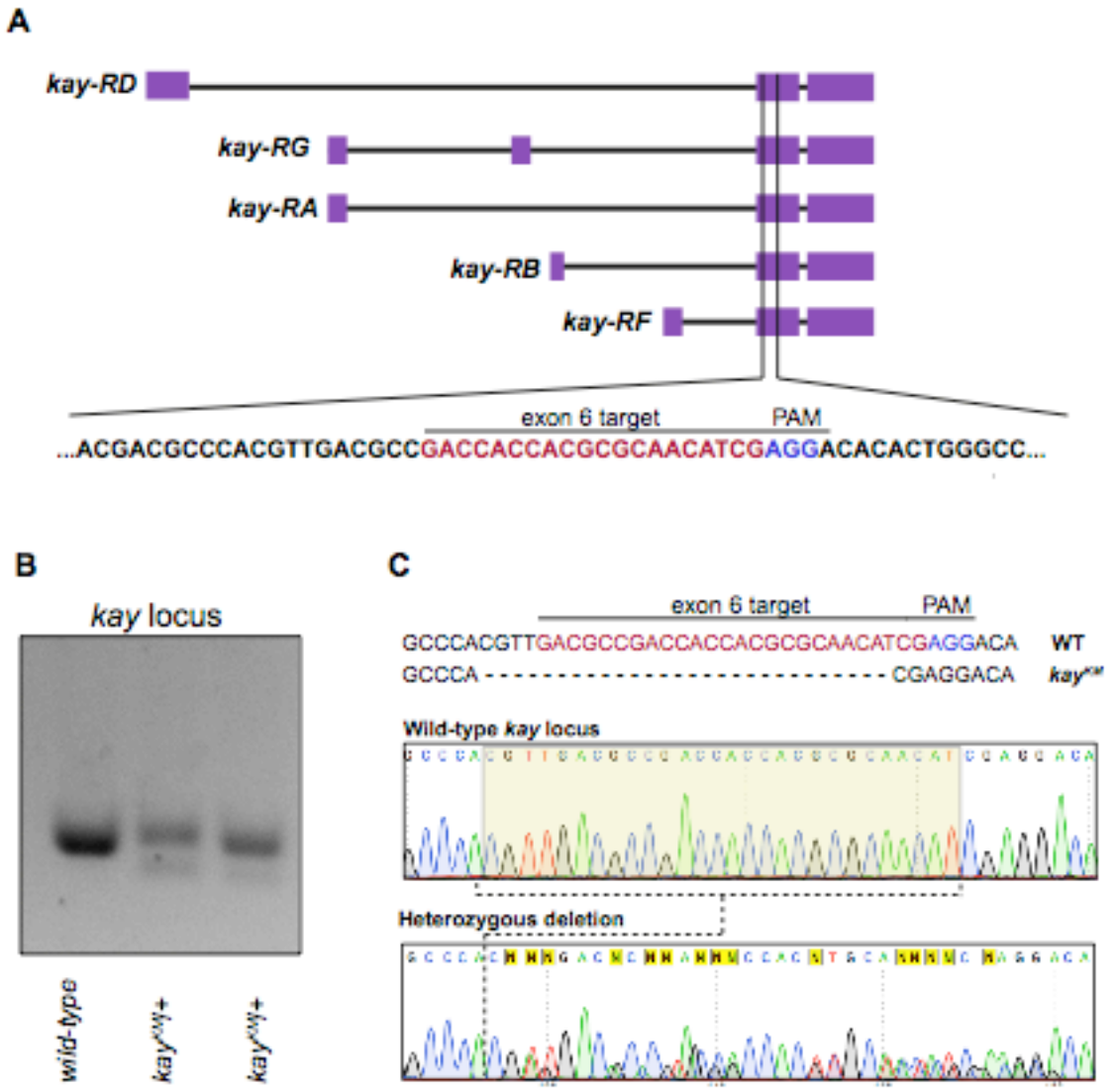
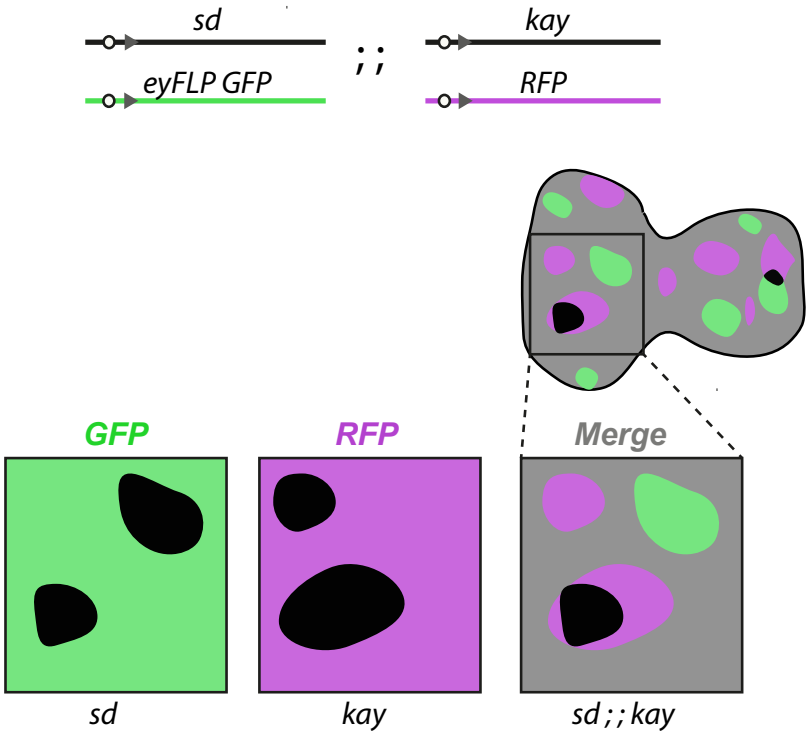

A
kay-RD
kay-RG
kay-RA
kay-RB
kay-RF
exon 6 target
PAM
...ACGACGCCCACGTTGACGCCGACCACCACGCGCAACATCGAGGACACACTGGGCC...
B
kay locus
wild-type
kay^KM+
kay^KM+
C
exon 6 target
PAM
GCCCACGTTGACGCCGACCACCACGCGCAACATCGAGGACA
WT
GCCCA - CGAGGACA
kay^KM
Wild-type kay locus
Heterozygous deletion
D
Target site
WT Kay
bZIP
Kay^KM
STOP
Predicted truncated protein

Figure 5.11: Schematic diagrams of scalloped-based genetic epistasis tests.

A. Double mutant clones can be generated in *Drosophila* eye discs by combining an eye-specific flippase (FLP) with double flippase recognition target (FRT) chromosomes containing *sd* and *kay* mutations in trans to double FRT chromosomes containing GFP and RFP markers. This creates an eye disc containing double heterozygous background clones (grey), *sd* single mutant clones (green), *kay* single mutant clones (magenta), and *sd;;kay* double-mutant clones (black). This schematic was adapted from Yu & Pan 2018.

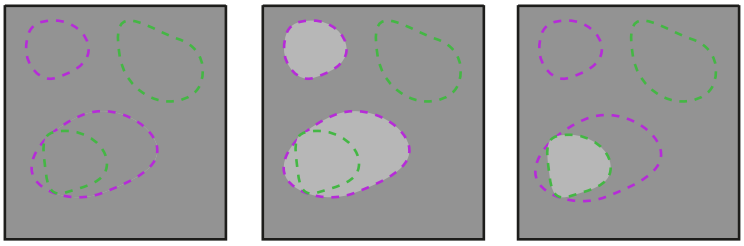
B. The role of Kay can be determined by analysing the expression of Hippo pathway target gene expression, for example *Diap1*, in *sd*-based epistasis experiments. The three possible results from these experiments are shown, which can determine if Kay regulates Hippo pathway target gene expression, and whether this is independent or dependent on Sd.

A



B

Target gene expression (e.g. *Diap-1*)



kay is not a regulator of target gene
Or
sd-dependent
Loss of kay has no effect on target gene expression

sd-independent
Loss of kay decreases target genes expression

sd-independent
Loss of both genes decreases target gene expression

Figure 5.12: Diap1 protein expression is not altered in *sd*, *kay* double mutant clones in the developing *Drosophila* eye disc.

A-C'''. Confocal micrographs of *Drosophila* larval eye-antennal discs from third instar larvae. *eyFLP* was used to create clones in the eye-antennal disc that contained the sd^{47M} mutant allele, kay^{KM} mutant allele, or the respective FRT controls. In all panels, mutant clones of sd^{47M} and kay^{KM} were marked by the loss of GFP (shown in green) and RFP (shown in magenta), respectively. Eye discs were stained for Diap1 protein (shown in greyscale) using a Diap1 antibody. Regions outlined by yellow dashed boxes are enlarged below. Clones that are double mutant for sd^{47M}, kay^{KM} or the controls are traced with a dashed yellow. Anterior is to the right in all panels. eyFLP induced clones containing sd^{47M} (GFP negative) and FRT82B (RFP negative) (A-A''), FRT19A (GFP negative) and kay^{KM} (RFP negative) (B-B'''), and sd^{47M} mutant (GFP negative) and kay^{KM} (RFP negative) (C-C'''). For A-B, n=4, and for C, n = 8. Scale bar is 100μm.

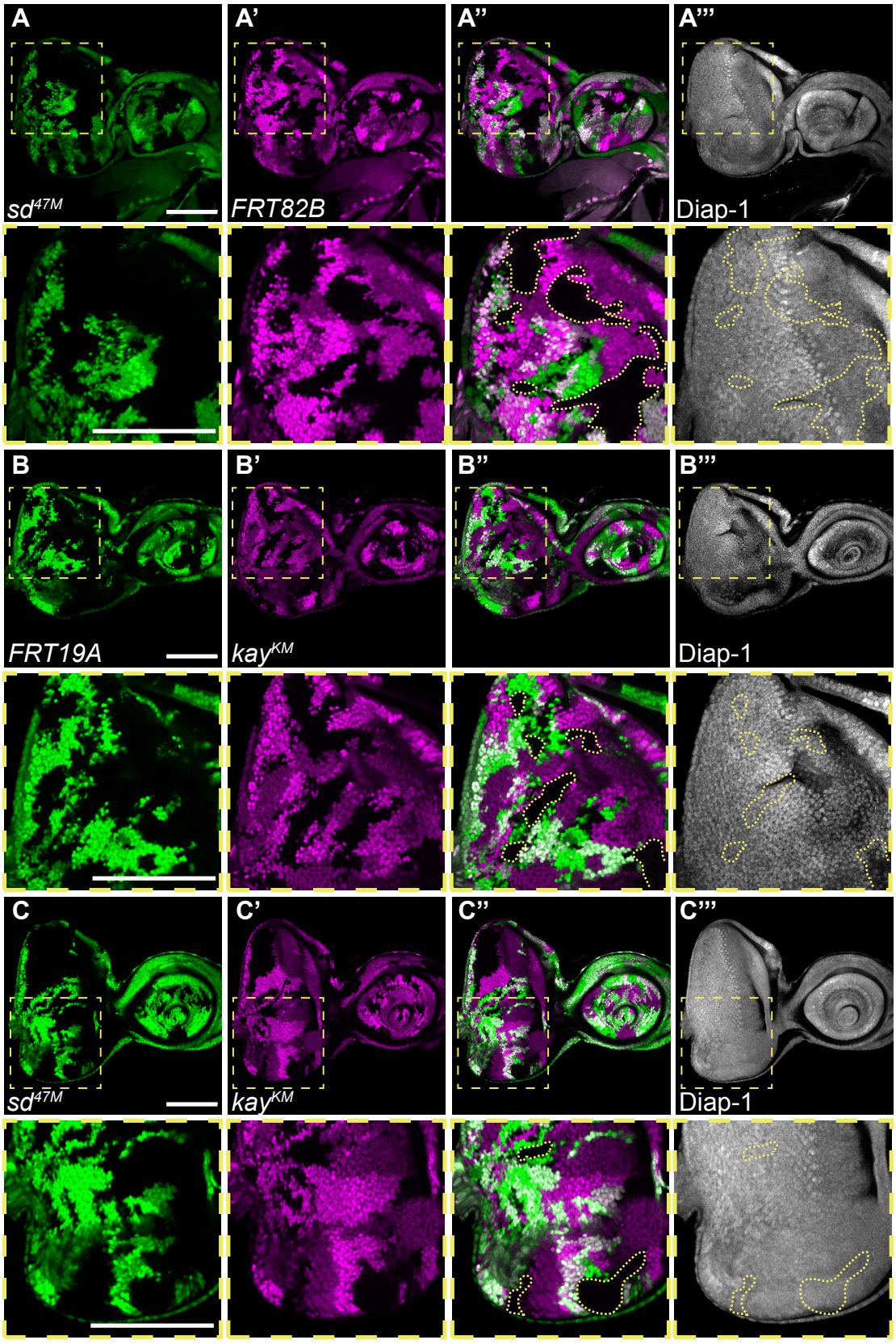

A
A'
A''
A'''
sd47M
FRT82B
Diap-1
B
B'
B''
B'''
FRT19A
kayKM
Diap-1
C
C'
C''
C'''
sd47M
kayKM
Diap-1

Figure 5.13: Loss of *scalloped* and *kayak* does not affect *Drosophila* adult eye size.

A-C'. Images of female adult *Drosophila* heads in which *eyFLP* was used to create clones in the eye-antennal disc that contained the sd^{47M} mutant allele, kay^{KM} mutant allele, or the respective FRT controls. In all panels, homozygous mutant clones are represented by white tissue, heterozygous clones are represented by orange tissue, and homozygous wild type clones are represented by red tissue. A-F shows the lateral view of the adult heads, while A'-F' shows the anterior view. eyFLP induced clones containing sd^{47M} and FRT82B (A-A''), FRT19A and kay^{KM} (B-B'''), and sd^{47M} and kay^{KM} (C-C'''). Scale bar is 25mm.

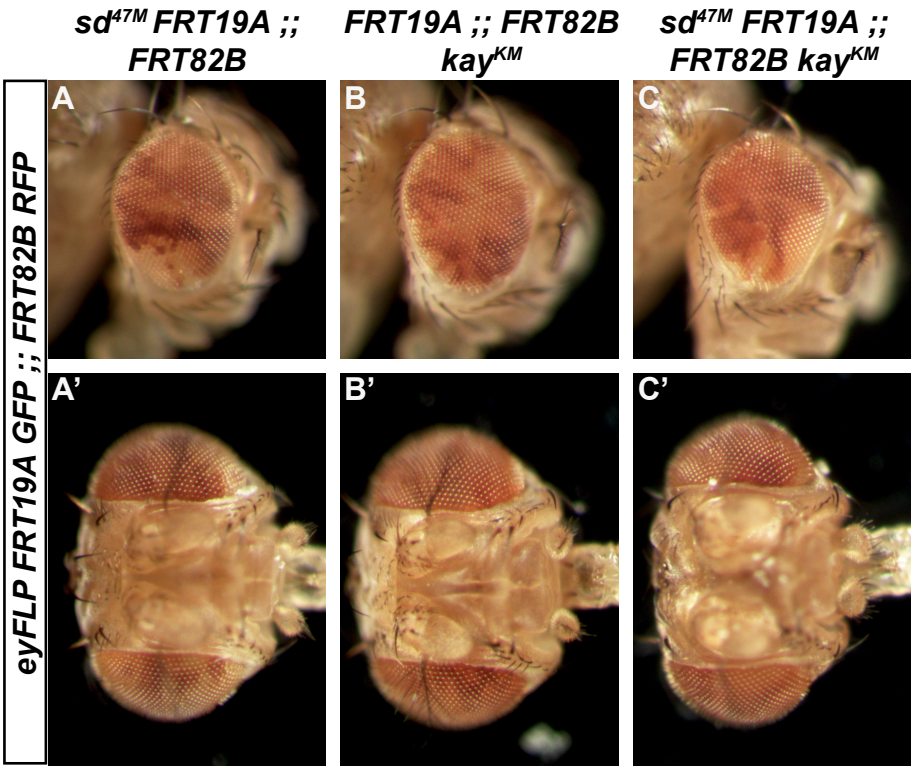

sd47M FRT19A ;; FRT82B
FRT19A ;; FRT82B kayKM
sd47M FRT19A ;; FRT82B kayKM
eyFLP FRT19A GFP ;; FRT82B RFP
A
B
C
A'
B'
C'

Figure 5.14: Loss of *kayak* does not suppress Yorkie-induced eye overgrowth.

A-C'. Images of female adult *Drosophila* heads in which *eyFLP* was used to create clones in the eye-antennal disc that contained the sav^{3} kay^{KM} double mutant or the respective controls. In all panels, homozygous mutant clones are represented by white tissue, heterozygous clones are represented by orange tissue, and homozygous wild type clones are represented by red tissue. A-C shows the lateral view of the adult heads, while A'-C' shows the anterior view.

eyFLP induced clones containing *FRT82B kay* mutant (A-A''), sav^{3} *FRT82B* mutant (B-B'''), and sav^{3} *kay* double mutant (C-C'''). Scale bar is 25mm.

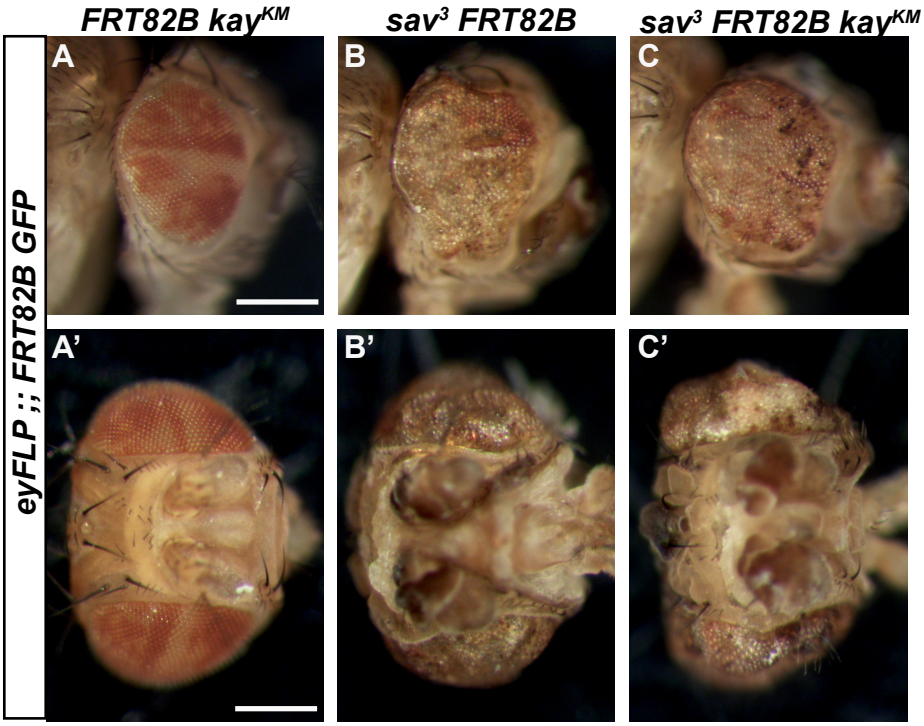
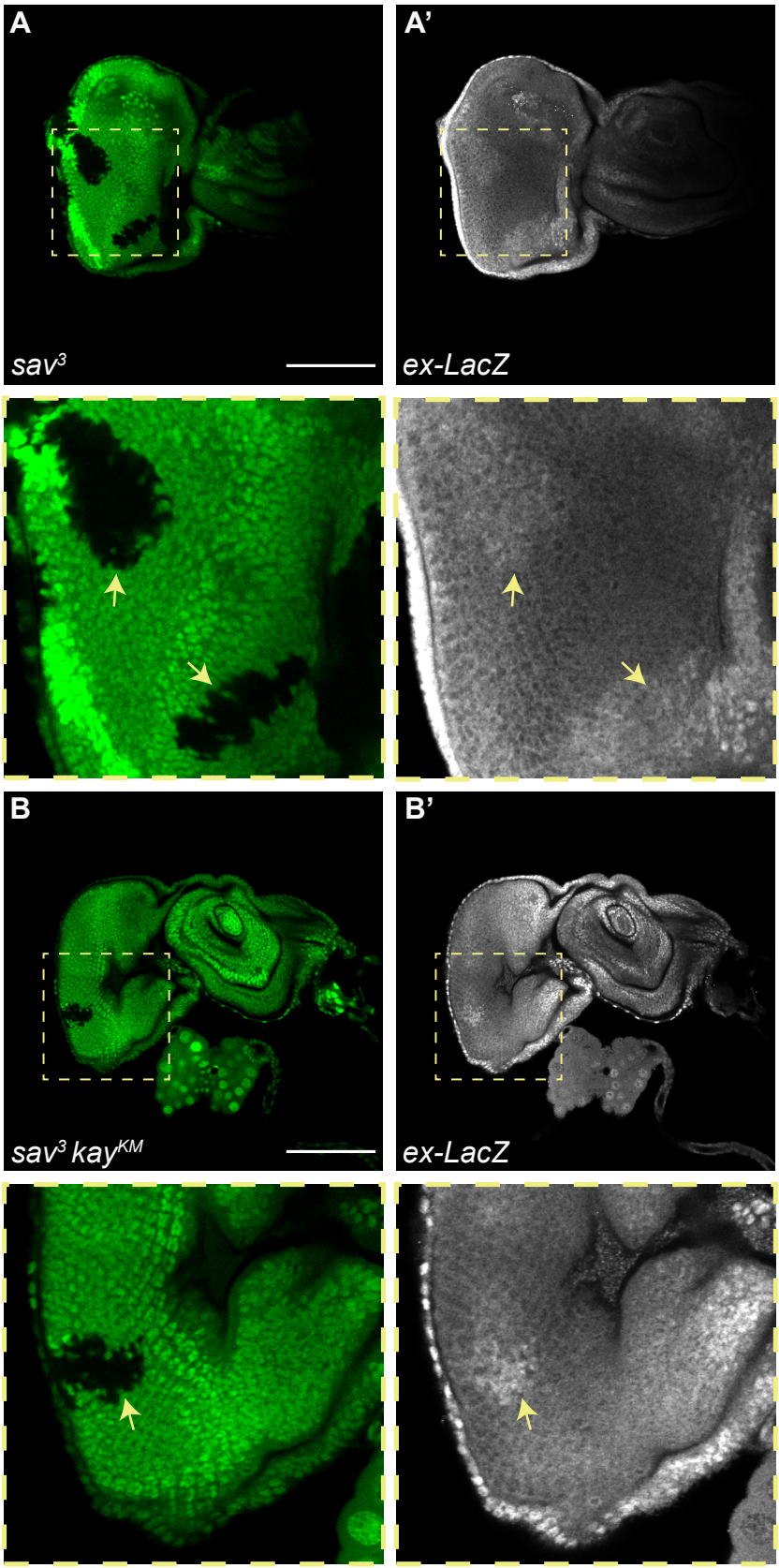

FRT82B kay^KM
sav^3 FRT82B
sav^3 FRT82B kay^KM
eyFLP ;; FRT82B GFP
A
B
C
A'
B'
C'

Figure 5.15: Loss of *kayak* does not influence Yorkie hyperactivity in eye discs.

A-B'. Confocal micrographs of *Drosophila* larval eye-antennal discs from third instar larvae. *hsFLP* was used to create clones in the eye-antennal disc that contained sav^{3} or sav^{3} kay^{KM} double mutation. In all panels, mutant clones of sav^{3} (A and A') and sav^{3} kay^{KM} (B and B') were marked by the loss of GFP. *ex-LacZ* was used as a reporter for Yki activity and was visualised using a βGal antibody (A' and B'). Regions outlined by yellow dashed boxes are enlarged below. Anterior is to the right in all panels. Scale bar is 100μm.


A
A'
sav3
ex-LacZ
B
B'
sav3 kayKM
ex-LacZ

Chapter 6: General Discussion

In this thesis, I identified the transcriptional targets of the *D. melanogaster* Hippo pathway, which is an essential regulator of organ growth and cell fate. In order to develop a deeper understanding of how the Hippo pathway controls these important biological processes, I investigated the following three key questions: 1) What genomic loci are bound by Yki, Sd, and Tgi in growing organs? 2) What is the gene expression program that is regulated by the Hippo pathway in growing organs? 3) What is the role of the AP-1 transcription factor complex in regulation of Hippo pathway target gene expression and organ growth? Prior to my thesis, Hippo pathway target genes had been investigated by others performing ChIP-seq for Yki and Sd (Ikmi et al., 2014; Oh et al., 2013; Slattery et al., 2013). These studies did not reveal key unknown genes that are targeted by the Hippo pathway for organ growth and there was also some uncertainty on their value given that the putative target genes identified by the different ChIP-seq studies only partially overlapped. Furthermore, the AP-1 transcription factors Jra and Kayak (Jun and Fos) had been implicated in neoplastic tissue control together with Yki and Sd (Liu et al. 2016; Pascual et al. 2017; Zanconato et al. 2015) but their genome binding profile has never before been reported in *D. melanogaster*.

Previous ChIP-seq experiments have revealed that the target genes of Yki and Sd vary between different studies, and Yki target genes only partially overlap with Sd target genes (Ikmi et al., 2014; Slattery et al., 2013). Additionally, despite Tgi playing a key role in transcriptional output of Hippo signalling (Guo et al., 2013; Koontz et al., 2013), its target genes have remained unknown. Therefore, I investigated the target genes of the three major transcriptional effectors of Hippo signalling using targeted DamID: the co-activator protein Yki, the co-repressor protein Tgi, and the DNA-binding transcription factor Sd. I found a very high degree of overlap between Yki, Sd, and Tgi target genes, indicating that Sd is the key transcription factor that both Yki and Tgi interact with to bind target gene loci. In the future, it will be important to test this hypothesis definitively by performing targeted DamID of Yki and Tgi in *sd* mutant

tissue, which will identify whether Yki and Tgi rely solely on Sd for interacting with their target genes in the growing eye disc.

Previous studies have focused on the transcriptional targets of Yki and Sd involved in proliferation and cell survival, as well as genes that form a negative feedback loop (Hamaratoglu et al. 2006b; Genevet et al. 2010; Goulev et al. 2008; Oh and Irvine 2011; Peng, Slattery, and Mann 2009; Tapon et al. 2002; S. Wu et al. 2003). In this thesis, comprehensive analyses of Yki, Sd, and Tgi target genes, and the transcriptional profile and accompanying changes in chromatin accessibility caused by Yki hyperactivation, has highlighted key biological processes that the Hippo pathway regulates during eye growth. I extend the list of Yki and Sd target genes to also include genes from the MAPK pathway, glutathione S-transferases (GSTs), cuticle development, and extracellular matrix (ECM), as well as retinal determination and differentiation. The future validation of these genes as Yki, Sd and Tgi target genes will allow us to extend the current list of bona fide Hippo pathway target genes beyond the handful that are currently known. Validation experiments can include verifying changes in gene expression and protein levels in tissues with Hippo signalling deregulation, for example using *Drosophila* enhancer trap lines and antibodies, respectively. Interestingly, some putative Hippo pathway target genes identified in this thesis, such as *Glutathione S transferase D3* and *Larval serum protein 1 γ,* display greater expression increases in hyperactive Yki eye tissue than frequently used reporter genes including *Diap1* and *expanded*. Therefore, these novel target genes could also provide improved and sensitive reporters of Hippo pathway activity, which could be useful for screening new members of the Hippo pathway.

My findings suggest that the Hippo pathway could control the fate of the eye disc, as hyperactive Yki promoted upregulation of cuticle and head cell fate genes, and downregulation of retinal determination genes. This is consistent with previous research showing that Yki activates *homothorax* (*hth*) in the peripodial epithelium (PE), which is a squamous epithelium layer that develops into the cuticle of the adult fly head (T. Zhang, Zhou, and Pignoni 2011). The activation of *hth* promotes PE fate, and represses the fate of the columnar epithelium layer that contributes to the adult eye,

called the disc proper (DP) (T. Zhang, Zhou, and Pignoni 2011). A striking number of genes involved in retinal cell fate were putative target genes of Yki, Sd and Tgi, suggesting that the repression of these genes may not just be a downstream result of *hth* activation. Interestingly, Hth is able to bind to Yki in eye progenitor cells to regulate the target gene *bantam* (Peng, Slattery, and Mann 2009). Therefore, it is possible that the activation of *hth* by Yki leads to association of Yki/Sd with Hth and the direct repression of retinal determination genes by this complex.

In the future, it will be essential to identify the target genes of the Hippo pathway in the PE and DP layers of the eye disc to determine whether Yki/Sd/Tgi target genes differ between these two cell types. This can be carried out using targeted DamID in the PE and DP and comparing Yki, Sd, and Tgi binding between these two epithelial cell layers. Future studies could also focus on identifying transcriptional co-factors of Yki and Sd in each cell type that might drive the different outcomes of Hippo signalling. This could be addressed by investigating the interactome of Yki and Sd in the PE and DP, for example through mass-spectrometry. It has been hypothesised that the less extensive cell-cell contacts in the squamous PE cell layer compared to the columnar DP cell layer stimulate an increased pool of nuclear Yki in these cells (T. Zhang, Zhou, and Pignoni 2011). In the DP, relatively higher activity of the Hippo pathway is thought to limit Yki activity so that it promotes a proliferation and survival gene expression program without promoting the PE cell fate (T. Zhang, Zhou, and Pignoni 2011). On the other hand, lower levels of Hippo signalling in the PE are thought to allow Yki to not only regulate proliferation and survival but also repress retinal determination genes (T. Zhang, Zhou, and Pignoni 2011). Interestingly, in the early mouse embryo, the Yki and Sd orthologues YAP and TEAD specify the first cell fate decision, i.e. the trophoectoderm, which gives rise to the placenta, from the inner cell mass, which forms the embryo proper and yolk sac (Manzanares and Rodriguez 2013; Nishioka et al. 2009). Evidence suggests that the extensive cell-cell contacts activate Hippo signalling in the inner cell mass to repress YAP and TEAD, while cell polarity proteins limit Hippo signalling in the trophoectoderm and allow YAP and TEAD activation in these cells (Cockburn et al. 2013; Hirate et al. 2013). It would be fascinating to investigate whether similar mechanisms also play a role in fate determination in the PE and DP of the eye disc.

A striking result I discovered was the number of Yki putative target genes that were downregulated upon Yki hyperactivation. This was unexpected because Yki has only been described as a transcriptional co-activator of genes in *Drosophila* tissues (Huang et al. 2005b; Oh et al. 2014a; Qing et al. 2014). This result could be explained by several possibilities, which are not mutually exclusive: 1) Direct repression of these target genes by Yki/Sd/Tgi 2) Loss of a co-activator at these target genes or 3) Gain of a repressor at these target genes. Several lines of evidence suggest that YAP represses target genes in mammalian cells. For example, YAP and TEAD interact with the nucleosome-remodelling and deacetylase (NuRD) repressor complex to repress mesendoderm cell fate genes in human embryonic stem cells (Beyer et al. 2013) and anti-proliferative and apoptosis-inducing genes in human cancer cell lines (M. Kim et al. 2015). An interesting topic for future studies would be to investigate the role of Yki as a repressor of target genes in *Drosophila* tissues. Focusing on target genes that are downregulated in Yki hyperactive tissues, and assessing these loci for histone modifications, as well as the presence of chromatin-modifying complexes and additional co-factors, could help to establish the mechanism of Yki-mediated target gene repression.

In this thesis, I found that the *Drosophila* AP-1 protein Jra, orthologous to mammalian JUN, shared 71% of its target genes with Yki and/or Sd. These shared target genes were highly enriched with predicted AP-1 and Sd DNA binding motifs, as indicated by Homer analysis. This is consistent with multiple reports showing the co-existence of AP-1 and Sd/TEAD DNA binding motifs in *Drosophila* tissues (Pascual et al. 2017), and human cancer cell lines (Stein et al. 2015; Verfaillie et al. 2015; Zanconato et al. 2015). As Sd is dispensable for eye growth (Koontz et al. 2013), I posited that the AP-1 complex and Yki/Sd regulate a set of common target genes in a genetically redundant fashion such that loss of either AP-1 or Yki/Sd does not impede eye growth. However, my experiments revealed this hypothesis to be incorrect, as experiments using a *kay* mutant allele revealed no obvious role for Kay in co-operatively regulating target genes and growth with Yki and Sd. So far, the only clear example of Yki and AP-1 converging to co-regulate numerous genes in *Drosophila* occurs in the heart, where Yki, Sd, and AP-1 synergistically activate target genes to

initiate alary muscle dedifferentiation (Schaub et al., 2019). In this context, Yki and Sd act downstream of the Hippo pathway, and AP-1 activity is regulated by JNK signalling (Schaub, Rose, and Frasch 2019a), which implies that both activation of JNK and inhibition of Hippo signalling are needed for the synergistic activation of shared target genes.

Given my finding that AP-1 and Yki/Sd share 1,737 genes in the developing eye disc, future experiments should address under what cellular contexts do these two different signalling inputs converge to regulate their shared target genes. One such context could be during tumorous growth of imaginal discs that possess hyperactive Ras and *scribble* loss. *Ras scrib* clones cause massive invasive tumours and increased levels of Yki target genes, such as *ex* and *fj* (Doggett et al. 2011; Enomoto et al. 2015). In this context, F-actin is accumulated through the activity of the LIM domain protein Ajuba, resulting in the inhibition of Wts and thereby Yki activation (Enomoto et al. 2015). It's been shown that the elevation of Yki target genes is, at least partly, due to increased Yki activity, however knockdown of Yki does not completely suppress tumour overgrowth (Doggett et al. 2011; Enomoto et al. 2015). In the future, it will be interesting to investigate whether AP-1 and Yki co-operate at Yki target genes in this context. It is also possible that Yki and AP-1 converge at shared target genes in tissues with apico-basal polarity defects (Bunker et al. 2015). Indeed, this has been reported in *Drosophila* wings that are entirely mutant for *scrib*. In this context, Yki and AP-1 are activated by aPKC and JNK, respectively. Yki and AP-1 converge to activate a polarity-responsive enhancer in the JAK-STAT ligand gene upd3, resulting in overgrown wing discs (Bunker et al. 2015). In the future, it will be interesting to investigate whether additional shared target genes are also activated by Yki and AP-1 in tissues with apico-basal polarity defects.

In summary, this thesis deepens our understanding of the downstream target genes and biological processes that the Hippo pathway regulates during organ growth. I have identified many putative novel target genes regulated of Yki, Sd, and Tgi in the developing eye disc, including genes involved in cuticle/head cell fate, retinal determination, glutathione S-transferases (GSTs), and extracellular matrix (ECM).

Interestingly, my results provide support for a role for the Hippo pathway in determining an early cell fate decision of the eye disc, i.e. peripodial epithelium from disc proper. I also identified shared putative targets genes of the AP-1 complex and Yki/Sd, which can help us to unravel the cooperative role that these complexes play in transcriptional regulation. Given the links between Hippo signalling and AP-1 in driving tumourigenesis (Galli et al. 2015; Koo et al. 2020; Maglic et al. 2018; Zanconato et al. 2015), understanding this complex interaction may help to identify therapeutic strategies for cancer patients as well as different developmental processes.

Chapter 7: References

Alarcón, Claudio, Alexia Ileana Zaromytidou, Qiaoran Xi, Sheng Gao, Jianzhong Yu, Sho Fujisawa, Afsar Barlas, et al. 2009. "Nuclear CDKs Drive Smad Transcriptional Activation and Turnover in BMP and TGF-β Pathways." *Cell*. doi:10.1016/j.cell.2009.09.035.

Alfonso-Gonzalez, Carlos, and Juan Rafael Riesgo-Escovar. 2018. "Fos Metamorphoses: Lessons from Mutants in Model Organisms." *Mechanisms of Development* 154 (March). Elsevier: 73–81. doi:10.1016/j.mod.2018.05.006.

Amoyel, Marc, Abigail M. Anderson, and Erika A. Bach. 2014. "JAK/STAT Pathway Dysregulation in Tumors: A Drosophila Perspective." *Seminars in Cell and Developmental Biology*. doi:10.1016/j.semcdb.2014.03.023.

Andersen, D. S., J. Colombani, and P. Léopold. 2013. "Coordination of Organ Growth: Principles and Outstanding Questions from the World of Insects." *Trends in Cell Biology* 23 (7): 336–44. doi:10.1016/j.tcb.2013.03.005.

Antonescu, Cristina R., Francois Le Loarer, Juan Miguel Mosquera, Andrea Sboner, Lei Zhang, Chun Liang Chen, Hsiao Wei Chen, et al. 2013. "Novel YAP1-TFE3 Fusion Defines a Distinct Subset of Epithelioid Hemangioendothelioma." *Genes Chromosomes and Cancer*. doi:10.1002/gcc.22073.

Asthagiri, Ashok R., Dilys M. Parry, John A. Butman, H. Jeffrey Kim, Ekaterini T. Tsilou, Zhengping Zhuang, and Russell R. Lonser. 2009. "Neurofibromatosis Type 2." *The Lancet*. doi:10.1016/S0140-6736(09)60259-2.

Aughey, Gabriel N., Alicia Estacio Gomez, Jamie Thomson, Hang Yin, and Tony D. Southall. 2018. "CATaDa Reveals Global Remodelling of Chromatin Accessibility during Stem Cell Differentiation in Vivo." *ELife* 7: 1–22. doi:10.7554/eLife.32341.

Badouel, Caroline, Laura Gardano, Nancy Amin, Ankush Garg, Robyn Rosenfeld, Thierry Le Bihan, and Helen McNeill. 2009. "The FERM-Domain Protein Expanded Regulates Hippo Pathway Activity via Direct Interactions with the Transcriptional Activator Yorkie." *Developmental Cell* 16 (3). Elsevier Ltd: 411–20. doi:10.1016/j.devcel.2009.01.010.

Baena-Lopez, Luis Alberto, Isabel Rodríguez, and Antonio Baonza. 2008. "The Tumor Suppressor Genes Dachsous and Fat Modulate Different Signalling Pathways by Regulating Dally and Dally-Like." *Proceedings of the National Academy of Sciences of the United States of America* 105 (28): 9645–50. doi:10.1073/pnas.0803747105.

Bainbridge, S. P., and M. Bownes. 1981. "Staging the Metamorphosis of Drosophila Melanogaster." *Journal of Embryology and Experimental Morphology* Vol.66 (1967): 57–80.

Barrio, Rosa, and Jose F. De Celis. 2004. "Regulation of Spalt Expression in the Drosophila Wing Blade in Response to the Decapentaplegic Signaling Pathway." *Proceedings of the National Academy of Sciences of the United States of America* 101 (16): 6021–26. doi:10.1073/pnas.0401590101.

Bassett, Andrew, and Ji Long Liu. 2014. "CRISPR/Cas9 Mediated Genome Engineering in Drosophila." *Methods* 69 (2). Elsevier Inc.: 128–36. doi:10.1016/j.ymeth.2014.02.019.

Bayarmagnai, Battuya, Brandon N. Nicolay, Abul B.M.M.K. Islam, Nuria Lopez-Bigas, and Maxim V. Frolov. 2012. "Drosophila GAGA Factor Is Required for Full Activation of the DE2f1-Yki/Sd Transcriptional Program." *Cell Cycle* 11 (22): 4191–4202. doi:10.4161/cc.22486.

Betz, Aurel, Don Ryoo Hyung, Hermann Steller, and James E. Darnell. 2008. "STAT92E Is a Positive Regulator of Drosophila Inhibitor of Apoptosis 1 (DIAP/1) and Protects against Radiation-Induced Apoptosis." *Proceedings of the National Academy of Sciences of the United States of America* 105 (37): 13805–10. doi:10.1073/pnas.0806291105.

Beyer, Tobias A., Alexander Weiss, Yuliya Khomchuk, Kui Huang, Abiodun A. Ogunjimi, Xaralabos Varelas, and Jeffrey L. Wrana. 2013. "Switch Enhancers Interpret TGF-β and Hippo Signaling to Control Cell Fate in Human Embryonic Stem Cells." *Cell Reports* 5 (6). The Authors: 1611–24. doi:10.1016/j.celrep.2013.11.021.

Boedigheimer, M., and A. Laughon. 1993. "Expanded: A Gene Involved in the Control of Cell Proliferation in Imaginal Discs." *Development* 118 (4): 1291–1301.

Boone, Emilie, Julien Colombani, Ditte S. Andersen, and Pierre Leópold. 2016. "The Hippo Signalling Pathway Coordinates Organ Growth and Limits Developmental Variability by Controlling Dilp8 Expression." *Nature Communications* 7: 1–8. doi:10.1038/ncomms13505.

Brennecke, Julius, David R. Hipfner, Alexander Stark, Robert B. Russell, and Stephen M. Cohen. 2003. "Bantam Encodes a Developmentally Regulated MicroRNA That Controls Cell Proliferation and Regulates the Proapoptotic Gene Hid in Drosophila." *Cell* 113 (1): 25–36. doi:10.1016/S0092-8674(03)00231-9.

Broadie, Kendal, Stefan Baumgartner, and Andreas Prokop. 2011. "Extracellular Matrix and Its Receptors in Drosophila Neural Development." *Developmental Neurobiology* 71 (11): 1102–30. doi:10.1002/dneu.20935.

Bryant, Peter J., and Patrice Levinson. 1985. "Intrinsic Growth Control in the Imaginal Primordia of Drosophila, and the Autonomous Action of a Lethal Mutation Causing Overgrowth." *Developmental Biology* 107 (2): 355–63. doi:10.1016/0012-1606(85)90317-3.

Bueno, Raphael, Eric W. Stawiski, Leonard D. Goldstein, Steffen Durinck, Assunta De Rienzo, Zora Modrusan, Florian Gnad, et al. 2016. "Comprehensive Genomic Analysis of Malignant Pleural Mesothelioma Identifies Recurrent Mutations, Gene Fusions and Splicing Alterations." *Nature Genetics*. doi:10.1038/ng.3520.

Bulger, Michael, and Mark Groudine. 2010. "NIH Public Access." *Development Biology* 339 (2): 250–57.

Bunker, Brandon D., Tittu T. Nellimoottil, Ryan M. Boileau, Anne K. Classen, and David Bilder. 2015. "The Transcriptional Response to Tumorigenic Polarity Loss in Drosophila." *ELife* 2015 (4): 1–19. doi:10.7554/eLife.03189.

Cagan, Ross. 2009. "Principles of Drosophila Eye Differentiation Ross." *Curr Top Dev Biol.* 89: 115–35. doi:10.1038/jid.2014.371.

Calses, Philamer C., James J. Crawford, Jennie R. Lill, and Anwesha Dey. 2019. "Hippo Pathway in Cancer: Aberrant Regulation and Therapeutic Opportunities." *Trends in Cancer* 5 (5). The Authors: 297–307. doi:10.1016/j.trecan.2019.04.001.

Camargo, Fernando D., Sumita Gokhale, Jonathan B. Johnnidis, Dongdong Fu, George W. Bell, Rudolf Jaenisch, and Thijn R. Brummelkamp. 2007. "YAP1 Increases Organ Size and Expands Undifferentiated Progenitor Cells." *Current Biology* 17 (23): 2054–60. doi:10.1016/j.cub.2007.10.039.

Campbell, Shelagh, Maneesha Inamdar, Veronica Rodrigues, Vijay Raghavan, Michael Palazzolo, and Arthur Chovnick. 1992. "The Scalloped Gene Encodes a Novel,

Evolutionarily Conserved Transcription Factor Required for Sensory Organ Differentiation in Drosophila." *Genes and Development* 6 (3): 367–79. doi:10.1101/gad.6.3.367.

Celis, Jose F. De, and Rosa Barrio. 2009. "Regulation and Function of Spalt Proteins during Animal Development." *International Journal of Developmental Biology*. doi:10.1387/ijdb.072408jd.

Chan, Eunice H.Y., Marjaana Nousiainen, Ravindra B. Chalamalasetty, Anja Schäfer, Erich A. Nigg, and Herman H.W. Sillje. 2005. "The Ste20-like Kinase Mst2 Activates the Human Large Tumor Suppressor Kinase Lats1." *Oncogene* 24 (12): 2076–86. doi:10.1038/sj.onc.1208445.

Chang, Lei, Luca Azzolin, Daniele Di Biagio, Francesca Zanconato, Giusy Battilana, Romy Lucon Xiccato, Mariaceleste Aragona, et al. 2018. "The SWI/SNF Complex Is a Mechanoregulated Inhibitor of YAP and TAZ." *Nature* 563 (7730). Springer US: 265–69. doi:10.1038/s41586-018-0658-1.

Chatterjee, Nirmalya, and Dirk Bohmann. 2012. "A Versatile ΦC31 Based Reporter System for Measuring AP-1 and NRF2 Signaling in Drosophila and in Tissue Culture." *PLoS ONE* 7 (4). doi:10.1371/journal.pone.0034063.

Chen, Chiao Lin, Molly C. Schroeder, Madhuri Kango-Singh, Chunyao Tao, and Georg Halder. 2012. "Tumor Suppression by Cell Competition through Regulation of the Hippo Pathway." *Proceedings of the National Academy of Sciences of the United States of America* 109 (2): 484–89. doi:10.1073/pnas.1113882109.

Chen, Qian, Nailing Zhang, Rui Xie, Wei Wang, Jing Cai, Kyung Suk Choi, Karen Kate David, et al. 2015. "Homeostatic Control of Hippo Signaling Activity Revealed by an Endogenous Activating Mutation in YAP." *Genes and Development*. doi:10.1101/gad.264234.115.

Chen, Xiyan, Weiting Gu, Qi Wang, Xucheng Fu, Ying Wang, Xin Xu, and Yong Wen. 2018. "C-MYC and BCL-2 Mediate YAP-Regulated Tumorigenesis in OSCC." *Oncotarget*. doi:10.18632/oncotarget.23089.

Cho, Eunjoo, Yongqiang Feng, Cordelia Rauskolb, Sushmita Maitra, Rick Fehon, and Kenneth D. Irvine. 2006. "Delineation of a Fat Tumor Suppressor Pathway." *Nature Genetics* 38 (10): 1142–50. doi:10.1038/ng1887.

Choi, Wonyoung, Jeongsik Kim, Jaeoh Park, Da Hye Lee, Daehee Hwang, Jeong Hwan Kim, Hassan Ashktorab, et al. 2018. "YAP/TAZ Initiates Gastric Tumorigenesis via Upregulation of MYC." *Cancer Research* 78 (12): 3306–20. doi:10.1158/0008-5472.CAN-17-3487.

Ciapponi, Laura, and Dirk Bohmann. 2002. "An Essential Function of AP-1 Heterodimers in Drosophila Development." *Mechanisms of Development* 115 (1–2): 35–40. doi:10.1016/S0925-4773(02)00093-X.

Clements, Jason, Korneel Hensa, Srinivas Merugub, Dichtlb Beatriz, Gert de Couetc, and Callaerts Patrick. 2009. "Mutational Analysis of the Eyeless Gene and Phenotypic Rescue Reveal That an Intact Eyeless Protein Is Necessary for Normal Eye and Brain Development in Drosophila." *Development Biology* 334 (2): 503–12. doi:10.1016/j.ydbio.2009.08.003.

Cockburn, Katie, Steffen Biechele, Jodi Garner, and Janet Rossant. 2013. "The Hippo Pathway Member Nf2 Is Required for Inner Cell Mass Specification." *Current Biology*. doi:10.1016/j.cub.2013.05.044.

Davie, Kristofer, Jelle Jacobs, Mardelle Atkins, Delphine Potier, Valerie Christiaens, Georg Halder, and Stein Aerts. 2015. "Discovery of Transcription Factors and Regulatory Regions Driving In Vivo Tumor Development by ATAC-Seq and FAIRE-Seq Open Chromatin Profiling." *PLoS Genetics* 11 (2): 1–24. doi:10.1371/journal.pgen.1004994.

Deng, Xiaochong, and Lin Fang. 2018. "VGLL4 Is a Transcriptional Cofactor Acting as a Novel Tumor Suppressor via Interacting with TEADs." *American Journal of Cancer Research*.

Djiane, Alexandre, Alena Krejci, Frédéric Bernard, Silvie Fexova, Katherine Millen, and Sarah J. Bray. 2013. "Dissecting the Mechanisms of Notch Induced Hyperplasia." *EMBO Journal* 32 (1): 60–71. doi:10.1038/emboj.2012.326.

Dnasekaran, ha Danny N., and E. Premkumar Reddy. 2017. "JNK-Signaling: A Multiplexing Hub in Programmed Cell Death." *Genes and Cancer* 8 (9–10): 682–94. doi:10.18632/genesandcancer.155.

Doggett, Karen, Felix A. Grusche, Helena E. Richardson, and Anthony M. Brumby. 2011. "Loss of the Drosophila Cell Polarity Regulator Scribbled Promotes Epithelial Tissue Overgrowth and Cooperation with Oncogenic Ras-Raf through Impaired Hippo Pathway

Signaling." *BMC Developmental Biology*. doi:10.1186/1471-213X-11-57.

Domínguez, María, and Fernando Casares. 2005. "Organ Specification-Growth Control Connection: New in-Sights from the Drosophila Eye-Antennal Disc." *Developmental Dynamics* 232 (3): 673–84. doi:10.1002/dvdy.20311.

Dong, Jixin, Georg Feldmann, Jianbin Huang, Shian Wu, Nailing Zhang, A Sarah, Mariana F Gayyed, Robert A Anders, Anirban Maitra, and Duojia Pan. 2007. "Elucidation of a Universal Size-Control Mechanism in Drosophila and Mammals." *Cell* 130 (6): 1120–33. doi:10.1016/j.cell.2007.07.019.Elucidation.

Driscoll, Tristan P., Brian D. Cosgrove, Su Jin Heo, Zach E. Shurden, and Robert L. Mauck. 2015. "Cytoskeletal to Nuclear Strain Transfer Regulates YAP Signaling in Mesenchymal Stem Cells." *Biophysical Journal* 108 (12). Biophysical Society: 2783–93. doi:10.1016/j.bpj.2015.05.010.

Duffy, Joseph B. 2002. "GAL4 System in Drosophila: A Fly Geneticist's Swiss Army Knife." *Genesis* 34 (1–2): 1–15. doi:10.1002/gene.10150.

Dupont, Sirio, Leonardo Morsut, Mariaceleste Aragona, Elena Enzo, Stefano Giulitti, Michelangelo Cordenonsi, Francesca Zanconato, et al. 2011. "Role of YAP/TAZ in Mechanotransduction." *Nature* 474 (7350). Nature Publishing Group: 179–84. doi:10.1038/nature10137.

Ege, Nil, Anna M. Dowbaj, Ming Jiang, Michael Howell, Steven Hooper, Charles Foster, Robert P. Jenkins, and Erik Sahai. 2018. "Quantitative Analysis Reveals That Actin and Src-Family Kinases Regulate Nuclear YAP1 and Its Export." *Cell Systems* 6 (6). Elsevier Inc.: 692-708.e13. doi:10.1016/j.cels.2018.05.006.

Elosegui-Artola, Alberto, Ion Andreu, Amy E.M. Beedle, Ainhoa Lezamiz, Marina Uroz, Anita J. Kosmalska, Roger Oria, et al. 2017. "Force Triggers YAP Nuclear Entry by Regulating Transport across Nuclear Pores." *Cell* 171 (6). Elsevier: 1397-1410.e14. doi:10.1016/j.cell.2017.10.008.

Enomoto, Masato, Daisuke Kizawa, Shizue Ohsawa, and Tatsushi Igaki. 2015. "JNK Signaling Is Converted from Anti- to pro-Tumor Pathway by Ras-Mediated Switch of Warts Activity." *Developmental Biology* 403 (2). Elsevier: 162–71. doi:10.1016/j.ydbio.2015.05.001.

Fernandez-L, Africa, Paul A. Northcott, James Dalton, Charles Fraga, David Ellison, Stephane

Angers, Michael D. Taylor, and Anna Marie Kenney. 2009. "YAP1 Is Amplified and Up-Regulated in Hedgehog-Associated Medulloblastomas and Mediates Sonic Hedgehog-Driven Neural Precursor Proliferation." *Genes and Development*. doi:10.1101/gad.1824509.

Ferrigno, Olivier, François Lallemand, Franck Verrecchia, Sébastien L'hoste, Jacques Camonis, Azeddine Atfi, and Alain Mauviel. 2002. "Yes-Associated Protein (YAP65) Interacts with Smad7 and Potentiates Its Inhibitory Activity against TGF-β/Smad Signaling." *Oncogene*. doi:10.1038/sj.onc.1205623.

Fischer, Janice A, Edward Giniger, Tom Maniatis, and Mark Ptashne. 1988. "GAL4 Activates Transcription in Drosophila." *Nature* 332 (28): 853–56.

Franklin, J. Matthew, Rajarshi P Ghosh, Quanming Shi, and Jan T Liphardt. 2019. "Spatial Resets Modulate YAP-Dependent Transcription." *BioRxiv*, 539049. doi:10.1101/539049.

Freeman, Matthew. 1996. "Reiterative Use of the EGF Receptor Triggers Differentiation of All Cell Types in the Drosophila Eye." *Cell* 87 (4): 651–60. doi:10.1016/S0092-8674(00)81385-9.

Fury, Wen, Franak Batliwalla, Peter K Gregersen, and Wentian Li. 2006. "Overlapping Probabilities of Top Ranking Gene Lists , Hypergeometric Distribution , and Stringency of Gene Selection Criterion." *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 5531–34. doi:10.1109/IEMBS.2006.260828.

Galli, Giorgio G., Matteo Carrara, Wei Chien Yuan, Christian Valdes-Quezada, Basanta Gurung, Brian Pepe-Mooney, Tinghu Zhang, et al. 2015. "YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers." *Molecular Cell* 60 (2). Elsevier Inc.: 328–37. doi:10.1016/j.molcel.2015.09.001.

Genevet, Alice, Michael C. Wehr, Ruth Brain, Barry J. Thompson, and Nicolas Tapon. 2010. "Kibra Is a Regulator of the Salvador/Warts/Hippo Signaling Network." *Developmental Cell* 18 (2). Elsevier: 300–308. doi:10.1016/j.devcel.2009.12.011.

Giot, L., J. S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y. L. Hao, et al. 2003. "A Protein Interaction Map of Drosophila Melanogaster." *Science*. doi:10.1126/science.1090289.

Gise, Alexander Von, Zhiqiang Lin, Karin Schlegelmilch, Leah B. Honor, Gina M. Pan, Jessica N. Buck, Qing Ma, et al. 2012. "YAP1, the Nuclear Target of Hippo Signaling, Stimulates

Heart Growth through Cardiomyocyte Proliferation but Not Hypertrophy." *Proceedings of the National Academy of Sciences of the United States of America* 109 (7): 2394–99. doi:10.1073/pnas.1116136109.

Gokhale, Rewatee, and Cathie M. Pfleger. 2019. "The Power of Drosophila Genetics: The Discovery of the Hippo Pathway." *Methods in Molecular Biology* 1893: 3–26. doi:10.1007/978-1-4939-8910-2_1.

Golic, Kent G. 1991. "Site-Specific Recombination between Homologous Chromosomes in Drosophila." *Science* 252 (5008): 958–61. doi:10.1126/science.2035025.

Golic, Kent G., and Susan Lindquist. 1989. "The FLP Recombinase of Yeast Catalyzes Site-Specific Recombination in the Drosophila Genome." *Cell* 59 (3): 499–509. doi:10.1016/0092-8674(89)90033-0.

Goulev, Youlian, Jean Daniel Fauny, Beatriz Gonzalez-Marti, Domenico Flagiello, Joël Silber, and Alain Zider. 2008. "SCALLOPED Interacts with YORKIE, the Nuclear Effector of the Hippo Tumor-Suppressor Pathway in Drosophila." *Current Biology* 18 (6): 435–41. doi:10.1016/j.cub.2008.02.034.

Grannas, Karin, Linda Arngården, Peter Lönn, Magdalena Mazurkiewicz, Andries Blokzijl, Agata Zieba, and Ola Söderberg. 2015. "Crosstalk between Hippo and TGFβ: Subcellular Localization of YAP/TAZ/Smad Complexes." *Journal of Molecular Biology*. doi:10.1016/j.jmb.2015.04.015.

Guo, Tong, Yi Lu, Peixue Li, Meng-Xin Yin, Dekang Lv, Wenjing Zhang, Huizhen Wang, et al. 2013a. "A Novel Partner of Scalloped Regulates Hippo Signaling via Antagonizing Scalloped-Yorkie Activity." *Cell Research* 23: 1201–14. doi:10.1038/cr.2013.120.

Guo, Tong, Yi Lu, Peixue Li, Meng Xin Yin, Dekang Lv, Wenjing Zhang, Huizhen Wang, et al. 2013b. "A Novel Partner of Scalloped Regulates Hippo Signaling via Antagonizing Scalloped-Yorkie Activity." *Cell Research* 23 (10): 1201–14. doi:10.1038/cr.2013.120.

Halder, Georg, Patricia Polaczyk, Mary Ellen Kraus, Angela Hudson, Jaeseob Kim, Allen Laughon, and Scan Carroll. 1998. "The Vestigial and Scalloped Proteins Act Together to Directly Regulate Wing-Specific Gene Expression in Drosophila." *Genes and Development* 12 (24). Cold Spring Harbor Laboratory Press: 3900–3909. doi:10.1101/gad.12.24.3900.

Hamaratoglu, Fisun, Maria Willecke, Madhuri Kango-Singh, Riitta Nolo, Eric Hyun, Chunyao

Tao, Hamed Jafar-Nejad, and Georg Halder. 2006a. “The Tumour-Suppressor Genes NF2/Merlin and Expanded Act through Hippo Signalling to Regulate Cell Proliferation and Apoptosis.” *Nat Cell Biol* 8 (1): 27–36. doi:ncb1339 [pii] 10.1038/ncb1339.

———. 2006b. “The Tumour-Suppressor Genes NF2/Merlin and Expanded Act through Hippo Signalling to Regulate Cell Proliferation and Apoptosis.” *Nature Cell Biology* 8 (1): 27–36. doi:10.1038/ncb1339.

Han, Yanyan. 2019. “Analysis of the Role of the Hippo Pathway in Cancer.” *Journal of Translational Medicine* 17 (1). BioMed Central: 1–17. doi:10.1186/s12967-019-1869-4.

Harvey, Kieran F., Cathie M. Pfleger, and Iswar K. Hariharan. 2003. “The Drosophila Mst Ortholog, Hippo, Restricts Growth and Cell Proliferation and Promotes Apoptosis.” *Cell* 114 (4): 457–67. doi:10.1016/S0092-8674(03)00557-9.

Harvey, Kieran F., Xiaomeng Zhang, and David M. Thomas. 2013. “The Hippo Pathway and Human Cancer.” *Nature Reviews Cancer* 13 (4): 246–57. doi:10.1038/nrc3458.

Hauck, Bernd, Walter J. Gehring, and Uwe Walldorf. 1999. “Functional Analysis of an Eye Specific Enhancer of the Eyeless Gene in Drosophila.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (2): 564–69. doi:10.1073/pnas.96.2.564.

Hayes, John D., Jack U. Flanagan, and Ian R. Jowsey. 2005. “Glutathione Transferases.” *Annual Review of Pharmacology and Toxicology* 45 (1): 51–88. doi:10.1146/annurev.pharmtox.45.120403.095857.

Heallen, Todd, Min Zhang, Jun Wang, Margarita Bonilla-Claudio, Ela Klysik, Randy L. Johnson, and James F. Martin. 2011. “Hippo Pathway Inhibits Wnt Signaling to Restrain Cardiomyocyte Proliferation and Heart Size Todd.” *Science* 332 (6028): 458–61. doi:10.1038/jid.2014.371.

Heberlein, Ulrike, Ulrike Heberlein, Alvin Y. Luk, and Terrence J. Donohoe. 1995. “Growth and Differentiation in the Drosophila Eye Coordinated by Hedgehog.” *Nature* 373 (6516): 709–11. doi:10.1038/373709a0.

Heinz, Sven, Christopher Benner, Nathanael Spann, Eric Bertolino, Yin C. Lin, Peter Laslo, Jason X. Cheng, Cornelis Murre, Harinder Singh, and Christopher K. Glass. 2010. “Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-Regulatory Elements Required for Macrophage and B Cell Identities.” *Molecular Cell* 38

(4). Elsevier Inc.: 576–89. doi:10.1016/j.molcel.2010.05.004.

Hillmer, Ryan E., and Brian A. Link. 2019. "The Roles of Hippo Signaling Transducers Yap and Taz in Chromatin Remodeling." *Cells* 8 (5): 502. doi:10.3390/cells8050502.

Hirate, Yoshikazu, Shino Hirahara, Ken Ichi Inoue, Atsushi Suzuki, Vernadeth B. Alarcon, Kazunori Akimoto, Takaaki Hirai, et al. 2013. "Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos." *Current Biology*. doi:10.1016/j.cub.2013.05.014.

Hua, Hui, Minjing Li, Ting Luo, Yancun Yin, and Yangfu Jiang. 2011. "Matrix Metalloproteinases in Tumorigenesis: An Evolving Paradigm." *Cellular and Molecular Life Sciences* 68 (23): 3853–68. doi:10.1007/s00018-011-0763-x.

Huang, Jianbin, Shian Wu, Jose Barrera, Krista Matthews, and Duojia Pan. 2005. "The Hippo Signaling Pathway Coordinately Regulates Cell Proliferation and Apoptosis by Inactivating Yorkie, the Drosophila Homolog of YAP." *Cell* 122 (3): 421–34. doi:10.1016/j.cell.2005.06.007.

Ikmi, Aissam, Bjoern Gaertner, Christopher Seidel, Mansi Srivastava, Julia Zeitlinger, and Matthew C. Gibson. 2014. "Molecular Evolution of the Yap/Yorkie Proto-Oncogene and Elucidation of Its Core Transcriptional Program." *Molecular Biology and Evolution* 31 (6): 1375–90. doi:10.1093/molbev/msu071.

Irvine, Kenneth D., and Kieran F. Harvey. 2015. "Control of Organ Growth by Patterning and Hippo Signaling in Drosophila." *Cold Spring Harbor Perspectives in Biology* 7 (6): 1–16. doi:10.1101/cshperspect.a019224.

Jacobs, Jelle, Mardelle Atkins, Kristofer Davie, Hana Imrichova, Lucia Romanelli, Valerie Christiaens, Gert Hulselmans, et al. 2018. "The Transcription Factor Grainy Head Primes Epithelial Enhancers for Spatiotemporal Activation by Displacing Nucleosomes." *Nature Genetics* 50 (7): 1011–20. doi:10.1038/s41588-018-0140-x.

Jasper, Heinrich, Vladimir Benes, Christian Schwager, Silvia Sauer, Sandra Clauder-Münster, Wilhelm Ansorge, and Dirk Bohmann. 2001. "The Genomic Response of the Drosophila Embryo to JNK Signaling." *Developmental Cell* 1 (4): 579–86. doi:10.1016/S1534-5807(01)00045-4.

Jia, Jianhang, Wensheng Zhang, Bing Wang, Richard Trinko, and Jin Jiang. 2003. "The Drosophila Ste20 Family Kinase DMST Functions as a Tumor Suppressor by Restricting

Cell Proliferation and Promoting Apoptosis." *Genes and Development* 17 (20): 2514–19. doi:10.1101/gad.1134003.

Jin, Yunyun, Jinjin Xu, Meng-Xin Yin, Yi Lu, Lianxin Hu, Peixue Li, Peng Zhang, et al. 2013 "Brahma Is Essential for Drosophila Intestinal Stem Cell Proliferation and Regulated by Hippo Signaling." *ELife* 2: 1–21. doi:10.7554/elife.00999.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. 2012. "A Programmable Dual-RNA – Guided" 337 (August): 816–22.

Justice, Robin W., Olav Zilian, Daniel F. Woods, Markus Noll, and Peter J. Bryant. 1995. "The Drosophila Tumor Suppressor Gene Warts Encodes a Homolog of Human Myotonic Dystrophy Kinase and Is Required for the Control of Cell Shape and Proliferation." *Genes and Development* 9 (5): 534–46. doi:10.1101/gad.9.5.534.

Kango-Singh, Madhuri, and Amit Singh. 2009. "Regulation of Organ Size: Insights from the Drosophila Hippo Signaling Pathway." *Developmental Dynamics* 238 (7): 1627–37. doi:10.1002/dvdy.21996.

Karim, Felix D., and Gerald M. Rubin. 1998. "Ectopic Expression of Activated Ras1 Induces Hyperplastic Growth and Increased Cell Death in Drosophila Imaginal Tissues." *Development* 125 (1): 1–9.

Karouzou, Maria V., Yannis Spyropoulos, Vassiliki A. Iconomidou, R. S. Cornman, Stavros J. Hamodrakas, and Judith H. Willis. 2007. "Drosophila Cuticular Proteins with the R&R Consensus: Annotation and Classification with a New Tool for Discriminating RR-1 and RR-2 Sequences." *Insect Biochemistry and Molecular Biology* 37 (8 SPEC. ISS): 754–60. doi:10.1016/j.ibmb.2007.03.007.

Kim, Han Byul, and Seung Jae Myung. 2018. "Clinical Implications of the Hippo-YAP Pathway in Multiple Cancer Contexts." *BMB Reports* 51 (3): 119–25. doi:10.5483/BMBRep.2018.51.3.018.

Kim, Minchul, Taekhoon Kim, Randy L. Johnson, and Dae Sik Lim. 2015. "Transcriptional Co-Repressor Function of the Hippo Pathway Transducers YAP and TAZ." *Cell Reports* 11 (2). The Authors: 270–82. doi:10.1016/j.celrep.2015.03.015.

Klemm, Sandy L., Zohar Shipony, and William J. Greenleaf. 2019. "Chromatin Accessibility and the Regulatory Epigenome." *Nature Reviews Genetics* 20 (4). Springer US: 207–20.

doi:10.1038/s41576-018-0089-8.

Knoblich, Jürgen A., Karsten Sauer, Lynn Jones, Helena Richardson, Robert Saint, and Christian F. Lehner. 1994. "Cyclin E Controls S Phase Progression and Its Down-Regulation during Drosophila Embryogenesis Is Required for the Arrest of Cell Proliferation." *Cell*. doi:10.1016/0092-8674(94)90239-9.

Kockel, L., J. Zeitlinger, L. M. Staszewski, M. Mlodzik, and D. Bohmann. 1998. "Jun in Drosophila Development: Redundant and Nonredundant Functions and Regulation by Two MAPK Signal Transduction Pathways." *Genes and Development* 12 (3): 447. doi:10.1101/gad.12.3.447.

Kockel, Lutz, Jason G. Homsy, and Dirk Bohmann. 2001. "Drosophila AP-1: Lessons from an Invertebrate." *Oncogene* 20 (19 REV. ISS. 2): 2347–64. doi:10.1038/sj.onc.1204300.

Kockel, Lutz, Julia Zeitlinger, Lena M. Staszewski, Marek Mlodzik, and Dirk Bohmann. 1997. "Jun in Drosophila Development: Redundant and Nonredundant Functions and Regulation by Two Mapk Signal Transduction Pathways." *Genes and Development*. doi:10.1101/gad.11.13.1748.

Koenecke, Nina, Jeff Johnston, Qiye He, Samuel Meier, and Julia Zeitlinger. 2017. "Drosophila Poised Enhancers Are Generated during Tissue Patterning with the Help of Repression." *Genome Research* 27 (1): 64–74. doi:10.1101/gr.209486.116.

Kofler, Michael, Pam Speight, Darby Little, Caterina Di Ciano-Oliveira, Katalin Szászi, and András Kapus. 2018. "Mediated Nuclear Import and Export of TAZ and the Underlying Molecular Requirements." *Nature Communications* 9 (1). doi:10.1038/s41467-018-07450-0.

Kondo, Shu, and Ryu Ueda. 2013. "Highly Improved Gene Targeting by Germline-Specific Cas9 Expression in Drosophila." *Genetics* 195 (3): 715–21. doi:10.1534/genetics.113.156737.

Koo, Ja Hyun, Steven W. Plouffe, Zhipeng Meng, Da Hye Lee, Di Yang, Dae Sik Lim, Cun Yu Wang, and Kun Liang Guan. 2020. "Induction of AP-1 by YAP/TAZ Contributes to Cell Proliferation and Organ Growth." *Genes & Development* 34 (1–2): 72–86. doi:10.1101/gad.331546.119.

Koontz, Laura M., Yi Liu-Chittenden, Feng Yin, Yonggang Zheng, Jianzhong Yu, Bo Huang, Qian Chen, Shian Wu, and Duojia Pan. 2013. "The Hippo Effector Yorkie Controls

Normal Tissue Growth by Antagonizing Scalloped-Mediated Default Repression." *Developmental Cell* 25 (4). Elsevier: 388–401. doi:10.1016/j.devcel.2013.04.021.

Kulkarni, Aishwarya, Matthew T. Chang, Joseph H.A. Vissers, Anwesha Dey, and Kieran F. Harvey. 2020. "The Hippo Pathway as a Driver of Select Human Cancers." *Trends in Cancer*. doi:10.1016/j.trecan.2020.04.004.

Kulshammer, Eva, Juliane Mundorf, Merve Kilinc, Peter Frommolt, Prerana Wagle, and Mirka Uhlirova. 2015. "Interplay among Drosophila Transcription Factors Ets21c, Fos and Ftz-F1 Drives JNK-Mediated Tumor Malignancy." *DMM Disease Models and Mechanisms* 8 (10): 1279–93. doi:10.1242/dmm.020719.

Külshammer, Eva, and Mirka Uhlirova. 2013. "The Actin Cross-Linker Filamin/Cheerio Mediates Tumor Malignancy Downstream of JNK Signaling." *Journal of Cell Science* 126 (4): 927–38. doi:10.1242/jcs.114462.

Kumar, Justin P. 2018. "The Fly Eye: Through the Looking Glass." *Dev Dyn.* 247 (1): 111–23. doi:10.1002/dvdy.24585.

Laborde, E. 2010. "Glutathione Transferases as Mediators of Signaling Pathways Involved in Cell Proliferation and Cell Death." *Cell Death and Differentiation* 17 (9). Nature Publishing Group: 1373–80. doi:10.1038/cdd.2010.80.

Lai, Zhi Chun, Xiaomu Wei, Takeshi Shimizu, Edward Ramos, Margaret Rohrbaugh, Nikolas Nikolaidis, Li Lun Ho, and Ying Li. 2005. "Control of Cell Proliferation and Apoptosis by Mob as Tumor Suppressor, Mats." *Cell* 120 (5): 675–85. doi:10.1016/j.cell.2004.12.036.

Lee, Kwang Pyo, Joo Hyeon Lee, Tae Shin Kim, Tack Hoon Kim, Hee Dong Park, Jin Seok Byun, Min Chul Kim, et al. 2010. "The Hippo-Salvador Pathway Restrains Hepatic Oval Cell Proliferation, Liver Size, and Liver Tumorigenesis." *Proceedings of the National Academy of Sciences of the United States of America* 107 (18): 8248–53. doi:10.1073/pnas.0912203107.

Lei, Q.-Y., H. Zhang, B. Zhao, Z.-Y. Zha, F. Bai, X.-H. Pei, S. Zhao, Y. Xiong, and K.-L. Guan. 2008. "TAZ Promotes Cell Proliferation and Epithelial-Mesenchymal Transition and Is Inhibited by the Hippo Pathway." *Molecular and Cellular Biology* 28 (7): 2426–36. doi:10.1128/mcb.01874-07.

Lieberman-aiden, Erez, Nynke L Van Berkum, Louise Williams, Maxim Imakaev, Tobias

Ragoczy, Agnes Telling, Ido Amit, et al. 2009. "Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome." *Science* 33292 (October): 289–94.

Lihua J Zhu, Claude Gazin, Nathan D Lawson, Hervé Pagès, Simon M Lin, David S Lapointe, and Michael R Green. 2010. "ChIPpeakAnno: A Bioconductor Package to Annotate ChIP-Seq and ChIP-Chip Data." *BMC Bioinformatics* 11: 237–47. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-237.

Ling, Jinli, Raphaëlle Dubruille, and Patrick Emery. 2012. "KAYAK-α Modulates Circadian Transcriptional Feedback Loops in Drosophila Pacemaker Neurons." *Journal of Neuroscience* 32 (47): 16959–70. doi:10.1523/JNEUROSCI.1888-12.2012.

Liu, Xiangfan, Huapeng Li, Mihir Rajurkar, Qi Li, Jennifer L. Cotton, Jianhong Ou, Lihua J. Zhu, et al. 2016. "Tead and AP1 Coordinate Transcription and Motility." *Cell Reports* 14 (5): 1169–80. doi:10.1016/j.celrep.2015.12.104.

Llano, Elena, Geza Adam, Alberto M. Pendás, Víctor Quesada, Luis M. Sánchez, Iñigo Santamaría, Stéphane Noselli, Carlos López-Otín, and Universidad De Oviedo. 2002. "Structural and Enzymatic Characterization of Drosophila Dm2-MMP, a Membrane-Bound Matrix Metalloproteinase with Tissue-Specific Expression." *Journal of Biological Chemistry* 277 (26): 23321–29. doi:10.1074/jbc.M200121200.

Luo, Xi, Oscar Puig, Joogyung Hyun, Dirk Bohmann, and Heinrich Jasper. 2007. "Foxo and Fos Regulate the Decision between Cell Death and Survival in Response to UV Irradiation." *EMBO Journal* 26 (2): 380–90. doi:10.1038/sj.emboj.7601484.

Ma, Xianjue, Yujun Chen, Wenyan Xu, Nana Wu, Maoquan Li, Ying Cao, Shian Wu, Qiutang Li, and Lei Xue. 2015. "Impaired Hippo Signaling Promotes Rho1-JNK-Dependent Growth." *Proceedings of the National Academy of Sciences of the United States of America* 112 (4): 1065–70. doi:10.1073/pnas.1415020112.

Ma, Yiqin, Daniel J. McKay, and Laura Buttitta. 2019. "Changes in Chromatin Accessibility Ensure Robust Cell Cycle Exit in Terminally Differentiated Cells." *PLoS Biology* 17 (9): 1–29. doi:10.1371/journal.pbio.3000378.

Maglic, Dejan, Karin Schlegelmilch, Antonella FM Dost, Riccardo Panero, Michael T Dill, Raffaele A Calogero, and Fernando D Camargo. 2018. "YAP-TEAD Signaling Promotes Basal Cell Carcinoma Development via a C-JUN/AP1 Axis." *The EMBO Journal*.

doi:10.15252/embj.201798642.

Manning, Samuel A., Benjamin Kroeger, and Kieran F. Harvey. 2020. “The Regulation of Yorkie, YAP and TAZ: New Insights into the Hippo Pathway.” *Development (Cambridge, England)* 147 (8). doi:10.1242/dev.179069.

Manning, Samuel A, Lucas G Dent, Shu Kondo, Ziqing W Zhao, Nicolas Plachta, Kieran F Harvey, Samuel A Manning, et al. 2018. “Dynamic Fluctuations in Subcellular Localization of the Hippo Pathway Effector Yorkie In Vivo Report Dynamic Fluctuations in Subcellular Localization of the Hippo Pathway Effector Yorkie In Vivo.” *Current Biology* 28 (10). Elsevier Ltd.: 1651-1660.e4. doi:10.1016/j.cub.2018.04.018.

Manzanares, Miguel, and Tristan A. Rodriguez. 2013. “Development: Hippo Signalling Turns the Embryo inside Out.” *Current Biology*. doi:10.1016/j.cub.2013.05.064.

Mao, Yaopan, Cordelia Rauskolb, Eunjoo Cho, Wei Li Hu, Heather Hayter, Ginny Minihan, Flora N. Katz, and Kenneth D. Irvine. 2006. “Dachs: An Unconventional Myosin That Functions Downstream of Fat to Regulate Growth, Affinity and Gene Expression in Drosophila.” *Development*. doi:10.1242/dev.02427.

Marca, John E. La, and Helena E. Richardson. 2020. “Two-Faced: Roles of JNK Signalling During Tumourigenesis in the Drosophila Model.” *Frontiers in Cell and Developmental Biology* 8 (February): 1–20. doi:10.3389/fcell.2020.00042.

Markstein, Michele, Chrysoula Pitsouli, Christians Villalta, and Susan E Celniker. 2008. “Exploiting Position Effects and the Gypsy Retrovirus Insulator to Engineer Precisely Expressed Transgenes” 40 (4): 476–83.

Marshall, Owen J., and Andrea H. Brand. 2015. “Damidseq-Pipeline: An Automated Pipeline for Processing DamID Sequencing Datasets.” *Bioinformatics* 31 (20): 3371–73. doi:10.1093/bioinformatics/btv386.

Marshall, Owen J., Tony D. Southall, Seth W. Cheetham, and Andrea H. Brand. 2016. “Cell-Type-Specific Profiling of Protein-DNA Interactions without Cell Isolation Using Targeted DamID with next-Generation Sequencing.” *Nature Protocols* 11 (9). Nature Publishing Group: 1586–98. doi:10.1038/nprot.2016.084.

Martin-Blanco, Enrique, Alexandra Gampel, Jenny Ring, Kanwar Virdee, Nikolai Kirov, Aviva M Tolkovsky, and Alfonso Martinez-arias. 1998. “Puckered Encodes a Phosphatase That Mediates a Feedback Loop Regulating JNK Activity during Dorsal

Closure in Drosophila," 557–70.

Martin, Marcel. 2011. "Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads." *EMBnet.Journal* 17 (1): 10–12.

Maugeri-Saccà, Marcello, Maddalena Barba, Laura Pizzuti, Patrizia Vici, Luigi Di Lauro, Rosanna Dattilo, Ilio Vitale, Monica Bartucci, Marcella Mottolese, and Ruggero De Maria. 2015. "The Hippo Transducers TAZ and YAP in Breast Cancer: Oncogenic Activities and Clinical Implications." *Expert Reviews in Molecular Medicine*. doi:10.1017/erm.2015.12.

Maugeri-Saccà, Marcello, and Ruggero De Maria. 2018. "The Hippo Pathway in Normal Development and Cancer." *Pharmacology and Therapeutics* 186. Elsevier Inc.: 60–72. doi:10.1016/j.pharmthera.2017.12.011.

Mcewen, Donald G, and Mark Peifer. 2005. "Puckered , a Drosophila MAPK Phosphatase , Ensures Cell Viability by Antagonizing JNK-Induced Apoptosis," 3935–46. doi:10.1242/dev.01949.

Meng, Zhipeng, Toshiro Moroishi, and Kun Liang Guan. 2016. "Mechanisms of Hippo Pathway Regulation." *Genes and Development* 30 (1): 1–17. doi:10.1101/gad.274027.115.

Menut, Laurent, Thomas Vaccari, Heather Dionne, Joseph Hill, Geena Wu, and David Bilder. 2007. "A Mosaic Genetic Screen for Drosophila Neoplastic Tumor Suppressor Genes Based on Defective Pupation." *Genetics* 177 (3): 1667–77. doi:10.1534/genetics.107.078360.

Metcalf, D. 1963. "The Autonomous Behaviour of Normal Thymus Grafts." *The Australian Journal of Experimental Biology and Medical Science* 41 (August). Australia: SUPPL437-47. doi:10.1038/icb.1963.64.

Mirzoyan, Zhasmine, Manuela Sollazzo, Mariateresa Allocca, Alice Maria Valenza, Daniela Grifoni, and Paola Bellosta. 2019. "Drosophila Melanogaster: A Model Organism to Study Cancer." *Frontiers in Genetics* 10 (March): 1–16. doi:10.3389/fgene.2019.00051.

Misra, Jyoti R., and Kenneth D. Irvine. 2018. "The Hippo Signaling Network and Its Biological Functions." *Annual Review of Genetics* 52 (1): 65–87. doi:10.1146/annurev-genet-120417-031621.

Mohan, M., H.-M. Herz, E. R. Smith, Y. Zhang, J. Jackson, M. P. Washburn, L. Florens, J. C.

Eissenberg, and A. Shilatifard. 2011. "The COMPASS Family of H3K4 Methylases in Drosophila." *Molecular and Cellular Biology* 31 (21): 4310–18. doi:10.1128/mcb.06092-11.

Monroe, Tanner O., Matthew C. Hill, Yuka Morikawa, John P. Leach, Todd Heallen, Shuyi Cao, Peter H.L. Krijger, et al. 2019. "YAP Partially Reprograms Chromatin Accessibility to Directly Induce Adult Cardiogenesis In Vivo." *Developmental Cell* 48 (6). Elsevier Inc.: 765-779.e7. doi:10.1016/j.devcel.2019.01.017.

Moon, Sungho, So Yeon Park, and Hyun Woo Park. 2018. "Regulation of the Hippo Pathway in Cancer Biology." *Cellular and Molecular Life Sciences* 75 (13): 2303–19. doi:10.1007/s00018-018-2804-1.

Moroishi, Toshiro, Hyun Woo Park, Baodong Qin, Qian Chen, Zhipeng Meng, Steven W. Plouffe, Koji Taniguchi, et al. 2015. "A YAP/TAZ-Induced Feedback Mechanism Regulates Hippo Pathway Homeostasis." *Genes and Development* 29 (12): 1271–84. doi:10.1101/gad.262816.115.

Moye-Rowley, W. S., K. D. Harshman, and C. S. Parker. 1989. "Yeast YAP1 Encodes a Novel Form of the Jun Family of Transcriptional Activator Proteins." *Genes & Development* 3 (3): 283–92. doi:10.1101/gad.3.3.283.

Mukherjee, T., I. Choi, and Utpal Banerjee. 2012. "Genetic Analysis of Fibroblast Growth Factor Signaling in the Drosophila Eye." *G3: Genes, Genomes, Genetics* 2 (1): 23–28. doi:10.1534/g3.111.001495.

Nagaraj, Raghavendra, Shubha Gururaja-Rao, Kevin T. Jones, Matthew Slattery, Nicolas Negre, Daniel Braas, Heather Christofk, Kevin P. White, Richard Mann, and Utpal Banerjee. 2012. "Control of Mitochondrial Structure and Function by the Yorkie/YAP Oncogenic Pathway." *Genes and Development* 26 (18): 2027–37. doi:10.1101/gad.183061.111.

Neto-Silva, Ricardo M., Simon de Beco, and Laura A. Johnston. 2010. "Evidence for a Growth Stabilizing Regulatory Feedback Mechanism between Myc and Yorkie, the Drosophila Homolog of Yap." *Dev Cell* 19 (4): 507–20. doi:10.1016/j.devcel.2010.09.009.

Nevil, Markus, Eliana R. Bondra, Katharine N. Schulz, Tommy Kaplan, and Melissa M. Harrison. 2017. "Stable Binding of the Conserved Transcription Factor Grainy Head to Its Target Genes throughout Drosophila Melanogaster Development." *Genetics* 205 (2):

605–20. doi:10.1534/genetics.116.195685.

Newsome, Timothy P., Bengt Åsling, and Barry J. Dickson. 2000. “Analysis of Drosophila Photoreceptor Axon Guidance in Eye-Specific Mosaics.” *Development* 127 (4): 851–60.

Nicolay, Brandon N., Battuya Bayarmagnai, Abul B.M.M.K. Islam, Nuria Lopez-Bigas, and Maxim V. Frolov. 2011. “Cooperation between DE2F1 and Yki/Sd Defines a Distinct Transcriptional Program Necessary to Bypass Cell Cycle Exit.” *Genes and Development* 25 (4): 323–35. doi:10.1101/gad.1999211.

Nishioka, Noriyuki, Ken ichi Inoue, Kenjiro Adachi, Hiroshi Kiyonari, Mitsunori Ota, Amy Ralston, Norikazu Yabuta, et al. 2009. “The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass.” *Developmental Cell* 16 (3). Elsevier Ltd: 398–410. doi:10.1016/j.devcel.2009.02.003.

Nolo, Riitta, Clayton M. Morrison, Chunyao Tao, Xinwei Zhang, and Georg Halder. 2006. “The Bantam MicroRNA Is a Target of the Hippo Tumor-Suppressor Pathway.” *Current Biology* 16 (19): 1895–1904. doi:10.1016/j.cub.2006.08.057.

Nukuda, A., C. Sasaki, S. Ishihara, T. Mizutani, K. Nakamura, T. Ayabe, K. Kawabata, and H. Haga. 2015. “Stiff Substrates Increase YAP-Signaling-Mediated Matrix Metalloproteinase-7 Expression.” *Oncogenesis* 4 (April). Nature Publishing Group: 1–11. doi:10.1038/oncsis.2015.24.

Oh, Hyangyee, and Kenneth D. Irvine. 2008a. “In Vivo Regulation of Yorkie Phosphorylation and Localization.” *Development* 135 (6): 1081–88. doi:10.1242/dev.015255.

———. 2009. “In Vivo Analysis of Yorkie Phosphorylation Sites.” *Oncogene* 28 (17): 1916–27. doi:10.1038/jid.2014.371.

———. 2011. “Cooperative Regulation of Growth by Yorkie and Mad through Bantam.” *Developmental Cell* 20 (1). Elsevier: 109–22. doi:10.1016/j.devcel.2010.12.002.

Oh, Hyangyee, and Kenneth D Irvine. 2008b. “In Vivo Regulation of Yorkie Phosphorylation and Localization.” *Development* 135 (6): 1081–88. doi:10.1242/dev.015255.

Oh, Hyangyee, B.V.V.G. Reddy, and Kenneth D. Irvine. 2009. “Phosphorylation-Independent Repression of Yorkie in Fat-Hippo Signaling.” *Dev Biol* 335 (1): 1188–97. doi:10.1038/jid.2014.371.

Oh, Hyangyee, Matthew Slattery, Lijia Ma, Alex Crofts, Kevin P. White, Richard S. Mann, and Kenneth D. Irvine. 2013. "Genome-Wide Association of Yorkie with Chromatin and Chromatin-Remodeling Complexes." *Cell Reports* 3 (2). The Authors: 309–18. doi:10.1016/j.celrep.2013.01.008.

Oh, Hyangyee, Matthew Slattery, Lijia Ma, Kevin P. White, Richard S. Mann, and Kenneth D. Irvine. 2014a. "Yorkie Promotes Transcription by Recruiting a Histone Methyltransferase Complex." *Cell Reports* 8 (2). The Authors: 449–59. doi:10.1016/j.celrep.2014.06.017.

Oh, Hyangyee, Matthew Slattery, Lijia Ma, Kevin P. White, Richard S. Mann, and Kenneth D Irvine. 2014b. "Yorkie Promotes Transcription by Recruiting a Histone Methyltransferase Complex Hyangyee." *Cell Reports* 8: 449–59. doi:10.1371/journal.pgen.1003753.

Organista, María F., Mercedes Martín, Jesus M. de Celis, Rosa Barrio, Ana López-Varea, Nuria Esteban, Mar Casado, and Jose F. de Celis. 2015. "The Spalt Transcription Factors Generate the Transcriptional Landscape of the Drosophila Melanogaster Wing Pouch Central Region." *PLoS Genetics* 11 (8): 1–27. doi:10.1371/journal.pgen.1005370.

Pagliarini, Raymond A., and Tian Xu. 2003. "A Genetic Screen in Drosophila for Metastatic Behavior." *Science* 302 (5648): 1227–31. doi:10.1126/science.1088474.

Pan, Duojia. 2010. "The Hippo Signaling Pathway and Cancer." *The Hippo Signaling Pathway in Development and Cancer* 19 (4): 491–505. doi:10.1007/978-1-4614-6220-0.

Pantalacci, Sophie, Nicolas Tapon, and Pierre Léopold. 2003. "The Salvador Partner Hippo Promotes Apoptosis and Cell-Cycle Exit in Drosophila." *Nature Cell Biology* 5 (10): 921–27. doi:10.1038/ncb1051.

Pascual, Justine, Jelle Jacobs, Leticia Sansores-Garcia, Malini Natarajan, Julia Zeitlinger, Stein Aerts, Georg Halder, and Fisun Hamaratoglu. 2017. "Hippo Reprograms the Transcriptional Response to Ras Signaling." *Developmental Cell* 42 (6). Elsevier Inc.: 667-680.e4. doi:10.1016/j.devcel.2017.08.013.

Pearson, Joseph C., Michelle T. Juarez, Myungjin Kim, and William McGinnis. 2009. "Multiple Transcription Factor Codes Activate Epidermal Wound-Response Genes in Drosophila." *Proceedings of the National Academy of Sciences of the United States of America* 106 (7): 2224–29. doi:10.1073/pnas.0810219106.

Peng, H. Wayne, Matthew Slattery, and Richard S. Mann. 2009. "Transcription Factor Choice in the Hippo Signaling Pathway: Homothorax and Yorkie Regulation of the MicroRNA

Bantam in the Progenitor Domain of the Drosophila Eye Imaginal Disc." *Genes and Development* 23 (19): 2307–19. doi:10.1101/gad.1820009.

Perkins, K. K., A. Admon, N. Patel, and R. Tjian. 1990. "The Drosophila Fos-Related AP-1 Protein Is a Developmentally Regulated Transcription Factor." *Genes and Development* 4 (5): 822–34. doi:10.1101/gad.4.5.822.

Perra, Andrea, Marta Anna Kowalik, Elena Ghiso, Giovanna Maria Ledda-Columbano, Luca Di Tommaso, Maria Maddalena Angioni, Carlotta Raschioni, et al. 2014. "YAP Activation Is an Early Event and a Potential Therapeutic Target in Liver Cancer Development." *Journal of Hepatology*. doi:10.1016/j.jhep.2014.06.033.

Poças, Gonçalo M, Pedro M Domingos, and Christen K Mirth. 2020. "Quantification of Macronutrients Intake in a Thermogenetic Neuronal Screen Using Drosophila Larvae Quantification of Macronutrients Intake in a Thermogenetic Neuronal Screen Using Drosophila Larvae," no. May. doi:10.3791/61323.

Qing, Yun, Feng Yin, Wei Wang, Yonggang Zheng, Pengfei Guo, Frederick Schozer, Hua Deng, and Duojia Pan. 2014. "The Hippo Effector Yorkie Activates Transcription by Interacting with a Histone Methyltransferase Complex through Ncoa6." *ELife* 3 (July2014): 1–13. doi:10.7554/eLife.02564.

Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26 (6): 841–42. doi:10.1093/bioinformatics/btq033.

Raabe, Thomas. 2000. "The Sevenless Signaling Pathway: Variations of a Common Theme." *Biochimica et Biophysica Acta - Molecular Cell Research* 1496 (2–3): 151–63. doi:10.1016/S0167-4889(00)00020-3.

Rebers, John E., and Judith H. Willis. 2001. "A Conserved Domain in Arthropod Cuticular Proteins Binds Chitin." *Insect Biochemistry and Molecular Biology* 31 (11): 1083–93. doi:10.1016/S0965-1748(01)00056-X.

Reginensi, Antoine, Rizaldy P. Scott, Alex Gregorieff, Mazdak Bagherie-Lachidan, Chaeuk Chung, Dae Sik Lim, Tony Pawson, Jeff Wrana, and Helen McNeill. 2013. "Yap- and Cdc42-Dependent Nephrogenesis and Morphogenesis during Mouse Kidney Development." *PLoS Genetics*. doi:10.1371/journal.pgen.1003380.

Ren, Fangfang, Bing Wang, Tao Yue, Eun Young Yun, Y. Tony Ip, and Jin Jiang. 2010. "Hippo

Signaling Regulates Drosophila Intestine Stem Cell Proliferation through Multiple Pathways." *Proceedings of the National Academy of Sciences of the United States of America* 107 (49): 21064–69. doi:10.1073/pnas.1012759107.

Ren, Fangfang, Lei Zhang, and Jin Jiang. 2010. "Hippo Signaling Regulates Yorkie Nuclear Localization and Activity through 14-3-3 Dependent and Independent Mechanisms." *Developmental Biology* 337 (2). Elsevier Inc.: 303–12. doi:10.1016/j.ydbio.2009.10.046.

Riesgo-Escovar, Juan R., and Ernst Hafen. 1997a. "Common and Distinct Roles of DFos and DJun during Drosophila Development." *Science* 278 (5338): 669–72. doi:10.1126/science.278.5338.669.

———. 1997b. "Drosophila Jun Kinase Regulates Expression of Decapentaplegic via the Ets-Domain Protein Aop and the Ap-1 Transcription Factor Djun during Dorsal Closure." *Genes and Development* 11 (13): 1717–27. doi:10.1101/gad.11.13.1717.

Roignant, Jean-Yves, and Jessica E. Treisman. 2009. "Pattern Formation in Drosophila." *Seminars in Cell Biology* 53 (6): 795–804. doi:10.1387/ijdb.072483jr.Pattern.

Rosenbluh, Joseph, Deepak Nijhawan, Andrew G. Cox, Xingnan Li, James T. Neal, Eric J. Schafer, Travis I. Zack, et al. 2012. "β-Catenin-Driven Cancers Require a YAP1 Transcriptional Complex for Survival and Tumorigenesis." *Cell*. doi:10.1016/j.cell.2012.11.026.

Rouleau, Guy A., Philippe Merel, Mohini Lutchman, Marc Sanson, Jessica Zucman, Claude Marineau, Khé Hoang-Xuan, et al. 1993. "Alteration in a New Gene Encoding a Putative Membrane-Organizing Protein Causes Neuro-Fibromatosis Type 2." *Nature*. doi:10.1038/363515a0.

Rousset, Raphaël, Fabrice Carballès, Nadège Parassol, Sébastien Schaub, Delphine Cérézo, and Stéphane Noselli. 2017. "Signalling Crosstalk at the Leading Edge Controls Tissue Closure Dynamics in the Drosophila Embryo." *PLoS Genetics* 13 (2): 1–26. doi:10.1371/journal.pgen.1006640.

Sansores-Garcia, Leticia, Mardelle Atkins, Ivan M. Moya, Maria Shahmoradgoli, Chunyao Tao, Gordon B. Mills, and Georg Halder. 2013. "Mask Is Required for the Activity of the Hippo Pathway Effector Yki/YAP." *Current Biology* 23 (3). Elsevier Ltd: 229–35. doi:10.1016/j.cub.2012.12.033.

Schaub, Christoph, Marcel Rose, and Manfred Frasch. 2019. "Yorkie and JNK Revert

Syncytial Muscles into Myoblasts during Org-1-Dependent Lineage Reprogramming." *The Journal of Cell Biology* 218 (11): 3572–82. doi:10.1083/jcb.201905048.

Schlegelmilch, Karin, Morvarid Mohseni, Oktay Kirak, Jan Pruszak, J. Renato Rodriguez, Dawang Zhou, Bridget T. Kreger, et al. 2011. "Yap1 Acts Downstream of α-Catenin to Control Epidermal Proliferation." *Cell*. doi:10.1016/j.cell.2011.02.031.

Shi, Ming, Dan Liu, Huijun Duan, Caili Han, Bo Wei, Lu Qian, Changguo Chen, et al. 2010. "Catecholamine Up-Regulates MMP-7 Expression by Activating AP-1 and STAT3 in Gastric Cancer." *Molecular Cancer* 9 (1). BioMed Central Ltd: 269. doi:10.1186/1476-4598-9-269.

Shilo, Ben Zion. 2014. "The Regulation and Functions of MAPK Pathways in Drosophila." *Methods* 68 (1). Elsevier Inc.: 151–59. doi:10.1016/j.ymeth.2014.01.020.

Shreberk-Shaked, Michal, and Moshe Oren. 2019. "New Insights into YAP/TAZ Nucleo-Cytoplasmic Shuttling: New Cancer Therapeutic Opportunities?" *Molecular Oncology* 13 (6): 1335–41. doi:10.1002/1878-0261.12498.

Sidor, Clara M., Ruth Brain, and Barry J. Thompson. 2013. "Mask Proteins Are Cofactors of Yorkie/YAP in the Hippo Pathway." *Current Biology* 23 (3). Elsevier Ltd: 223–28. doi:10.1016/j.cub.2012.11.061.

Silva, Elizabeth, Yonit Tsatskis, Laura Gardano, Nic Tapon, and Helen McNeill. 2006. "The Tumor-Suppressor Gene Fat Controls Tissue Growth Upstream of Expanded in the Hippo Signaling Pathway." *Current Biology* 16 (21): 2081–89. doi:10.1016/j.cub.2006.09.004.

Simmonds, Andrew J., Xiaofeng Liu, Kelly H. Soanes, Henry M. Krause, Kenneth D. Irvine, and John B. Bell. 1998. "Molecular Interactions between Vestigial and Scalloped Promote Wing Formation in Drosophila." *Genes and Development* 12 (24). Cold Spring Harbor Laboratory Press: 3815–20. doi:10.1101/gad.12.24.3815.

Skibinski, Adam, Jerrica L Breindel, Aleix Prat, Patricia Galván, Elizabeth Smith, Andreas Rolfs, Piyush B Gupta, Joshua Labaer, and Charlotte Kuperwasser. 2014. "The Hippo Transducer TAZ Interacts with the SWI/SNF Complex to Regulate Breast Epithelial Lineage Commitment." *Cell Reports* 6 (6): 1059–72. doi:10.1016/j.celrep.2014.02.038.The.

Slattery, Matthew, Roumen Voutev, Lijia Ma, Nicolas Nègre, Kevin P. White, and Richard S. Mann. 2013. "Divergent Transcriptional Regulatory Logic at the Intersection of Tissue

Growth and Developmental Patterning." *PLoS Genetics* 9 (9): 22–25. doi:10.1371/journal.pgen.1003753.

Slattery, Matthew, Roumen Voutev, Lijia Ma, Nicolas Nègre, Kevin P. White, Richard S. Mann, Louis Gervais, et al. 2013. "Genome-Wide Association of Yorkie with Chromatin and Chromatin-Remodeling Complexes." *Developmental Cell* 3 (2). Elsevier Inc.: 667-680.e4. doi:10.1016/j.devcel.2017.08.013.

Snigdha, Kirti, Karishma Sanjay Gangwani, Gauri Vijay Lapalikar, Amit Singh, and Madhuri Kango-Singh. 2019. "Hippo Signaling in Cancer: Lessons from Drosophila Models." *Frontiers in Cell and Developmental Biology* 7 (May): 1–16. doi:10.3389/fcell.2019.00085.

Song, Hai, Kinglun Kingston Mak, Lilia Topol, Kangsun Yun, Jianxin Hu, Lisa Garrett, Yongbin Chen, et al. 2010. "Mammalian Mst1 and Mst2 Kinases Play Essential Roles in Organ Size Control and Tumor Suppression." *Proceedings of the National Academy of Sciences of the United States of America* 107 (4): 1431–36. doi:10.1073/pnas.0911409107.

Song, Shilin, Héctor Herranz, and Stephen M. Cohen. 2017. "The Chromatin Remodeling BAP Complex Limits Tumor-Promoting Activity of the Hippo Pathway Effector Yki to Prevent Neoplastic Transformation in Drosophila Epithelia." *DMM Disease Models and Mechanisms* 10 (10): 1201–9. doi:10.1242/dmm.030122.

Southall, Tony D., Katrina S. Gold, Boris Egger, Catherine M. Davidson, Elizabeth E. Caygill, Owen J. Marshall, and Andrea H. Brand. 2013. "Cell-Type-Specific Profiling of Gene Expression and Chromatin Binding without Cell Isolation: Assaying RNA Pol II Occupancy in Neural Stem Cells." *Developmental Cell* 26 (1). The Authors: 101–12. doi:10.1016/j.devcel.2013.05.020.

Srivastava, Ajay, Andrew J. Simmonds, Ankush Garg, Leif Fossheim, Shelagh D. Campbell, and John B. Bell. 2004. "Molecular and Functional Analysis of Scalloped Recessive Lethal Alleles in Drosophila Melanogaster." *Genetics* 166 (4): 1833–43. doi:10.1534/genetics.166.4.1833.

Stark, Mary B. 1918. "An Hereditary Tumor in the Fruit Fly Drosophila." *Journal of Cancer Research*. doi:10.1158/jcr.1918.279.

Stein, Claudia, Guglielmo Roma, Sebastian Bergling, Ieuan Clay, Alexandra Ruchti, Claudia

Agarinis, Tobias Schmelzle, Tewis Bouwmeester, Dirk Schübeler, and Andreas Bauer. 2015. "YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers," 1–29. doi:10.1371/journal.pgen.1005465.

Strano, Sabrina, Olimpia Monti, Natalia Pediconi, Alessia Baccarini, Giulia Fontemaggi, Eleonora Lapi, Fiamma Mantovani, et al. 2005. "The Transcriptional Coactivator Yes-Associated Protein Drives P73 Gene-Target Specificity in Response to DNA Damage." *Molecular Cell*. doi:10.1016/j.molcel.2005.04.008.

Sumabat, Taryn M., Melanie I. Worley, Brett J. Pellock, Justin A. Bosch, and Iswar K. Hariharan. 2019. "The Transcriptional Co-Repressor CtBP Is a Negative Regulator of Growth That Antagonizes the Yorkie and JNK/AP-1 Pathways." *BioRxiv*. doi:10.1017/CBO9781107415324.004.

Sun, Xiaohan, Yan Ding, Meixiao Zhan, Yan Li, Dongqing Gao, Guiping Wang, Yang Gao, et al. 2019. "Usp7 Regulates Hippo Pathway through Deubiquitinating the Transcriptional Coactivator Yorkie." *Nature Communications* 10 (1). Springer US: 1–16. doi:10.1038/s41467-019-08334-7.

Sun, Yubing, Koh Meng Aw Yong, Luis G. Villa-Diaz, Xiaoli Zhang, Weiqiang Chen, Renee Philson, Shinuo Weng, Haoxing Xu, Paul H. Krebsbach, and Jianping Fu. 2014. "Hippo/YAP-Mediated Rigidity-Dependent Motor Neuron Differentiation of Human Pluripotent Stem Cells." *Nature Materials* 13 (6): 599–604. doi:10.1038/nmat3945.

Szklarczyk, Damian, Andrea Franceschini, Michael Kuhn, Milan Simonovic, Alexander Roth, Pablo Minguez, Tobias Doerks, et al. 2011. "The STRING Database in 2011: Functional Interaction Networks of Proteins, Globally Integrated and Scored." *Nucleic Acids Research* 39 (SUPPL. 1): 561–68. doi:10.1093/nar/gkq973.

Tanas, M. R., S. Ma, F. O. Jadaan, C. K.Y. Ng, B. Weigelt, J. S. Reis-Filho, and B. P. Rubin. 2016. "Mechanism of Action of a WWTR1(TAZ)-CAMTA1 Fusion Oncoprotein." *Oncogene* 35 (7): 929–38. doi:10.1038/onc.2015.148.

Tanas, Munir R., Andrea Sboner, Andre M. Oliveira, Michele R. Erickson-Johnson, Jessica Hespelt, Philip J. Hanwright, John Flanagan, et al. 2011. "Identification of a Disease-Defining Gene Fusion in Epithelioid Hemangioendothelioma." *Science Translational Medicine*. doi:10.1126/scitranslmed.3002409.

Tapon, Nicolas, Kieran F. Harvey, Daphne W. Bell, Doke C.R. Wahrer, Taryn A. Schiripo,

Daniel A. Haber, and Iswar K. Hariharan. 2002. “Salvador Promotes Both Cell Cycle Exit and Apoptosis in Drosophila and Is Mutated in Human Cancer Cell Lines.” *Cell* 110 (4): 467–78. doi:10.1016/S0092-8674(02)00824-3.

Thompson, Barry J., and Stephen M. Cohen. 2006. “The Hippo Pathway Regulates the Bantam MicroRNA to Control Cell Proliferation and Apoptosis in Drosophila.” *Cell* 126 (4): 767–74. doi:10.1016/j.cell.2006.07.013.

Tolhuis, Bas, Inhua Muijrers, Elzo De Wit, Hans Teunissen, Wendy Talhout, Bas Van Steensel, and Maarten Van Lohuizen. 2006. “Genome-Wide Profiling of PRC1 and PRC2 Polycomb Chromatin Binding in Drosophila Melanogaster.” *Nature Genetics* 38 (6): 694–99. doi:10.1038/ng1792.

Treisman, Jessica E. 2013. “Retinal Differentiation in Drosophila.” *Wiley Interdiscip Rev Dev Biol.* 2 (4): 545–557.

Tu, Chen Pei D., and Bünyamin Akgül. 2005. “Drosophila Glutathione S-Transferases.” *Methods in Enzymology* 401: 204–26. doi:10.1016/S0076-6879(05)01013-X.

Twitty, Victor, and Joseph Schwind. 1931. “The Growth of Eyes and Limbs Transplanted Heteroplastically between Two Species of Amblystoma.” *Journal of Experimental Zoology* 59 (1): 61–86.

Udan, Ryan S., Madhuri Kango-Singh, Riitta Nolo, Chunyao Tao, and Georg Halder. 2003a. “Hippo Promotes Proliferation Arrest and Apoptosis in the Salvador/Warts Pathway.” *Nature Cell Biology* 5 (10): 914–20. doi:10.1038/ncb1050.

Udan, Ryan S, Madhuri Kango-Singh, Riitta Nolo, Chunyao Tao, and Georg Halder. 2003b. “Hippo Promotes Proliferation Arrest and Apoptosis in the Salvador/Warts Pathway.” *Nature Cell Biology* 5 (10): 914–20. doi:10.1038/ncb1050.

Uhlirova, Mirka, and Dirk Bohmann. 2006. “JNK- and Fos-Regulated Mmp1 Expression Cooperates with Ras to Induce Invasive Tumors in Drosophila.” *EMBO Journal* 25 (22): 5294–5304. doi:10.1038/sj.emboj.7601401.

Vassilev, Alex, Kotaro J Kaneko, Hongjun Shu, Yingming Zhao, and Melvin L Depamphilis. 2001. “TEAD / TEF Transcription Factors Utilize the Activation Domain of YAP65 , a Src / Yes-Associated Protein Localized in the Cytoplasm,” 1229–41. doi:10.1101/gad.888601.required.

Verfaillie, Annelien, Hana Imrichova, Zeynep Kalender Atak, Michael Dewaele, Florian

Rambow, Gert Hulselmans, Valerie Christiaens, et al. 2015. “Decoding the Regulatory Landscape of Melanoma Reveals TEADS as Regulators of the Invasive Cell State.” *Nature Communications*. doi:10.1038/ncomms7683.

Villegas, Santiago Nahuel. 2019. “One Hundred Years of Drosophila Cancer Research: No Longer in Solitude.” *DMM Disease Models and Mechanisms* 12 (4). doi:10.1242/dmm.039032.

Vissers, Joseph H.A., Lucas G. Dent, Colin M. House, Shu Kondo, and Kieran F. Harvey. 2020. “Pits and CtBP Control Tissue Growth in Drosophila Melanogaster with the Hippo Pathway Transcription Repressor Tgi.” *Genetics* 215 (1): 117–28. doi:10.1534/genetics.120.303147.

Vissers, Joseph H.A., Francesca Froldi, Jan Schröder, Anthony T. Papenfuss, Louise Y. Cheng, and Kieran F. Harvey. 2018. “The Scalloped and Nerfin-1 Transcription Factors Cooperate to Maintain Neuronal Cell Fate.” *Cell Reports* 25 (6): 1561-1576.e7. doi:10.1016/j.celrep.2018.10.038.

Vissers, Joseph H.A., Samuel A. Manning, Aishwarya Kulkarni, and Kieran F. Harvey. 2016. “A Drosophila RNAi Library Modulates Hippo Pathway-Dependent Tissue Growth.” *Nature Communications* 7. Nature Publishing Group: 1–6. doi:10.1038/ncomms10368.

Vogel, Maartje J., Daniel Peric-Hupkes, and Bas van Steensel. 2007. “Detection of in Vivo Protein - DNA Interactions Using DamID in Mammalian Cells.” *Nature Protocols* 2 (6): 1467–78. doi:10.1038/nprot.2007.148.

Wang, Chao, Meng-Xin Yin, Wei Wu, Liang Dong, Shimin Wang, Yi Lu, Jinjin Xu, et al. 2016. “Taiman Acts as a Coactivator of Yorkie in the Hippo Pathway to Promote Tissue Growth and Intestinal Regeneration.” *Cell Discovery* 2: 16006. doi:10.1038/celldisc.2016.6.

Wang, Shimin, Yi Lu, Meng Xin Yin, Chao Wang, Wei Wu, Jinhui Li, Wenqing Wu, et al. 2016. “Importin A1 Mediates Yorkie Nuclear Import via an N-Terminal Non-Canonical Nuclear Localization Signal.” *Journal of Biological Chemistry* 291 (15): 7926–37. doi:10.1074/jbc.M115.700823.

Wang, Susan L., Christine J. Hawkins, Soon Ji Yoo, H. Arno J. Müller, and Bruce A. Hay. 1999. “The Drosophila Caspase Inhibitor DIAP1 Is Essential for Cell Survival and Is Negatively Regulated by HID.” *Cell*. doi:10.1016/S0092-8674(00)81974-1.

Wang, Ting, Beibei Mao, Chi Cheng, Zhuangzhi Zou, Junling Gao, Yanglu Yang, Tong Lei, et al. 2018. "YAP Promotes Breast Cancer Metastasis by Repressing Growth Differentiation Factor-15." *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1864 (5). Elsevier: 1744–53. doi:10.1016/j.bbadis.2018.02.020.

Wei, Xiaomu, Takeshi Shimizu, and Zhi Chun Lai. 2007. "Mob as Tumor Suppressor Is Activated by Hippo Kinase for Growth Inhibition in Drosophila." *EMBO Journal* 26 (7): 1772–81. doi:10.1038/sj.emboj.7601630.

Weston, Claire R., and Roger J. Davis. 2007. "The JNK Signal Transduction Pathway." *Current Opinion in Cell Biology* 19 (2): 142–49. doi:10.1016/j.ceb.2007.02.001.

Willsey, Helen Rankin, Xiaoyan Zheng, Josécarlos Pastor-Pareja, A. Jeremy Willsey, Philip A. Beachy, and Tian Xu. 2016. "Localized JNK Signaling Regulates Organ Size during Development." *ELife* 5 (MARCH2016): 1–18. doi:10.7554/eLife.11491.

Winnebeck, Eva C., Craig D. Millar, and Guy R. Warman. 2010. "Why Does Insect RNA Look Degraded?" *Journal of Insect Science* 10 (159): 1–7. doi:10.1673/031.010.14119.

Wittkorn, Erika, Ankita Sarkar, Kristine Garcia, Madhuri Kango-Singh, and Amit Singh. 2015. "The Hippo Pathway Effector Yki Downregulates Wg Signaling to Promote Retinal Differentiation in the Drosophila Eye." *Development (Cambridge)* 142 (11): 2002–13. doi:10.1242/dev.117358.

Wolff, T., and D. F. Ready. 1991. "The Beginning of Pattern Formation in the Drosophila Compound Eye: The Morphogenetic Furrow and the Second Mitotic Wave." *Development* 113 (3): 841–50.

Wu, Hai, Meng C. Wang, and Dirk Bohmann. 2009. "JNK Protects Drosophila from Oxidative Stress by Trancriptionally Activating Autophagy." *Mechanisms of Development* 126 (8–9). Elsevier Ireland Ltd: 624–37. doi:10.1016/j.mod.2009.06.1082.

Wu, Ming, José Carlos Pastor-Pareja, and Tian Xu. 2010. "Interaction between RasV12 and Scribbled Clones Induces Tumour Growth and Invasion." *Nature* 463 (7280): 545–48. doi:10.1038/nature08702.

Wu, Shian, Jianbin Huang, Jixin Dong, and Duojia Pan. 2003. "Hippo Encodes a Ste-20 Family Protein Kinase That Restricts Cell Proliferation and Promotes Apoptosis in Conjunction with Salvador and Warts." *Cell* 114 (4): 445–56. doi:10.1016/S0092-8674(03)00549-X.

Wu, Shian, Y. Liu, Yonggang Zheng, Jixin Dong, and Duojia Pan. 2008. "The TEAD/TEF

Family Protein Scalloped Mediates Transcriptional Output of the Hippo Growth-Regulatory Pathway." *Developmental Cell* 14 (3): 388–98. doi:10.1016/j.devcel.2008.01.007.

Wu, Yen Chi, Kyu Sun Lee, Yan Song, Stephan Gehrke, and Bingwei Lu. 2017. "The Bantam MicroRNA Acts through Numb to Exert Cell Growth Control and Feedback Regulation of Notch in Tumor-Forming Stem Cells in the Drosophila Brain." *PLoS Genetics* 13 (5): 1–20. doi:10.1371/journal.pgen.1006785.

Xiao, Ling, Yuanhong Chen, Ming Ji, and Jixin Dong. 2011. "KIBRA Regulates Hippo Signaling Activity via Interactions with Large Tumor Suppressor Kinases." *Journal of Biological Chemistry*. doi:10.1074/jbc.M110.173468.

Xiao, Weifan, Jiayi Wang, Chao Ou, Yue Zhang, Lifang Ma, Wenhao Weng, Qiuhui Pan, and Fenyong Sun. 2013. "Mutual Interaction between YAP and C-Myc Is Critical for Carcinogenesis in Liver Cancer." *Biochemical and Biophysical Research Communications*. doi:10.1016/j.bbrc.2013.08.071.

Xin, Mei, Yuri Kim, Lillian B. Sutherland, Xiaoxia Qi, John McAnally, Robert J. Schwartz, James A. Richardson, Rhonda Bassel-Duby, and Eric N. Olson. 2011. "Development: Regulation of Insulin-like Growth Factor Signaling by Yap Governs Cardiomyocyte Proliferation and Embryonic Heart Size." *Science Signaling* 4 (196). doi:10.1126/scisignal.2002278.

Xu, Jiajie, Pamela J. Vanderzalm, Michael Ludwig, Ting Su, Sherzod A. Tokamov, and Richard G. Fehon. 2018. "Yorkie Functions at the Cell Cortex to Promote Myosin Activation in a Non-Transcriptional Manner." *Dev Cell* 46 (3): 271–84. doi:10.1016/j.devcel.2018.06.017.

Xu, M. Z., S. W. Chan, A. M. Liu, K. F. Wong, S. T. Fan, J. Chen, R. T. Poon, et al. 2011. "AXL Receptor Kinase Is a Mediator of YAP-Dependent Oncogenic Functions in Hepatocellular Carcinoma." *Oncogene*. doi:10.1038/onc.2010.504.

Xu, T., and G. M. Rubin. 1993. "Analysis of Genetic Mosaics in Developing and Adult Drosophila Tissues." *Development* 117 (4): 1223–37.

Xu, Tian, Weiyi Wang, Sheng Zhang, Rodney A Stewart, and Wan Yu. 1995. "Identifying Tumor Suppressors in Genetic Mosaics: The Drosophila Lats Gene Encodes a Putative Protein Kinase." *Development* 121: 1053–63.

Youle, Richard J., and Alexander M. Van Der Bliek. 2012. “Mitochondrial Fission, Fusion, and Stress.” *Science*. doi:10.1126/science.1219855.

Yu, Guangchuang, Li Gen Wang, and Qing Yu He. 2015. “ChIP Seeker: An R/Bioconductor Package for ChIP Peak Annotation, Comparison and Visualization.” *Bioinformatics* 31 (14): 2382–83. doi:10.1093/bioinformatics/btv145.

Yue, Beatrice. 2014. “Biology of the Extracellular Matrix: An Overview.” *Journal of Glaucoma* 23 (8): S20–23. doi:10.1097/IJG.0000000000000108.

Zaidi, Sayyed K., Andrew J. Sullivan, Ricardo Medina, Yoshiaki Ito, Andre J. Van Wijnen, Janet L. Stein, Jane B. Lian, and Gary S. Stein. 2004. “Tyrosine Phosphorylation Controls Runx2-Mediated Subnuclear Targeting of YAP to Repress Transcription.” *EMBO Journal*. doi:10.1038/sj.emboj.7600073.

Zanconato, Francesca, Michelangelo Cordenonsi, and Stefano Piccolo. 2016. “YAP/TAZ at the Roots of Cancer.” *Cancer Cell* 29 (6): 783–803. doi:10.1016/j.ccell.2016.05.005.

Zanconato, Francesca, Mattia Forcato, Giusy Battilana, Luca Azzolin, Erika Quaranta, Beatrice Bodega, Antonio Rosato, Silvio Bicciato, Michelangelo Cordenonsi, and Stefano Piccolo. 2015. “Genome-Wide Association between YAP/TAZ/TEAD and AP-1 at Enhancers Drives Oncogenic Growth.” *Nature Cell Biology* 17 (9). Nature Research: 1218–27. doi:10.1038/ncb3216.

Zeitlinger, Julia, Lutz Kockel, Fiorenzo A. Peverali, David B. Jackson, Marek Mlodzik, and Dirk Bohmann. 1997. “Defective Dorsal Closure and Loss of Epidermal Decapentaplegic Expression in Drosophila Fos Mutants.” *EMBO Journal* 16 (24): 7393–7401. doi:10.1093/emboj/16.24.7393.

Zemke, Nathan R., Dawei Gou, and Arnold J. Berk. 2019. “Dedifferentiation by Adenovirus E1A Due to Inactivation of Hippo Pathway Effectors YAP and TAZ.” *Genes and Development* 33 (13–14): 828–43. doi:10.1101/gad.324814.119.

Zender, Lars, Mona S. Spector, Wen Xue, Peer Flemming, Carlos Cordon-Cardo, John Silke, Sheung Tat Fan, et al. 2006. “Identification and Validation of Oncogenes in Liver Cancer Using an Integrative Oncogenomic Approach.” *Cell*. doi:10.1016/j.cell.2006.05.030.

Zhang, Can, Brian S. Robinson, Wenjian Xu, Liu Yang, Bing Yao, Heya Zhao, Phil K. Byun, Peng Jin, Alexey Veraksa, and Kenneth H. Moberg. 2015. “The Ecdysone Receptor Coactivator Taiman Links Yorkie to Transcriptional Control of Germline Stem Cell

Factors in Somatic Tissue." *Developmental Cell* 34 (2). Elsevier Inc.: 168–80. doi:10.1016/j.devcel.2015.05.010.

Zhang, Haiying, H. Amalia Pasolli, and Elaine Fuchs. 2011. "Yes-Associated Protein (YAP) Transcriptional Coactivator Functions in Balancing Growth and Differentiation in Skin." *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.1019603108.

Zhang, Jie, Christina Grek, Zhi-wei Ye, Yefim Manevich, Kenneth D Tew, M Townsend, Experimental Therapeutics, et al. 2014. "Pleiotropic Functions of Glutathione S-Transferase P." *Adv Cancer Res* 122: 143–75. doi:10.1016/B978-0-12-420117-0.00004-9.Pleiotropic.

Zhang, Lei, Fangfang Ren, Qing Zhang, Yongbin Chen, Bing Wang, and Jiang Jian. 2008. "The TEAD/TEF Family of Transcription Factor Scalloped Mediates Hippo Signaling in Organ Size Control." *Dev Cell* 14 (3): 377–87. doi:10.1016/j.devcel.2008.01.006.The.

Zhang, Lei, Fangfang Ren, Qing Zhang, Yongbin Chen, Bing Wang, and Jin Jiang. 2008. "Article The TEAD / TEF Family of Transcription Factor Scalloped Mediates Hippo Signaling in Organ Size Control," no. March: 377–87. doi:10.1016/j.devcel.2008.01.006.

Zhang, Peng, Chunli Pei, Xi Wang, Jinyi Xiang, Bao Fa Sun, Yongsheng Cheng, Xiaolong Qi, et al. 2017. "A Balance of Yki/Sd Activator and E2F1/Sd Repressor Complexes Controls Cell Survival and Affects Organ Size." *Developmental Cell* 43 (5). Elsevier Inc.: 603-617.e5. doi:10.1016/j.devcel.2017.10.033.

Zhang, Tianyi, Qingxiang Zhou, and Francesca Pignoni. 2011. "Yki/YAP, Sd/TEAD and Hth/MEIS Control Tissue Specification in the Drosophila Eye Disc Epithelium." *PLoS ONE* 6 (7). doi:10.1371/journal.pone.0022278.

Zhao, Bin, Li Li, Qunying Lei, and Kun-Liang Guan. 2010. "The Hippo–YAP Pathway in Organ Size Control and Tumorigenesis: An Updated Version." *Genes and Development* 24238: 862–74.

Zhao, Bin, Li Li, Lloyd Wang, Cun Yu Wang, Jindan Yu, and Kun Liang Guan. 2012. "Cell Detachment Activates the Hippo Pathway via Cytoskeleton Reorganization to Induce Anoikis." *Genes and Development* 26 (1): 54–68. doi:10.1101/gad.173435.111.

Zhao, Bin, Karen Tumaneng, and Kun Liang Guan. 2011. "The Hippo Pathway in Organ Size Control, Tissue Regeneration and Stem Cell Self-Renewal." *Nature Cell Biology* 13 (8).

Nature Publishing Group: 877–83. doi:10.1038/ncb2303.

Zhao, Bin, Xiaomu Wei, Weiquan Li, Ryan S. Udan, Qian Yang, Joungmok Kim, Joe Xie, et al. 2007. "Inactivation of YAP Oncoprotein by the Hippo Pathway Is Involved in Cell Contact Inhibition and Tissue Growth Control." *Genes and Development* 21 (21): 2747–61. doi:10.1101/gad.1602907.

Zhao, Bin, Xin Ye, Jindan Yu, Li Li, Weiquan Li, Siming Li, Jianjun Yu, et al. 2008. "TEAD Mediates YAP-Dependent Gene Induction and Growth Control." *Genes and Development* 22 (14): 1962–71. doi:10.1101/gad.1664408.

Zhou, Dawang, Claudius Conrad, Fan Xia, Ji Sun Park, Bernhard Payer, Yi Yin, Gregory Y. Lauwers, et al. 2009. "Mst1 and Mst2 Maintain Hepatocyte Quiescence and Suppress Hepatocellular Carcinoma Development through Inactivation of the Yap1 Oncogene." *Cancer Cell* 16 (5). Elsevier Ltd: 425–38. doi:10.1016/j.ccr.2009.09.026.

Zhou, Yuan Yuan, Ying Li, Wei Qin Jiang, and Lin Fu Zhou. 2015. "MAPK/JNK Signalling: A Potential Autophagy Regulation Pathway." *Bioscience Reports* 35 (3): 1–10. doi:10.1042/BSR20140141.

Zhu, Ye, Dong Li, Yadong Wang, Chunli Pei, Song Liu, Lei Zhang, Zengqiang Yuan, and Peng Zhang. 2015. "Brahma Regulates the Hippo Pathway Activity through Forming Complex with Yki-Sd and Regulating the Transcription of Crumbs." *Cellular Signalling* 27 (3). The Authors: 606–13. doi:10.1016/j.cellsig.2014.12.002.

Ziosi, Marcello, Luis Alberto Baena-López, Daniela Grifoni, Francesca Froldi, Andrea Pession, Flavio Garoia, Vincenzo Trotta, Paola Bellosta, Sandro Cavicchi, and Annalisa Pession. 2010. "DMyc Functions Downstream of Yorkie to Promote the Supercompetitive Behavior of Hippo Pathway Mutant Cells." *PLoS Genetics* 6 (9). doi:10.1371/journal.pgen.1001140.

Chapter 8: Appendix

8.1 Bioinformatics Analysis

8.1.1 DamID-seq analysis

Code for DamIDseq analysis

Installation and loading of packages required for DamID analysis pipeline:

```
```{r}

#installing packages:
if (!requireNamespace("BiocManager", quietly = TRUE))
install.packages("BiocManager")
BiocManager::install(c("tidyverse", "edgeR", "Glimma", "gplots", "RColorBrewer", "org.Mm.eg.db",
"org.Dm.eg.db", "BiasedUrn", "statmod", "Glimma", "reshape2", "ggplot2", "dplyr", "rprojroot",
"AnnotationDbi", "ChIPpeakAnno","dplyr","GO.db","tibble","EDASeq","eulerr","RUVSeq","svglite"))


#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":") %in%
search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("dplyr", "ggplot2", "ggsignif", "tidyr", "reshape2", "tibble", "scales", "ggpubr",
"readr", "stringr", "cowplot", "purrr", "PANTHER.db", "AnnotationDbi", "org.Dm.eg.db", "statmod",
"edgeR", "limma", "rprojroot"))


```
```

The code below was created by Dr Jan Schröder with edits from Katrina Mitchell and Jonathan Pojer.

The DamIDseq sequencing data was aligned to the Drosophila dm6 genome using Subread in Terminal.

The reads were then assigned to GATC sequences (subsequently referred to as 'tags') using FeatureCounts in Terminal. This generated a counts table with the number of reads ('counts') for each tag.

Generating counts file in terminal:

```bash
cd /Volumes/Shared-2/Research/Cluster\ 5/Harvey/Lab\ Shared/Everything\ else/Kat/DamID_bamfiles

featureCounts -Q 10 -T 15 -t gatc --read2pos 5 -a
Drosophila_melanogaster.BDGP6.GATC_stranded_cutadapt.gff -o counts.normal
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Dam-1_S49_s.bam
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Dam-2_S51_s.bam
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Dam-3_S53_s.bam
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Yki-1_S50_s.bam
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Yki-2_S52_s.bam
../Normal\ growth\ data/YkiDam/Bam\ files/KM-Yki-3_S54_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/Dam1_S1_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/Dam2_S4_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/Dam3_S7_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamSd1_S2_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamSd2_S5_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamSd3_S8_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamTgi1_S3_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamTgi2_S6_s.bam
../Normal\ growth\ data/DamSd\ and\ DamTgi_April_2017/Bam\ files_JS/DamTgi3_S9_s.bam

```

The count data was imported and analysed in RStudio using the EdgeR and limma packages. This analysis generated a list of differentially methylated tags of the Dam-fusion compared to the Dam-alone control.

```{r Differential methylation analysis}
setwd( /Volumes/Shared-2/Research/Cluster\ 5/Harvey/Lab\ Shared/Everything\ else/Kat/Dam_IDfiles)

# Read the sample information into R
sampleinfo <- read.delim("data/sampleinfo.normal.txt")
#read in the count file:
data <- read.delim("data/counts.normal", header = TRUE, comment.char = "#")

#renaming the sample names
names(data) <- c("GeneID", "Chr", "Start", "End", "Strand", "Length","Dam1", "Dam2", "Dam3", "Yki1",
"Yki2", "Yki3", "Dam4", "Dam5", "Dam6", "Sd1", "Sd2", "Sd3", "Tgi1", "Tgi2", "Tgi3")
samples = names(data)[6:14]


#creating the design matrix to analyse the data using group (genotype), batch (biological replicate)
and seq (sequencing batch)
group <- c("Dam", "Dam", "Dam", "Yki", "Yki", "Yki", "Dam", "Dam", "Dam", "Sd", "Sd", "Sd", "Tgi",
"Tgi", "Tgi")
batch <- c(1, 2, 3, 1, 2, 3, 4, 5, 6, 4, 5, 6, 4, 5, 6)
seq <- c(1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2)
design <- model.matrix(~ batch + group + seq)

# Storing the count data in a format that can be analysed using edgeR
dge <- DGEList(data.normal[,7:21], group = group.normal, genes = data.normal$GeneID)
```

```
# Filter and normalise the data

## counts per million has to be greater than 0.5 for at least 3 samples - 3 because there is 3
replicates per each sample:
keep <- rowSums(cpm(dge) >= 0.5) >= 3
## below states the number of counts that are kept after filtering:
summary(keep)
              #   Mode     FALSE    TRUE
              #   logical  436291   331030
## elminates lower rows; updates libary size for the new dataset:
dge <- dge[keep, , keep.lib.sizes = FALSE]
## uses TMM method to work out the composition bias in the data:
dge <- calcNormFactors(dge)
## calculate within group dispersion - variablitiy between replicates:
dge <- estimateDisp(dge, robust = TRUE, design = design)

# Diagnostic plots
## plots counts after filtering
dge$samples$lib.size
barplot(dge.normal$samples$lib.size,names=colnames(dge.normal),las=2, main = "Sample library size"))

## plots variablility between replicates
plotBCV(dge)

## Multidimension scaling (MDS) plot
levels(sampleinfo$SeqRun)
col.gen <- c("grey32","dodgerblue3","springgreen3","magenta2")[sampleinfo$Genotype]
pch.seq <- c(15,17)[sampleinfo$SeqRun]
data.frame(sampleinfo$SeqRun,col.seq)
svg('output/MDS_plot.svg')
par(xpd=T, mar=par()$mar+c(0,0,0,6))

plotMDS(removeBatchEffect(dge$counts, batch = batch),col= col.gen, pch=pch.seq, cex = 1.5,cex.lab = 1,
cex.axis = 1)
legend(5000, 0,fill=c("gray32","dodgerblue3","springgreen3","magenta2"),legend=levels(sampleinfo$Genot
ype),cex=0.8)

# Differental methylation analysis
## create a linear model
fit <- glmFit(dge, design = design)

lrt_Yki <- glmLRT(fit, coef = 5)
lrt_Sd <- glmLRT(fit, coef = 3)
lrt_Tgi <- glmLRT(fit, coef = 4)
```

```
## summary of results for differential methylation
summary(de.Yki <- decideTestsDGE(lrt_Yki))
        #Down       9083
        #NotSig   290115
        #Up        31832

summary(de.Sd <- decideTestsDGE(lrt_Sd))
        #Down     45707
        #NotSig  246686
        #Up       38637

summary(de.Tgi <- decideTestsDGE(lrt_Tgi))
        #Down      44393
        #NotSig   251858
        #Up        34779
```

```
# Exporting the datasets to use for peak calling

names(keep) <- data$GeneID
write.table(keep, file = 'data/keep', quote = FALSE, col.names = FALSE)

row.names(lrt_Yki$table) <- data[keep, "GeneID"]
lrt_Yki$table %>%
  tibble::rownames_to_column("coord") %>%
  mutate(sig = de.Yki) %>%
  write.table(., file = "data/lrt_Yki", quote = FALSE, row.names = FALSE)

row.names(lrt_Sd$table) <- data[keep, "GeneID"]
lrt_Sd$table %>%
  tibble::rownames_to_column("coord") %>%
  mutate(sig = de.Sd) %>%
  write.table(., file = "data/lrt_Sd", quote = FALSE, row.names = FALSE)

row.names(lrt_Tgi$table) <- data[keep, "GeneID"]
lrt_Tgi$table %>%
  tibble::rownames_to_column("coord") %>%
  mutate(sig = de.Tgi) %>%
  write.table(., file = "data/lrt_Tgi", quote = FALSE, row.names = FALSE)

```

Peak calling was then carried out in Terminal. The peak calling pipeline was created by Dr Jan Schröder (unpublished). Peak calling generated 2 files:

1) filtered peak (bed file) - these are significant peaks that had at least one differentially methylated tag, more than two significant tags (p value) or all the tags were differentially methylated within a 1kb of the centre of the peak.

2) Peaks assigned to the closest gene (bed file) - these are filtered peaks that were assigned to the nearest transcriptional start site (TSS) within a 5kb range. This was carried out using the Bedtools 'closestBed' function.

<u>The peak files contained the folowing information:</u>

Column 1 Chromosome arm

Column 2 Start of the peak

Column 3 End of the peak

Column 4 Number of tags in a peak

Column 5 1kb methylation rate

Column 6 Peak log fold change (logFC)

Column 7 p-value

8.1.2 RNA-seq analysis

RNAseq analysis code for examining differential gene expression in overgrowth (wtsX1/wtsLacZ) compared to wild type growth (FRT82B).

Analysis was carried out according to COMBINE pipeline, which can be found at:

http://combine-australia.github.io/RNAseq-R/

Sequencing data was aligned to the reference genome using Subread tool in Terminal.

Installing packages and loading libraries:

```{r}

if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("tidyverse", "edgeR", "Glimma", "gplots", "RColorBrewer", "org.Mm.eg.db",
"org.Dm.eg.db", "BiasedUrn", "statmod", "Glimma", "reshape2", "ggplot2", "dplyr", "rprojroot",
"AnnotationDbi", "RColorBrewer", "Rsubread", "limma", "svglite","KEGG.db", "EnhancedVolcano"))

#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":") %in%
search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("tidyverse", "edgeR", "Glimma", "gplots", "RColorBrewer", "org.Mm.eg.db",
"org.Dm.eg.db", "BiasedUrn", "statmod", "Glimma", "reshape2", "ggplot2", "dplyr", "rprojroot",
"AnnotationDbi", "RColorBrewer", "Rsubread", "limma", "svglite","KEGG.db", "EnhancedVolcano"))

```

FeatureCounts was then used to count how many sequencing reads are at each genomic position, which generated a counts table.

```{r}

bam <- list.files(path = "data", pattern = "*.bam$")
# count file to generate how many reads are at each genomic position
counts <- featureCounts(bam, annot.ext="dmel-all-r6.27.gtf", isGTFAnnotationFile=TRUE,
              GTF.featureType="exon", GTF.attrType="gene_id", useMetaFeatures=TRUE,
              allowMultiOverlap=TRUE, isPairedEnd=TRUE)

# made a new matrix containing only the counts, the row names are the flybase ID
countdata <- counts$counts
# renaming the sample names
colnames(countdata) <- c("FRT82B1", "FRT82B2", "FRT82B3", "FRT82B4", "FRT82B5", "WTS1", "WTS2",
"WTS3", "WTS4", "WTS5")
# reading in the sample info file
sampleinfo <- read.csv("data/sampleinfo.csv")
table(colnames(countdata)==sampleinfo$SampleName)
write.csv(countdata, file="output/count.table.csv", row.names=TRUE, col.names=TRUE)

```

Filtering the data:

```r
# filtering to remove lowly expressed genes
myCPM <- cpm(countdata)
head(myCPM)
# greater than 0.5 counts per million as the threshold
thresh <- myCPM > 0.5

# Summary: there are 9,338 that have TRUE in all 10 samples (that are above 0.5 cpm)
# filter to keep genes that have at least 5 TRUES in each row of thresh, because i have 5 replicates
for each genotype
keep <- rowSums(thresh) >= 5
counts.keep <- countdata[keep,]

# plotting the countdata
plot(myCPM[,1],countdata[,1],ylim=c(0,50),xlim=c(0,3), abline(v=0.5))
# a cpm of 0.5 corresponds to a count of 15 for these samples. Any lower than 15, the gene is likely
not expressed in the sample and is filtered out from the analysis
```

Creating a DGEList object and looking at library sizes:

```r

# Converting counts to DGEList object which is an object used by edgeR to store count data
y <- DGEList(counts.keep)

# Quality control
# looking at how many reads is in each sample in y
y$samples$lib.size
barplot(y$samples$lib.size,names=colnames(y),las=2)

# we can see that the library sizes are all quite similar and large

# count data is not normall distributed, so we will look at the distribution on a log scale

# boxplot showing the distribution of the read counts on the log2 scale
# Used the cpm function to get log2 counts per million, which are corrected for the different library
sizes

logcounts <- cpm(y,log=TRUE)
boxplot(logcounts, xlab="", ylab="Log2 counts per million",las=2)
abline(h=median(logcounts),col="blue")
title("Boxplots of logCPMs (unnormalised)")

# summary: the boxplots show that the distributions of raw log intensities are quite similar across
the samples, although not identical

```

```r
```{r Multidimensional scaling plots}


#determine the greatest sources of variation in the data using a MDS plot
plotMDS(y)

levels(sampleinfo$Status)
col.cell <- c("magenta2","dodgerblue3")[sampleinfo$Status]
data.frame(sampleinfo$Status,col.cell)
pch <- c(19)

svglite(file = "MDSplotRNAseq.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
par(xpd=T, mar=par()$mar+c(0,0,0,6))  # Expand right side of clipping rect to make room for the legend
plotMDS(y, col= col.gen, pch=pch.seq, cex = 1.5,cex.lab = 1, cex.axis = 1)
legend(5000, 0,fill=c("magenta2","dodgerblue3"),legend=levels(sampleinfo$Status),cex=0.8)
```
```

Normalisation for composition bias:

```r
```{r}
# TMM normalization is performed to eliminate composition biases between libraries
# A normalization factor below one indicates that the library size will be scaled down, as there is
more suppression (i.e., composition bias)

y <- calcNormFactors(y)
y$samples

#before normalisation
par(mfrow=c(1,2))
plotMD(logcounts,column = 5)
abline(h=0,col="grey")
plotMD(logcounts,column = 10)
abline(h=0,col="grey")
#after normalisation
par(mfrow=c(1,2))
plotMD(y,column = 5)
abline(h=0,col="grey")
plotMD(y,column = 10)
abline(h=0,col="grey")

```
```

Differential expression with limma-voom:

```{r}
# The limma package (Ritchie et al. 2015) (since version 3.16.0) offers the voom function, which transforms the read counts into logCPMs while taking into account the mean-variance relationship in the data

y$samples
group <- c("FRT82B", "FRT82B", "FRT82B", "FRT82B", "FRT82B", "WTS", "WTS", "WTS", "WTS", "WTS")
design <- model.matrix(~ 0 + group)
colnames(design) <- c("FRT82B", "WTS")

# Voom transform the data

par(mfrow=c(1,1))
v <- voom(y,design,plot = TRUE)
dev.off()
png(file="Mean variance.png")
voom(y,design,plot = TRUE)
dev.off()


par(mfrow=c(1,2))
boxplot(logcounts, xlab="", ylab="Log2 counts per million",las=2,main="Unnormalised logCPM")
## Adding a blue horizontal line that corresponds to the median logCPM
abline(h=median(logcounts),col="blue")
boxplot(v$E, xlab="", ylab="Log2 counts per million",las=2,main="Voom transformed logCPM")

```

Testing for differential expression:

```{r}
## Fit the linear model
fit <- lmFit(v)
names(fit)
# Matrix which compares genes that are differentially expressed between WildType and Mutant
cont.matrix <- makeContrasts(MutantvsWildtype= WTS - FRT82B,levels=design)

##apply the contrasts matrix to the fit object to get the statistics and estimated parameters of our comparison that we are interested in
fit.cont <- contrasts.fit(fit, cont.matrix)

##perform empirical Bayes shrinkage on the variances, and estimates moderated t-statistics and the associated p-values
fit.cont <- eBayes(fit.cont)

summa.fit <- decideTests(fit.cont)
summary(summa.fit)

#Summary results:
#MutantvsWildtype
#Down               2336
#NotSig             5457
#Up                 2364

#top 10 according to p value
top_pvalue <- topTable(fit.cont,coef="MutantvsWildtype",sort.by="p")
write.csv(top_pvalue, file = "output/top_mutantvswildtype_pvalue.csv", row.names = FALSE)
```

```
#get the full table:
#this table is all genes listed (significant and non significant)
limma.res <- topTable(fit.cont,coef="MutantvsWildtype",sort.by="p",n="Inf")
#these are the ones that are significant
limma.res.significant <- (topTable(fit.cont,coef="MutantvsWildtype",p.val=0.05,n=Inf))
#if cut off is 0.01:
limma.res.p0.01 <- (topTable(fit.cont,coef="MutantvsWildtype",p.val=0.01,n=Inf))


##SUBSETTING THE DIFFERENTIAL GENE EXPRESSION DATA

#subsetting the data based on a LogFC of 0.5 and an adjusted value of 0.01
limma.res.p0.01.logfc0.5 <- limma.res[which(abs(limma.res$logFC) > 0.5 & limma.res$adj.P.Val < 0.01),]
write.csv(limma.res.p0.01.logfc0.5, file="output/lists/limma.res.p0.01.logfc0.5.csv")
#list of upregulated genes
upgenes <- read.csv("output/lists/limma.res.p0.01.logfc0.5_upregulated.csv")
gene_list_up <- as.vector(upgenes)
#list of downregulated genes
downgenes <- read.csv("output/lists/limma.res.p0.01.logfc0.5_downregulated.csv")
gene_list_down <- as.vector(downgenes)

```

Visualising differential gene expression using a Volcano plot, and also integrating the DamIDseq data:

```
```{r Volcano plot integrating DamID target genes}

###volcano plot for just looking at upregulated and downregulated genes
svglite(file = "Volcano_Plot.svg", width = 10, height = 8, bg = "white",
pointsize = 12, standalone = TRUE, system_fonts = list(),
user_fonts = list())
#basic volcano plot
with(limma.res, plot(logFC, -log(limma.res$P.Value, base=10), pch=20, col="gray54"))
#highlight genes that are logfc >0.5 and p value <0.01
with(subset(limma.res, abs(logFC)>0.5), points(logFC, -log10(P.Value), pch=20, col="gray53"))
with(subset(limma.res, abs(-log10(P.Value))>2.5), points(logFC, -log10(P.Value), pch=20,
col="gray53"))
#highlight the upregulated genes in red
with(subset(limma.res, rownames(limma.res) %in% gene_list_up$FLYBASE), points(logFC, -log10(P.Value),
pch=20, col= "red3")) +
#highlight the downregulated genes in green
  with(subset(limma.res, rownames(limma.res) %in% gene_list_down$FLYBASE), points(logFC,
-log10(P.Value), pch=20, col= "royalblue3"))
  abline(h=2.5, v=0.5, col="black", lty = 5, lwd=2)
    abline(v=-0.5, col="black", lty = 5, lwd=2)


## Integrating DamID data
# Figure 4.2
svglite(file = "Volcanoplot_normalgrowthDamID.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
#basic volcano plot
with(limma.res, plot(logFC, -log(limma.res$P.Value, base=10), pch=21, bg = "gray54", col="gray54"))
```
```

```
#highlighting the genes that are yki/sd DamID target genes in purple
#higlighting the genes that are DamID targets and LogFC > 0.5
with(subset(limma.res, rownames(limma.res) %in% ykitargets_ID$Gene_ID & abs(logFC)>0.5 &
abs(-log10(P.Value))>2.5), points(logFC, -log10(P.Value), pch=21, bg = "magenta2", col="magenta4"))
with(subset(limma.res, rownames(limma.res) %in% sdtargets_ID$Gene_ID & abs(logFC)>0.5 &
abs(-log10(P.Value))>2.5), points(logFC, -log10(P.Value), pch=21, bg = "dodgerblue",
col="dodgerblue"))
with(subset(limma.res, rownames(limma.res) %in% tgitargets_ID$Gene_ID & abs(logFC)>0.5 &
abs(-log10(P.Value))>2.5), points(logFC, -log10(P.Value), pch=21, bg = "springgreen2",
col="springgreen2"))
with(subset(limma.res, rownames(limma.res) %in% sharedtargets$Gene_ID & abs(logFC)>0.5 &
abs(-log10(P.Value))>2.5), points(logFC, -log10(P.Value), pch=21, bg = "darkorchid3",
col="darkorchid3"))
abline(h=2.5, v=0.5, col="black", lty = 5, lwd=2)
abline(v=-0.5,col="black", lty = 5, lwd=2)
dev.off()

```

8.1.3 Comparisons between datasets

Intersect function was used to identify genes that overlap between datasets:

```{r}

#intersections of FlyBaseID between two datasets
ListA <- GeneList1$FlyBaseID
ListB <- GeneList2$FlyBaseID
intersection <- intersect(ListA, ListB)

#save this as a csv file
write.csv(intersection, file = "intersection.csv", row.names = FALSE)

```

Venn diagrams to visualise overlapping genes/features between datasets:

```{r}
BiocManager::install("eulerr")
library(eulerr)

#Venn diagrams were created using the Eulur package in RStudio
Dataset1 = numberofgenes
Dataset2 = numberofgenes
Overlap = number of genes overlapping Dataset1 and Dataset2

Colour pallete <- (yki = "magenta2", sd = "dodgerblue3", tgi = "springgreen1")
venn.Overlapdatasets <- euler(c(Dataset1, Dataset2, Overlap))

svglite(file = "venn.Overlapdatasets.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
plot(venn.Overlapdatasets,
     fills = c("magenta2","dodgerblue3", "springgreen1"),
     edges = TRUE,
     alpha = 0.85,
     fontsize = 12,
     lty = c("solid", "dotted", "dashed"),
     lwd = 2,
     col= c("Black","Black", "Black"),
     quantities = list(fontsize = 10))
dev.off()
```

Bedtools was used to intersect DamID-seq bed peak files in Terminal and reported the overlap between two sets of genomic features

Input files = bed peak files

As input, I used the significant peaks (not assigned to genes yet), these were generated from the filtered.bed files and the bed files were rearranged to a format that bedtools can read

format for bedtools:

1 - chrom (e.g. chr2)

2 - start (e.g. 9)

3 - end (e.g. 20)

```
```{r getting the peak files into the correct format }
yki.peaks %>%
  mutate(chrom = paste0("chr", Chr)) %>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End) %>%
  write_delim(., path = "...", col_names = FALSE, delim = "\t")

sd.peaks %>%
  mutate(chrom = paste0("chr", Chr)) %>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End) %>%
  write_delim(., path = "...", col_names = FALSE, delim = "\t")

tgi.peaks %>%
  mutate(chrom = paste0("chr", Chr)) %>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End) %>%
  write_delim(., path = "...", col_names = FALSE, delim = "\t")

```
```

Intersecting the peaks:

```
```{bash}
bedtools intersect -a A.bed -b B.bed > overlapping.AB.bed
```
```

Hypergeometric distribution:

```
```{r}
phyper(overlap, sampleb, totala - sampleb, samplec)
```
```

8.1.4 Gene Profiles

The code for visualising genes profiles and peaks was created by Dr Jonathan Pojer with edits by Katrina Mitchell.

```
```{r installing packages and libraries}

#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":")
%in% search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("cowplot", "tidyr", "dplyr", "gridExtra", "reshape2", "ggplot2", "scales",
"GGally", "grid","gggenes"))

```
```

Marshall & Brand (2015) code for generating profiles (bedgraphs)

Using Marshall & Brand (2015) pipeline in Terminal to generate bedgraphs for each biological replicate:

```
```{bash}
cd /Volumes/Shared-2/Research/Cluster\ 5/Harvey/Lab\ Shared/Everything\ else/Kat/

damidseq_pipeline --gatc_frag_file=Dmel_BDGP6.GATC.gff
--bowtie2_genome_dir=Drosophila_melanogaster_Ensembl_BDGP6/Drosophila_melanogaster/Ensembl/BDG
P6/Sequence/Bowtie2Index/genome --save_defaults

# Creating bedgraphs of Dam-fusion over Dam-alone control
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=KM-Dam-1_S49_s.bam
KM-Yki-1_S50_s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=KM-Dam-2_S51_s.bam
KM-Yki-2_S52_s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=KM-Dam-3_S53_s.bam
KM-Yki-3_S54_s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam1_S1_s.bam DamSd1_S2_s.bam
DamTgi1_S3_s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam2_S4_s.bam DamSd2_S5_s.bam
DamTgi2_S6_s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam3_S7_s.bam DamSd3_S8_s.bam
DamTgi3_S9_s.bam

#Jra data
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam-1_S4.s.bam
Dam-Jra-1_S6.s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam-2_S7.s.bam
Dam-Jra-2_S9.s.bam
damidseq_pipeline --bamfiles --full_data_files --ps_debug --dam=Dam-3_S10.s.bam
Dam-Jra-3_S12.s.bam
```
```

```
# Averaging bedgraphs
## Yki
bedtools unionbedg -i KM-Yki-1_S50_s-vs-Dam.bedgraph KM-Yki-2_S52_s-vs-Dam.bedgraph
KM-Yki-3_S54_s-vs-Dam.bedgraph | awk '{sum=0; for (col=4; col<=NF; col++) sum += $col; print
$1"\t"$2"\t"$3"\t"sum/(NF-3); }' > Yki_mean.bedgraph
## Sd
bedtools unionbedg -i DamSd1_S2_s-vs-Dam.bedgraph DamSd2_S5_s-vs-Dam.bedgraph
DamSd3_S8_s-vs-Dam.bedgraph | awk '{sum=0; for (col=4; col<=NF; col++) sum += $col; print
$1"\t"$2"\t"$3"\t"sum/(NF-3); }' > Sd_mean.bedgraph
## Tgi
bedtools unionbedg -i DamTgi1_S3_s-vs-Dam.bedgraph DamTgi2_S6_s-vs-Dam.bedgraph
DamTgi3_S9_s-vs-Dam.bedgraph | awk '{sum=0; for (col=4; col<=NF; col++) sum += $col; print
$1"\t"$2"\t"$3"\t"sum/(NF-3); }' > Tgi_mean.bedgraph
## Jra
bedtools unionbedg -i Dam-Jra-1_S6.s-vs-Dam.bedgraph Dam-Jra-2_S9.s-vs-Dam.bedgraph
Dam-Jra-3_S12.s-vs-Dam.bedgraph | awk '{sum=0; for (col=4; col<=NF; col++) sum += $col; print
$1"\t"$2"\t"$3"\t"sum/(NF-3); }' > Jra_mean.bedgraph

```

Code for visualising gene profiles and peaks

The code below was written by Jonathan Pojer, with edits by Katrina Mitchell

This code allows visualisation of the gene profiles, peaks, and gene view for each Dam-fusion.

Import the annotated genome, downloaded from:

ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/gtf/

```
```{r Dmel genome features}

setwd

dmel.r6.27 <- read.delim("dmel-all-r6.27.gtf",
                         header = FALSE,
                         col.names=c("Chr", "Source", "Feature", "Start", "End",
                                     "Score", "Strand", "Frame", "Attribute")) %>%
  dplyr::select(-Source, -Score) %>%
  separate(Attribute, into = c("Gene_ID", "Gene_Symbol", "Transcript_ID",
                               "Transcript_Symbol", "Notes"), sep = "; ") %>%
  mutate(Gene_ID = word(str_remove_all(Gene_ID, ";"), 2),
         Gene_Symbol = word(str_remove_all(Gene_Symbol, ";"), 2),
         Transcript_ID = word(str_remove_all(Transcript_ID, ";"), 2),
         Transcript_Symbol = word(str_remove_all(Transcript_Symbol, ";"), 2))
```
```

Importing profile and peak files:

```
```{r Importing files}

yki.profile <- read.delim("Bedgraphs/Normal growth/Yki_mean.bedgraph", header = FALSE,
                          col.names = c("Chr", "Start", "End", "logFC"))
sd.profile <- read.delim("Bedgraphs/Normal growth/Sd_mean.bedgraph", header = FALSE,
                         col.names = c("Chr", "Start", "End", "logFC"))
tgi.profile <- read.delim("Bedgraphs/Normal growth/Tgi_mean.bedgraph", header = FALSE,
                         col.names = c("Chr", "Start", "End", "logFC"))
jra.profile <- read.delim("Bedgraphs/Normal growth/Jra_mean.bedgraph", header = FALSE,
                         col.names = c("Chr", "Start", "End", "logFC"))

yki.peaks <- read.delim("Peak files/yki_normal_peaks_filtered.bed",
                        header = FALSE, comment.char = "#",
                        col.names = c("Chr", "Start", "End", "Tags",
                              "Methylation_Rate", "logFC", "Sig_Tags", "pVal"))
sd.peaks <- read.delim("Peak files/sd_normal_peaks_filtered.bed",
                         header = FALSE, comment.char = "#",
                         col.names = c("Chr", "Start", "End", "Tags",
                               "Methylation_Rate", "logFC", "Sig_Tags", "pVal"))
tgi.peaks <- read.delim("Peak files/tgi_normal_peaks_filtered.bed",
                         header = FALSE, comment.char = "#",
                         col.names = c("Chr", "Start", "End", "Tags",
                               "Methylation_Rate", "logFC", "Sig_Tags", "pVal"))
jra.profile <- read.delim("Bedgraphs/Normal growth/Jra_mean.bedgraph", header = FALSE,
                         col.names = c("Chr", "Start", "End", "logFC"))
```
```

Plot the profile and peaks

Code shown for Yki, Sd and Tgi profiles

```
```{r Profile - peaks separate}
damid.profile <- function(name, chr, start, end,
                          dam.gene = c("yki", "sd", "tgi")) {
  col.palette <- c(yellow = "#FED600", magenta = "#EE42E0", green = "#19C103",
                   blue = "#0026CB", orange = "#F69303", cyan = "#1FACE0",
                   red = "#FF0000", grey = "#808080")
  dam.gene <- match.arg(dam.gene)
  gene <- switch(dam.gene,
                 yki = "Yorkie",
                 sd = "Scalloped",
                 tgi = "Tgi"
                )
  profile.col <- switch(dam.gene,
                        yki = "#FF00FF",
                        sd = "#1E90FF",
                        tgi = "#00FA9A"
                        )
  peaks.col <- switch(dam.gene,
                      yki = "#C71585",
                      sd = "#0000CD",
                      tgi = "#3CB371")
```
```

```
  label <- paste(gene)
  print(paste("Generating", gene, "profile for", name))

  # Get profile and peaks data
  profile <- switch(dam.gene,
                    yki = yki.profile,
                    sd= sd.profile,
                    tgi=tgi.profile
                                    ) %>%
    mutate(End = End + 1, Factor = label) %>%
    filter(Chr == chr, Start >= start, End <= end)

  peaks <- switch(dam.gene,
                  yki = yki.peaks,
                  sd = sd.peaks,
                  tgi = tgi.peaks
                                  ) %>%
    mutate(Factor = label) %>%
    filter(Chr == chr, Start >= start, End <= end)

  # Make a plot
profile.breaks <- c(ifelse(min(profile$logFC) > 0, 0, min(profile$logFC)),
                    ifelse(max(profile$logFC) < 3, 3, max(profile$logFC)))
  plot.profile <- ggplot() +
    geom_rect(data = profile, aes(xmin = Start, xmax = End, ymin = 0, ymax = logFC),
              fill = profile.col) +
    geom_hline(yintercept = 0) +
    scale_x_continuous(limits = c(start, end), expand = c(0, 0)) +
    scale_y_continuous(expand = c(0, 0), breaks = profile.breaks,
                       limits = profile.breaks,
                       labels = number_format(accuracy = 0.01)) +
    labs(y = "log(Dam-fusion / Dam-only)") +
    theme_classic(base_family = "HelveticaNeue") +
    theme(text = element_text(size = 10 * 4/3), legend.position = "none",
          plot.background = element_blank(), panel.background = element_blank(),
          axis.text.y = element_text(colour = "black"),
          axis.ticks.y = element_line(colour = "black"),
          axis.text.x = element_blank(), axis.ticks.x = element_blank(),
          axis.title.x = element_blank(), axis.line.x = element_blank())
```

```
  # Add data for peaks, if present
  if(nrow(peaks) != 0) {
    plot.peaks <- ggplot() +
      geom_segment(data = peaks, aes(x = Start, xend = End, y = 0, yend = 0),
                   size = 4, colour = peaks.col) +
      scale_y_continuous(expand = c(0,0), limits = c(0, 0)) +
      scale_x_continuous(limits = c(start, end), expand = c(0, 0)) +
      theme_classic() +
      theme(plot.background = element_blank(), panel.background = element_blank(),
            axis.text = element_blank(), axis.ticks = element_blank(),
            axis.title = element_blank(), axis.line = element_blank())
    plots <- list(plot.peaks, plot.profile)
  }
  else {plots <- list(ggplot() + theme_nothing(), plot.profile)}

  return(plots)
}
print("damid.profile() has been loaded.")
```

Generate a plot for the gene view:

```
```{r Geneview - whole gene}

damid.geneview <- function(name, chr, start, end) {
  col.palette <- c(yellow = "#FED600", magenta = "#EE42E0", green = "#19C103",
                   blue = "#0026CB", orange = "#F69303", cyan = "#1FACE0",
                   red = "#FF0000", grey = "#808080")
  label.spacing <- (end - start) / 200
  print(paste0("Generating gene view for ", name, "..."))

  # Load gene and exon information
  gene <- dmel.r6.27 %>%
    filter(Chr == chr, Start >= start, End <= end, Feature == "gene") %>%
    dplyr::select(Symbol = Gene_Symbol, Start, End, Strand) %>%
    mutate(Width = End - Start,
           Label = ifelse(nchar(Symbol) >= Width / (end - start) * 40, FALSE, TRUE))

  # Generate gene view plot
  plot.geneview <- ggplot(data = gene, aes(xmin = Start, xmax = End, y = Symbol,
                                           forward = ifelse(Strand == "+", 1, -1))) +
    geom_gene_arrow(arrow_body_height = unit(4, "mm"),
                    arrowhead_width = unit(4, "mm"),
                    arrowhead_height = unit(4, "mm"),
                    colour = "black", fill = "black") +
    geom_text(data = filter(gene, Label == TRUE),
              aes(x = ifelse(Strand == "+", Start + label.spacing,
                             End - label.spacing), label = Symbol,
                  hjust = ifelse(Strand == "+", 0, 1)),
              colour = "white", fontface = "italic") +
    geom_text(data = filter(gene, Label == FALSE),
              aes(x = ifelse(Strand == "+", Start - label.spacing,
                             End + label.spacing), label = Symbol,
                  hjust = ifelse(Strand == "+", 1, 0)),
              colour = "black", fontface = "italic") +
    scale_x_continuous(limits = c(start, end), expand = c(0, 0)) +
    labs(x = paste("Position on chromosome", chr)) +
    theme_classic(base_family = "HelveticaNeue") +
    theme(text = element_text(size = 10 * 4/3), axis.text.y = element_blank(),
          axis.title.y = element_blank(), axis.ticks.y = element_blank(),
          axis.line.y = element_blank(), axis.text.x = element_text(colour = "black"),
          axis.ticks.x = element_line(colour = "black"),
          plot.background = element_blank(), panel.background = element_blank())

  return(plot.geneview)
}
```
```

Graph the profile:

```r
damid.graph <- function(name, chr, plots) {
  n.plots <- length(plots)
  print(paste0("Compiling profile for ", name, "..."))

  # Generate plot
  plot.damid <- ggmatrix(plots = plots, nrow = n.plots, ncol = 1, byrow = FALSE,
                         xlab = paste("Position on chromosome", chr)) +
    theme(text = element_text(family = "HelveticaNeue"),
          plot.margin = unit(unit = "mm", x = c(0, 10, 0, 0)),
          panel.spacing = unit(2, units = "mm"),
          axis.title.y = element_blank())

  return(plot.damid)
}
print("damid.graph() has been loaded.")

```

Function that calls all the other functions and saves the plot:

```
```{r Plot - peaks separate}
damid.plot <- function(name, chr, start, end, file.name = NULL,
                       path = "Plots/",
                       dam.gene = c("both", "yki", "sd", "tgi"), sd.binding = FALSE,
                       gene.view = NULL, ext = c("png", "svg"), save = TRUE,
                       show = TRUE, dpi = 300, width = NULL, height = NULL) {
  col.palette <- c(yellow = "#FED600", magenta = "#EE42E0", green = "#19C103",
                   blue = "#0026CB",  orange = "#F69303", cyan = "#1FACE0",
                   red = "#FF0000", grey = "#808080")
  ext <- match.arg(ext)
  if(ext == "svg") {library(svglite, quietly = TRUE)}
  dam.gene <- match.arg(dam.gene)

  # Make gene view and profiles
  if(is.null(gene.view)) gene.view <- damid.geneview(name = name, chr = chr,
                                                     start = start, end = end)
  else {
    if(!is.ggplot(gene.view)) stop("Gene view input is not a ggplot object.")
  }
{
    if(dam.gene == "both") {
      profile.yki <- damid.profile(name = name, chr = chr,
                                   start = start, end = end,
                                   dam.gene = "yki")
      profile.sd <- damid.profile(name = name, chr = chr,
                                  start = start, end = end,
                                  dam.gene = "sd")
      profile.tgi <- damid.profile(name = name, chr = chr,
                                   start = start, end = end,
                                   dam.gene = "tgi")
      plot.list <- c(profile.yki, profile.sd, profile.tgi, list(gene.view))
    } }
 {
    if(dam.gene == "both") {
      profile.yki <- damid.profile(name = name, chr = chr,
                                   start = start, end = end,
                                   dam.gene = "yki")
      profile.sd <- damid.profile(name = name, chr = chr,
                                  start = start, end = end,
                                  dam.gene = "sd")
      profile.tgi <- damid.profile(name = name, chr = chr,
                                   start = start, end = end,
                                   dam.gene = "tgi")
      plot.list <- c(profile.yki, profile.sd, profile.tgi, list(gene.view))
    }
    else {
      profile <- damid.profile(name = name, chr = chr,
                               start = start, end = end,
                               dam.gene = dam.gene)
      plot.list <- c(profile, list(gene.view))
    }
  }
```
```

```
  plot.damid <- damid.graph(name = name, chr = chr,
                            plots = plot.list)

  # Save graph
  if(save) {
    if(is.null(width)) {width <- 7.5}
    if(is.null(height)) {height <- 5}
    path <- paste0(path, format(Sys.Date(), "%y%m%d"), "/")
    if(!dir.exists(path)) dir.create(path)
    print(paste0("Saving profile for ", name, "..."))
    if(is.null(file.name)) {
      save.file <- paste0(path, name, "_profile.", ext)
    }
    else {save.file <- paste0(path, file.name, ".", ext)}
    save_plot(filename = save.file, plot = plot.damid,
              bg = "transparent", base_width = width * 4/3,
              base_height = height * 4/3, units = "cm", dpi = dpi)
  }
  print(paste("Profile for", name, "is complete."))
  if(show) {return(plot.damid)}
}
print("damid.plot() has been loaded.")

```

Generating gene profiles used in this thesis:

```{r}
# Figure 3.1: Hippo pathway target genes
damid.plot(name = "Diap1", chr = "3L", start = 16037000, end = 16053000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ex", chr = "2L", start = 422000, end = 449000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "CycE", chr = "2L", start = 15727000, end = 15749000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ban", chr = "3L", start = 627000, end = 650000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.8: Spalt related genes
damid.plot(name = "salm", chr = "2L", start = 11431000, end = 11449500,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "salr", chr = "2L", start = 11354000, end = 11376000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "brk", chr = "X", start = 7304000, end = 7314000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.9: GST genes
damid.plot(name = "GstD gene cluster", chr = "3R", start = 12360000, end = 12389000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "GstE gene cluster", chr = "2R", start = 18394000, end = 18410000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "GstT3", chr = "X", start = 20464000, end = 20471000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "GstT4", chr = "X", start = 13401000, end = 13411000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
```

```
# Figure 4.10: Cuticle related genes
damid.plot(name = "TwdlE", chr = "2L", start = 7854000, end = 7865000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "CG1136", chr = "3L", start = 4042000, end = 4061000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "Lcp and cpr gene cluster 1", chr = "3L", start = 6130000, end = 6145000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "Lcp and cpr gene cluster 2", chr = "3L", start = 6145000, end = 6160000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.11: Early retinal determination genes
damid.plot(name = "ey", chr = "4", start = 694000, end = 725000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "toy", chr = "4", start = 987000, end = 1009500,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "hth", chr = "3R", start = 10504000, end = 10641000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "wg", chr = "2L", start = 7304000, end = 7319000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.13: Key retinal determination genes
damid.plot(name = "hh", chr = "3R", start = 23126000, end = 23145000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "dpp", chr = "2L", start = 2426000, end = 2460000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "eya", chr = "2L", start = 6526000, end = 6549800,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "so", chr = "2R", start = 7415000, end = 7437000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "dac", chr = "2L", start = 16462000, end = 16486000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.15: ECM related genes
damid.plot(name = "wb", chr = "2L", start = 14260000, end = 14329500,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "Mmp2", chr = "2R", start = 9610000, end = 9687000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 4.16: MAPK genes
damid.plot(name = "pnt", chr = "3R", start = 23288000, end = 23349500,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "RasGAP1", chr = "3L", start = 9835000, end = 9859000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "aos", chr = "3L", start = 16467000, end = 16486000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "kek1", chr = "2L", start = 12814000, end = 12827000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
```

```
# Figure 4.19: CATADA
damid.plot(name = "Diap1", chr = "3L", start = 16037000, end = 16053000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ex", chr = "2L", start = 422000, end = 449000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "kibra", chr = "3R", start = 14695000, end = 14728000,
          dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ban", chr = "3L", start = 627000, end = 650000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
```

```
#Figure 4.20: CATADA
damid.plot(name = "GstD gene cluster", chr = "3R", start = 12360000, end = 12389000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "GstE gene cluster", chr = "2R", start = 18394000, end = 18410000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "GstT4", chr = "X", start = 13401000, end = 13411000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ey", chr = "4", start = 694000, end = 725000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "hth", chr = "3R", start = 10504000, end = 10641000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

# Figure 5.6:
damid.plot(name = "puc", chr = "3R", start = 8102000, end = 8126000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "msn", chr = "3L", start = 2548000, end = 2590000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "cher", chr = "3R", start = 17085000, end = 17130000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "upd3", chr = "X", start = 18270000, end = 18296000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")

#Figure 5.7
damid.plot(name = "Diap1", chr = "3L", start = 16037000, end = 16053000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ex", chr = "2L", start = 422000, end = 449000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "kibra", chr = "3R", start = 14695000, end = 14728000,
           dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
damid.plot(name = "ban", chr = "3L", start = 627000, end = 650000,
            dam.gene = "both", save = TRUE, width = 7, height = 7, ext = "svg")
```

8.1.5 Distribution of genomic features

The below code reads in the *Drosophila* annotated genome and then splits it according to feature. The code then intersects the DamID-seq peaks with the annotated feature to calculate the distribution of the DamID-seq peak throughout the genome.

Installing and loading the packages:

```
```{r}
#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":")
%in% search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("ggplot2", "dplyr", "reschape2", "scales"))

```
```

Reading in the files and splitting the genomic features into separate files:

```{r}
dmel <- read.delim("Genomic_features/dmel-all-r6.27.s.gtf", header = FALSE,
                   col.names = c("Chr", "Source", "Feature", "Start", "End", "Score",
"Strand", "Frame", "Attribute")) %>%
  separate(Attribute, into = c("Gene_ID", "Gene_Symbol", "Transcript_ID", "Transcript_Symbol",
"Notes"), sep = "; ") %>%
  mutate(Gene_ID = word(str_remove_all(Gene_ID, ";"), 2),
         Gene_Symbol = word(str_remove_all(Gene_Symbol, ";"), 2),
         Transcript_ID = word(str_remove_all(Transcript_ID, ";"), 2),
         Transcript_Symbol = word(str_remove_all(Transcript_Symbol, ";"), 2))
dmel.split <- split(dmel, dmel$Feature)
write_delim(dmel.split$`3UTR`, path = "Genomic_features/dmel-3UTR-r6.27.s.gtf", col_names =
FALSE, delim = "\t")
write_delim(dmel.split$`5UTR`, path = "Genomic_features/dmel-5UTR-r6.27.s.gtf", col_names =
FALSE, delim = "\t")
write_delim(dmel.split$exon, path = "Genomic_features/dmel-exon-r6.27.s.gtf", col_names =
FALSE, delim = "\t")
write_delim(dmel.split$gene, path = "Genomic_features/dmel-gene-r6.27.s.gtf", col_names =
FALSE, delim = "\t")
```

Calculating the genomic feature (carried out in Terminal):

```{bash}
cd /Volumes/Shared-2/Research/Cluster\ 5/Harvey/Lab\ Shared/Everything\
else/Kat/GenomicFeatures

bedtools sort -i dmel-3UTR-r6.29.gtf > dmel-3UTR-r6.29.s.gtf

bedtools sort -i dmel-3UTR-r6.27.s.gtf > dmel-3UTR-r6.27.sorted.gtf
bedtools merge -i dmel-3UTR-r6.27.s.gtf > dmel-3UTR-r6.27.merge.bed
bedtools sort -i dmel-5UTR-r6.27.gtf > dmel-5UTR-r6.27.s.gtf
bedtools merge -i dmel-5UTR-r6.27.s.gtf > dmel-5UTR-r6.27.merge.bed
bedtools sort -i dmel-exon-r6.27.gtf > dmel-exon-r6.27.s.gtf
bedtools merge -i dmel-exon-r6.27.s.gtf > dmel-exon-r6.27.merge.bed
bedtools sort -i dmel-gene-r6.27.gtf > dmel-gene-r6.27.s.gtf
bedtools merge -i dmel-gene-r6.27.s.gtf > dmel-gene-r6.27.merge.bed

bedtools subtract -a dmel-gene-r6.27.merge.bed -b dmel-exon-r6.27.merge.bed >
dmel-intron-r6.27.merge.bed
bedtools subtract -a dmel-exon-r6.27.merge.bed -b dmel-3UTR-r6.27.merge.bed >
dmel-exon_no_3UTR-r6.27.merge.bed
bedtools subtract -a dmel-exon-r6.27.merge.bed -b dmel-5UTR-r6.27.merge.bed >
dmel-exon_no_5UTR-r6.27.merge.bed
bedtools intersect -a dmel-exon_no_3UTR-r6.27.merge.bed -b dmel-exon_no_5UTR-r6.27.merge.bed >
dmel-exon_no_UTR-r6.27.merge.bed
```

Calculating the distribution of the genome:

```{r}

dmel.3utr <- read.delim("Drosophilagenome_annotated/dmel-3UTR-r6.27.merge.bed", header =
FALSE)
dmel.5utr <- read.delim("Drosophilagenome_annotated/dmel-5UTR-r6.27.merge.bed", header =
FALSE)
dmel.exon <- read.delim("Drosophilagenome_annotated/dmel-exon-r6.27.merge.bed", header =
FALSE)
dmel.intron <- read.delim("Drosophilagenome_annotated/dmel-intron-r6.27.merge.bed", header =
FALSE)
dm6.chromsizes <- read.delim("Drosophilagenome_annotated/dm6.chrom.sizes.tsv", header = FALSE,
stringsAsFactors = FALSE, col.names = c("Chrom", "Size"))

dist <- data.frame(Feature = factor(c("3UTR", "5UTR", "Exon", "Intron", "Intergenic"),
                                    levels = c("3UTR", "5UTR", "Intron", "Exon",
"Intergenic")),
                   Genome = c(sum(dmel.3utr$V3 - dmel.3utr$V2), sum(dmel.5utr$V3 -
dmel.5utr$V2),
                              sum(dmel.exon$V3 - dmel.exon$V2), sum(dmel.intron$V3 -
dmel.intron$V2),
                              sum(dm6.chromsizes$Size) - (sum(dmel.3utr$V3 - dmel.3utr$V2) +
sum(dmel.5utr$V3 - dmel.5utr$V2) +
                                                            sum(dmel.exon$V3 - dmel.exon$V2) +
sum(dmel.intron$V3 - dmel.intron$V2)))) %>%
  melt("Feature", variable.name = "Dataset", value.name = "Size")
dist <- dist %>%
  group_by(Dataset) %>%
  mutate(Percent = Size / sum(dist$Size[dist$Dataset == Dataset])) %>%
  ungroup()

```

Calculating the distribution of the DamIDseq peaks (carried out in Terminal):

```{bash}

bedtools intersect -a  data/yki_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-3UTR-r6.27.merge.bed > Yki-3UTR.bed
bedtools intersect -a  data/yki_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-5UTR-r6.27.merge.bed > Yki-5UTR.bed
bedtools intersect -a  data/yki_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-exon-r6.27.merge.bed > Yki-exon.bed
bedtools intersect -a  data/yki_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-intron-r6.27.merge.bed > Yki-intron.bed
bedtools subtract -a  data/yki_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-gene-r6.27.merge.bed > Yki-intergenic.bed

bedtools intersect -a  data/sd_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-3UTR-r6.27.merge.bed > Sd-3UTR.bed
bedtools intersect -a  data/sd_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-5UTR-r6.27.merge.bed > Sd-5UTR.bed
bedtools intersect -a  data/sd_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-exon-r6.27.merge.bed > Sd-exon.bed
bedtools intersect -a  data/sd_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-intron-r6.27.merge.bed > Sd-intron.bed
bedtools subtract -a  data/sd_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-gene-r6.27.merge.bed > Sd-intergenic.bed
```

```
bedtools intersect -a  data/tgi_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-3UTR-r6.27.merge.bed > Tgi-3UTR.bed
bedtools intersect -a  data/tgi_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-5UTR-r6.27.merge.bed > Tgi-5UTR.bed
bedtools intersect -a  data/tgi_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-exon-r6.27.merge.bed > Tgi-exon.bed
bedtools intersect -a  data/tgi_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-intron-r6.27.merge.bed > Tgi-intron.bed
bedtools subtract -a  data/tgi_normal_peaks_filtered.bed -b
Drosophilagenome_annotated/dmel-gene-r6.27.merge.bed > Tgi-intergenic.bed

bedtools sort -i data/Dmel_BDGP6.GATC.gff > Dmel_BDGP6.GATC.s.gff
bedtools merge -i Dmel_BDGP6.GATC.s.gff > Dmel_BDGP6.GATC.bed
bedtools intersect -a Dmel_BDGP6.GATC.bed -b dmel-3UTR-r6.27.merge.bed > GATC-3UTR.bed
bedtools intersect -a Dmel_BDGP6.GATC.bed -b dmel-5UTR-r6.27.merge.bed > GATC-5UTR.bed
bedtools intersect -a Dmel_BDGP6.GATC.bed -b dmel-exon-r6.27.merge.bed > GATC-exon.bed
bedtools intersect -a Dmel_BDGP6.GATC.bed -b dmel-intron-r6.27.merge.bed > GATC-intron.bed
bedtools subtract -a Dmel_BDGP6.GATC.bed -b dmel-gene-r6.27.merge.bed > GATC-intergenic.bed

```
```

Importing the files into RStudio:

```
```{r}

dmel.3utr <- read.delim("Genomic_features/dmel-3UTR-r6.27.merge.bed", header = FALSE)
dmel.5utr <- read.delim("Genomic_features/dmel-5UTR-r6.27.merge.bed", header = FALSE)
dmel.exon <- read.delim("Genomic_features/dmel-exon-r6.27.merge.bed", header = FALSE)
dmel.intron <- read.delim("Genomic_features/dmel-intron-r6.27.merge.bed", header = FALSE)
dm6.chromsizes <- read.delim("Genomic_features/dm6.chrom.sizes.tsv", header = FALSE,
stringsAsFactors = FALSE, col.names = c("Chrom", "Size"))

yki.3utr <- read.delim("Genomic_features/Yki-3UTR.bed", header = FALSE)
yki.5utr <- read.delim("Genomic_features/Yki-5UTR.bed", header = FALSE)
yki.exon <- read.delim("Genomic_features/Yki-exon.bed", header = FALSE)
yki.intergenic <- read.delim("Genomic_features/Yki-intergenic.bed", header = FALSE)
yki.intron <- read.delim("Genomic_features/Yki-intron.bed", header = FALSE)

sd.3utr <- read.delim("Genomic_features/Sd-3UTR.bed", header = FALSE)
sd.5utr <- read.delim("Genomic_features/Sd-5UTR.bed", header = FALSE)
sd.exon <- read.delim("Genomic_features/Sd-exon.bed", header = FALSE)
sd.intergenic <- read.delim("Genomic_features/Sd-intergenic.bed", header = FALSE)
sd.intron <- read.delim("Genomic_features/Sd-intron.bed", header = FALSE)

tgi.3utr <- read.delim("Genomic_features/Tgi-3UTR.bed", header = FALSE)
tgi.5utr <- read.delim("Genomic_features/Tgi-5UTR.bed", header = FALSE)
tgi.exon <- read.delim("Genomic_features/Tgi-exon.bed", header = FALSE)
tgi.intergenic <- read.delim("Genomic_features/Tgi-intergenic.bed", header = FALSE)
tgi.intron <- read.delim("Genomic_features/Tgi-intron.bed", header = FALSE)

gatc.3utr <- read.delim("Genomic_features/GATC-3UTR.bed", header = FALSE)
gatc.5utr <- read.delim("Genomic_features/GATC-5UTR.bed", header = FALSE)
gatc.exon <- read.delim("Genomic_features/GATC-exon.bed", header = FALSE)
gatc.intergenic <- read.delim("Genomic_features/GATC-intergenic.bed", header = FALSE)
gatc.intron <- read.delim("Genomic_features/GATC-intron.bed", header = FALSE)

```

```
dist <- data.frame(Feature = factor(c("3UTR", "5UTR", "Exon", "Intron", "Intergenic"),
                                    levels = c("3UTR", "5UTR", "Intron", "Exon",
"Intergenic")),
                   Genome = c(sum(dmel.3utr$V3 - dmel.3utr$V2), sum(dmel.5utr$V3 -
dmel.5utr$V2),
                              sum(dmel.exon$V3 - dmel.exon$V2), sum(dmel.intron$V3 -
dmel.intron$V2),
                              sum(dm6.chromsizes$Size) - (sum(dmel.3utr$V3 - dmel.3utr$V2) +
sum(dmel.5utr$V3 - dmel.5utr$V2) +
                                                           sum(dmel.exon$V3 - dmel.exon$V2) +
sum(dmel.intron$V3 - dmel.intron$V2))),
                   GATC = c(sum(gatc.3utr$V3 - gatc.3utr$V2), sum(gatc.5utr$V3 -
gatc.5utr$V2),
                            sum(gatc.exon$V3 - gatc.exon$V2), sum(gatc.intron$V3 -
gatc.intron$V2),
                            sum(gatc.intergenic$V3 - gatc.intergenic$V2)),
                   Yki = c(sum(yki.3utr$V3 - yki.3utr$V2), sum(yki.5utr$V3 - yki.5utr$V2),
                              sum(yki.exon$V3 - yki.exon$V2), sum(yki.intron$V3 -
yki.intron$V2),
                              sum(yki.intergenic$V3 - yki.intergenic$V2)),
                   Sd = c(sum(sd.3utr$V3 - sd.3utr$V2), sum(sd.5utr$V3 - sd.5utr$V2),
                               sum(sd.exon$V3 - sd.exon$V2), sum(sd.intron$V3 - sd.intron$V2),
                               sum(sd.intergenic$V3 - sd.intergenic$V2)),
                   Tgi = c(sum(tgi.3utr$V3 - tgi.3utr$V2), sum(tgi.5utr$V3 - tgi.5utr$V2),
                              sum(tgi.exon$V3 - tgi.exon$V2), sum(tgi.intron$V3 -
tgi.intron$V2),
                              sum(tgi.intergenic$V3 - tgi.intergenic$V2))) %>%
  melt("Feature", variable.name = "Dataset", value.name = "Size")
dist <- dist %>%
  group_by(Dataset) %>%
  mutate(Percent = Size / sum(dist$Size[dist$Dataset == Dataset])) %>%
  ungroup()

```

Plotting the genomic features of peaks:

```{r}

ggplot()+
  geom_bar(aes(y = Percent, x=Dataset, fill = Feature), data = dist,
           stat="identity", position = "stack") + #scale_fill_brewer(palette="") +
 coord_flip()+ facet_grid(Dataset~., scale="free_y", switch="y") +
  theme_bw()+
  theme(axis.title.y=element_blank(), axis.text.y=element_blank(),
        axis.ticks.y=element_blank()) +
scale_fill_manual(values = c("dodgerblue3","seagreen2", "purple1","magenta","turquoise3"))
ggsave("output/genomic_distribution2.svg", width = 8, height = 5)

```

8.1.6 Proximity to the transcriptional start site (TSS)

Installation and loading libraries:

```
```{r}
# installing packages
install.packages("ChIPseeker")
# loading packages
library(ChIPseeker)
library(TxDb.Dmelanogaster.UCSC.dm6.ensGene)
library(clusterProfiler)

```
```

Loading the peak files and the Drosophila genome:

```
```{r}
# peak files:
peakfiles <- list.files("data/peak files/", pattern= ".bed", full.names=T)
peakfiles <- as.list(peakfiles)
names(peakfiles) <- c("GATC", "Sd", "Tgi", "Yki")

#reading in the drosophila genome file
txdb <- TxDb.Dmelanogaster.UCSC.dm6.ensGene

```
```

Plotting the coverage of peaks over the Drosophila genome:

```
```{r}
#coverage of the peaks over genome
covplot(peakfiles$GATC, weightCol="V5")
covplot(peakfiles$Yki, weightCol="V5")
covplot(peakfiles$Sd, weightCol="V5")
covplot(peakfiles$Tgi, weightCol="V5")
```
```

Calculating and plotting the distance of peaks from transcriptional start sites:

```
```{r}

promoter <- getPromoters(TxDb=txdb, upstream=3000, downstream=3000)
gatc.tagMatrix <- getTagMatrix(peakfiles, windows=promoter)
lapply(peakfiles, annotatePeak, TxDb=txdb,
       tssRegion=c(-1000, 1000), verbose=FALSE)
tagMatrix <- lapply(peakfiles, getTagMatrix, windows=promoter)

tagMatrix <- list.files("data/peak files/", pattern= ".bed", full.names=T)

#below is the plot with statistics on the graph - confidence 95%
TSS.plot2 <- plotAvgProf(tagMatrix, xlim=c(-3000, 3000),
            xlab="Genomic Region (5'->3')", ylab = "Read Count Frequency", facet ="row", conf =
0.95, resample = 1000) +
  scale_colour_manual(values=c("Gray", "dodgerblue1", "darkorchid2", "deeppink1"))

svg(filename="TSSproximity_withconfidenceintervals.svg", units="in",width=5, height=4,
pointsize=12, res=72)

```
```

8.1.7 KEGG pathway enrichment analysis

This code was used to carry out KEGG pathway analysis on all datasets presented in this thesis, including the DamID-seq data and RNA-seq data. The KEGG analysis was visualised either

using a bar plot (used for single dataset), or using a dot plot (used when visualising several datasets).

Installing packages and loading libraries:

```
```{r}
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("AnnotationDbi","org.Dm.eg.db","ggplot2", "limma", "dplyr"))

#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":")
%in% search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("AnnotationDbi","org.Dm.eg.db","ggplot2", "limma", "dplyr"))

```
```

KEGG analysis:

```
``` {r KEGG analysis and graphing }

# KEGG pathway analysis:
keglink = getGeneKEGGLinks(species.KEGG = "dme", convert = TRUE)
kegnames = getKEGGPathwayNames(species.KEGG = "dme", remove.qualifier = FALSE)

# Retrieving Entrez ID:
map.yki = AnnotationDbi::select(org.Dm.eg.db, as.character(yki$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
yki.kegg <- topKEGG(kegga(map.yki$ENTREZID, species = "Dm",gene.pathway = keglink,
pathway.names = kegnames))
write.csv(yki.kegg, file = "Lists/yki.kegg.csv", row.names = TRUE)

map.sd = AnnotationDbi::select(org.Dm.eg.db, as.character(sd$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
sd.kegg <- topKEGG(kegga(map.sd$ENTREZID, species = "Dm",gene.pathway = keglink, pathway.names
= kegnames))
write.csv(sd.kegg, file = "Lists/sd.kegg.csv", row.names = TRUE)

map.tgi = AnnotationDbi::select(org.Dm.eg.db, as.character(tgi$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
tgi.kegg <- topKEGG(kegga(map.tgi$ENTREZID, species = "Dm",gene.pathway = keglink,
pathway.names = kegnames))
write.csv(tgi.kegg, file = "Lists/tgi.kegg.csv", row.names = TRUE)

YkiSdTgi_overlap_GeneID <- data.frame(YkiSdTgi_overlap_GeneID)
names(YkiSdTgi_overlap_GeneID) <- c("Gene_ID")
```
```

```
#for shared target genes:
map = AnnotationDbi::select(org.Dm.eg.db, as.character(YkiSdTgi_overlap_GeneID$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
shared.kegg <- topKEGG(kegga(map$ENTREZID, species = "Dm",gene.pathway = keglink,
pathway.names = kegnames))
write.csv(shared.kegg , file = "Lists/shared.kegg .csv", row.names = TRUE)

# makes a data frame with the Pathways (without '- fly - Drosophila', etc.), P values, and
GeneRatio (how many enriched genes over the number of genes in that category)
yki.kegg.p <- data.frame(Pathway = sub(" - .*", "", yki.kegg$Pathway), P.value =
yki.kegg$P.DE, GeneRatio = yki.kegg$DE/yki.kegg$N)

sd.kegg.p <- data.frame(Pathway = sub(" - .*", "", sd.kegg$Pathway), P.value = sd.kegg$P.DE,
GeneRatio = sd.kegg$DE/sd.kegg$N)

tgi.kegg.p <- data.frame(Pathway = sub(" - .*", "", tgi.kegg$Pathway), P.value =
tgi.kegg$P.DE, GeneRatio = tgi.kegg$DE/tgi.kegg$N)

# removes duplicated Pathways (keeping the most significant one)
yki.kegg.p <- data.frame(Pathway = yki.kegg.p$Pathway[!duplicated(yki.kegg.p$Pathway)],
                         P.value = yki.kegg.p$P.value[!duplicated(yki.kegg.p$Pathway)],
                         GeneRatio =
yki.kegg.p$GeneRatio[!duplicated(yki.kegg.p$Pathway)])

sd.kegg.p <- data.frame(Pathway = sd.kegg.p$Pathway[!duplicated(sd.kegg.p$Pathway)],
                         P.value = sd.kegg.p$P.value[!duplicated(sd.kegg.p$Pathway)],
                         GeneRatio =
sd.kegg.p$GeneRatio[!duplicated(sd.kegg.p$Pathway)])

tgi.kegg.p <- data.frame(Pathway = tgi.kegg.p$Pathway[!duplicated(tgi.kegg.p$Pathway)],
                         P.value = tgi.kegg.p$P.value[!duplicated(tgi.kegg.p$Pathway)],
                         GeneRatio =
tgi.kegg.p$GeneRatio[!duplicated(tgi.kegg.p$Pathway)])

# makes the pathway as a factor to retain the order; this order is from least to most
significant
yki.kegg.p$Pathway <- factor(yki.kegg.p $Pathway, levels = rev(yki.kegg.p $Pathway))

sd.kegg.p$Pathway <- factor(sd.kegg.p $Pathway, levels = rev(sd.kegg.p $Pathway))

tgi.kegg.p$Pathway <- factor(tgi.kegg.p $Pathway, levels = rev(tgi.kegg.p $Pathway))

#take only the significant kegg terms
yki.kegg.p <- subset(yki.kegg.p, P.value <= 0.05)
sd.kegg.p <- subset(sd.kegg.p, P.value <= 0.05)
tgi.kegg.p <- subset(tgi.kegg.p, P.value <= 0.05)

```

Plotting a bar graph the KEGG analysis results:

```{r}

yki.kegg_plot <- ggplot(data = subset(yki.kegg.p, P.value <= 0.05), aes(x = Pathway, y =
-log10(P.value), fill =  Pathway)) +
  coord_flip() +
geom_col(stat = "identity", width = 0.5) +
scale_fill_hue(l=70) +
theme_classic() +
guides(fill = FALSE) +
labs(x = "", y = "-Log(P value)") +
ggtitle("") +
theme(plot.title=element_text(size=16, face= "bold")) +
theme(axis.text=element_text(size=12), axis.title=element_text(size=12))
ggsave("output/yki.kegg_plot.png")

sd.kegg_plot <- ggplot(data = subset(sd.kegg.p, P.value <= 0.05), aes(x = Pathway, y =
-log10(P.value), fill =  Pathway)) +
  coord_flip() +
geom_col(stat = "identity", width = 0.5) +
scale_fill_hue(l=70) +
theme_classic() +
guides(fill = FALSE) +
labs(x = "", y = "-Log(P value)") +
ggtitle("") +
theme(plot.title=element_text(size=16, face= "bold")) +
theme(axis.text=element_text(size=12), axis.title=element_text(size=12))
ggsave("output/sd.kegg_plot.png")

tgi.kegg_plot <- ggplot(data = subset(tgi.kegg.p, P.value <= 0.05), aes(x = Pathway, y =
-log10(P.value), fill =  Pathway)) +
  coord_flip() +
geom_col(stat = "identity", width = 0.5) +
scale_fill_hue(l=70) +
theme_classic() +
guides(fill = FALSE) +
labs(x = "", y = "-Log(P value)") +
ggtitle("") +
theme(plot.title=element_text(size=16, face= "bold")) +
theme(axis.text=element_text(size=12), axis.title=element_text(size=12))
ggsave("output/tgi.kegg_plot.png")

```

Plotting a dot plot the KEGG analysis results:

```r
# place all of the dataframes into a list
kegg.list <- list(yki.kegg.p, sd.kegg.p,tgi.kegg.p)
names(kegg.list) <- c("Yki", "Sd", "Tgi")

# make into a merged dataframe
kegg <- plyr::ldply(kegg.list, rbind)

# make the pvalue a minus log to the 10 p value
kegg$P.value <- -log10(kegg$P.value)
names(kegg) <- c("Protein", "Pathway", "Adjusted.Pvalue", "GeneRatio")

# plotting the data:
ggplot(data = kegg, aes(x=Protein , y=reorder(Pathway, Adjusted.Pvalue), color=
Adjusted.Pvalue, size=GeneRatio)) +
    geom_point() + scale_color_gradient(name = "-Log10(P-value)", low="blue", high ="red") +
  scale_size_continuous(range = c(1,6)) +
  labs(x="", y = "", colour ="Adjusted.Pvalue") +
  theme_bw() +
  scale_x_discrete(limits=c("Yki", "Sd","Tgi"))

ggsave("output/kegg_plot.svg", width = 8, height = 7)
```

8.1.8 GO enrichment analysis

This code was used to carry out GO analysis on all datasets presented in this thesis, including the DamID-seq data and RNA-seq data. The GO analysis was visualised using a dot plot (used when visualising several datasets).

Installing packages and loading libraries:

```r
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("AnnotationDbi","org.Dm.eg.db","ggplot2", "limma", "dplyr"))
#loading libraries:
libraries <- function(packages) { for(p in packages) { if(!(paste("package", p, sep = ":")
%in% search())) { lapply(p, require, character.only = TRUE, quietly = TRUE)
}}}
print(paste(p, "package has been loaded."))
libraries(list("AnnotationDbi","org.Dm.eg.db","ggplot2", "limma", "dplyr"))
```

GO analysis:

```r
map = AnnotationDbi::select(org.Dm.eg.db, as.character(YkiSdTgi_overlap_GeneID$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
GO.shared <- topGO(goana(as.character(map$ENTREZID), species = "Dm"))
write.csv(GO.shared, file = "Lists/GO.csv", row.names = TRUE)

map.yki = AnnotationDbi::select(org.Dm.eg.db, as.character(yki$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
GO.yki <- topGO(goana(as.character(map.yki$ENTREZID), species = "Dm"))
write.csv(GO.yki, file = "Lists/GO.yki.csv", row.names = TRUE)

map.sd = AnnotationDbi::select(org.Dm.eg.db, as.character(sd$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
GO.sd <- topGO(goana(as.character(map.sd$ENTREZID), species = "Dm"))
write.csv(GO.sd, file = "Lists/GO.sd.csv", row.names = TRUE)

map.tgi = AnnotationDbi::select(org.Dm.eg.db, as.character(tgi$Gene_ID),
keytype="ENSEMBL",columns=c("ENTREZID","SYMBOL"))
GO.tgi <- topGO(goana(as.character(map.tgi$ENTREZID), species = "Dm"))
write.csv(GO.tgi, file = "Lists/GO.tgi.csv", row.names = TRUE)


# makes a data frame with the Pathways (without '- fly - Drosophila', etc.), P values, and
GeneRatio (how many enriched genes over the number of genes in that category)
GO.yki.p <- data.frame(Term = sub(" - .*", "", GO.yki$Term), P.value = GO.yki$P.DE, GeneRatio
= GO.yki$DE/GO.yki$N)
GO.sd.p <- data.frame(Term = sub(" - .*", "", GO.sd$Term), P.value = GO.sd$P.DE, GeneRatio =
GO.sd$DE/GO.yki$N)
GO.tgi.p <- data.frame(Term = sub(" - .*", "", GO.tgi$Term), P.value = GO.tgi$P.DE, GeneRatio
= GO.tgi$DE/GO.yki$N)

# removes duplicated Pathways (keeping the most significant one)
GO.yki.p <- data.frame(Term = GO.yki.p$Term[!duplicated(GO.yki.p$Term)],
                       P.value =  GO.yki.p$P.value[!duplicated(GO.yki.p$Term)],
                       GeneRatio =
GO.yki.p$GeneRatio[!duplicated(GO.yki.p$Term)])

GO.sd.p <- data.frame(Term = GO.sd.p$Term[!duplicated(GO.sd.p$Term)],
                       P.value =  GO.sd.p$P.value[!duplicated(GO.sd.p$Term)],
                       GeneRatio =
GO.sd.p$GeneRatio[!duplicated(GO.sd.p$Term)])

GO.tgi.p <- data.frame(Term = GO.tgi.p$Term[!duplicated(GO.tgi.p$Term)],
                       P.value =  GO.tgi.p$P.value[!duplicated(GO.tgi.p$Term)],
                       GeneRatio =
GO.tgi.p$GeneRatio[!duplicated(GO.tgi.p$Term)])
```

```
# makes the pathway as a factor to retain the order; this order is from least to most
significant...
GO.yki.p$Term <- factor(GO.yki.p$Term, levels = rev(GO.yki.p$Term))
GO.sd.p$Term <- factor(GO.sd.p$Term, levels = rev(GO.sd.p$Term))
GO.tgi.p$Term <- factor(GO.tgi.p$Term, levels = rev(GO.tgi.p$Term))

#take only the significant kegg terms
GO.yki.p <- subset(GO.yki.p, P.value <= 0.05)
GO.sd.p <- subset(GO.sd.p, P.value <= 0.05)
GO.tgi.p <- subset(GO.tgi.p, P.value <= 0.05)

```

Plotting a dot plot the KEGG analysis results:

```{r}
#place all of the dataframes into a list
go.list <- list(GO.yki.p, GO.sd.p, GO.tgi.p)
names(go.list) <- c("Yki", "Sd", "Tgi")

#make into a merged dataframe
go <- plyr::ldply(go.list, rbind)

#make the pvalue a minus log to the 10 p value
go$P.value <- -log10(go$P.value)
names(go) <- c("Protein", "Term", "Adjusted.Pvalue", "GeneRatio")

# plotting the results
ggplot(data = go, aes(x=Protein , y=reorder(Term, Adjusted.Pvalue), color= Adjusted.Pvalue,
size=GeneRatio)) +
    geom_point() + scale_color_gradient(name = "-Log10(P-value)", low="blue", high ="red") +
  scale_size_continuous(range = c(1,6)) +
  labs(x="", y = "", colour ="Adjusted.Pvalue") +
  theme_bw() +
  scale_x_discrete(limits=c("Yki", "Sd","Tgi"))
ggsave("output/go_plot.svg", width = 8, height = 7)

```

8.1.9 Motif enrichment analysis

Motif enrichment analysis was completed using HOMER in Terminal

The following code takes peak bed files and converts it to a better .bed file structure (that works with HOMER analysis). The order of columns it likes is: chromosome, start of the peak, end of the peak, name (here, 'chr:start-end'), score (here, logFC), strand (here, '.' = no strand)

```{r}

yki.peaks %>%
  mutate(name = paste0(Chr, ":", Start, "-", End), chrom = paste0("chr", Chr), strand = ".")
%>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End, name, score = logFC, strand) %>%
  write_delim(., path = "Yki_peaks_motif.bed", col_names = FALSE, delim = "\t")

sd.peaks %>%
  mutate(name = paste0(Chr, ":", Start, "-", End), chrom = paste0("chr", Chr), strand = ".")
%>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End, name, score = logFC, strand) %>%
  write_delim(., path = "sd_peaks_motif.bed", col_names = FALSE, delim = "\t")

tgi.peaks %>%
  mutate(name = paste0(Chr, ":", Start, "-", End), chrom = paste0("chr", Chr), strand = ".")
%>%
  dplyr::select(chrom, chromStart = Start, chromEnd = End, name, score = logFC, strand) %>%
  write_delim(., path = "tgi_peaks_motif.bed", col_names = FALSE, delim = "\t")

read.delim("overlapping.peaks.yki.sd.tgi.bed", header = FALSE, col.names = c("Chr", "Start",
"End", "", "", "")) %>%
  mutate(name = paste0(Chr, ":", Start, "-", End), score = "", strand = ".") %>%
  dplyr::select(Chr, chromStart = Start, chromEnd = End, name) %>%
  write_delim(., path = "bed files/overlap_all_peaks_merge_motif.bed", col_names = FALSE, delim
= "\t")


read.delim("overlapping.peaks.yki.sd.tgi.bed", header = FALSE, col.names = c("Chr", "Start",
"End", "", "", "")) %>%
  mutate(name = paste0(Chr, ":", Start, "-", End), score = "", strand = ".") %>%
  dplyr::select(Chr, chromStart = Start, chromEnd = End, name) %>%
  write_delim(., path = "bed files/overlap_all_peaks_merge_motif.bed", col_names = FALSE, delim
= "\t")
```

Homer analysis in Terminal:

```{bash}
cd

findMotifsGenome.pl yki.peaks.bed dm6 190509_yki_fdr100/ -fdr 100 -p 5
findMotifsGenome.pl sd.normal.peaks.bed dm6 190509_sd_fdr100/ -fdr 100 -p 5
findMotifsGenome.pl tgi.normal.peaks.bed dm6 190509_tgi_fdr100/ -fdr 100 -p 5
findMotifsGenome.pl jra_peaks.bed  dm6 190509_jra_fdr100/ -fdr 100 -p 5
```

8.1.10 Heatmaps

The code below subsets, orders, and creates a matrix which is then used to generate a heatmap of the selected genes. This was carried out on the RNAseq dataset.

Installing packages and loading libraries:

```
```{r}
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("gplots","RColorBrewer"))
library("gplots")
library("RColorBrewer")
```
```

Subsetting genes and logcounts and plotting on a heatmap:

```
```{r}
#Figure 4.3: Known Hippo pathway target genes
## subset gene symbols
knowntargets.list <- list("ft", "ex", "Diap1", "CycE", "kibra", "Ilp8")
knowntargets <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% knowntargets.list)
knowntargets.matrix <- data.matrix(knowntargets, rownames.force = NULL)
## save as svg
svglite(file = "output/heatmap_knowntargets_eye.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(knowntargets.matrix,col=rev(morecols(50)),trace="none", main="Known target
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )

# Figure 4.8: SPALT genes
## subset gene symbols
spalt.genes.list <- list("Gadd45", "GstE6", "GstE8", "CG3008", "mthl8", "Ilp8", "Irbp",
"mre11", "Irbp18", "ImpL2", "ltl", "Arc1","Arc2","Ady43A", "Hen1", "Cul2", "Cpr11A",
"CG31875", "CG32581", "CG32625", "CG3448", "CG15784", "AGO3", "Snap25", "Cyp6d2", "salm",
"salr", "brk")
spalt.genes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% spalt.genes.list)
spalt.genes.matrix <- data.matrix(spalt.genes, rownames.force = NULL)
row.reorder <- c("Gadd45", "GstE6", "GstE8", "CG3008", "mthl8", "Ilp8", "Irbp", "mre11",
"Irbp18", "ImpL2", "ltl", "Arc1","Arc2","Ady43A", "Hen1", "Cul2", "Cpr11A", "CG31875",
"CG32581", "CG32625", "CG3448", "CG15784", "AGO3", "Snap25", "Cyp6d2", "salm", "salr", "brk")
spalt.genes.matrix <- spalt.genes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_Spaltgenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(spalt.genes.matrix,col=rev(morecols(50)),trace="none", main="Spalt
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )
```
```

```
# Figure 4.9: GST genes
## subset gene symbols
GST.genes.list <- list("GstD3", "GstD10", "GstE1", "GstE3", "GstE6", "GstE8", "GstE10",
"GstT4")
GST.genes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% GST.genes.list)
GST.genes.matrix <- data.matrix(GST.genes, rownames.force = NULL)
row.reorder <- c("GstD3", "GstD10", "GstE1", "GstE3", "GstE6", "GstE8", "GstE10", "GstT4")
GST.genes.matrix <- GST.genes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_GSTgenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(GST.genes.matrix,col=rev(morecols(50)),trace="none", main="GST
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )
```

```
# Figure 4.10: cuticle related genes
## subset gene symbols
cuticle.genes.list <- list("TwdlE", "obst-E", "Lcp65Ag1", "Lcp65Ag2", "Lcp65Ag3", "Lcp65Ae",
"Lcp65Af", "Cpr78E", "Cpr66D", "Cpr92A", "Cpr56F", "CG1136" )
cuticle.genes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% cuticle.genes.list)
cuticle.genes.matrix <- data.matrix(cuticle.genes, rownames.force = NULL)
row.reorder <- c("TwdlE", "obst-E", "Lcp65Ag1", "Lcp65Ag2", "Lcp65Ag3", "Lcp65Ae", "Lcp65Af",
"Cpr78E", "Cpr66D", "Cpr92A", "Cpr56F", "CG1136")
cuticle.genes.matrix <- cuticle.genes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_cuticlegenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(cuticle.genes.matrix,col=rev(morecols(50)),trace="none", main="cuticle
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )

# Figure 4.11: Early retinal determinant genes
## subset gene symbols
retinal.earlyeyegenes.list <- list("ey", "tsh", "toy", "hth", "wg", "tio")
retinal.earlyeyegenes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in%
retinal.earlyeyegenes.list)
retinal.earlyeyegenes.matrix <- data.matrix(retinal.earlyeyegenes, rownames.force = NULL)
row.reorder <- c("ey", "tsh", "toy", "hth", "wg", "tio")
retinal.earlyeyegenes.matrix <- retinal.earlyeyegenes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_earlyeyegenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(retinal.earlyeyegenes.matrix,col=rev(morecols(50)),trace="none", main="Early eye
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )
```

```
# Figure 4.12: Later retinal determinant genes
## subset gene symbols
retinal.determinationgenes.list <- list("hh", "eya", "so", "dac", "dan", "danr", "shf",
"ato","Optix", "stg", "dpp", "N", "h", "emc", "ro", "da")
retinal.determinationgenes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in%
retinal.determinationgenes.list)
retinal.determinationgenes.matrix <- data.matrix(retinal.determinationgenes, rownames.force =
NULL)
row.reorder <- c("hh", "eya", "so", "dac", "dan", "danr", "shf", "ato","Optix", "stg", "dpp",
"N", "h", "emc", "ro", "da")
retinal.determinationgenes.matrix <- retinal.determinationgenes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_retinaldeterminationgenes.svg", width = 10, height = 8, bg =
"white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(retinal.determinationgenes,col=rev(morecols(50)),trace="none", main="Retinal
determinant genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )

# Figure 4.14: Differentiation related genes
## subset gene symbols
retinal.differentiationgenes.list <- list("sens", "rho", "ru", "pnt", "spi", "Dl", "sca",
"sev", "boss", "pros", "svp", "lz", "gl", "phyl", "salm", "oc", "ss", "melt", "elav", "gogo",
"jeb", "retn", "Rh6", "rdgC", "inaC", "Ekar", "sls", "pes", "cd", "ry", "E(spl)mdelta-HLH",
"E(spl)mgamma-HLH", "E(spl)m4-BFM", "E(spl)m2-BFM", "E(spl)m5-HLH", "E(spl)m7-HLH",
"E(spl)m8-HLH")
retinal.differentiationgenes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in%
retinal.differentiationgenes.list)
retinal.differentiationgenes.matrix <- data.matrix(retinal.differentiationgenes,
rownames.force = NULL)
row.reorder <- c("sens", "rho", "ru", "pnt", "spi", "Dl", "sca", "sev", "boss", "pros", "svp",
"lz", "gl", "phyl", "salm", "oc", "ss", "melt", "elav", "gogo", "jeb", "retn", "Rh6", "rdgC",
"inaC", "Ekar", "sls", "pes", "cd", "ry", "E(spl)mdelta-HLH", "E(spl)mgamma-HLH",
"E(spl)m4-BFM", "E(spl)m2-BFM", "E(spl)m5-HLH", "E(spl)m7-HLH", "E(spl)m8-HLH")
retinal.differentiationgenes.matrix <- retinal.differentiationgenes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_retinaldifferentiaitongenes.svg", width = 10, height = 8, bg =
"white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(retinal.differentiationgenes.matrix,col=rev(morecols(50)),trace="none",
main="Retinal differentiation genes",ColSideColors=col.cell,scale="row", Colv=FALSE,
Rowv=FALSE )
```

```
# Figure 4.15: ECM genes
## subset gene symbols
ecm.genes.list <- list("wb", "LanA", "trol", "Col4a1", "vkg", "SPARC", "mspo", "Ppn", "Mp",
"frac", "Tig", "Pxn", "Mmp2")
ecm.genes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% ecm.genes.list)
ecm.genes.matrix <- data.matrix(ecm.genes, rownames.force = NULL)
row.reorder <- c("wb", "LanA", "trol", "Col4a1", "vkg", "SPARC", "mspo", "Ppn", "Mp", "frac",
"Tig", "Pxn", "Mmp2")
ecm.genes.matrix <- ecm.genes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_ecmgenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(ecm.genes.matrix,col=rev(morecols(50)),trace="none", main="ECM
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )

# Figure 4.16: MAPK genes
## subset gene symbols
mapk.genes.list <- list("pnt","aos", "kek1", "kek2", "kek5", "sty", "RasGAP1", "RapGAP1",
"phyl", "Pvr", "Dsor1", "egr", "peb", "boss", "csw", "btl", "bnl", "ths", "pyr")
mapk.genes <- subset(logcounts_symbol, rownames(logcounts_symbol) %in% mapk.genes.list)
mapk.genes.matrix <- data.matrix(mapk.genes, rownames.force = NULL)
row.reorder <- c("pnt","aos", "kek1", "kek2", "kek5", "sty", "RasGAP1", "RapGAP1", "phyl",
"Pvr", "Dsor1", "egr", "peb", "boss", "csw", "btl", "bnl", "ths", "pyr")
mapk.genes.matrix <- mapk.genes.matrix[row.reorder,]
## save as svg
svglite(file = "output/heatmap_mapkgenes.svg", width = 10, height = 8, bg = "white",
        pointsize = 12, standalone = TRUE, system_fonts = list(),
        user_fonts = list())
## create the heatmap
heatmap.2(mapk.genes.matrix,col=rev(morecols(50)),trace="none", main="MAPK
genes",ColSideColors=col.cell,scale="row", Colv=FALSE, Rowv=FALSE )


```

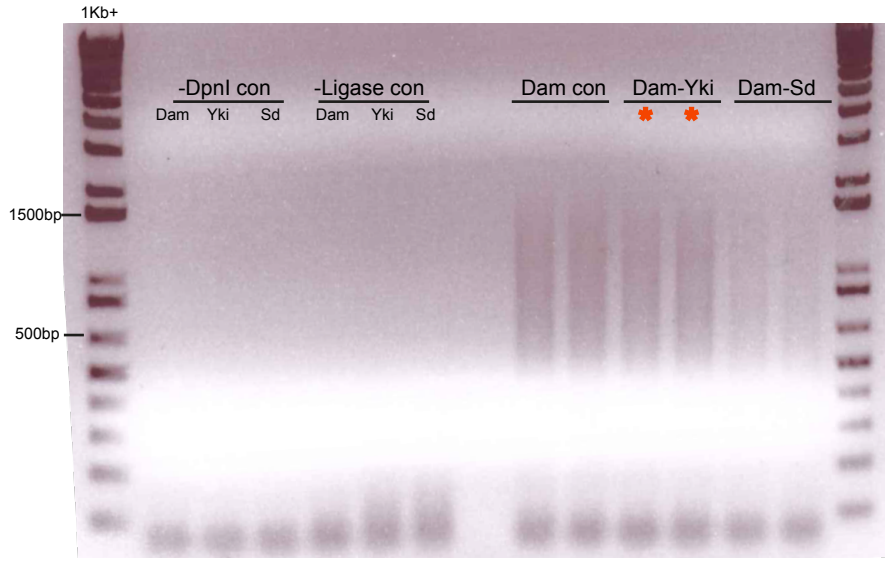
8.2 Figures and Tables

Figure 8.2.1: Targeted DamID DNA methylation and gene profiles for Dam-Yki transgene

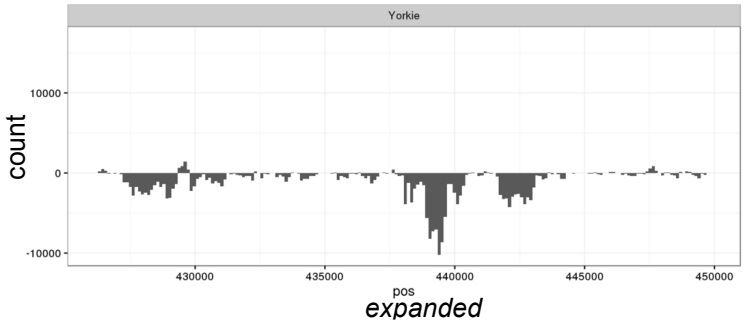
A. Representative PCR amplified adenine methylated DNA fragments for Dam-Yki, Dam-Sd, and Dam-alone control. The red asterix indicates the Dam-Yki replicates. An aliquot of 5µl from each DamID PCR sample was separated on a 0.8% agarose gel. The DNA ladder is shown in the 1kb+ lane. A smear concentrated between 200bp and 2kb can be seen in each experimental sample (+), whereas no smears were observed in the negative control lanes (-DpnI and -Lig). The lower band shows primer dimer binding, which was subsequently eliminated at the library preparation stage, prior to sequencing.

B-D. Gene binding profiles showing adenine methylation by Dam-Yki at the *expanded* (B), *Diap1* (C), and *kibra* (D) genes. Gene binding profiles for Dam-Yki represent differential methylation in which GATC sequences were differentially methylated higher in the Dam-fusion protein compared to the Dam-alone control. The gene binding profiles for Dam-Yki are shown on the y-axis and represent the subtraction of the Dam-fusion counts from the Dam-alone control counts. The x-axis shows the genomic region of each gene locus.

A

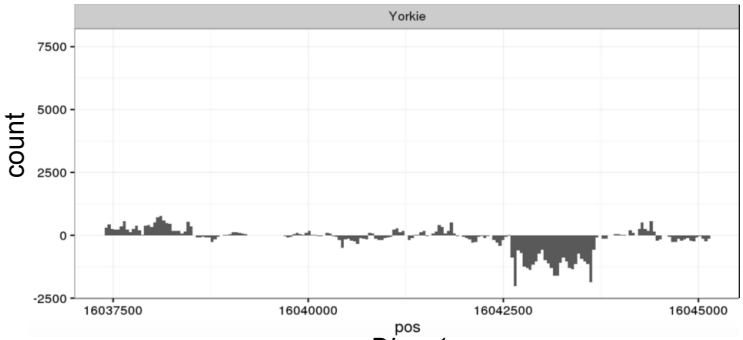


B



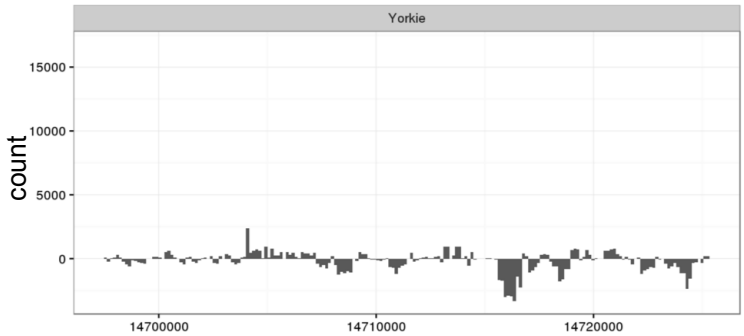
expanded

C



Diap-1

D



kibra

Table 8.2.1: KEGG analysis of Yorkie, Scalloped, and Tgi target genes

A. KEGG analysis of shared Yorkie, Scalloped and Tgi target genes in normal growth

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	32	3.53E-11
path:dme04391	Hippo signaling pathway	61	24	4.84E-10
path:dme04310	Wnt signaling pathway	106	26	5.23E-06
path:dme04512	ECM-receptor interaction	12	6	0.00045489
path:dme04320	Dorso-ventral axis formation	28	9	0.00091633
path:dme04214	Apoptosis	63	14	0.00233937
path:dme04341	Hedgehog signaling pathway	38	9	0.00901902
path:dme04150	mTOR signaling pathway	96	17	0.01005294
path:dme04144	Endocytosis	122	20	0.01284963
path:dme04350	TGF-beta signaling pathway	41	9	0.01494057
path:dme04330	Notch signaling pathway	24	6	0.02423316
path:dme00340	Histidine metabolism	8	3	0.03543856
path:dme00600	Sphingolipid metabolism	28	6	0.04880132

B. KEGG analysis of Yorkie target genes in normal growth

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	40	9.47E-12
path:dme04391	Hippo signaling pathway	61	30	7.09E-11
path:dme04310	Wnt signaling pathway	106	33	5.95E-06
path:dme04512	ECM-receptor interaction	12	7	0.00050319
path:dme04320	Dorso-ventral axis formation	28	10	0.00397468
path:dme04330	Notch signaling pathway	24	9	0.00428516
path:dme04214	Apoptosis	63	17	0.00609077
path:dme04144	Endocytosis	122	27	0.01233546
path:dme04341	Hedgehog signaling pathway	38	11	0.01493723
path:dme00532	Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	8	4	0.0183541
path:dme01230	Biosynthesis of amino acids	66	16	0.02177706

path:dme04350	TGF-beta signaling pathway	41	11	0.02625473
path:dme00230	Purine metabolism	101	21	0.04776357

C. KEGG analysis of Scalloped target genes in normal growth

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	43	4.77E-12
path:dme04391	Hippo signaling pathway	61	32	4.49E-11
path:dme04310	Wnt signaling pathway	106	33	7.54E-05
path:dme04624	Toll and Imd signaling pathway	67	23	0.00018877
path:dme04214	Apoptosis	63	21	0.00056366
path:dme04512	ECM-receptor interaction	12	7	0.00105747
path:dme04341	Hedgehog signaling pathway	38	14	0.00158573
path:dme04350	TGF-beta signaling pathway	41	13	0.00999182
path:dme00250	Alanine, aspartate and glutamate metabolism	30	10	0.01586992
path:dme02010	ABC transporters	22	8	0.01747153
path:dme04144	Endocytosis	122	29	0.01751632
path:dme00564	Glycerophospholipid metabolism	63	17	0.01934326
path:dme04320	Dorso-ventral axis formation	28	9	0.02758268
path:dme00230	Purine metabolism	101	24	0.02924022
path:dme01040	Biosynthesis of unsaturated fatty acids	25	8	0.03793639
path:dme04711	Circadian rhythm	13	5	0.04559352

D. KEGG analysis of Tgi target genes in normal growth

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	38	4.24E-09
path:dme04391	Hippo signaling pathway	61	27	1.12E-07
path:dme04310	Wnt signaling pathway	106	33	5.26E-05
path:dme04214	Apoptosis	63	21	0.00044189
path:dme01212	Fatty acid metabolism	53	17	0.00245929
path:dme04341	Hedgehog signaling pathway	38	13	0.00418949

path:dme00520	Amino sugar and nucleotide sugar metabolism	47	15	0.00459622
path:dme04512	ECM-receptor interaction	12	6	0.00618546
path:dme00564	Glycerophospholipid metabolism	63	18	0.00739946
path:dme04320	Dorso-ventral axis formation	28	10	0.00829615
path:dme01040	Biosynthesis of unsaturated fatty acids	25	9	0.01142881
path:dme04150	mTOR signaling pathway	96	24	0.01314463
path:dme01100	Metabolic pathways	1138	202	0.02058649
path:dme04350	TGF-beta signaling pathway	41	12	0.02199771
path:dme00340	Histidine metabolism	8	4	0.02604452
path:dme00760	Nicotinate and nicotinamide metabolism	20	7	0.02918379
path:dme00650	Butanoate metabolism	16	6	0.03046775
path:dme00920	Sulfur metabolism	9	4	0.04109307

Table 8.2.2: GO analysis of Yorkie, Scalloped, and Tgi target genes

A. GO analysis of shared Yorkie, Scalloped and Tgi target genes in normal growth

GO ID	Term	Ont	N	DE	P.DE
GO:0048731	system development	BP	1874	475	4.55E-113
GO:0007275	multicellular organism development	BP	2577	547	5.67E-99
GO:0009653	anatomical structure morphogenesis	BP	1616	413	8.52E-97
GO:0009888	tissue development	BP	1138	338	2.30E-96
GO:0048513	animal organ development	BP	1312	366	2.65E-96
GO:0060429	epithelium development	BP	1045	317	2.14E-92
GO:0032502	developmental process	BP	3087	597	5.22E-92
GO:0048856	anatomical structure development	BP	2988	582	3.25E-90
GO:0035295	tube development	BP	758	259	7.59E-87
GO:0009887	animal organ morphogenesis	BP	802	264	1.79E-84
GO:0048729	tissue morphogenesis	BP	651	232	1.99E-81

GO:0002009	morphogenesis of an epithelium	BP	633	224	8.93E-78
GO:0048869	cellular developmental process	BP	1842	413	3.25E-77
GO:0030154	cell differentiation	BP	1793	402	9.22E-75
GO:0007444	imaginal disc development	BP	608	214	1.26E-73
GO:0032501	multicellular organismal process	BP	3949	652	7.93E-71
GO:0007399	nervous system development	BP	1143	299	1.83E-69
GO:0035239	tube morphogenesis	BP	532	194	1.95E-69
GO:0065007	biological regulation	BP	4238	677	2.94E-68
GO:0060562	epithelial tube morphogenesis	BP	493	183	8.66E-67

B. GO analysis of Yorkie target genes in normal growth

GO ID	Term	Ont	N	DE	P.DE
GO:0048731	system development	BP	1874	609	3.49E-115
GO:0007275	multicellular organism development	BP	2577	727	5.10E-106
GO:0009653	anatomical structure morphogenesis	BP	1616	533	2.78E-101
GO:0009888	tissue development	BP	1138	427	1.02E-99
GO:0032502	developmental process	BP	3087	800	1.92E-97
GO:0048856	anatomical structure development	BP	2988	781	1.79E-96
GO:0060429	epithelium development	BP	1045	397	5.34E-94
GO:0048513	animal organ development	BP	1312	456	1.34E-93
GO:0035295	tube development	BP	758	320	1.82E-88
GO:0009887	animal organ morphogenesis	BP	802	323	6.77E-83
GO:0048729	tissue morphogenesis	BP	651	285	8.73E-83
GO:0065007	biological regulation	BP	4238	947	1.45E-79
GO:0002009	morphogenesis of an epithelium	BP	633	274	4.43E-78
GO:0048869	cellular developmental process	BP	1842	535	8.84E-78
GO:0007399	nervous system development	BP	1143	390	5.39E-76
GO:0030154	cell differentiation	BP	1793	518	5.61E-74
GO:0007444	imaginal disc development	BP	608	257	2.12E-70
GO:0022008	neurogenesis	BP	909	326	1.21E-68
GO:0032501	multicellular organismal process	BP	3949	874	3.29E-68
GO:0050789	regulation of biological process	BP	3781	847	5.47E-68

C. GO analysis of Scalloped target genes in normal growth

GO ID	Term	Ont	N	DE	P.DE
GO:0048731	system development	BP	1874	626	1.45E-104
GO:0007275	multicellular organism development	BP	2577	756	2.52E-97
GO:0009653	anatomical structure morphogenesis	BP	1616	551	4.79E-94
GO:0032502	developmental process	BP	3087	837	8.52E-90
GO:0048856	anatomical structure development	BP	2988	818	1.62E-89
GO:0048513	animal organ development	BP	1312	474	3.17E-89
GO:0009888	tissue development	BP	1138	432	1.37E-88
GO:0060429	epithelium development	BP	1045	400	1.03E-82
GO:0035295	tube development	BP	758	326	1.09E-81
GO:0009887	animal organ morphogenesis	BP	802	337	2.46E-81
GO:0048729	tissue morphogenesis	BP	651	285	4.48E-73
GO:0007399	nervous system development	BP	1143	406	1.15E-72
GO:0048869	cellular developmental process	BP	1842	557	2.27E-72
GO:0007444	imaginal disc development	BP	608	272	3.91E-72
GO:0030154	cell differentiation	BP	1793	544	5.44E-71
GO:0002009	morphogenesis of an epithelium	BP	633	276	3.37E-70
GO:0065007	biological regulation	BP	4238	989	4.73E-69
GO:0022008	neurogenesis	BP	909	342	1.23E-67
GO:0032501	multicellular organismal process	BP	3949	932	5.09E-66
GO:0048699	generation of neurons	BP	869	327	1.66E-64

D. GO analysis of Tgi target genes in normal growth

GO ID	Term	Ont	N	DE	P.DE
GO:0048731	system development	BP	1874	620	1.27E-102
GO:0007275	multicellular organism development	BP	2577	752	5.65E-97
GO:0009653	anatomical structure morphogenesis	BP	1616	541	1.01E-89
GO:0032502	developmental process	BP	3087	830	2.38E-88
GO:0048856	anatomical structure development	BP	2988	809	4.16E-87
GO:0048513	animal organ development	BP	1312	465	4.64E-85

GO:0009888	tissue development	BP	1138	423	4.24E-84
GO:0035295	tube development	BP	758	322	1.03E-79
GO:0009887	animal organ morphogenesis	BP	802	332	8.93E-79
GO:0060429	epithelium development	BP	1045	390	1.48E-77
GO:0048729	tissue morphogenesis	BP	651	285	9.71E-74
GO:0007444	imaginal disc development	BP	608	268	5.97E-70
GO:0002009	morphogenesis of an epithelium	BP	633	274	1.84E-69
GO:0007399	nervous system development	BP	1143	396	5.23E-68
GO:0022008	neurogenesis	BP	909	337	1.89E-65
GO:0048869	cellular developmental process	BP	1842	539	4.02E-65
GO:0032501	multicellular organismal process	BP	3949	924	9.35E-65
GO:0065007	biological regulation	BP	4238	972	1.44E-64
GO:0035239	tube morphogenesis	BP	532	238	2.13E-63
GO:0030154	cell differentiation	BP	1793	524	8.06E-63

Table 8.2.3: Comparisons of targeted DamID and Ikmi et al (2014) ChIP-seq data

A. KEGG analysis of putative Yki target genes shared between DamID and Ikmi ChIP seq

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	35	1.35E-13
path:dme04391	Hippo signaling pathway	61	24	3.98E-10
path:dme04310	Wnt signaling pathway	106	29	9.40E-08
path:dme04341	Hedgehog signaling pathway	38	11	6.30E-04
path:dme04330	Notch signaling pathway	24	8	1.28E-03
path:dme04144	Endocytosis -	122	22	2.54E-03
path:dme04512	ECM-receptor interaction	12	5	3.64E-03
path:dme04320	Dorso-ventral axis formation	28	8	3.82E-03
path:dme04214	Apoptosis	63	13	6.02E-03
path:dme04150	mTOR signaling pathway	96	16	1.98E-02
path:dme00534	Glycosaminoglycan biosynthesis - heparan sulfate / heparin	13	4	3.02E-02
path:dme00340	Histidine metabolism	8	3	3.46E-02
path:dme04350	TGF-beta signaling pathway	41	8	3.90E-02
path:dme04213	Longevity regulating pathway -	56	10	3.91E-02
path:dme00515	Mannose type O-glycan biosynthesis	4	2	4.89E-02
path:dme04624	Toll and Imd signaling pathway	67	11	5.36E-02
path:dme00250	Alanine, aspartate and glutamate metabolism	30	6	6.29E-02
path:dme00760	Nicotinate and nicotinamide metabolism -	20	4	1.20E-01
path:dme04711	Circadian rhythm	13	3	1.23E-01

B. KEGG analysis of putative Sd target genes shared between DamID and Ikmi ChIP seq

KEGG ID	Pathway	N	DE	P.DE
path:dme04391	Hippo signaling pathway	61	16	3.40E-10
path:dme04013	MAPK signaling pathway -	93	16	2.29E-07
path:dme04310	Wnt signaling pathway	106	14	3.39E-05
path:dme04341	Hedgehog signaling pathway	38	7	4.56E-04
path:dme04214	Apoptosis	63	8	2.31E-03

path:dme04144	Endocytosis	122	11	5.92E-03
path:dme04512	ECM-receptor interaction	12	3	9.16E-03
path:dme04330	Notch signaling pathway	24	4	1.17E-02
path:dme04624	Toll and Imd signaling pathway	67	7	1.26E-02
path:dme04150	mTOR signaling pathway -	96	8	2.79E-02
path:dme00565	Ether lipid metabolism	24	3	6.06E-02
path:dme04350	TGF-beta signaling pathway -	41	4	6.82E-02
path:dme04711	Circadian rhythm	13	2	8.50E-02
path:dme04320	Dorso-ventral axis formation	28	3	8.80E-02
path:dme04068	FoxO signaling pathway	65	5	9.89E-02
path:dme00650	Butanoate metabolism	16	2	1.22E-01
path:dme00590	Arachidonic acid metabolism	17	2	1.35E-01
path:dme00515	Mannose type O-glycan biosynthesis	4	1	1.44E-01
path:dme04213	Longevity regulating pathway - multiple species	56	4	1.62E-01

C. KEGG analysis of putative Yki target genes only identified using targeted DamID

KEGG ID	Pathway	N	DE	P.DE
path:dme00900	Terpenoid backbone biosynthesis	27	5	0.00772021
path:dme01230	Biosynthesis of amino acids	66	7	0.03469909
path:dme00981	Insect hormone biosynthesis	27	4	0.03625676
path:dme00983	Drug metabolism - other enzymes	97	9	0.03820788
path:dme00982	Drug metabolism - cytochrome P450	68	7	0.03995049
path:dme00980	Metabolism of xenobiotics by cytochrome P450	69	7	0.04276136
path:dme00230	Purine metabolism	101	9	0.0476068

D. KEGG analysis of putative Yki target genes only identified using ChIPseq

KEGG ID	Pathway	N	DE	P.DE
path:dme00562	Inositol phosphate metabolism	47	23	0.00173507
path:dme03040	Spliceosome	128	51	0.0021139
path:dme04933	AGE-RAGE signaling pathway in diabetic complications	31	16	0.00443704

E. KEGG analysis of putative Sd target genes only identified using targeted DamID

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	27	1.09E-05
path:dme04391	Hippo signaling pathway	61	16	2.32E-03
path:dme02010	ABC transporters	22	8	3.43E-03
path:dme04624	Toll and Imd signaling pathway	67	16	6.34E-03
path:dme01040	Biosynthesis of unsaturated fatty acids	25	8	8.28E-03
path:dme00230	Purine metabolism	101	21	1.05E-02
path:dme01100	Metabolic pathways	1138	161	1.54E-02
path:dme00564	Glycerophospholipid metabolism	63	14	1.95E-02
path:dme00520	Amino sugar and nucleotide sugar metabolism	47	11	2.54E-02
path:dme01212	Fatty acid metabolism	53	12	2.56E-02
path:dme00250	Alanine, aspartate and glutamate metabolism	30	8	2.57E-02
path:dme00440	Phosphonate and phosphinate metabolism	6	3	2.84E-02
path:dme04214	Apoptosis	63	13	4.18E-02
path:dme00471	D-Glutamine and D-glutamate metabolism	3	2	4.22E-02
path:dme01200	Carbon metabolism	123	22	4.55E-02
path:dme04512	ECM-receptor interaction	12	4	5.11E-02
path:dme01230	Biosynthesis of amino acids	66	13	5.80E-02
path:dme04310	Wnt signaling pathway	106	19	5.93E-02
path:dme04350	TGF-beta signaling pathway -	41	9	5.93E-02
path:dme00532	Glycosaminoglycan biosynthesis	8	3	6.58E-02

Table 8.2.4: Comparisons of targeted DamID and Slattery et al (2013) ChIP-ChIP data

A. KEGG analysis of putative Yki target genes shared between DamID and Slattery ChIP ChIP

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	30	2.36E-11
path:dme04391	Hippo signaling pathway	61	16	2.66E-05
path:dme04310	Wnt signaling pathway	106	20	4.11E-04
path:dme04330	Notch signaling pathway	24	7	2.79E-03
path:dme04144	Endocytosis	122	18	1.26E-02
path:dme00760	Nicotinate and nicotinamide metabolism	20	5	2.23E-02
path:dme04320	Dorso-ventral axis formation	28	6	2.64E-02
path:dme04214	Apoptosis	63	10	3.61E-02
path:dme04341	Hedgehog signaling pathway	38	7	3.68E-02
path:dme00515	Mannose type O-glycan biosynthesis	4	2	3.80E-02
path:dme00440	Phosphonate and phosphinate metabolism	6	2	8.49E-02
path:dme04711	Circadian rhythm	13	3	9.04E-02
path:dme00534	Glycosaminoglycan biosynthesis	13	3	9.04E-02
path:dme04150	mTOR signaling pathway	96	12	1.07E-01
path:dme00514	Other types of O-glycan biosynthesis	14	3	1.08E-01
path:dme00072	Synthesis and degradation of ketone bodies	7	2	1.12E-01
path:dme04140	Autophagy	100	12	1.33E-01
path:dme00565	Ether lipid metabolism	24	4	1.39E-01
path:dme00532	Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate -	8	2	1.42E-01

B. KEGG analysis of putative Sd target genes shared between DamID and Slattery ChIP ChIP

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	33	1.85E-13

path:dme04391	Hippo signaling pathway	61	21	1.11E-08
path:dme04310	Wnt signaling pathway	106	23	2.12E-05
path:dme04341	Hedgehog signaling pathway	38	11	2.35E-04
path:dme04350	TGF-beta signaling pathway	41	10	1.95E-03
path:dme04214	Apoptosis	63	13	2.25E-03
path:dme04144	Endocytosis	122	20	3.32E-03
path:dme04711	Circadian rhythm	13	5	3.33E-03
path:dme00440	Phosphonate and phosphinate metabolism	6	3	1.04E-02
path:dme01230	Biosynthesis of amino acids	66	11	2.38E-02
path:dme04624	Toll and Imd signaling pathway	67	11	2.64E-02
path:dme01200	Carbon metabolism	123	17	3.30E-02
path:dme04330	Notch signaling pathway	24	5	5.02E-02
path:dme00030	Pentose phosphate pathway	24	5	5.02E-02
path:dme00450	Selenocompound metabolism	11	3	6.23E-02
path:dme04320	Dorso-ventral axis formation	28	5	8.78E-02
path:dme04130	SNARE interactions in vesicular transport	21	4	1.01E-01
path:dme04150	mTOR signaling pathway	96	12	1.20E-01

C. KEGG analysis of putative Yki target genes in DamID only

KEGG ID	Pathway	N	DE	P.DE
path:dme04512	ECM-receptor interaction	12	7	1.46E-06
path:dme04391	Hippo signaling pathway	61	14	8.17E-06
path:dme01230	Biosynthesis of amino acids	66	11	1.48E-03
path:dme00981	Insect hormone biosynthesis	27	6	4.24E-03
path:dme02010	ABC transporters	22	5	8.12E-03
path:dme04310	Wnt signaling pathway	106	13	9.18E-03
path:dme04350	TGF-beta signaling pathway	41	7	9.44E-03
path:dme04213	Longevity regulating pathway	56	8	1.64E-02
path:dme00600	Sphingolipid metabolism	28	5	2.27E-02
path:dme00380	Tryptophan metabolism	21	4	3.27E-02
path:dme04013	MAPK signaling pathway	93	10	4.74E-02
path:dme00520	Amino sugar and nucleotide sugar metabolism	47	6	5.76E-02

path:dme00230	Purine metabolism	101	10	7.50E-02
path:dme00532	Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	8	2	7.78E-02
path:dme04214	Apoptosis	63	7	7.78E-02
path:dme00564	Glycerophospholipid metabolism	63	7	7.78E-02
path:dme04320	Dorso-ventral axis formation	28	4	8.13E-02
path:dme00260	Glycine, serine and threonine metabolism	28	4	8.13E-02
path:dme04068	FoxO signaling pathway	65	7	8.88E-02

D. KEGG analysis of putative Yki target genes in Slattery ChIP only

KEGG ID	Pathway	N	DE	P.DE
path:dme03013	RNA transport	148	74	0.00033005
path:dme00562	Inositol phosphate metabolism	47	27	0.00235714
path:dme03030	DNA replication	35	21	0.00349273
path:dme00510	N-Glycan biosynthesis	38	22	0.00518212
path:dme04141	Protein processing in endoplasmic reticulum	133	62	0.00804319
path:dme04144	Endocytosis	122	57	0.01023118
path:dme03060	Protein export	23	14	0.0139629
path:dme03050	Proteasome	52	27	0.01452314
path:dme03040	Spliceosome	128	57	0.03089049
path:dme03440	Homologous recombination	23	13	0.03759825
path:dme03420	Nucleotide excision repair	39	20	0.03854802
path:dme00513	Various types of N-glycan biosynthesis	30	16	0.04135334
path:dme04120	Ubiquitin mediated proteolysis	100	45	0.04278473
path:dme03015	mRNA surveillance pathway	73	34	0.04377964

E. KEGG analysis of putative Sd target genes in DamID only

KEGG ID	Pathway	N	DE	P.DE
path:dme02010	ABC transporters	22	8	0.00012091
path:dme04512	ECM-receptor interaction	12	5	0.00121452
path:dme00561	Glycerolipid metabolism	42	9	0.00342
path:dme04624	Toll and Imd signaling pathway)	67	12	0.00378768
path:dme04391	Hippo signaling pathway	61	11	0.00521198
path:dme00981	Insect hormone biosynthesis	27	6	0.01353252
path:dme00600	Sphingolipid metabolism	28	6	0.01616217
path:dme00471	D-Glutamine and D-glutamate metabolism	3	2	0.01626873
path:dme01212	Fatty acid metabolism	53	9	0.01637067
path:dme00250	Alanine, aspartate and glutamate metabolism	30	6	0.02243845
path:dme01100	Metabolic pathways	1138	100	0.03215409
path:dme01040	Biosynthesis of unsaturated fatty acids	25	5	0.03623261
path:dme00230	Purine metabolism -	101	13	0.03900538
path:dme00564	Glycerophospholipid metabolism	63	9	0.04530033

F. KEGG analysis of putative Sd target genes in Slattery ChIP only

KEGG ID	Pathway	N	DE	P.DE
path:dme03013	RNA transport	148	65	5.01E-07
path:dme04141	Protein processing in endoplasmic reticulum	133	54	7.34E-05
path:dme03040	Spliceosome	128	49	8.14E-04
path:dme03050	Proteasome	52	24	9.50E-04
path:dme03015	mRNA surveillance pathway	73	28	1.01E-02
path:dme04140	Autophagy	100	36	1.20E-02
path:dme04120	Ubiquitin mediated proteolysis	100	36	1.20E-02
path:dme03420	Nucleotide excision repair	39	16	2.41E-02
path:dme00513	Various types of N-glycan biosynthesis	30	13	2.51E-02
path:dme03022	Basal transcription factors	40	16	3.09E-02
path:dme04933	AGE-RAGE signaling pathway in diabetic complications	31	13	3.34E-02

path:dme03450	Non-homologous end-joining	6	4	4.04E-02
path:dme00510	N-Glycan biosynthesis	38	15	4.06E-02
path:dme03030	DNA replication	35	14	4.21E-02
path:dme03018	RNA degradation	58	21	4.56E-02
path:dme03460	Fanconi anemia pathway	26	11	4.56E-02
path:dme03060	Protein export	23	10	4.61E-02
path:dme04068	FoxO signaling pathway	65	23	4.85E-02
path:dme04144	Endocytosis	122	39	6.26E-02
path:dme01210	2-Oxocarboxylic acid metabolism	19	8	8.56E+00

Table 8.2.5: Comparisons of targeted DamID in the eye (this study) and in the wing (Zhang et al., 2017)

A. KEGG analysis of putative Yki target genes shared between DamID in eye and wing

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	28	2.64E-10
path:dme04391	Hippo signaling pathway	61	22	5.10E-10
path:dme04310	Wnt signaling pathway	106	29	1.56E-09
path:dme04330	Notch signaling pathway	24	9	5.78E-05
path:dme04214	Apoptosis	63	13	1.26E-03
path:dme04150	mTOR signaling pathway	96	17	1.46E-03
path:dme04512	ECM-receptor interaction	12	5	1.65E-03
path:dme04350	TGF-beta signaling pathway	41	9	4.54E-03
path:dme04341	Hedgehog signaling pathway	38	8	9.55E-03
path:dme00532	Glycosaminoglycan biosynthesis	8	3	2.17E-02
path:dme04320	Dorso-ventral axis formation	28	6	2.20E-02
path:dme04144	Endocytosis	122	16	3.50E-02
path:dme00515	Mannose type O-glycan biosynthesis	4	2	3.51E-02
path:dme00600	Sphingolipid metabolism	28	5	7.07E-02
path:dme04213	Longevity regulating pathway	56	8	7.90E-02
path:dme00520	Amino sugar and nucleotide sugar metabolism	47	7	8.13E-02
path:dme04711	Circadian rhythm	13	3	8.19E-02
path:dme00514	Other types of O-glycan biosynthesis	14	3	9.82E-02
path:dme00565	Ether lipid metabolism	24	4	1.24E-01

B. KEGG analysis of putative Yki target genes in the eye only

KEGG ID	Pathway	N	DE	P.DE
path:dme00981	Insect hormone biosynthesis	27	6	0.00559188
path:dme00230	Purine metabolism	101	13	0.00970691
path:dme04013	MAPK signaling pathway - fly - Drosophila melanogaster (fruit fly)	93	12	0.01253681
path:dme01230	Biosynthesis of amino acids	66	9	0.02088651
path:dme04391	Hippo signaling pathway	61	8	0.03532207
path:dme02010	ABC transporters	22	4	0.04554343

C. KEGG analysis of putative Yki target genes in the wing only

KEGG ID	Pathway	N	DE	P.DE
path:dme04215	Apoptosis	23	6	0.00153795
path:dme04933	AGE-RAGE signaling pathway in diabetic complications	31	6	0.00761502
path:dme04391	Hippo signaling pathway	61	9	0.00763891
path:dme04350	TGF-beta signaling pathway	41	7	0.0081993
path:dme00230	Purine metabolism	101	11	0.03002894
path:dme00750	Vitamin B6 metabolism	6	2	0.04292063
path:dme00620	Pyruvate metabolism	46	6	0.04731887

Table 8.2.6: KEGG analysis of differentially expressed genes in hyperplastic eye growth compared to normal eye growth.

A. KEGG analysis of upregulated genes in wts mutant eye discs (hyperplastic growth) compared to FRT82B eye discs (normal growth)

KEGG ID	Pathway	N	DE	P.DE
path:dme00983	Drug metabolism - other enzymes	97	12	1.35E-06
path:dme00982	Drug metabolism - cytochrome P450	68	10	2.26E-06
path:dme00980	Metabolism of xenobiotics by cytochrome P450	69	10	2.60E-06
path:dme03450	Non-homologous end-joining	6	3	0.00022347
path:dme00480	Glutathione metabolism	78	7	0.00184795
path:dme00053	Ascorbate and aldarate metabolism	33	4	0.00640632
path:dme00830	Retinol metabolism	34	4	0.00713571
path:dme04214	Apoptosis	63	5	0.01408864
path:dme04068	FoxO signaling pathway	65	5	0.01597914
path:dme01230	Biosynthesis of amino acids	66	5	0.01698516
path:dme01100	Metabolic pathways	1138	35	0.0222688
path:dme00260	Glycine, serine and threonine metabolism	28	3	0.02544794
path:dme00230	Purine metabolism	101	6	0.0274549
path:dme00250	Alanine, aspartate and glutamate metabolism	30	3	0.03052485
path:dme00010	Glycolysis / Gluconeogenesis	55	4	0.03655732
path:dme00071	Fatty acid degradation	33	3	0.03906487
path:dme00220	Arginine biosynthesis	15	2	0.04530411
path:dme00290	Valine, leucine and isoleucine biosynthesis	2	1	0.04555344

B. KEGG analysis of downregulated genes in wts mutant eye discs (hyperplastic growth) compared to FRT82B eye discs (normal growth)

KEGG ID	Pathway	N	DE	P.DE
path:dme04013	MAPK signaling pathway	93	12	0.00041699
path:dme04320	Dorso-ventral axis formation	28	6	0.00086496
path:dme04512	ECM-receptor interaction	12	3	0.01208541
path:dme00564	Glycerophospholipid metabolism	63	7	0.01535414

path:dme01230	Biosynthesis of amino acids	66	7	0.01947773
path:dme00330	Arginine and proline metabolism	53	6	0.02230969
path:dme00260	Glycine, serine and threonine metabolism	28	4	0.027901
path:dme00561	Glycerolipid metabolism	42	5	0.02969051
path:dme01100	Metabolic pathways	1138	58	0.03895653

Table 8.2.7: GO analysis of differentially expressed genes in hyperplastic eye growth compared to normal eye growth.

A. Top 50 GO terms enriched in upregulated genes in wts mutant eye discs (hyperplastic growth) compared to FRT82B eye discs (normal growth)

GO ID	Term	Ont	N	DE	P.DE
GO:0004364	glutathione transferase activity	MF	39	7	1.69E-05
GO:0071480	cellular response to gamma radiation	BP	11	4	6.14E-05
GO:0044421	extracellular region part	CC	919	38	7.67E-05
GO:0006749	glutathione metabolic process	BP	51	7	0.00010193
GO:0006575	cellular modified amino acid metabolic process	BP	111	10	0.00013791
GO:0006790	sulfur compound metabolic process	BP	162	12	0.0001991
GO:0031012	extracellular matrix	CC	212	14	0.0002014
GO:0005576	extracellular region	CC	1215	45	0.00021231
GO:0042302	structural constituent of cuticle	MF	142	11	0.00024795
GO:0016765	transferase activity, transferring alkyl or aryl (other than methyl) groups	MF	60	7	0.00028836
GO:0010332	response to gamma radiation	BP	17	4	0.00039978
GO:0051186	cofactor metabolic process	BP	257	15	0.00045599
GO:0043564	Ku70:Ku80 complex	CC	2	2	0.00046097
GO:0003693	P-element binding	MF	2	2	0.00046097
GO:0008010	structural constituent of chitin-based larval cuticle	MF	106	9	0.00046137
GO:0071479	cellular response to ionizing radiation	BP	18	4	0.00050536
GO:0006303	double-strand break repair via nonhomologous end joining	BP	8	3	0.0005089

GO:0015562	efflux transmembrane transporter activity	MF	8	3	0.0005089
GO:0009066	aspartate family amino acid metabolic process	BP	32	5	0.00055595
GO:0005214	structural constituent of chitin-based cuticle	MF	135	10	0.00067093
GO:0071481	cellular response to X-ray	BP	9	3	0.00075122
GO:0070192	chromosome organization involved in meiotic cell cycle	BP	55	6	0.00112464
GO:0042594	response to starvation	BP	145	10	0.00116419
GO:0040003	chitin-based cuticle development	BP	198	12	0.00120699
GO:0061330	Malpighian tubule stellate cell differentiation	BP	3	2	0.00136322
GO:0036375	Kibra-Ex-Mer complex	CC	3	2	0.00136322
GO:0007380	specification of segmental identity, head	BP	11	3	0.00142918
GO:0010165	response to X-ray	BP	11	3	0.00142918
GO:0000726	non-recombinational repair	BP	11	3	0.00142918
GO:0007178	transmembrane receptor protein serine/threonine kinase signaling	BP	81	7	0.00177032
GO:0035326	enhancer binding	MF	42	5	0.00197131
GO:0016840	carbon-nitrogen lyase activity	MF	14	3	0.00300559
GO:0031667	response to nutrient levels	BP	167	10	0.00330628
GO:0042335	cuticle development	BP	254	13	0.00343973
GO:0009991	response to extracellular stimulus	BP	169	10	0.00359978
GO:0048730	epidermis morphogenesis	BP	5	2	0.00441588
GO:2001270	regulation of cysteine-type endopeptidase activity involved in execution phase of apoptosis	BP	5	2	0.00441588
GO:0000784	nuclear chromosome, telomeric region	CC	16	3	0.00447924
GO:0009897	external side of plasma membrane	CC	16	3	0.00447924
GO:0006302	double-strand break repair	BP	98	7	0.00517605
GO:0002778	antibacterial peptide production	BP	33	4	0.00524694
GO:0002780	antibacterial peptide biosynthetic process	BP	33	4	0.00524694
GO:0002786	regulation of antibacterial peptide production	BP	33	4	0.00524694

GO:0002808	regulation of antibacterial peptide biosynthetic process	BP	33	4	0.00524694
GO:0010043	response to zinc ion	BP	17	3	0.00535341
GO:0098595	perivitelline space	CC	17	3	0.00535341
GO:0007379	segment specification	BP	35	4	0.00649246
GO:0010212	response to ionizing radiation	BP	35	4	0.00649246
GO:0033353	S-adenosylmethionine cycle	BP	6	2	0.00652999
GO:0046498	S-adenosylhomocysteine metabolic process	BP	6	2	0.00652999

B. Top 50 GO terms enriched in downregulated genes in wts mutant eye discs (hyperplastic growth) compared to FRT82B eye discs (normal growth)

GO ID	Term	Ont	N	DE	P.DE
GO:0005886	plasma membrane	CC	1251	177	3.99E-33
GO:0071944	cell periphery	CC	1396	186	1.97E-31
GO:0048731	system development	BP	1874	221	6.09E-30
GO:0007399	nervous system development	BP	1143	158	3.66E-28
GO:0009653	anatomical structure morphogenesis	BP	1616	192	5.69E-26
GO:0065007	biological regulation	BP	4238	369	5.19E-25
GO:0007275	multicellular organism development	BP	2577	258	3.61E-24
GO:0048513	animal organ development	BP	1312	163	7.63E-24
GO:0022008	neurogenesis	BP	909	128	2.31E-23
GO:0048856	anatomical structure development	BP	2988	283	4.44E-23
GO:0048699	generation of neurons	BP	869	123	1.24E-22
GO:0032502	developmental process	BP	3087	288	1.59E-22
GO:0007154	cell communication	BP	1675	188	2.16E-22
GO:0023052	signaling	BP	1630	183	9.33E-22
GO:0050789	regulation of biological process	BP	3781	329	3.11E-21
GO:0030182	neuron differentiation	BP	769	111	4.28E-21
GO:0032501	multicellular organismal process	BP	3949	339	4.68E-21
GO:0050794	regulation of cellular process	BP	3476	307	2.05E-20
GO:0007423	sensory organ development	BP	517	85	7.69E-20
GO:0009887	animal organ morphogenesis	BP	802	111	1.20E-19
GO:0001654	eye development	BP	427	74	1.08E-18

GO:0048880	sensory system development	BP	427	74	1.08E-18
GO:0150063	visual system development	BP	427	74	1.08E-18
GO:0044459	plasma membrane part	CC	816	108	1.15E-17
GO:0050896	response to stimulus	BP	2982	266	1.22E-17
GO:0048592	eye morphogenesis	BP	342	63	2.14E-17
GO:0090596	sensory organ morphogenesis	BP	342	63	2.14E-17
GO:0001754	eye photoreceptor cell differentiation	BP	186	45	2.24E-17
GO:0032990	cell part morphogenesis	BP	469	75	6.70E-17
GO:0030154	cell differentiation	BP	1793	182	6.97E-17
GO:0048858	cell projection morphogenesis	BP	460	74	7.84E-17
GO:0001751	compound eye photoreceptor cell differentiation	BP	177	43	9.91E-17
GO:0048666	neuron development	BP	635	90	1.16E-16
GO:0007409	axonogenesis	BP	345	62	1.35E-16
GO:0016020	membrane	CC	2718	245	1.45E-16
GO:0007267	cell-cell signaling	BP	479	75	2.21E-16
GO:0031226	intrinsic component of plasma membrane	CC	479	75	2.21E-16
GO:0048869	cellular developmental process	BP	1842	184	2.31E-16
GO:0000904	cell morphogenesis involved in differentiation	BP	501	77	2.49E-16
GO:0000902	cell morphogenesis	BP	579	84	3.32E-16
GO:0005887	integral component of plasma membrane	CC	465	73	4.97E-16
GO:0046530	photoreceptor cell differentiation	BP	209	46	4.97E-16
GO:0048749	compound eye development	BP	395	66	5.98E-16
GO:0120039	plasma membrane bounded cell projection morphogenesis	BP	458	72	7.43E-16
GO:0061564	axon development	BP	358	62	8.42E-16
GO:0031224	intrinsic component of membrane	CC	1457	154	1.04E-15
GO:0001745	compound eye morphogenesis	BP	321	58	1.05E-15
GO:0016021	integral component of membrane	CC	1436	152	1.48E-15
GO:0032989	cellular component morphogenesis	BP	710	94	1.91E-15
GO:0048812	neuron projection morphogenesis	BP	457	71	2.24E-15

Table 8.2.8: KEGG analysis of putative Jra target genes in normal eye growth

KEGG ID	Pathway	N	DE	P.DE
path:dme04144	Endocytosis	122	54	2.22E-09
path:dme04013	MAPK signaling pathway	93	44	5.99E-09
path:dme04391	Hippo signaling pathway	61	33	7.57E-09
path:dme04310	Wnt signaling pathway	106	39	8.49E-05
path:dme04341	Hedgehog signaling pathway	38	16	0.00225682
path:dme04214	Apoptosis	63	23	0.0027169
path:dme00240	Pyrimidine metabolism	40	16	0.0041995
path:dme04150	mTOR signaling pathway	96	31	0.00503065
path:dme04120	Ubiquitin mediated proteolysis	100	32	0.00514233
path:dme00564	Glycerophospholipid metabolism	63	22	0.00619223
path:dme04350	TGF-beta signaling pathway	41	15	0.0140522
path:dme00565	Ether lipid metabolism	24	10	0.01621712
path:dme04320	Dorso-ventral axis formation	28	11	0.01929982
path:dme04512	ECM-receptor interaction	12	6	0.02318033
path:dme04140	Autophagy	100	29	0.03037259
path:dme04624	Toll and Imd signaling pathway	67	20	0.04954198