IDENTIFICATION AND CHARACTERIZATION OF NOVEL TUMOR SUPPRESSORS USING *IN VIVO* TUMOR MODEL IN *DROSOPHILA MELANOGASTER*

A Thesis

Submitted in partial fulfillment of requirements

Of the degree of

Doctor of Philosophy

By

Sanket Nagarkar

Roll no. 20153357

Indian Institute of Science Education and Research

(2020)

CERTIFICATE

Certified that the work incorporated in the thesis entitled “Identification and characterization of novel tumor suppressors using *in vivo* tumor model in *Drosophila melanogaster*.” submitted by Sanket Nagarkar was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

(Supervisor)

Date: 20/04/2020

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Signature

Sanket Nagarkar

Roll No. 20153357

Date: 20/04/2020

Acknowledgements

Firstly, I would like to express my gratitude towards my thesis supervisor Prof LS Shashidhara. He has been supportive throughout the course of PhD and I am also grateful for the training I have received from him. The freedom of thought and research are immense contributors to a successful research project and Shashi has consistently provided these. He has been an inspiration during my work and will continue to be so later. He has provided the vision and critical guidance during thesis writing and manuscript preparations.

I would also like to extent my gratitude to Prof Stephen Cohen for hosting me at Copenhagen as well for the invaluable inputs and directions he has given for my research. Interactions with him have always been an enriching experience. He has been instrumental in shaping this thesis and manuscripts resulted from this work. Dr. Hector Herranz for being the contact point in Copenhagen and promptly arranging for reagents when I required them. Interactions with Hector have been a motivation through and through. I would also like to thank Dr. Teresa Eichenlaube for setting up the screen procedure here at IISER, Pune and training me in ways of the screening, Additionally, she has been my local guide in Copenhagen, without her the stay at Copenhagen would not have been possible. I am also grateful to Dr. Nguyen Hung, who mentored me in tissue culture and molecular work at Copenhagen, without his guidance it would not have been possible to conduct mammalian cell based experiments. I am also thankful to SC lab members at Copenhagen for stimulating interactions during my stay.

My sincere thanks to my Thesis advisory committee members Dr. Nagaraj Subramaniam, Dr. Kundan Sengupta and Dr. Jomon Joseph for bringing a new perspective during meetings and prompt communication of ideas and otherwise. Their advice and suggestions have been critical in shaping the work presented in this thesis. I am grateful for their support and precious inputs for my work.

I am also grateful to IISER faculty members for making scientific research at IISER a fun and fulfilling activity. A special thanks to Dr Girish Ratnaparkhi and Dr Richa Rikhy for taking charge of fly lab and making the *Drosophila* research smooth sailing journey. Dr Girish Deshpande has been a great support. I would like to thank him for the discussions about my work, which he made

a point to have always and also for the constant encouragement he has provided in every discussion.

I thank my fellow labmates- Soumen, Madhumita, Dilsha for making themselves available for discussions of science as well as non-scientific thought exchange. I wish them all the best for teir PhDs. Nelchi, Pooja, Aditi have been great team members during the screen. We have been each other's support sytem during the screen and have covered each other's back always. I am also grateful to friends at IISER, for the constant support, treats and celebrations which made these years memorable.

A special thanks to non-academic staff at IISER with special mention of IISER biology staff- Mahesh, Piyush, Kalpesh, Dr Mrinalini who have made administrative work a breeze for us. A sincere thanks to fly facility staff- Snehal, Ashwini, Bhargavi and Yashwant without whom the screen and fly work would not have happened.

Last but not the least, I would like to thank my family and friends for standing by me during the course of PhD.

Contents

Abstract

Cancer is a disease that hitch hikes growth regulatory network and causes uncontrollable cell proliferation and metastasis which is usually fatal. Genetic alterations in cancer cells enable such hitch hiking of regulatory networks. These alterations allow cancer cells to acquire abilities to grow and metastasize, which together with other enabling characteristics of cancer form the hallmarks of cancer. Cancer cells gradually acquire multiple hallmark characteristics defining cancer progression as a multi-step process. Disease progression and accumulating genetic alterations suggest cooperative mechanisms underlying progression of disease which is also supported by experimental evidence. Thus, it is of fundamental importance to identify causal genetic alterations in different cancer types. This has motivated a large number of genome wide omics studies, which have revealed a plethora of genomic, transcriptomic changes in cancer tissues. Resultant large data has presented a bigger challenge of identifying relevant causal factors.

To understand the cooperative mechanisms underlying tumorigenesis we carried out a genome-wide screen in *Drosophila* tumor model for epithelial cancers. We used epithelial growth factor (EGFR) and Yorkie (Yki) as oncogenic drivers and depleted one gene at a time using RNAi mediated knock down in each of these contexts. We have identified several novel putative tumor suppressors, depletion of which enhance effects of EGFR and Yki resulting in massive overgrowth of wing imaginal discs.

Interestingly, we identified components of the Negative Elongation Factor Complex (NELF) as tumor suppressors specifically in context of Yki. These findings were particularly intriguing because, Yki is a co-activator of transcription, while the NELF complex is required for promoter proximal pausing (PPP). PPP, as the name suggests, occurs in region proximal to transcription start site and has been shown to be a critical regulatory mechanism in transcription of genes in response to stimuli before elongation phase and after promoter escape by RNA Pol II. Characterization of the overgrown wing disc tissue formed by combination of depletion of NELF complex components and Yki overexpression showed neoplastic transformation. Additionally, we observed that CDK9, the kinase component of Positive transcription elongation factor (P-TEFb) is necessary as well as cooperative with Yki in tumorigenesis.

Our work shows that PPP functions to limit tumorigenic potential of Yki activity. I will also present evidence from RNA-seq of tumor discs which indicate possible mechanisms that underlie tumorigenesis observed in combination of NELF depleted and Yki overexpressing wing discs. These findings shed light upon mechanisms that might regulate oncogenic outcome of Yes associated protein 1 (YAP1), mammalian ortholog of well conserved Hippo pathway.

Introduction

Tissue growth is a fundamental biological process that involves increase in cell number and cell size. These two facets of tissue growth are influenced by signaling modules that interface between extrinsic and intrinsic factors governing growth. Intuitively, there are robust mechanisms of regulation of such a critical process via fine tuning the signaling pathways. Signaling pathways essentially bring about coordination between cell proliferation and cell death such that the resultant growth or tissue pruning meets the requirements (Hipfner and Cohen, 2004). Remarkably, a handful of signaling modules are involved in regulating growth and homeostasis throughout the lifespan of an organism and additionally are highly conserved across evolution. Broadly, there are two types of regulators of growth, ones which promote cell proliferation, oncogenes and others that suppress growth either by limiting oncogene function or by inducing cell death, tumor suppressors. Our knowledge of how a handful of signaling pathways are regulated and crosstalk with each other is derived from animal models for organ and tissue development (Barolo and Posakony, 2002; Hariharan, 2015). In retrospect, it wasn't surprising that process of tumorigenesis was observed to hitch-hike on these growth regulatory mechanisms leading to uncontrolled proliferation and acquiring survival advantage (Hanahan and Weinberg, 2011).

Findings from initial investigations about the nature of cancer which were only strengthened later, showed that genetic alterations are at the root of multiple types of cancers. Thus, was defined the genetic nature of cancer. Original observations regarding this were reported more than a century ago by David von Hansemann on carcinoma samples (Hansemann, 1890). These observations were validated by the nature of different characteristics exhibited by cancer cells and also explained how cancer cells pass on such characteristics in a tumor. Additional evidence of role carcinogens in tumor causation via mutating genes has made genetic nature of neoplasia a cornerstone of modern cancer research.

Cancer research over past century has used this central genetic tenet of oncogenesis and has developed it further by addition of specific causative factors such as detailed mechanisms of mutagenesis by different carcinogens, identification of genes that impart special abilities to cancer cells etc. Most seminal of the principals that has become a milestone in cancer research is the idea of progressive accumulation of mutations is a necessity for development of cancers. This principle

is based on a popular hypothesis termed as the 'Two-hit hypothesis' proposed by Knudson in a study published in 1971 (Knudson, 1971). The two-hit hypothesis has marked the beginning from where cancer research took off and advanced into an era where understanding of individual genes that are causally linked with cancer as well as a large catalogue of mutations and other genomic alterations that are present in the variety of cancers is now available to researchers globally. Despite such tremendous advance, cancer presents a diversity of challenges in terms of understanding causal mechanisms as well as mechanisms that bring about progression of disease. Such detailed understanding is critical for accurate prognosis and treatment of the disease.

Hallmarks of cancer

One of the reasons cancer treatment is challenging is that the diversity and variations it presents at each stage of progression. Cancer initiation and progression is characterized by acquisition of variety of capabilities by the cancer cells. These have been referred to as 'Hallmarks of cancers' (Hanahan and Weinberg, 2011). Hallmarks of cancer represent the enabling characteristics that allow cancer to initiate, sustain and progress.

Although, it is not necessary that all cancer types exhibit all of these hallmarks, most show acquisition of multiple the hallmarks. Considering the functional distinction between the hallmark capabilities, intuitively, more than one change in genomic and epigenomic landscape of cancer cells is causative of appearance of hallmark capabilities. Interestingly, progression of cancer shows remarkable parallels with process of evolution (McGranahan and Swanton, 2017), which implicates process of selection, active or passive, by which cells that possess favorable characteristics are selected for. the cells that show hallmark capabilities have certain degree of advantage over other cells with genomic alterations that did not result in an advantageous characteristic. Such cells are thus selected. This school of thought finds support from cell competition studies which showed that cancerous cells are effectively eliminated in a milieu of normal cells, unless they are conferred with specific advantage (Tabassum and Polyak, 2015). How cancer cells develop hallmark capabilities? Mutagenesis has been believed to be the most common cause of developing cancerous properties. Based on outcome, mutations are classified in two categories- Driver mutations which confer a specific advantage to cancer cells and Passenger mutation which likely do not have a function in process of tumorigenesis. Thus, several studies

have been done to identify how driver mutations contribute to progression of tumorigenesis. Following are a few examples-

Progression of colorectal cancer: Jones and colleagues used colorectal cancer as a model to draw a correlation between time (patient age) and mutational events that led to metastatic disease. Based on mutation rates in colorectal cancer cells, sequencing of tissues derived from mouse xenografts, they were able to predict key mutational events in a tumor that led to a metastatic form of the disease (Jones *et al.*, 2008).

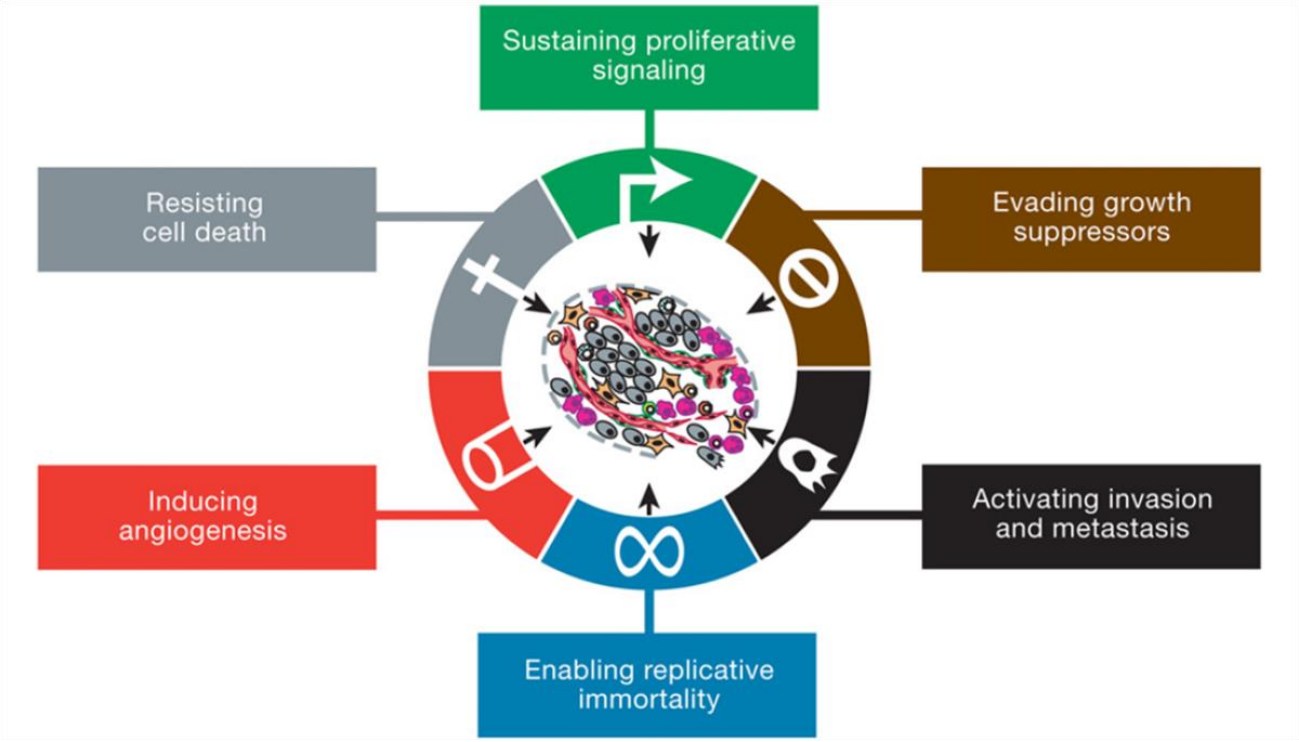


Fig. 1: Hallamark capabilities of cancer (Hanahan and Weinberg, 2011)

Progression of liver cancer: Progression of liver cancer is understood to a lesser extent. Commonly, many liver diseases or stress condition lead to liver cirrhosis. Later, molecular events that cause mutations in oncogenes lead to progression of stressed liver to dysplastic nodules. These, later progress to hepatocellular carcinoma, due to sequential acquisition of mutations (Marquardt, Andersen and Thorgeirsson, 2015). Transcriptome analysis of low and high grade hepatocellualar

carcinomas (HCCs) has revealed progressive acquisition of activated oncogenic signaling modules such as *TGF-β, WNT, NOTCH*. Early stage dysplastic nodules, unlike high grade HCC samples, had relatively uniform transcriptome (Marquardt *et al.*, 2014). Similarly, exome sequencing of HCCs, in stage wise manner has identified driver mutations in Teleomerase reverse transcriptase (*TERT)* gene promoter to be an early event (Schulze *et al.*, 2015).

Pancreatic cancer: In case of pancreatic cancer Yachida and colleagues have identified progressive occurrence of somatic mutations in pancreatic cancers using 8 patient samples. Tissue samples from progressive stages and metastases of patients were used to identify mutation signatures in progressive stages of the pancreatic cancers. This study provides a time line representing gap between mutational events that lead to metastasis. Additionally, it also revealed that, the cells metastasizing to distinct organs (eg, lung, liver) originate from spatially separated populations of cancer cells in the original tumor based on the mutation signatures (Yachida *et al.*, 2010).

Therefore, it can be seen across all cancers that different mutations mark distinct stages of cancer and such progression of disease inherently demands crosstalk between at least two or more alteration in cancer cells, direct or indirect. Thus, it imperative to understand this cross-talk to understand and tackle cancer.

Cooperative mechanisms of tumorigenesis

"The whole is more than the sum of its parts" – Aristotle

This quotation from the Greek philosopher Aristotle, aptly shows the gap between the number of mutations present in cancer cells versus the causal relationship with the disease. It is reflected in the fact that the number of cancer genes (genes causally linked with cancer) we know about is little more than 2% of total protein coding genes and it is speculated that a minimum of 5 mutations of cancer genes is required for transformation of cells into cancerous cells (Stratton, 2011).

First step in identifying cause-effect relationship is to catalogue all the changes alterations. Over past few years several omics studies have revealed millions of mutations that are present in different kinds of cancers and are available as resource at different places. One such resource available in public domain is headed by The Cancer Genome Atlas (TCGA). TCGA has recently took their omics studies to next level with publication of Pan-Cancer data. These datasets have tremendously advanced characterization of individual caner types.

Despite the advantage, delineating causal mechanisms with an added complexity of cooperativity between two or more changes remains a daunting task. Thus, augmenting bioinformatic repertoire of data with studies from cancer cell lines and animal models is essential. Investigations of different signaling modules have already revealed cross-talk between them.

Epidermal Growth factor receptor and NF-κB:

Epidermal growth factor receptor (EGFR) signaling module is one of the tyrosine receptor kinase module and is frequently hyperactivated in carcinomas ((Normanno *et al.*, 2006). Canonically, EGFR signaling functions via RAS-MAPK axis. EGFR targeting drugs are frequently used in treatment of epithelial tumors (Hynes and Lane, 2006). Interestingly, tumors develop resistance to EGFR targeting drugs and resistance develops because of EGFR independent reasons. One of such reasons is NF-κB pathway activation (Shostak and Chariot, 2015). NF-κB is a family of transcription factors with a transactivator domain which are activated downstream of interleukin signaling. These are known to be elevated in different cancer cells and coincide with activated EGFR in EGFR positive breast cancer patients (Biswas and Iglehart, 2006). Such coincidence is explained by the interdependency of these two pathways. EGFR activation leads to activation of NF-κB via PKCε and IKK complex. It is interesting that EGFR dependent tumorigenesis requires activation of NF-κB targets in glioblastoma cells, lung and breast cancer cells in culture and xenografts (Jiang *et al.*, 2011; Yang *et al.*, 2012).

Hippo pathway crosstalk with other pathways:

Hippo pathway was originally discovered in *Drosophila*. Later, several components of this pathway which function as tumor suppressors were identified and were found to be evolutionarily conserved in mammals (Camargo *et al.*, 2007). Components of Hpo pathway ultimately impinge on activity of its effector YAP/TAZ. YAP/TAZ seem to take central stage in regulation of cell proliferation and cell survival as several other growth regulatory pathways such as Wnt, EGFR, BMP cross-talk with Hpo pathway influencing YAP/TAZ activity (Hansen, Moroishi and Guan, 2015).

APC, a tumor suppressor, which is frequently mutated in gastro-intestinal tumors is a component of Wnt pathway. APC regulates Hpo pathway by serving as a scaffold for kinase module containing Lats1 and Sav1 (Cai *et al.*, 2015). Interestingly, β-catenin, the effector of Wnt signaling

is known to bind YAP (Rosenbluh *et al.*, 2012). Depletion of YAP has been shown to decrease β-catenin target gene expression in gastric cancer cells while increased YAP expression in mice showed remarkable increase in β-catenin and its target gene expression. Additionally YAP is necessary for β-catenin dependent cancers in cell lines and animal models (Rosenbluh *et al.*, 2012).

Cancer cells have an ability to survive adverse treatments (radiotherapy) and evade apoptosis. YAP/TAZ play a role in giving survival advantage to cancer cells by increasing expression of IGF2 in medulloblastoma (Fernandez-L *et al.*, 2012). Increased expression of IGF2 in turn activates Akt. This was an interesting finding as IGF2 is indispensable for Shh driven subclass of medulloblastomas and therefore links YAP with Shh signaling module through IGF2. Although direct links are not yet clearly established.

Similarly, Activator protein 1 (AP-1, a dimer of Fos and Jun) is required for YAP/TAZ/TEAD mediated tumorigenesis in cancer cells, xenografts and mouse tumor model (Zanconato *et al.*, 2015). This adds another known tumorigenic module to cooperative mode of tumorigenesis involving Hpo pathway.

Cross-talk with Ras in tumorigenesis

Ras is a small GTPase known to be activated downstream of growth promoting signaling pathways such as EGFR, PI3K pathway etc. Activating mutations of Ras cause in transformation of cells resulting in cancerous growth. Occurrence of such mutations and increased activity of Ras across cancer types is abundant and mechanisms of Ras mediated tumorigenesis are also well studied. Yet, it requires presence of cooperating factors that enhance tumorigenesis.

One of such examples is cell polarity protein scribbled. It was identified as a tumor suppressor in a screen using *Drosophila* tumor model to function in cooperation with active mutant of Ras, Ras^{V12} (Brumby and Richardson, 2003). This interaction is interesting because, Ras^{V12} on one hand requires loss of Scribbled for neoplastic transformation, on the other hand loss of scribbled alone leads to elimination of cells (Wu, Pastor-Pareja and Xu, 2010). Similar cooperative function has been observed in human cells, where, loss of scribbled in combination activated Ras led to invasion in 3D cell culture models (Dow *et al.*, 2008). Reinforcing cooperative mechanisms of tumorigenesis involving Ras. Interestingly, loss of other tumor suppressors linked to cell polarity such as Lgl, Dlg also cooperate with Ras in transformation of cells.

Another example of cooperative tumorigenesis involving Ras is with Akt. This mechanism is unique as it involves cooperation between Ras and Akt at the level of translation and not only transcription (Parsa and Holland, 2004). The oncogenic combination of Ras and Akt signaling increased transcription of many genes involved in growth but the increase in mRNAs that are being actively translated was much more than the extent of change in transcription (Rajasekhar *et al.*, 2003).

Integrated analysis of co-occurring alterations in signaling pathways in cancers

Described previously are specific examples of cooperative mechanisms in tumorigenesis. Cancers are diverse and alterations are plenty thus there is a need to analyze alterations signaling modules in an integrated manner to give us an idea of which alterations are concurrent in cancers and which are mutually exclusive. The TCGA PanCancer initiative has made headways in this direction with analysis of 10 signaling pathways in >9000 tumors (Sanchez-Vega *et al.*, 2018). This analysis covers 10 signaling modules excluding alterations in DNA damage repair pathways, epigenetic modifications and splicing due to confounding genomic instability that accompanies these modifications. Bioinformatic analysis yielded significant co-occurrence of multiple different pathway modules with p53 alterations. Among others are, co-occurring alterations in PI3K-Nrf2 pathway, Hippo-p53 pathway. Similar analysis of master regulators of metabolism in tumors has been conducted in a parallel study of PanCancer initiative (Peng *et al.*, 2018). Such analysis paves the way for identifying novel crosstalk between different specific modules in specific cancers types and develop therapeutic strategies, diagnostic and prognostic mmeasures.

There many more examples of independent signaling modalities getting deregulated and inturn cooperating with each other aiding and advancing cancer. Additionally, we also need to appreciate the cellular heterogeneity in a tumor tissue and thus resulting autonomous and non-autonomous cross talk within the tumor tissue. Authors Tabassum and Polyak aptly summarize the complexity as, “Tumorigenesis takes a village” in their review (Tabassum and Polyak, 2015).

Diversity and complexity of disease and obvious limitation of studying patients of cancer demanded development of model systems that can be used to simulate oncogenic mechanisms as well as to discover novel mechanisms of tumorigenesis through cooperativity of genetic alterations. Most commonly used are Mouse based tumor models and *Drosophila* based tumor models.

Modelling tumorigenesis in animal systems

Mice provide a mammalian model to study process of tumorigenesis *in vivo.* Advancement of genetic tools has eased creation of transgenic mice and thus, studies in mice are no longer limited only to xenograft models. Several genetic approaches of generation of transgenic mice have been used such as, conditional mice to express or knock out genes in tissue specific manner, virus based generation of transgenic mice and so on (Cheon and Orsulic, 2011).

Similarly, a transposon based approach has shown success in identifying novel cancer genes remarkable of which is Sleeping beauty (SB) transposon (Copeland and Jenkins, 2010). The SB based tumor models have identified genes cooperating with *Arf* in sarcomas (Collier *et al.*, 2005), colorectal carcinoma (Dupuy *et al.*, 2009) and several other types of carcinomas (Starr *et al.*, 2009). Additionally, stepwise process of tumorigenesis that implies gradual accumulation of genetic alterations has also been modelled in mice. Wherein sequential application of mutagens has been used to mimic stepwise tumorigenesis process (Wu and Pandolfi, 2001). Such models have been used to reinforce cell line based evidence of cooperative tumorigenesis (Zanconato *et al.*, 2015). Also, SB based models have led to identification of cancer genes and in generating mice models of therapeutics where known alterations in cancer genes are used as sensitized background to carry out screens (Keng *et al.*, 2009; Takeda *et al.*, 2015).

On the other hand, *Drosophila melanogaster* has also been utilized to develop tumor models (Miles, Dyson and Walker, 2011). This invertebrate model system provides distinctly sophisticated genetic tools which also allow for large scale testing *in vivo*. Historically, *Drosophila* has been a model system that in which majority of signaling pathways were first elucidated and characterized. Till date, high level of conservation of molecular mechanisms from flies to humans make it an attractive system of study. Short lifespan and plethora advanced genetic tools allow for large scale and rapid screening in *Drosophila*, which is particularly advantageous for cancer gene discovery (Gonzalez, 2013).

The thesis work presented here envisioned to identify novel tumor suppressors, loss of which in context of known oncogenic drivers could result in neoplasia. We did a screen to identify novel tumor suppressors using *Drosophila* tumor model. We conducted screens with EGFR and Yki expression as sensitized backgrounds and KK RNAi collection from VDRC for combining gene

depletion with the oncogenic drivers. These were based on concept of cooperative tumorigenesis in a genetic combination of an oncogene and depletion of a tumor suppressor.

During the screens, we identified novel tumor suppressor role of two transcription regulatory complexes in Yki background. Transcription, is one of the fundamental processes in cells and is at the critical juncture between information content in DNA and the functional proteins. RNA Polymerase II is the enzyme that brings about transcription of genes encoding proteins and this process of transcription is highly regulated and sensitive to a variety of stimuli. Transcription factors, which are master regulators of transcription have great influence on outcome of developmental processes in an organism as well as are determinants of disease phenotypes (Levine, 2011). The evidence showing importance of regulated transcription in cancer initiation and progression is one of the many highlighting link between transcription regulation and diseases (Kandoth *et al.*, 2013; Bradner, Hnisz and Young, 2017). Thus, intrigued by discovery of two complexes that regulate transcription step of promoter proximal pausing, especially in the background of an oncogenic transcriptional co-activator, Yki, we present characterization of these candidate complexes from tumor perspective. Further, we provide evidence of a potential novel mechanism that links this transcription regulatory event with oncogenic potential of Yki along with bioinformatic analysis of cancer patient sample data.

Chapter 1: *In vivo* screen to identify novel tumor suppressors using *Drosophila* tumor model

In the field of cancer research the diversity and complexity the disease presents are gradually forming a quantitative picture owing omics studies over past decade. We now know about the mutation load, genome wide methylation status, transcriptomic differences etc. Bioinformatic toolset at our disposal despite being remarkable relies largely on frequencies of alteration and context specificity is also challenging to consider (Copeland and Jenkins, 2010). Additionally, the aspect of clonal heterogeneity, tumor and non-tumor cell interaction and role of systemic factors such as immune system are being appreciated greatly and have become inseparable part of diagnostic and therapeutic process in cancer treatment (Tabassum and Polyak, 2015).

On one hand omics studies have revealed signatures across all types of cancers and in cases of specific cancer types and also have laid down map of progression of disease. While on the other hand *in vivo* systemic complexity has made it imperative to study findings of bioinformatics analysis in a complex set up of tissues and organs. In such cases, development and study of animal tumor models have again come to fore and are being newly appreciated for advancing understanding of disease and therapeutics.

Modelling tumorigenesis in *Drosophila*

Drosophila as a model system is credited with discovery and exploration of most of the major signaling pathways involved in growth and development. *Drosophila* is best known for studying process of development and growth control. As is mentioned earlier here, most of the mechanisms that regulate growth and development are also found to be responsible for disease when deregulated. Modelling of cancer in *Drosophila* has come up as promising tool in cancer biology particularly because of tools of genetic manipulation that have been developed for more than century in *Drosophila*, shorter lifespan of flies, simple husbandry and most importantly high level of conservation of molecular mechanisms and genes from *Drosophila* to humans.

Of course, like other model systems *Drosophila* has limitations when it comes to studying cancer. Some of these are presence of an open circulatory system unlike closed system in mammals, absence of tissues such as bone, cartilage and four chromosomes in *Drosophila* limit study of

aneuploidy. Despite such limitations, the wealth sophisticated genetic and other tools in *Drosophila* allow for creating and studying tumorigenesis in great detail.

Past century has marked identification of 50 tumor suppressor genes in *Drosophila*, most of which have mammalian orthologs, many had already been characterized to molecular level (Watson, Justice and Bryant, 1994). Notably, last few years have taken a step further to model cooperative tumorigenesis with emergence of a variety of tumor models that utilize distinct developing larval tissue to study relevant type of tumor (Fig. 2). The growing larval tissues serve as responsive field to genetic manipulation that affects activity of well-known growth regulators. Genome wide screens which typically use a combination of knockdown and expression of genes are used to create tumors, rightly referred to as 'Tumors a la carte' by Gonzalez.

Notable was a screen at the beginning of 21st century identified *scribbled* as a tumor suppressor along with other cell polarity regulators which were known to exhibit tumor suppressor activity. (Pagliarini and Xu, 2003). Loss of which caused metastatic tumors in larvae that also expressed constitutively active mutant form of Ras, Ras^{V12} (Brumby and Richardson, 2003; Pagliarini and Xu, 2003). The findings were remarkable as human scribbled protein also showed tumor suppressor function (Dow *et al.*, 2008). Similar to such findings is glioblastoma model in larval glia which mimics demonstrated limiting role of *Rbf1* (one of the RB genes in fly) in glioma model overexpressing EGFR-PI3K (Read *et al.*, 2009). Modelling of Rhabdomyosarcoma in fly muscle tissue, based on expression of fused human PAX7-FOXO1 transcription factors, identified *rolling pebbles (rols)* as a suppressor of neoplasia. Consistent with these findings was the tumor suppressor function of *TANC1*, the mammalian ortholog of *rols* (Avirneni-Vadlamudi *et al.*, 2012). These are but a few examples of potential of *Drosophila* as a tumor model.

Imaginal disc tissue also serves as a model for most commonly occurring type of tumors called carcinomas. These are of epithelial in origin and are commonly modelled in wing imaginal tissue as shown in schematic. Of particular interest is that activation of known oncogenic drivers such as Notch, EGFR, Myc YAP in imaginal disc tissue causes hyperplastic growth in the tissue. Such a characteristic phenotype sets a platform to identify genes which could cooperate with these drivers causing neoplasia in the imaginal disc tissue (Herranz, Eichenlaub and Cohen, 2016). Pioneering studies that utilized such feature are mentioned earlier come from Richardson's group and Paliagrini and Xu (Pagliarini and Xu, 2003; Yachida *et al.*, 2010).

Wing imaginal disc is an epithelial tissue and we have used genetic manipulation in wing disc tissue for modelling carcinoma. This *in vivo* tumor model is primarily based on neoplastic tumor

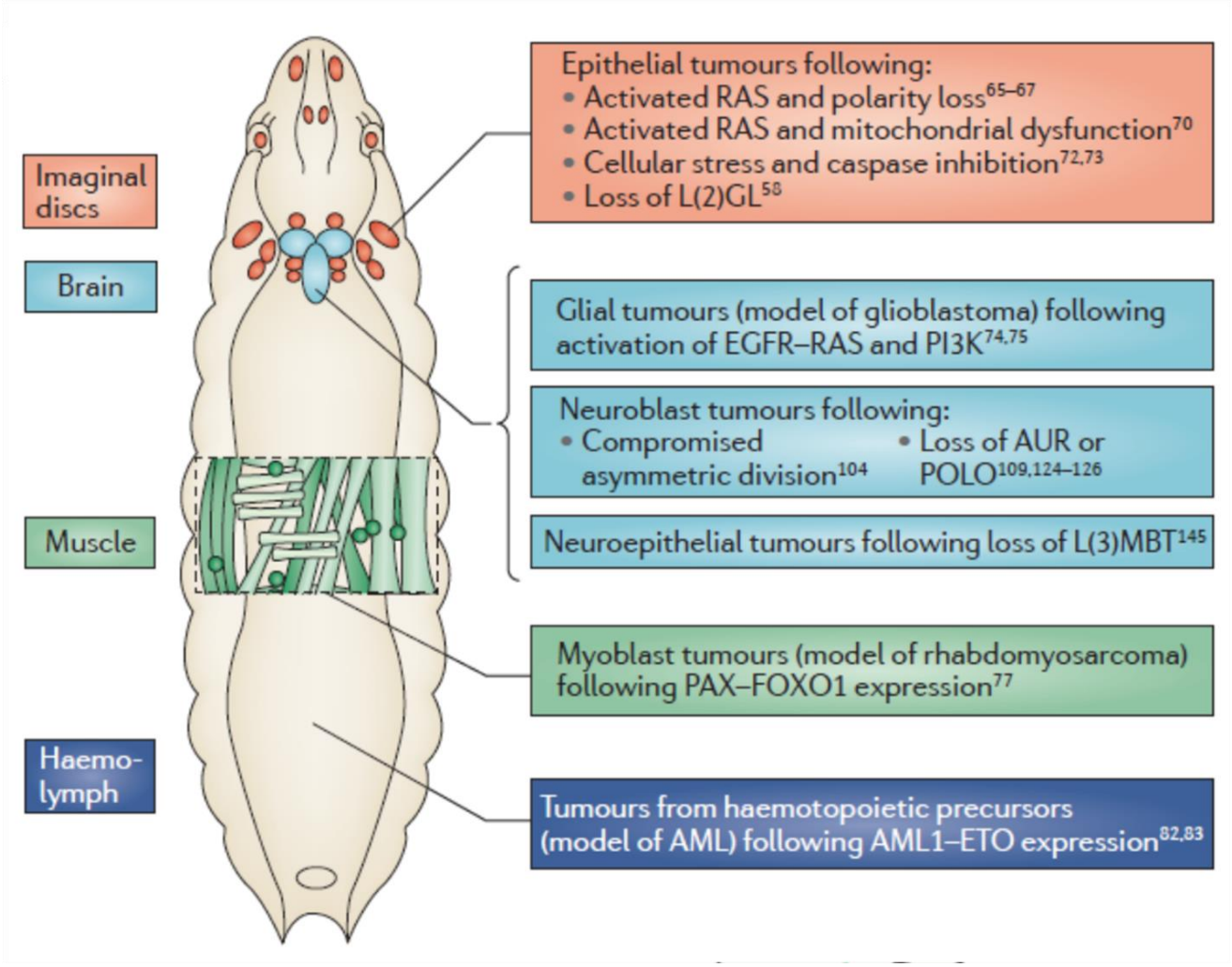


Figure 2: Schematic representing different parts of *Drosophila* larva used to model different types of tumors (Gonzalez, 2013). It can be observed here that cognate type of tissue is used to model the type of cancer. For eg. Epithelial tissue of imaginal discs represents a model of epithelial tumors.

growth observed upon combining EGFR as oncogenic driver and depletion of *Socs36E* using RNAi (Herranz *et al.*, 2012) and oncogenic cooperation of Yki and BAP complex (Song, Herranz and Cohen, 2017). Our work, envisaged to combine the oncogenic drivers EGFR and Yki with

depletion of genes (one gene at a time) using KK RNAi library from VDRC. EGFR and Yki are well known oncogenic drivers and are placed critically in their respective signaling modules. Following is a brief summary of EGFR and Yki in context of tumorigenesis.

EGFR pathway in tumorigenesis

Epidermal growth factor receptor (EGFR) is a member of receptor tyrosine kinases. As the name suggests it is a responsive to epidermal growth factor family of ligands. EGFR activation leads to activation of RAS-MEK axis as well as STAT pathway (Gschwind, Fischer and Ullrich, 2004) ultimately promoting cell growth, survival. Due to its receptor activity and potential to influence a very broad cascade of growth promoting factors in a cell, EGFR provides a pivotal point in translating extracellular cues into growth promoting signals. Primary function of EGFR is in development and patterning of tissues in mammals as well as *Drosophila* (Citri and Yarden, 2006; Harden, 2017). From 1980s EGFR has been documented to be highly expressed in multiple types of tumors, consistent with these, several mutations that lead to increased or constitutive activation of EGFR have been deemed as drivers of tumorigenesis, particularly of carcinomas (Hynes and Lane, 2006). Consistent with these findings, agents that suppress EGFR activity such as antibodies against EGFR and drugs that inhibit EGFR activity have been used as treatment measure.

In *Drosophila* EGFR has been used to study process of tumorigenesis mimicking scenario in cancer patients by overexpression of EGFR. Such expression of EGFR in wing disc tissue results in hyperplastic growth in wing imaginal disc tissue (Herranz *et al.*, 2012). Thus, enabling studies of cooperation using EGFR overexpression as a sensitized genetic background. Using this model, *Socs36E* was discovered as tumor suppressor (Herranz *et al.*, 2012) as well as tumor suppressive function of *Psq* was studied (Herranz, Weng and Cohen, 2014). EGFR overexpression as background also gave insights into how Lactate Dehydrogenase (LDH) expression synergizes with EGFR leading to neoplastic transformation of wing disc tissue (Eichenlaub *et al.*, 2018).

Hippo pathway and its effector Yki in tumorigenesis

Hippo pathway components were identified as tumor suppressors in *Drosophila melanogaster* as a result of mutant screens (Pan, 2010). These include the core kinase cassette of NDR family kinase Warts (Wts) (Justice *et al.*, 1995; Xu *et al.*, 1995), Ste-like kinase Hippo (Hpo) (Jia *et al.*, 2003; Udan *et al.*, 2003) with scaffolding proteins Salvador (Sav) (Tapon *et al.*, 2002) and Mob as Tumor

suppressor (Mats) (Lai *et al.*, 2005). Interestingly all the components of the core kinase pathways were identified to regulate the delicate balance between cell proliferation and apoptosis which in turn is necessary to determine tissue and organ size. This kinase cassette was later shown to regulate activity of a transcriptional coactivator protein Yorkie (Yki) which when activated mimics effects of loss of function of Hpo, Wts, Sav and Mats (Huang *et al.*, 2005a). Yki is known to promote expression of a repertoire of genes that promote cell proliferation such as *CycE* and promote cell survival by inhibition of apoptosis such as *Diap1*. Since these pioneering studies the expanse of Hpo pathway has been greatly studied both in regulation of growth and organ size regulation.

Hpo pathway is highly conserved from flies to humans and functionality of its components is consistent across the systems (Harvey, Zhang and Thomas, 2013). Unique ability of Hpo pathway effector Yki to promote cell proliferation and survival without causing senescence which is a common feature of most of the oncogenes made it an intuitive candidate driver cancer gene. Evidence supported this notion when, Yes associated protein (YAP), a mammalian orthologue of Yki, was discovered to be amplified in multiple cancer types as well as in mouse tumor models (Overholtzer *et al.*, 2006; Zender *et al.*, 2006; Lin, Park and Guan, 2018). Additionally, deregulation of Hpo pathway components has been studied in several mouse models and cancer cell lines for driving growth and implicated thus in tumorigenesis (Harvey and Tapon, 2007).

Overexpression of Yki in imaginal disc tissue of developing larva causes massive overgrowth (Huang *et al.*, 2005a). Additionally, overactivating Yki by depletion of core Hpo pathway components or expressing mutant form of Yki that is constitutively active also leads to massive overgrowth, showing immense growth promoting ability of Yki. Similar to EGFR overexpression of Yki also provides a sensitized background for identification of context dependent tumor suppressor candidates.

Screen methodology

Overexpression of a growth promoter such as EGFR and Yki causes overgrowth in wing disc tissue however, combining it with an additional genetic manipulation can lead to a neoplastic transformation (Herranz *et al.*, 2012; Herranz, Weng and Cohen, 2014; Song, Herranz and Cohen, 2017; Eichenlaub *et al.*, 2018). These observations held a potential to use the tumor model system to screen for novel genes that are context dependent growth regulators and are potentially tumor

suppressors. With that vision we designed the screen which sought to combine the overexpression of EGFR or Yki with a panel of RNAi lines that will cover a large portion of *Drosophila* genome.

The choice of KK library from VDRC was based on rationale that all the insertions of KK RNAi library are using phiC31-based transgene which has single specific insertion site in the *Drosophila* genome. Such method provides an inherent consistency to genome wide screen. The KK RNAi library contains about 9822 transgenic lines covering about 69% of *Drosophila* genome. More information about KK RNAi library can be found here (https://stockcenter.vdrc.at/control/library_rnai;jsessionid=7DF44DACA69BBCDCE93426533325B86C.jvm1).

We used *UAS-Gal4* system to drive expression of transgenes. *apterous-Gal4* (apGal4) which is expressed in dorsal compartment of wing imaginal disc was used along with $Gal80^{ts}$. $Gal80^{ts}$ is a temperature sensitive version of Gal80 which is active at restrictive temperature of 18^{o}C while gets degraded at 29^{o}C resulting in expression of transgenes. UAS-GFP was used to visualize the tissue expressing EGFR/Yki and transgenic brought in by the RNAi line. Schema of the screen is depicted in Fig.3.

It has been reported that upon induction of overgrowth resulting in tumor formation in imaginal disc tissue, larval phase of *Drosophila* life cycle gets prolonged ending in prepupal lethality at much delayed time than in larvae with non-perturbed imaginal disc growth (Gateff, 1978). In such cases, larvae keep growing resulting development of 'Giant larvae', a phenotype characteristic of tumor growth in imaginal disc tissue (Gateff, Löffler and Wismar, 1993). We utilized this characteristic feature of giant larval phenotype (depicted in Fig. 4) as a primary scoring visual. Followed by observation of larvae under fluorescence microscope to visually gauge the growth of GFP positive wing imaginal disc tissue. Our points of comparison have been the oncogene driver alone control and positive control of EGFR/Yki in combination with depletion of pTEN using RNAi. pTEN is a very well documented tumor suppressor (Lee, Chen and Pandolfi, 2018) that also showed massive growth in wing disc tissue when combined with EGFR as well as Yki expression background (Fig.5).

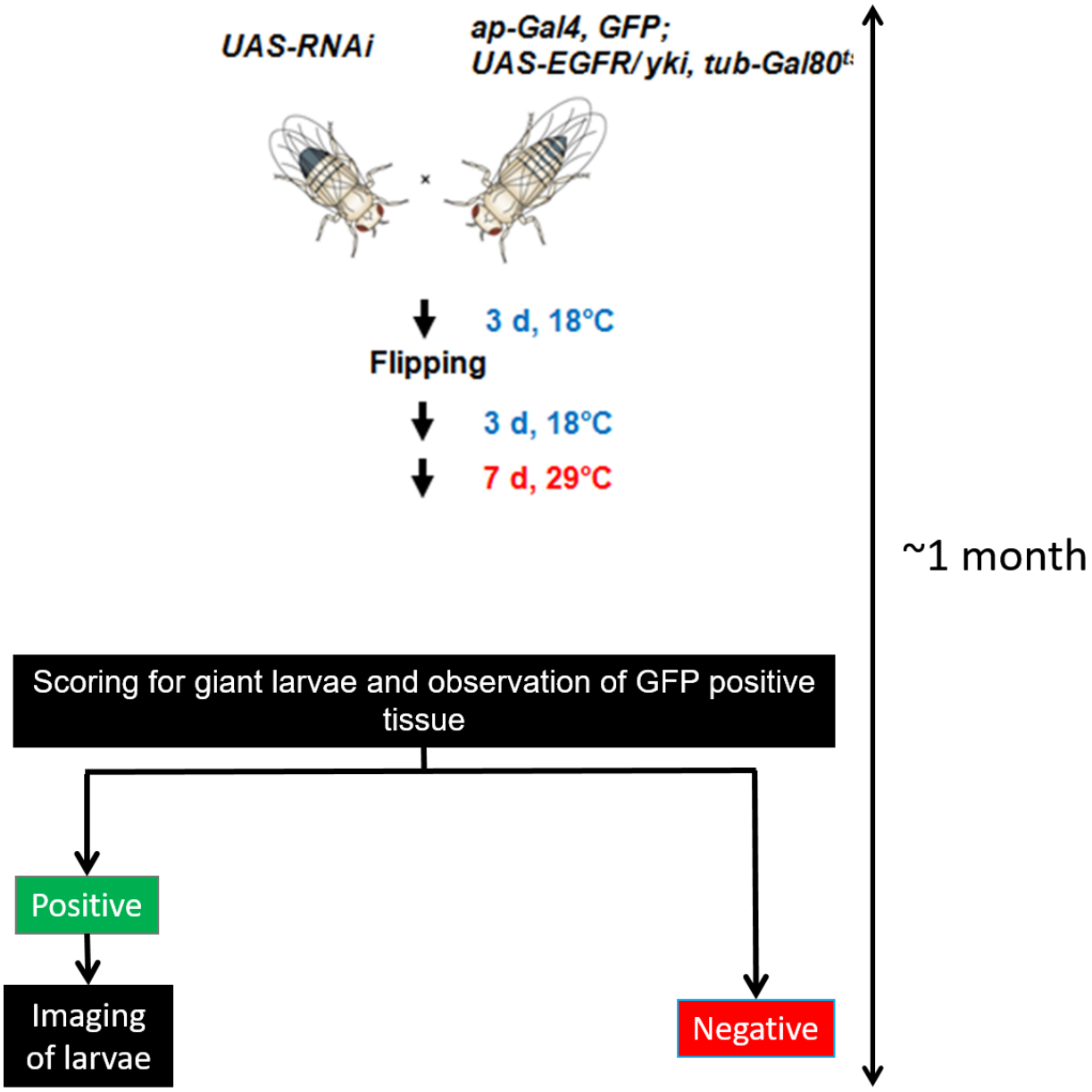


Figure 3.: A schematic representing workflow of the screen. Parental stock containing transgene for EGFR and/or Yki is crossed with an RNAi line and is grown at restrictive temperature at first followed by induction of transgenes at permissive temperature. Time of induction was 7th day after egg laying such that the transgenes start expressing at L1 stage of larval development .

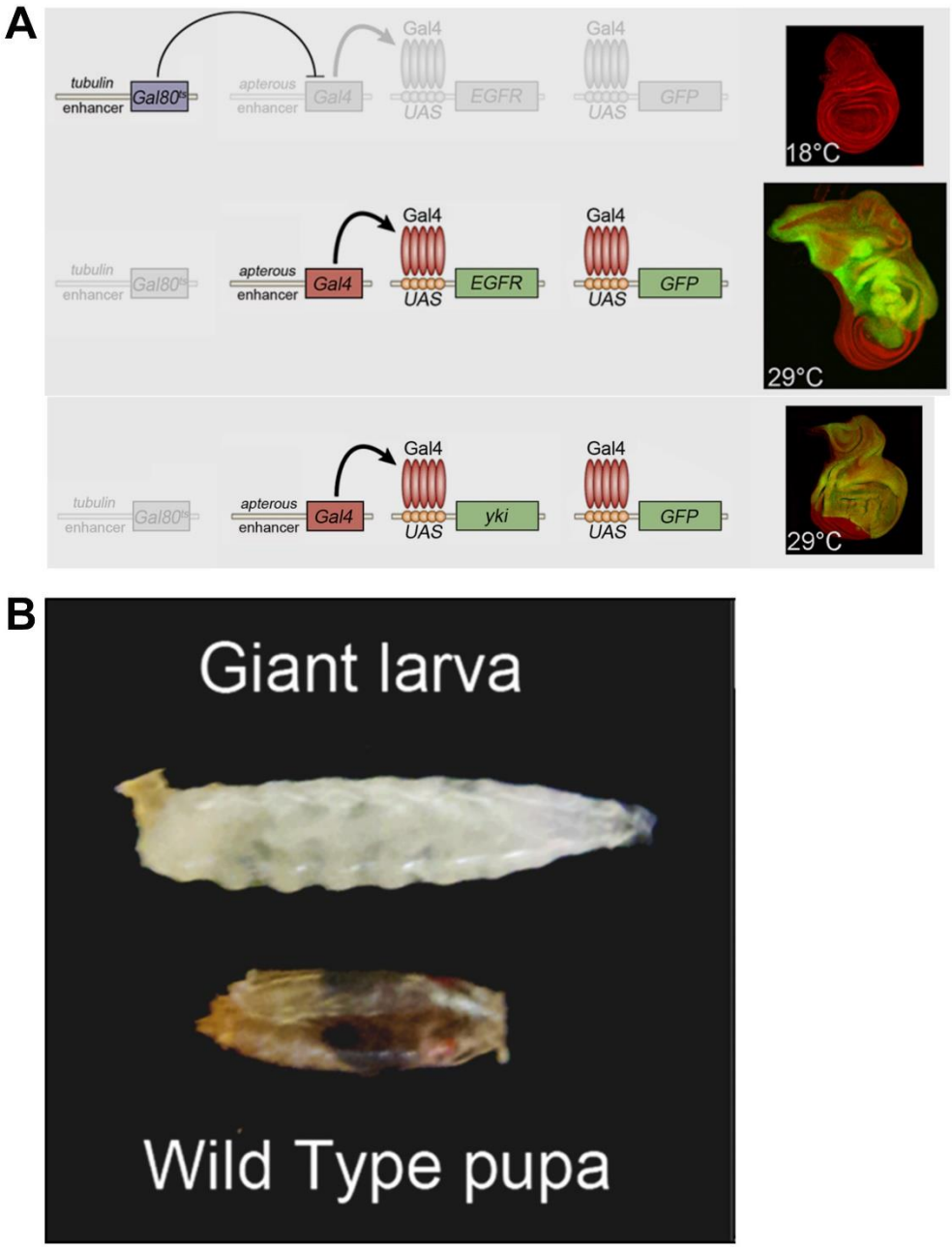


Figure 4.: (A) Schematic representing *Gal4-Gal80^{ts}* system used for the screen. The resultant expression of GFP and hyperplastic overgrowth shown by expression of EGFR and Yki alone is depicted at permissive temperature. Adapted from (Herranz *et al.*, 2012)

(B) The giant larval phenotype along with puparium which formed by wild type larvae in the same time scale in which giant larval phenotype occurs.

While the EGFR and Yki expression backgrounds provided an opportunity to screen for enhancement of growth when combined with depletion of genes, the neoplastic tumors induced by combination of EGFR and Socs36 RNAi (Herranz *et al.*, 2012) could be used for screening of suppression/reversal of neoplastic growth. Thus, we introduced the third screening background with the EGFR+Socs36 RNAi (will be referred to as SOCS screen).

Results and discussion

The screening was carried out together for the three backgrounds and crosses were repeated to conform the consistency of positives observed during the screen. Results here are presented as a comparison with the complete screens, which were performed with a screening team at IISER Pune. Thus, following results are for a sub set of KK RNAi lines and are juxtaposed with the results of the genome wide screen. We have recently published the whole genome wide screen with rich analysis making it a valuable resource (Groth *et al.*, 2020).

EGFR Screen:

EGFR background was used to test 862 KK RNAi lines. From this repertoire, about 19 RNAi lines were found to be positives. We repeated these 19 positives and observed that only 5 of these 19 show the giant larval phenotype and large wing discs consistently. We refer to these as confirmed positives (Table 1).

	YKI	EGFR	SOCS
Screened RNAi lines	940	862	1063
Positives	40	19	4
Confirmed Positives	14	5	4

Table 1: Result summary of screen for each of the three backgrounds with numbers showing KK RNAi lines subjected to screen and the positives observed. Confirmed positives are the KK RNAi lines that show overgrown discs with giant larval phenotype consistently.

These observations showed that EGFR screen yielded about 2.2% positives while we observed that only 0.6% of the total stocks screened were consistently showing a positive phenotype. These percentages of confirmed positives do not differ greatly when compared to the complete screening effort which has 9137 lines screened for EGFR background, with total confirmed positives being 67 (0.73%) and. These results are being published as a resource (Groth *et al.*, 2020).

The positives obtained in the screen consist of well-known tumor suppressors such as *wrts, hpo* which are tumor suppressor kinases of Hpo pathway. And also highlight the crosstalk between two pathways in context of growth and tumorigenesis. The screen was done blinded such that the screeners did not know the genes we were depleting. Thus, identifying known tumor suppressors in the screen re-confirmed the potential of screen to identify tumor suppressors *in vivo.*

Among the novel candidates identified are *elonginB, atrophin*/*gug* of which *atrophin* was also a positive in Yki screen. These are promising candidates for further validation and understanding in context of EGFR pathway and tumorigenesis.

SOCS screen:

SOCS screen was an interesting endeavor since the positives in this screen were the genes that could revert the neoplastic tumors in EGFR and *SOCS36E* RNAi combination. As a proof of principle, we found that RNAi line for EGFR (KK107130) was found to be positive in SOCS screen. Overall, we observed that the number of positives in the SOCS screen were few. We identified 4 positives from the 1063 lines screened, a percentage of 0.4%. In the full screen the percentage of positives is very similar as 0.31% (32 positives in 9332 lines screened).

SOCS screen positives could be of particular interest given the very specific context in which they were shown to revert neoplasia upon depletion and thus hold a great potential for therapeutics.

Yki screen:

940 KK RNAi lines were screened in the Yki expression background. Positives of the Yki screen were numerous compared to EGFR and SOCS screen. We identified 40 positives from Yki screen out of which 14 consistently showed giant larval phenotype and massive overgrowth in visual analysis of positives. The percentages were 4.25% positives and 1.5% of conformed positives. These stats differed yet again when the full screening effort is considered, wherein 9032 KK RNAi

lines were screened and confirmed positives were 950 (10.5%) which was a very large number compared to EGFR screen.

While the screen was being conducted a study regarding KK RNAi library in context of Yki background was published which showed that about 25% of KK RNAi lines have an unintended landing site at 40D on second chromosome and is causing expression of tiptop (tio) leading to enhancement of Yki driven phenotype (Vissers *et al.,* 2016). This study made it imperative that we test sensitivity of our screen to this 40D insertion. Based on UAS-40D line (an empty UAS construct at 40D site) control with our genetic background and Gal4 driver showed that Yki screen we conducted is not affected by 40D insertion present in KK RNAi lines. These results are consistent with results published in Vissers *et al.* where they have tested multiple Gal4 drivers for sensitivity to additional 40D insertion showing *ap-Gal4* is not affected by this. Detailed analysis is presented in Groth *et al.*2020

Yki screen presented a large set target and thus several genes which are novel and have not been reported with growth regulatory function earlier. The overlap of Yki positives with EGFR provides another subset of interesting genes. In the smaller set of screened lines presented here the overlap was minimal (3 positives shared between the two screens). However, the subset of shared positives between EGFR and Yki screens, from all of screen positives, presents an interesting group of positive candidates. Although, many of these are known for their growth regulatory function, there are a few candidates that could be promising to examine in detail.

From the subset reported in this thesis, there were unique candidates of Yki screen such as- *No child left behind (NCLB), Und, Graf* and *NelfA*. *NCLB, Graf* are genes that encode proteins about which very little is known. However, most intriguing of all the positives was identification of *NelfA* as candidate in Yki screen. Primarily, this piqued our interest because, we also identified *NelfB* and *NelfD* as positives. These are components of one single Negative transcription elongation factor (NELF) complex. This discovery prompted us to study function of these putative tumor suppressors in detail. At the same time, screen team identified *bin3, hexim* as positives in Yki screen. These are components of the 7SK snRNP complex. Interestingly, these complexes regulate promoter proximal pausing which is a rate limiting step before productive elongation begins. We will elaborate on this in next chapter.

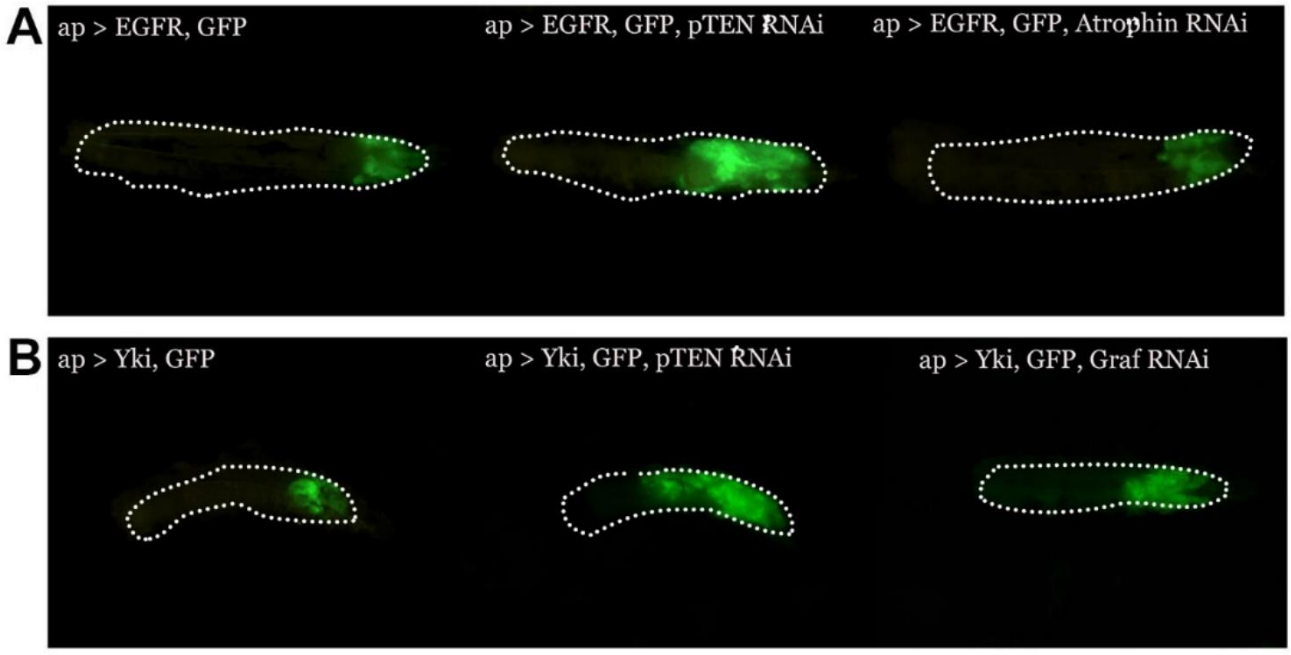


Figure 5.: (A) Panel shows larval phenotype of EGFR screen visualized under GFP fluorescence microscope. (From L to R) EGFR expression alone, EGFR expression combined with pTEN RNAi (the positive control) and a positive identified in screen.

(B) Panel shows larval phenotype of Yki screen visualized under GFP fluorescence microscope. (From L to R) Yki expression alone, Yki expression combined with pTEN RNAi (the positive control) and a positive identified in screen.

Detailed screen data along with larval images of positives can be seen in a free database hosted by IISER, Pune at http://www.iiserpune.ac.in/rnai/home.

Chapter 2: Characterization of 7SK snRNP and NELF complex components as tumor suppressors

Transcription is a process by which messenger RNA is made from DNA. mRNA is translated in a protein which brings about the intended function. Thus, regulation of transcription forms one of the most fundamental and commonly associated with regulation of cellular function. Consistently, dysregulated transcription is invariably associated with various diseases and disorders. Thus, understanding of transcription process has taken a central place in studies of growth, development and diseases. Importance of transcription in development and disease was appealing prospect particularly in context of the NELF and 7SK snRNA components being identified as tumor suppressor candidates in the screen (Yki screen- Chapter 1). With this perspective, it important that we elaborate the regulation of transcription and its function in cancer.

Transcription is an interplay of well-choreographed events. Primarily, the regulation occurs at the step of initiation of transcription by assembly of preinitiation complex (PIC). PIC is made up by RNA Polymerase II, general transcription factors such as TFIID, TFIIH, GAGA factor, several chromatin modifiers etc. which bind to specific sequences of promoter DNA (Levine, Cattoglio and Tjian, 2014). Assembly of PIC is triggered by binding of transcription factor/s (TF). Interestingly, these transcription factors are respondents of signaling pathways that orchestrate growth and development of an organism. Thus, it does not come as surprise that most of the morphogens that direct early development or factors that determine cell fate, cell identity are TFs that are expressed in a cell or tissue specific manner with temporal regulation (Levine, 2011).

Transcription is no exception to the recurrent theme that cancer relies on perturbation of normal regulatory processes. Extensive genome wide studies from cancer patient samples have led to identification of several somatic mutations and a large number of these mutations result in perturbation of transcriptional program (Kandoth *et al.*, 2013; Vogelstein *et al.*, 2013). It has to be noted that transcriptional processes are altered in multiple ways such as alterations of *cis*-regulatory elements on DNA (enhancers, promoter elements), mutations of transcription factors that can constitutively activate or repress expression, alterations in DNA methylation and other chromatin modifiers that can indirectly alter expression of genes (Bywater *et al.*, 2013; Garraway

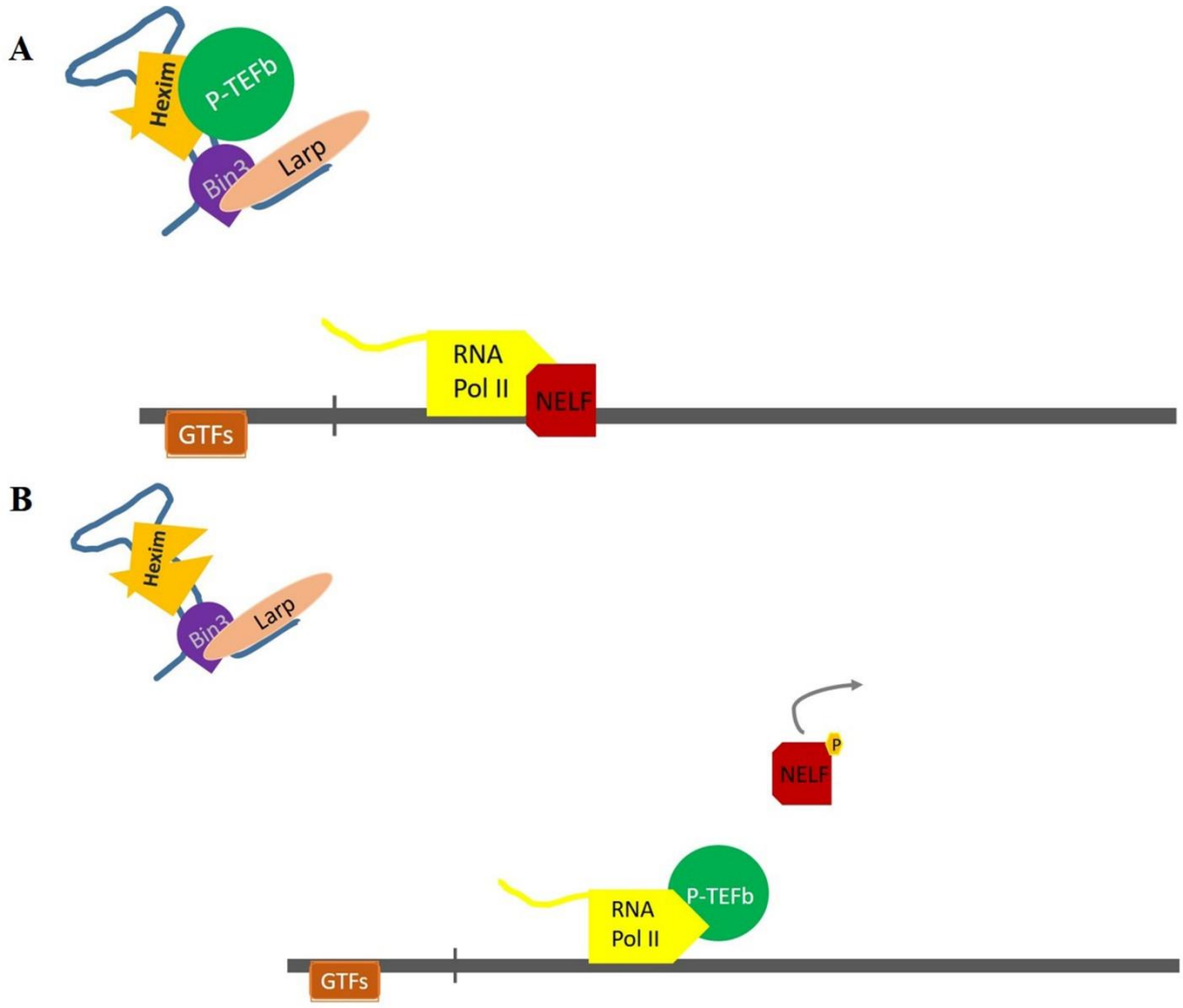
and Lander, 2013; Sur and Taipale, 2016). Alterations leading to changes in gene expression have cascading effects ultimately leading to characteristic cancer traits, such is the importance of transcription making it a premier process to target therapeutically.

In context of the screens discussed previously, we identified a few regulators that impinge directly on process of transcription. We were intrigued particularly by discovery of a set of tumor suppressor candidates- Bin3 RNAi, Hexim RNAi, NELFA RNAi, NELFB RNAi and NELFD RNAi in Yki screen. These candidates are components of *Drosophila* 7SK snRNP complex (Bin3, Hexim) and Negative transcription elongation factor (NELF) (NELFA, B and D) complex. The 7SK snRNP and NELF complex are involved in promoter proximal pausing (PPP) of RNA Pol II.

Promoter Proximal Pausing of RNA Pol II

Historically it was believed that RNA Pol II proceeds into elongation phase after 'clearing the promoter'. However, in higher organisms Pol II was observed to be present at the inducible gene promoters even in absence of a stimulus and was still transcriptionally engaged (Gilmour and Lis, 1986; Rougvie and Lis, 1988). Along with earlier evidence suggesting that post-initiation of transcription there is a rate limiting step before elongation takes place it became increasingly evident that in case of many of the metazoan genes RNA Pol II is paused and accumulates along template DNA after clearing the promoter (Jonkers and Lis, 2015). Pausing has been observed in the range of 30-60bp downstream from transcription start site (TSS) (Kwak and Lis, 2013) and thus is termed as promoter proximal pausing (PPP). Evidence suggest that PPP serves as a regulatory step dependent on stimulus as has been shown with *Hsp70* in *Drosophila* (Rasmussen and Lis, 1993) or determines synchronous vs asynchronous expression of genes during development (Lagha *et al.*, 2013) or has a role as a quality check point before elongation (Core and Adelman, 2019), making PPP a critical step for spatio-temporal regulation of gene expression.

The phenomenon of PPP involves complex interplay of stimuli and molecular players involved in regulation of gene expression. The most critical players of pausing are three complexes that are directly involved in regulation of PPP. These are the 7SK snRNP complex, the NELF complex and the Positive transcription elongation factor complex (P-TEFb). Since, components of the 7SK snRNP complex and the NELF complex were identified as a tumor suppressor in Yki screen we were keenly interested in the transcription associated function of these complexes reason being



transcription co-activator function of Yki. Following is a brief description of 7SK snRNP complex and NELF complex and their function in regulating PPP (Fig. 6).

Figure. 6: Schematic representing the 7SK snRNP and NELF complex invovled in Promoter proximal pausing (20-60bp downstream of TSS). The PPP is rate limiting step before productive elongation begins. The 7SK snRNP sequesters P-TEFb mantaining the paused state while the NELF complex hinder mobility of RNA Pol II by binding to RNA Pol II establishing the pasue. P-TEFb upon release from 7SK snRNP phosphorylates NELF complex evicting it and releasing the pause.

7SK snRNP complex

The ribonucleoprotein complex consists of non-cosing 7SK snRNA, 5'-RNA methyl capping protein MePCE and 3'-RNA stability protein Larp7. Hexim-1 and 2 are the kinase inhibitors that

are bound to P-TEFb and bring it in the 7SK snRNP complex. The discovery of 7SK small nuclear RNA is attributed to screen for RNA factors that are involved in inhibition of P-TEFb induced transcription (Nguyen *et al.*, 2001). 7SK snRNA forms a complex with MePCE and Larp7 and formation of RNA-protein complex stabilizes the 7SK snRNA which in absence of either MePCE and Larp7 gets degraded (Krueger *et al.*, 2008; Xue *et al.*, 2009; Quaresma, Bugai and Barboric, 2016). Hexims are reversibly associated with 7SK snRNP and with P-TEFb resulting in sequestration of P-TEFb (Yik *et al.*, 2003).

7SK snRNP function is sequestering P-TEFb resulting in reduced unavailability of P-TEFb to release the paused Pol II. Such characteristic contribution of 7SK snRNA to regulation of transcription provides an additional layer of restricting transcription until a suitable stimulus is received. It has been reported that many transcription factors such as NFκB, CTIP, TRIM28 provide the stimulus that leads to release of paused Pol II (Cherrier *et al.*, 2013; McNamara *et al.*, 2013). In such cases the 7SK with P-TEFb is recruited to the paused gene resulting release of P-TEFb. It is interesting to note here that, in this scenario large multi-protein complexes such as super elongation complex (SEC) accommodate 7SK snRNP which cycles repeatedly between P-TEFb sequestered vs P-TEFb released form (D'Orso, 2016; McNamara, Bacon and D'Orso, 2016; Quaresma, Bugai and Barboric, 2016). This provides a local sequestration and release mechanism for regulating transcription possibly regulating rate of transcription as well.

Drosophila 7SK snRNP complex has been shown to be functionally conserved with d7SK snNRA, Bin3, Hexim and Larp being the orthologues of 7Sk snRNA, MePCE, Heximes 1and 2 and Larp7 respectively (Nguyen *et al.*, 2012). Conserved function is also evidenced by developmental defects observed in *Drosophila* Hexim mutant. Overall 7Sk snRNP is poised at a juncture where it is critical to the decision of releasing promoter proximally paused Pol II.

Negative elongation factor complex (NELF)

NELF complex is composed of four proteins NELFA (WHSC2), NELFB (COBRA1), NELFC/D (TH1), NELFE (RD). Interestingly all the components of the NELF complex were originally discovered independently and was then identified as components of a complex that is imparts DRB sensitivity to *in vitro* (Yamaguchi *et al.*, 1999, 2002). The *Drosophila* NELF was also observed to pause Pol II, evidenced by the Pol II and NELF complex component supplemented by release of paused Pol II from *hsp70* (Wu *et al.*, 2003). Further molecular characterization of NELF complex

revealed that RNA Pol II exhibits pause based regulation of house-keeping genes in *Drosophila* (Wu *et al.*, 2005).

NELF complex functions with another complex of Spt-4 and 5 constituting DRB sensitivity inducing factor (DSIF). Loss of NELF results in release of Pol II from paused sites these observations led to questions such as extent of PPP and the extent to which PPP affects transcription output. Genome wide analysis revealed that paused genes are present throughout development and in *Drosophila* to Human cells (Guenther *et al.*, 2007; Zeitlinger *et al.*, 2007). Loss of NELF led to profound change in Pol II profile despite there is not a large difference in overall transcriptome profile (Muse *et al.*, 2007). *In vitro* and *in vivo* studies both consistently showed that depletion of single component of NELF complex led to loss of NELF complex localization and release of paused Pol II (Narita *et al.*, 2003).

To understand the process and regulation of PPP it is important to understand the interplay between the 7SK snRNP, P-TEFb and NELF complex together. Although, deregulation of pausing by perturbation in NELF function does not show significant change in global transcription, it must be noted that perturbations of pausing have gene specific or context specific role as shown in case of *hsp70* and other heat shock genes (Wu *et al.*, 2003), context dependent regulation of ER target genes (Aiyar *et al.*, 2004) or regulation of genes involved in differentiation and maintaining of stem cell status (Amleh *et al.*, 2009; Min *et al.*, 2011). Additional recent evidence also pinpoints that regulation of pausing is how gene transcription is regulated rather than via regulating recruitment of Pol II (Bartman *et al.*, 2019).

Discovery of the 7Sk snRNP and NELF complexes as tumor suppressors in specific context of a transcription coactivator-Yki along with emerging importance of PPP in regulating gene expression, addiction of cancer cells to transcription, were compelling factors for detailed characterization of overgrown wing discs we observed.

Results

Depletion 7SK snRNP complex components cooperates with Yki in causing tumorous growth

In continuation with identification of 7SK snRNP complex components as candidates in Yki screen we sought to analyze the phenotype in detail. We first compared the larval phenotypes. We found that depletion of 7SK snRNP components, *bin3* and *Hexim,* did not produce overgrowth on their

own (Fig.7A) in contrast to RNAi-mediated depletion of *bin3* and *Hexim,* in combination with Yki overexpression which led to massive overgrowth in wing disc tissue and giant larval phenotype (Fig.8A). Larval phenotype in case of Yki expression alone showed only moderate overgrowth phenotype, and larvae eventually pupate as was the case with larvae with depletion of 7SK snRNP complex components alone (Fig.8A). Occurrence of larval phenotype was 100% in GFP expressing larvae. We observed no GFP expressing pupae in case tumorigenic genotypes.

To understand the extent of growth induced by the combination of Yki and depletion of 7SK snRNP components we quantified area of wing disc in each of the genotype. As the *ap*-GAL4 drives expression of transgenes only the dorsal wing disc, we wanted to restrict our measurements to the domain in which *ap*-Gal4 drives expression. For this purpose, we measured area of wing discs expressing GFP, which served as a marker of *ap*-Gal4 domain. We found that, the depletion of 7SK snRNP alone did not show difference in wing disc area expressing GFP when compared to wing discs with *ap*-GAL4 driven GFP expression, which is a wild type control (Fig.8C). These observations conformed with larval phenotype described above. On the other hand, combination of the 7SK snRNP complex depletion with Yki expression showed manifold increase in wing disc area in comparison to Yki expression alone as well as GFP expressing WT wing discs (Fig.8D). Compared to wing discs with GFP expression alone, Yki and 7SK snRNP RNAi combinations showed an increase in wing disc area in range of 15-20 fold. Interestingly, Yki expression alone, although led to moderate growth in wing disc tissue compared to GFP expressing WT wing discs (7.5 fold), it was not statistically significant. Though it is important to note here that, the area measurements do not represent the phenotype we observed accurately. Because, the tissue shows folding, such measurements likely underestimate the extent of growth.

PPP is believed to be a generic phenomenon and specificity of its function is not clearly understood (Core and Adelman, 2019). Thus, we questioned whether the growth regulatory role of 7SK snRNP was specific to Yki or other growth promoters are also regulated via the 7SK snRNP complex. To test this hypothesis, we used expression two other growth promoting factors EGFR and Notch intracellular domain in combination with depletion of 7SK snRNP. We did not observe wing disc overgrowth when depletion of 7SK snRNP components in combination with overexpression of other well-known onco-proteins such as EGFR or Notch intracellular domain (NICD) (Fig.7B,C)**.** These observations suggest that, *Drosophila* 7SK snRNP complex may exclusively function to repress tumorigenic potential of Yki *in vivo* in an epithelial tumor model.

To validate the tumorigenic phenotype, we examined effect of depletion of 7SK snRNP components using RNAi lines from an independent resource. We observed that RNAi-mediated depletion using Bloomington RNAi resource stocks for *bin3 and Hexim* showed massive overgrowth in wing disc tissue similar to KK RNAi phenotypes. Thus, validating tumor suppressor function of 7SK snRNP complex in context of Yki.

Components of the NELF complex function as tumor suppressors.

The NELF complex is composed of four sub-units- NELF-A, B, C/D and E. As described earlier, the NELF complex is also crucial for PPP establishment and has to be actively removed from promoter proximally paused Pol II (Adelman and Lis, 2012). To extend the findings in the Yki screen, we examined larval phenotypes in case of depletion of the NELF complex components without Yki expression. Depletion of the NELF components on their own did not cause such giant larval phenotype or overgrowth of the wing disc tissue (Fig.7A). In contrast, depletion of each of these NELF components using RNAi in combination with Yki produced a giant larval phenotype (Fig.8B) and massively overgrown wing disc tissue compared to the larvae only over-expressing Yki (Fig.8A).

In concordance with larval phenotype, quantitation of GFP expressing area of the wing disc tissue, showed no difference in wing disc area when the NELF complex components were depleted alone (Fig.8C). Manifold increase in area was observed in the wing discs with combination of Yki expression and NELF complex depletion (Fig.8D). Magnitude of this increase was in the range of 15 to 17.5 fold compared to WT GFP expressing wing discs and significantly higher compared to wing discs with Yki expression alone.

Similar to observations with depletion of 7SK snRNP, depletion of the NELF complex components too, did not show any growth or tumor phenotype in the context of over-expressed EGFR or NICD (Fig.7B,C). Additional confirmation of tumor suppressor function was made using Bloomington RNAi resource stocks for NELF complex components, which showed giant larval phenotype with massive overgrowth of wing disc tissue.

Neoplastic transformation induced by Yki combined with depletion of 7SK snRNP or NELF complexes

The 7SK snRNP and the NELF complexes could regulate growth, given the specific context of Yki expression. Interestingly, Yki on its own, is known to promote cell proliferation and cell

survival (Pan, 2010). Thus, it is possible that larger size of the wing disc tissue observed upon loss of either 7SK snRNP or NELF complex is a result of enhancement of growth and survival effect of Yki rather than neoplastic transformation. To distinguish between the two possibilities, we analyzed the tumor tissue using markers that can distinguish between neoplastic and non-neoplastic growth.

First, we examined epithelial cell polarity. Wing disc tissue is an epithelial tissue and thus exhibits the characteristic apico-basal polarity. The polarity is defined by presence of distinct groups of proteins that are exclusive to apical or basal side (Martin-belmonte and Perez-moreno, 2012). Neoplastic transformation of an epithelial tissue is often accompanied by the loss of their characteristic apico-basal cell polarity (Ellenbroek, Iden and Collard, 2012), thus making is it a reliable indicator of neoplastic transformation versus the WT epithelium. There are multiple polarity determinants which have been implicated as tumor suppressors such as *Dlg, lgl scribbled* to name a few. E-cadherin (E-Cad) is a part of adherens junctions which are sub-apically localized in *Drosophila* and thus it provides a convenient marker for examining epithelial polarity (Tanos and Rodriguez-Boulan, 2008). Wing discs overexpressing Yki alone showed localization of E-Cad, in a pattern similar to the wild type wing discs, although former discs are much larger (Fig.9A). This indicated that Yki overexpression caused overgrowth of the epithelium without perturbation of epithelial cell polarity. In contrast, when Yki overexpression was combined with depletion of a component of the 7SK snRNP complex or the NELF complex, sub-apical localization of E-cad was lost or perturbed (Fig.9A). The perturbation in E-cad localization can be visualized in the optical sections of all tumor discs (Fig.9A(a'-e')). It should be noted that this phenotype was consistently observed in a patchy format across a wing discs tissue, wherein a region of the tissue showed complete absence of E-cad while another region showed mis-localization of E-cad. Interestingly, we observed the severity of such perturbation was higher in wing discs with Yki and NELF depletion compared to Yki and 7SK snRNP depletion.

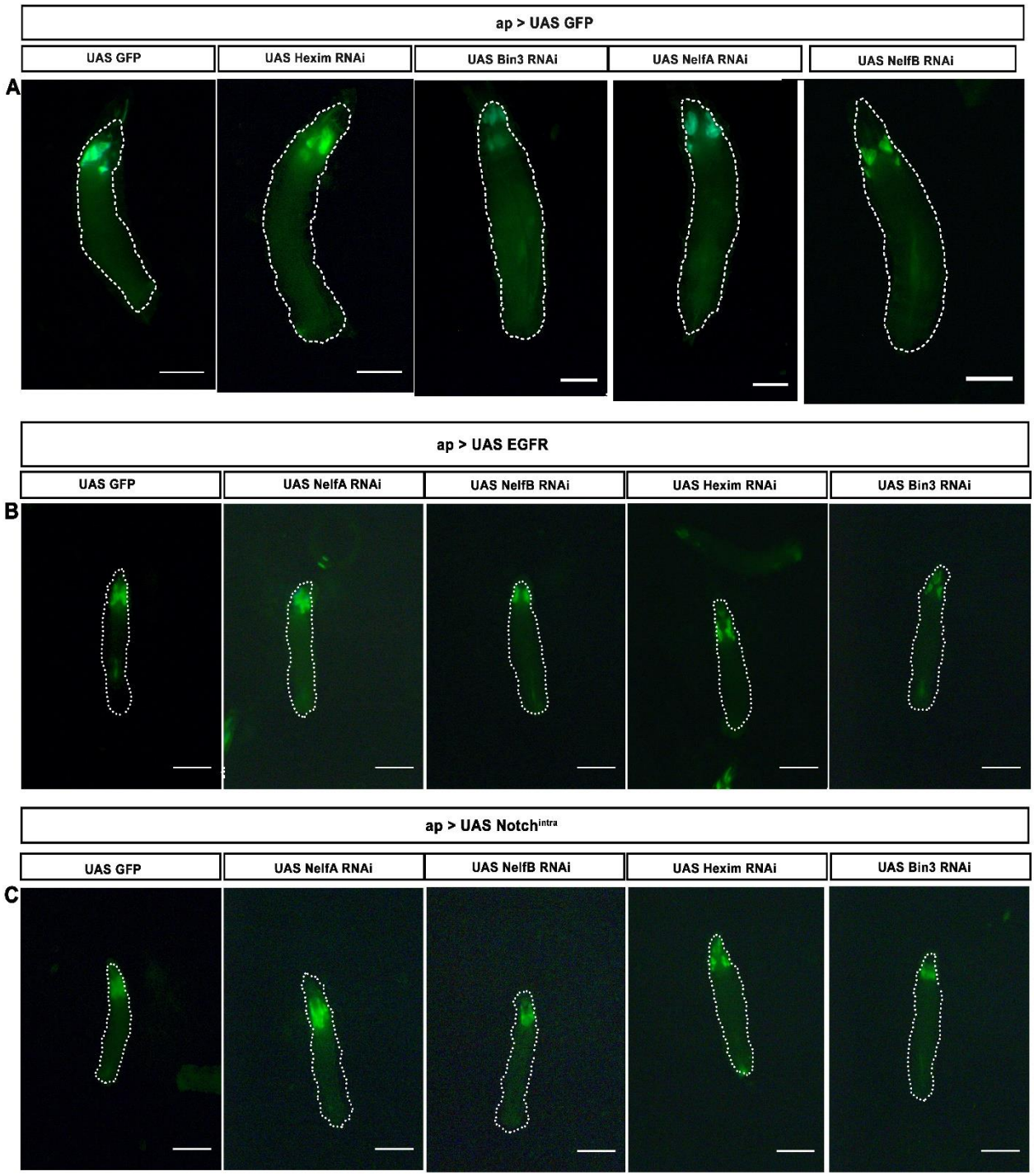


Figure 7.: The larval images of RNAi mediated depletion of 7SK snRNP and depletion of NELF complex components compared to GFP alone with the *apGal4* driver which is the wild type control. A) Depletion of 7SK snRNP and NELF complex shows phenotype similar to larvae expressing GFP alone. B) Depletion of 7SK snRNP and NELF complex combined with EGFR overexpression does not show massive overgrowth. C) Depletion of 7SK snRNP and NELF complex combined with Notch^{intra} overexpression does not show massive overgrowth. (Scale bar: 0.77mm)

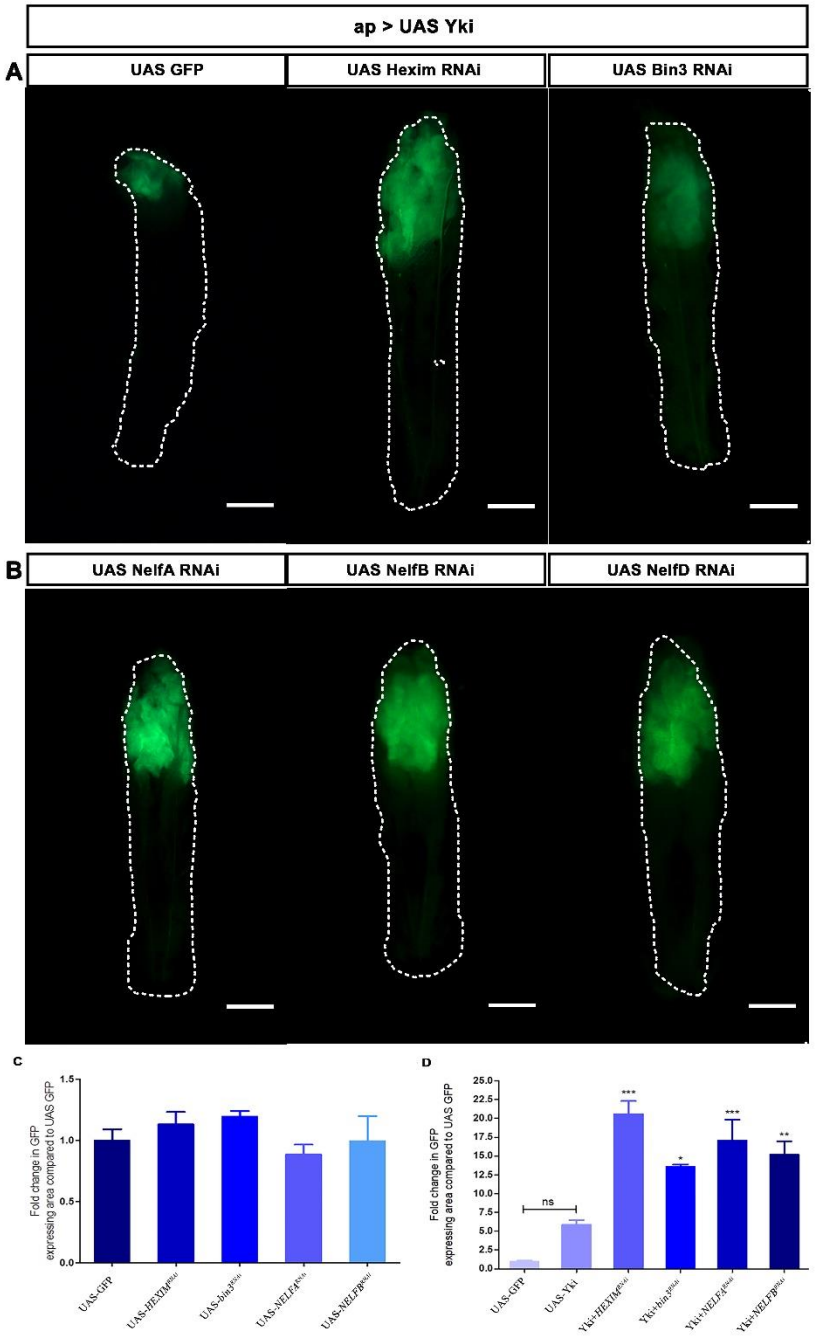


Figure 8.: The giant larval phenotype A) Massive overgrowth was observed when Yki overexpression combined with RNAi mediated depletion of 7SK snRNP. B) Depletion of NELF complex components combined with Yki overexpression also showed giant larval phenotype with massive overgrowth. Compared to milder growth observed in Yki expression alone. (Scale bar: 0.77mm). C) Fold change in the size of wing disc tissue caused by the depletion of the 7SK snRNP and the NELF complex components. There is no significant difference between these genotypes and in the control at $p< 0.05$. D) Fold change in the size of wing disc tissue caused by the overexpression of Yki alone or in the background of the depletion of the NELF and 7SK snRNP components (*$p< 0.05$; **$p< 0.01$; *** $p< 0.001$; One-way ANNOVA). Comparison between UAS-GFP and tumor tissue is significant at $p< 0.001$ and is not mentioned in the graphs. N=5 wing discs of each genotype.

Additionally, we analyzed F-actin, which localizes near the apical junctions of the wing disc epithelial cells, using Rhodamine labelled Phalloidin. As with E-Cad, we observed loss of apical localization of F-Actin in the Yki expressing tissue depleted of a component of the 7SK snRNP or the NELF complex, but not in wing disc tissue expressing Yki alone (Fig.9C). Here also, we observed the severity of such perturbation was higher in wing discs with Yki and NELF depletion. We did not observe any change in cell polarity as indicated by E-Cad and F-actin localization in wing discs with depletion of components of 7SK snRNP and NELF complexes alone (Fig.10A).

The matrix metallo-protease MMP1 has been used as a marker of epithelial to mesenchymal transition (EMT) and neoplastic transformation in *Drosophila* tumor models. MMP1expression is elevated in tumor models and its depletion by RNAi has been reported to block metastasis (Uhlirova and Bohmann, 2006; Miles, Dyson and Walker, 2011). We examined the effects of depleting components of 7SK snRNP and NELF complexes in Yki-expressing tissue on the levels of MMP1 by immunohistochemistry. We observed elevated levels of MMP1 in wing discs over-expressing Yki and depleted for Bin3, Hexim or the NELF complex (Fig.9B). We observed only mild increase in MMP1 levels in wing discs expressing Yki alone (Fig.9B), while in wing discs with depletion of components of 7SK snRNP and NELF complexes alone did not show MMP1 expression (Fig.10B). These observations were supported by quantitation of mean intensity of MMP1 in wing disc tissue. Wing discs with depletion of the 7SK snRNP and NELF components did show difference in MMP1 intensity compared to GFP expressing WT wing discs (Fig.10C). Yki expression alone showed a mildly elevated MMP1 levels (Fig.10D) in contrast to 3.5-fold to 6-fold increase in MMP1 levels observed in wing discs with combination of the 7SK snRNP or NELF complex depletion with Yki expression ((Fig.10D). Interestingly, in case of MMP1 levels as well we observed depletion of NELF complex gave much robust increase in MMP1 levels (6 and ~4-fold increase with NELFA and NELFB depletion respectively) compared to depletion of the 7SK snRNP components. Such differences combined with the severity of perturbation in epithelial cell polarity might indicate a stronger role of the NELF complex in regulating neoplastic activity of Yki in wing imaginal disc tissue.

Depletion of NELF complex components regulate YAP target gene expression in mammalian cells

The 7SK snRNP and NELF complex components show conservation across *Drosophila* and mammals (Peterlin and Price, 2006). Therefore, we were curious to examine the effects of depletion of 7SK snRNP and NELF complex components on YAP driven transcription. Firstly, we tested effect of siRNA mediated depletion of the two complexes on Luciferase system responsive to YAP (Nguyen *et al.*, 2014). siRNAs from SMARTpool (SIGMA-Aldrich) against NELFA, HEXIM1, HEXIM2 and MEPCE (orthologue of Bin3) we determined the extent of knockdown, which was >95% with siNELFA, 60% with siHEXIM1, siHEXIM2 and siMEPCE (Fig.12). Preliminary observations suggested that depletion of NELFA or HEXIM1 and 2 or MEPCE alone did not change Luciferase reporter activity except a slight dip in case of MEPCE depletion (Fig.11A).

However, when we combined siRNAs for these factors with YAP (S127A/S397A), a constitutively active form of YAP, expression vector we observed interesting trends. The combination of siNELFA and YAP led to an increased luciferase activity over and above the increase resulting from YAP expression alone (Fig.11B). However, depletion of MEPCE or co-depletion of HEXIMs with YAP showed a trend of decrease in luciferase activity compared to that of YAP expression alone (Fig.11B).

Based on Luciferase reporter assay we sought to determine if depletion of the 7SK snRNP complex and the NELF complex affects YAP target genes. To test this, we examined changes in expression of levels of YAP target genes in 293T cells. To study outcome and function of the NELF and 7SK snRNP in mammalian cells, we induced conditions similar to the screen in *Drosophila*, wherein constitutively active YAP is expressed in combination with siRNA mediated depletion of the NELFA and 7SK snRNP complex components, MEPCE and HEXIMs. Preliminary results indicate a similar trend as observed in case of Luciferase assay. Depletion of NELFA in combination with YAP led to an enhancement in expression of YAP target genes such as *CTGF, ANKRD1, Cyr61* over and above the change caused by YAP expression alone (Fig. 12A). On the other hand, non-target gene expression levels did not get altered (Fig. 11A). Depletion of 7SK snRNP complex components either did not change expression of Yap target genes or caused a slight dip in expression (Fig. 12B,C). However, these are preliminary findings and need further validation of the trends we observed.

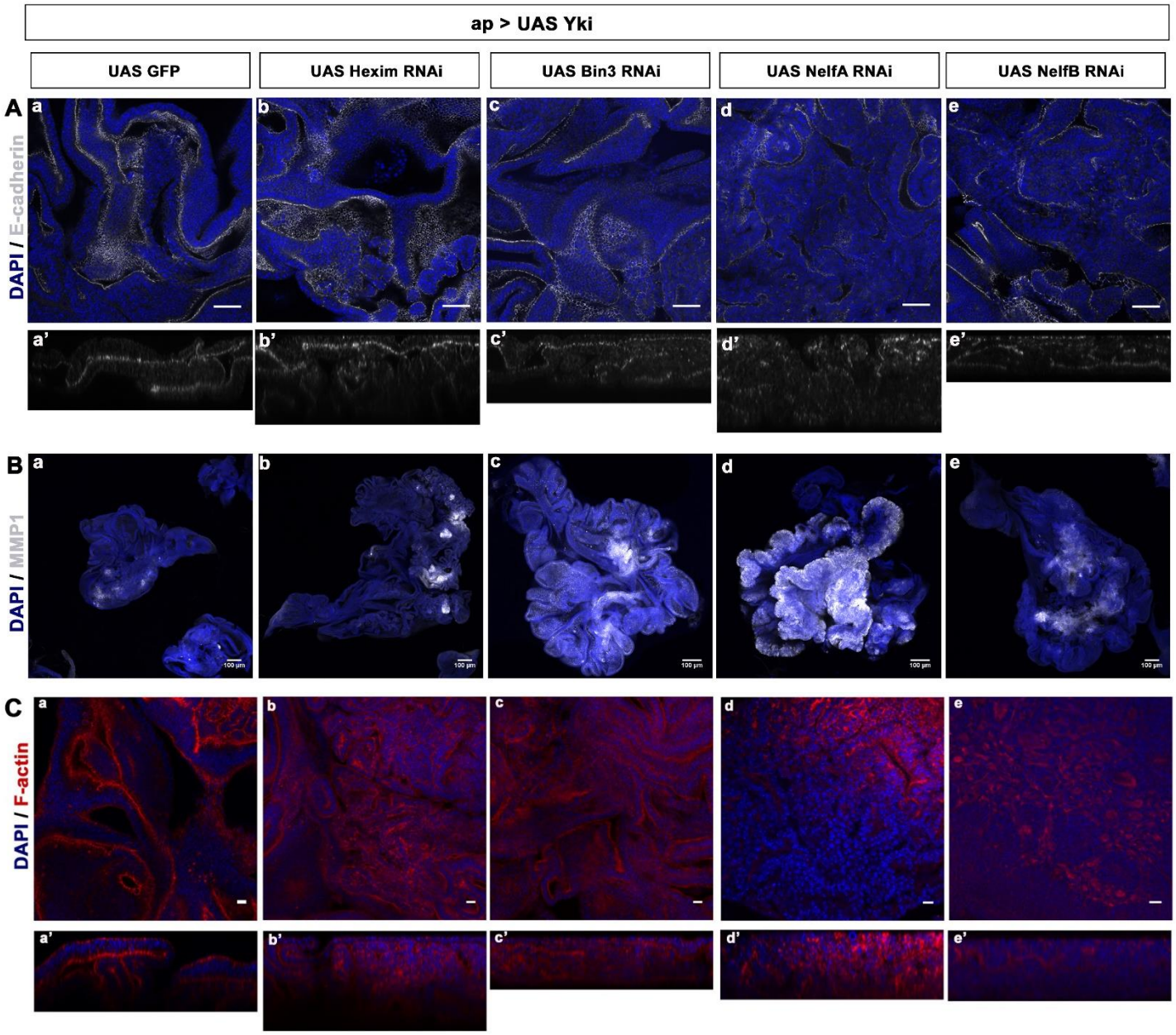
Taken together, tumors formed upon depletion of the 7SK snRNP or NELF complex components in combination with Yki exhibit neoplastic characters. As neither genetic change alone produced these results, it appears that they act in combination to promote neoplasia, a classical mechanism of cooperative tumorigenesis known in mammals. The observations in 293T cells provide support to these findings of tumor suppressor activity of the 7SK snRNP and NELF complexes may have, in the context of elevated Yki activity.

Clinical data analysis

Given the premise of tumorigenesis driven by combination of Yki overexpression and deregulation of PPP our findings established, discovering that the NELF complex and 7SK snRNP complex likely function in mammalian cells as in *Drosophila* was encouraging. These parallels motivated us to investigate cancer patient data with respect to the NELF and the 7SK snRNP complexes in conjunction with YAP. The Hpo pathway effector YAP has already been implicated in multiple cancer types as a cancer gene and based on our findings we wanted to ask if the NELF and 7SK snRNP complex components were important in context of cancer patient who have elevated YAP expression or activity. We were also curious about connection of these PPP regulators with clinical outcome across cancers. To test that we asked whether expression mammalian orthologues of NELF complex components and 7SK snRNP components have correlation with survival of cancer patients across a range of cancers.

For analysis of patient data, we made use of UCSC Xena browser. Xena browser accesses patient data from public platform of GDC-TCGA and allows users to perform categorization of patient data based on gene expression, clinical parameters and several other genomic and phenotypic characters (Goldman *et al.*, 2019).

To examine correlation of NELF complex and 7SK snRNP complex components with patient survival, we chose copy number alterations, mutations and gene expression levels as parameters for classification of patient samples. Copy number alterations and gene mutations were observed to occur infrequently in genes encoding the NELF and the 7SK snRNP complex components, ruling out these two criteria for further analysis.



<u>Figure 9.</u>: Characterization of tumors induced in the wing disc. (A) Disruption of characteristic epithelial apico-basal polarity in tumor discs. Images of wing discs over-expressing Yki alone (crossed to UAS-GFP as control) (a) or in combination RNAi-mediated knowckdown of *Hexim, bin3, NelfA* or *NelfB* using GAL80^{TS}; *ap*-GAL4; UAS-GFP (From a-e respectively) (scale bar = 50µm). Discs are stained for E-cad (white). Bottom panel (a'-e') of each image shows orthogonal optical section of respective genotype. Note delocalization of E-cad in tumorous tissues caused by the depletion of a component of PPP and Yki over-expression. All discs are also stained with DAPI (blue) to visualize nuclei.

(B) Increased expression of MMP1 is observed in tumor discs. Images of wing discs over-expressing Yki alone (crossed to UAS-GFP as control) or in combination RNAi-mediated knowckdown of *Hexim, bin3, NelfA* or *NelfB* (From a-e respectively) (scale bar = 100µm). Wing disc tissues are stained for MMP1 (white)

(C) Discs are stained with Phalloidin (red), which stain F-Actin. Bottom panel of each image shows orthogonal optical section of respective genotype. Note delocalization of F-actin in tumorous tissues caused by the depletion of a component of 7SK snRNP and NELF complexes, *Hexim, bin3, NelfA* or *NelfB* and Yki over-expression (From a-e respectively). All discs are also stained with DAPI (blue) to visualize nuclei (scale bar = 10µm

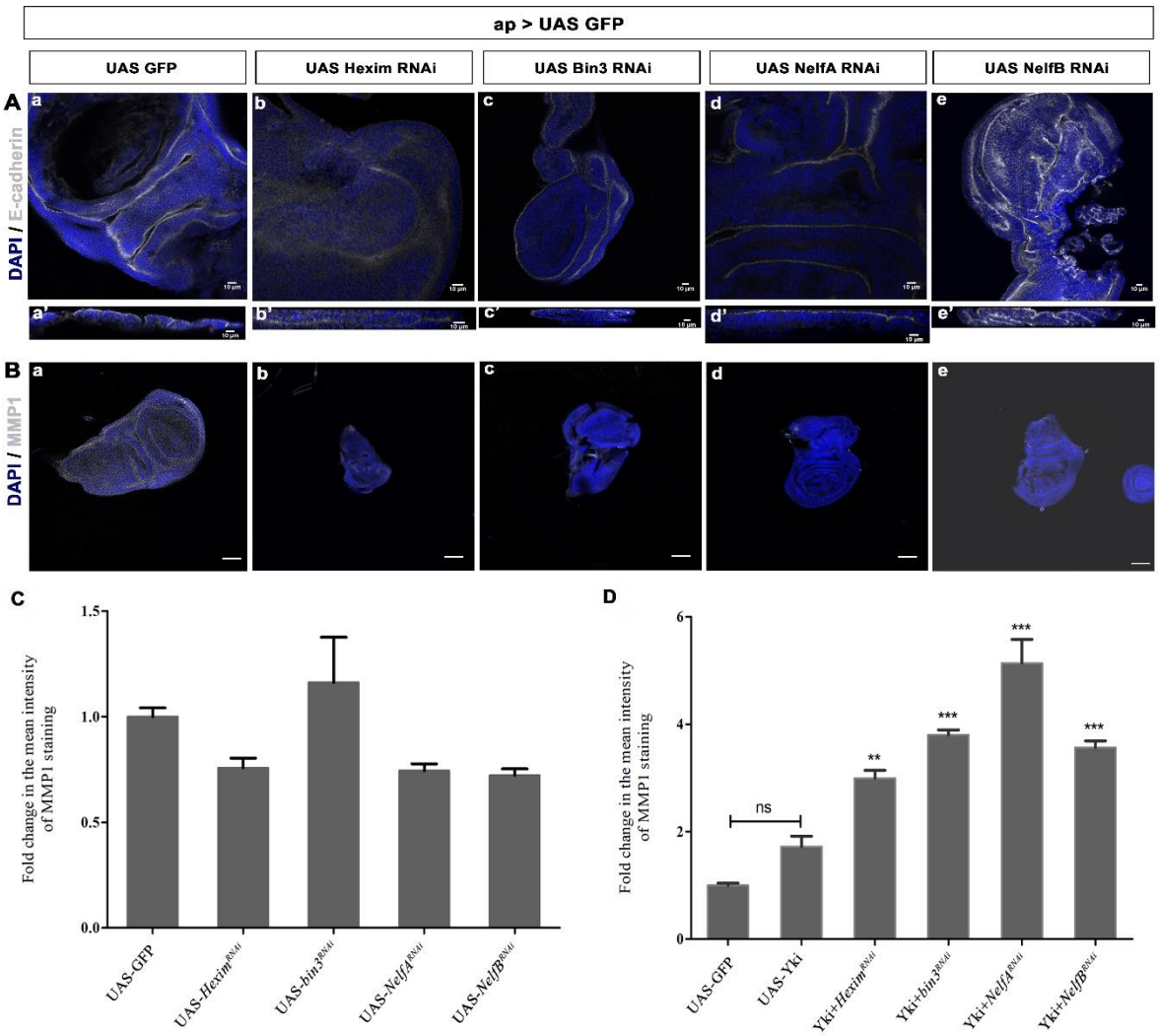
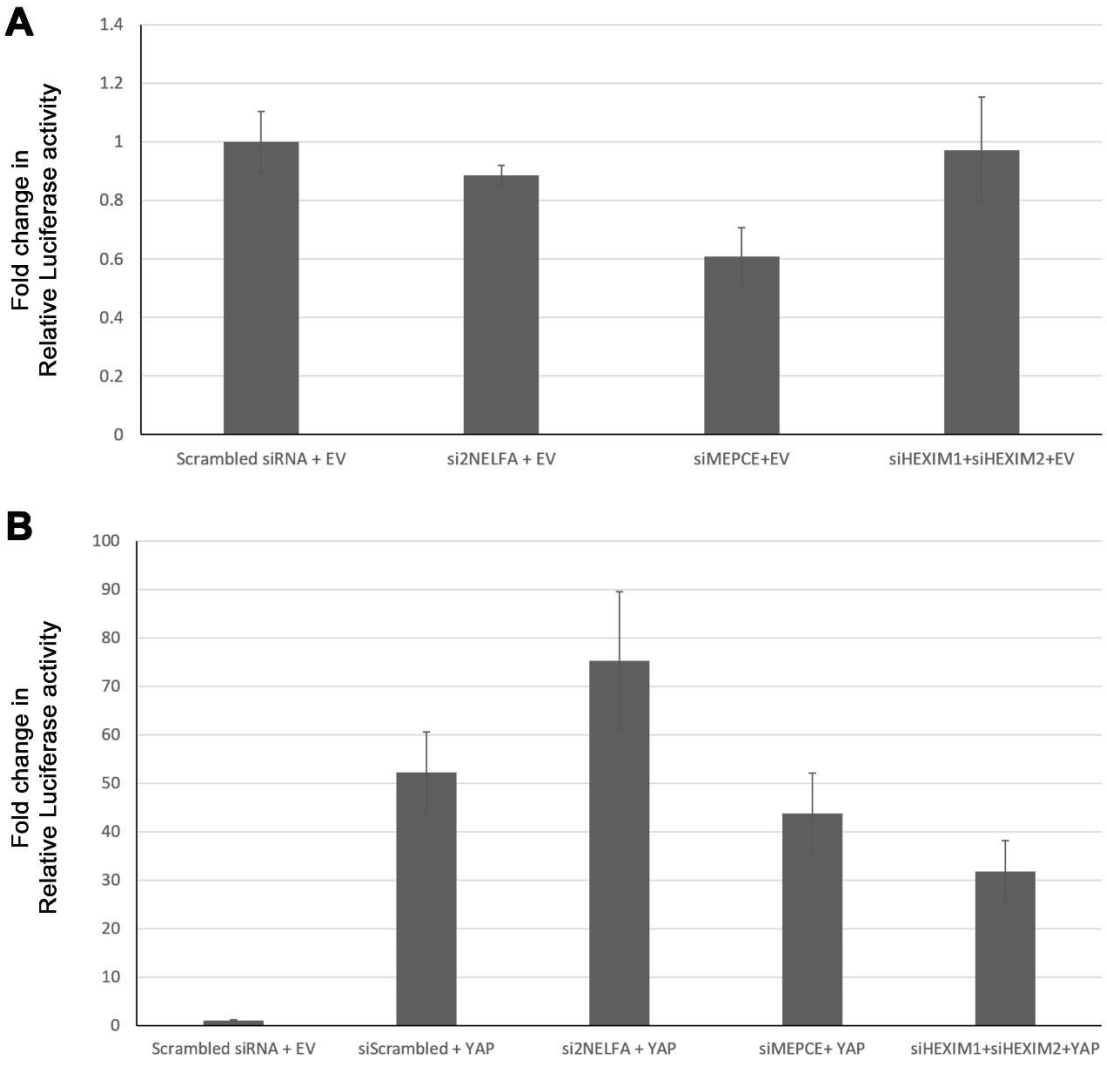


Figure 10.: A) Ecadherin staining in 7SK snRNP (b and c) and NELF depleted tissue (d and e) compared to GFP expressing WT wing disc tissue (a). B) MMP1 expression in 7SK snRNP (b and c) and NELF RNAi alone (d and e) wing disc tissue compared to GFP expressing WT wing disc tissue (a) (Scale bar = 100µm). Ecad is localized in apical region while we did not observe any MMP1 expression in these discs. C) Quantification of fold change in mean intensity of MMP1 staining in wing imaginal discs shows no significant difference in MMP1 staining intensity between wing discs of RNAi-mediated depletion of the 7SK snRNP and NELF complex components. Mean intensity measurements are normalized with mean intensity of MMP1 in GFP expressing WT wing discs. D) Quantification of fold change in mean intensity of MMP1 staining in wing imaginal discs shows significant difference in MMP1 staining intensity between wing discs over-expressing Yki alone and wing disc tissue of Yki in combination with depletion of the 7SK snRNP and NELF complex components. Mean intensity measurements are normalized with mean intensity of MMP1 in GFP expressing WT wing discs. (*p< 0.05; **p< 0.01; *** p< 0.001; One-way ANNOVA). N=5 wing discs of each genotype.



<u>Figure 11.:</u> Effect of depletion of the NELF and 7SK snRNP components on YAP responsive luciferase reporter. (A) Shows effect of depletion of NELFA, MEPCE and HEXIMs alone using siRNA on luciferase reporter compared to control siRNA, siScrambled (N=2, n=6). (B) Effect of combination of siNELFA, siMEPCE and siHEXIMs with YAP (S127A/S397A) (N=2, n=6). Error bars represent SEM. EV is the empty vector control.

Gene encoding these two complex components, did not show any particular trend, such as higher or lower expression of individual genes, across cancer types. Rather, it presented a diverse pattern, thus we reasoned that, grouping of patients-based gene expression level based on quartiles of expression range, in individual cancer types, could refine the cohort revealing trends. low expression group contained lower than first quartile referred as and greater than equal to fourth quartile as high expression of a gene. The rationale behind such criterion was to analyze a situation similar to the screen in *Drosophila*, wherein RNA interference lowered expression of NELF and 7SK snRNP complex components. Apart from analysis of survival in cohort of samples with lower expression of individual components of NELF and 7SK snRNP complex, we combined high expression of YAP as criterion.

We examined all cancer datasets available on UCSC Xena browser and observed a few of these showed correlation of patients with lower expression of NELF, 7SK snRNP complex components and/or with higher expression of YAP compared to remaining samples. High or low expression categories were defined based on quartiles, such that samples with expression lower than first quartile were categorized as low expressing cohort while, samples with expression greater than fourth quartile were termed high expressing cohort. The Following are the results of our analysis.

Bladder cancer (BLCA)

According to definition by National cancer institute (NCI), BLCA originates in urinary bladder tissue and is a carcinoma i.e. cancer of epithelial origin. TCGA has a total of 430 patient samples with expression data.

High and low expression categories of YAP1 expression each had 105 patient samples. The KM plots for survival showed that patients in the high expression cohort had significantly poorer survival compared to those in low expression category (p= 0.008391) (Fig.13A). In case of *NELFA* on the other hand, patients in low expression group had a significantly poorer survival (p= 0.009658), with each cohort containing ~105 patients (Fig.13B). Thus, we sought to test if we combined these two sets of patients with poorer survival. And interestingly we observed significant difference in survival of patients with combination of lower expression and *NELFA* and higher expression of YAP1 (n=40) compared to remaining patients ((p= 0.01062) (Fig.13F). This is

supportive evidence suggestive of PPP being deregulated via depleted NELF complex and 7SKsn RNP complex in the cohort of interest.

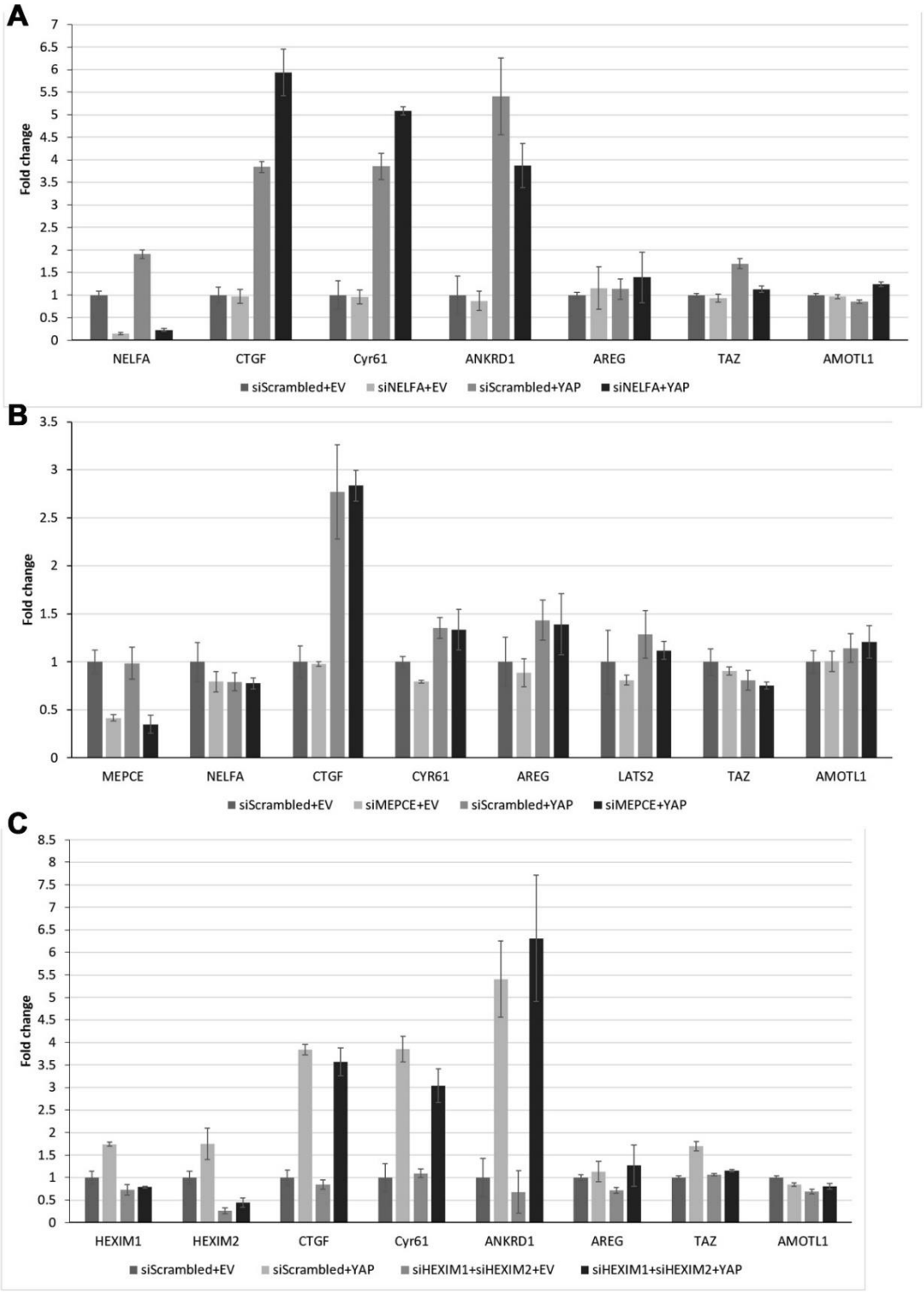
We observed similar results with poorer survival of patients with lower expression of HEXIM2 (n=104) compared to high expressing cohort (n=103) (p= 0.03737) (Fig.13C). However, the statistical significance was lost when high YAP1 expression criterion was applied in combination with low HEXIM2 expression (Fig.13G).

Other, 7SK snRNP complex components except Larp7 (Fig.13D) and NELF complex components did not show significant difference in survival of patients, alone or in combination with high YAP1 expression (except HEXIM1 (Fig.13E)).

Breast cancer (BRCA)

BRCA is another type of carcinoma originating in breast tissue. BRCA dataset is the largest dataset in terms of number of patients. Analysis of individual genes showed the following.

Survival of high YAP1 expressing cohort (n=298) is significantly poorer compared to low YAP1 expressing cohort (n=293) (p=0.0028) (Fig.14A). In case of NELF complex components, none of the NELF components showed significant difference in survival of low vs high expressing cohorts (Fig.14B-D). In case of 7SK snRNP complex components, patients with lower expression of HEXIM1 (n=296) and HEXIM2 (n=295) show poorer survival compared to patients with higher expression of HEXIM1 and 2 (n=294 each) (p=0.002985, p=0.00999 respectively) (Fig.15A-C). Next, we examined survival of patients with combination of higher YAP1 expression and lower expression of individual components of NELF and 7SK snRNP complexes. Cohort of patients with low expression of NELF and high expression of YAP1 (n=136) showed a highly significant poorer survival compared to remaining patient samples (p=2.649X10^{-6}) (Fig.14E-G). Interestingly, the p value in this case is much lower than in case of YAP1 high expression as mentioned above. This cohort also had significantly lower expression of remaining NELF complex components and MEPCE, HEXIM1 and HEXIM2 of 7SK snRNP complex suggestive of depletion of PPP regulation in cancer tissue.



<u>Figure 12.:</u> Effect of NELF an 7SK snRNP depletion on YAP target genes. (A) Effect of NELFA knockdown alone and in combination with constitutively active YAP on YAP target genes and non-target genes (*AMOTL1*) tested using qRTPCR on total RNA isolated from 293T cells (N=2, n=3). (B),(C) Effect of MEPCE and HEXIM1 and 3 knockdown on Yap target genes (N=2, =3). Error bars depict SEM

Along with NELFA, high YAP1+low NELFB cohort (p=0.0174) and high YAP1+low NELFE cohort (p=0.01397) also showed poorer survival. On the other hand, 7SK snRNP component, HEXIM2 which had showed poorer survival correlation also had significantly lower p value when YAP1 high expression criterion was added (p=0.0001457) (Fig.15D). Low expression of MEPCE showed poorer survival when combined with high expression of YAP1 (p=0.011) (Fig.15E).

Pancreatic cancer (PAAD)

PAAD is an adenocarcinoma arising from pancreatic tissue. It is commonly found at an aggressive stage. We probed the TCGA patient data for this cancer type using criterion same as above. In a comparison of survival between patients of Ist quartile with patients of IVth quartile of YAP1 expression, we observed that high expressing cohort of IVth quartile had significantly poorer survival (p=0.000255) (Fig.16A). However, examination of NELF complex components and 7SKsn RNP complex components revealed that none of these showed significant difference in survival when we compared high expressing group with low expressing group.

Next, we examined a cohort made by combination of high expression and low NELF or 7SK snRNP component expression. In this analysis we observed that high YAP1+low NELFA cohort (p=0.0000148) (Fig.16F) and high YAP1+low NELFB cohort (p=0.002459) (Fig.16G) showed significant difference in survival such that, patients of these cohort fared poorer. High YAP1+low MEPCE (p=0.0008199) (Fig.16H) and high YAP1+low HEXIM2 (p=0.008179) (Fig.16I) cohorts of patients also showed poorer survival.

In addition to these cancer types we observed Kidney clear cell carcinoma showing poorer survival in case of high expressing cohort of NELFA (p=0.00005273) and MEPCE (p=0.003137). Other cancer types- Ovarian, Prostrate, Colon cancer did not show significant difference in survival correlating with lower or higher expression of YAP1, components of NELF and 7SK snRNP.

Discussion

PPP is emerging as a widespread and a critical step in regulation of transcription as described earlier. Despite the fact that PPP is observed at about 30-40% of the genes in mammals as well as in *Drosophila* the functional significance of the phenomenon remains underappreciated (Core *et al.*, 2012; Kwak, 2013; Core and Adelman, 2019). Here we present a unique study that emphasizes

on previously unknown function of PPP in limiting oncogenic output of Yki, a well-known oncogene.

Our findings reiterate the fact that Yki a known promoter of cell proliferation and survival results in an overgrown wing disc tissue. However, we find that this growth is strictly hyperplastic and the tissue does not exhibit any perturbations that indicate neoplastic character. However, depletion of transcription regulatory complex, 7SK snRNP or NELF augments the activity of Yki leading to neoplastic transformation. The indicators of neoplastic transformation we tested are epithelial cell polarity and MMP1 expression. Firstly, the epithelial cell polarity is a crucial factor for wing disc epithelium development. This has been evidenced by several studies that highlight importance of cell polarity in cell to cell communication and how it ensures appropriate growth, signaling and patterning events in a developing wing disc epithelium (Pallavi and Shashidhara, 2005; Campbell *et al.,* 2018). In addition to that, perturbation of epithelial polarity is seen in progressing tumors and is accompanied by epithelial to mesenchymal transition typical of neoplastic, invasive tumors in *Drosophila* tissue (Campbell *et al.,* 2018). Interestingly, such perturbation in apico-basal polarity induce apoptosis in wing disc epithelium unless an inhibited by molecular players such as p35 (Jezowska *et al.,* 2011). The genetic conditions we have examined provide a sustainable proliferation and apoptosis inhibition via Yki expression thus, perturbation of epithelial polarity which appear very typical of an wing disc epithelium undergoing EMT (Campbell *et al.,* 2018) provide a strong support to occurrence of neoplastic transformation and might explain absence of metamorphosis event of puparium formation. Secondly, it has been observed that wing disc tissue that overgrows undergoes folding as we have also seen in tumor discs as well as Yki expressing wing disc tissue. However, such tissue is known to secrete excess of basal membrane proteins that allow sustenance of tissue integrity which is what we have observed in case of Yki expressing wing discs which are grown larger than WT wing discs. Interestingly, MMP1 expression which is the second marker we have tested for neoplastic transformation is elevated significantly in only tumor wing discs, which can allow ECM to be digested and thus assisting the neoplastic transformation. Thus, the evidence we provide here strongly indicates an invasive neoplastic tissue as a result of cooperative mechanisms between depletion of PPP machinery and Yki expression.

Additionally, previous studies have shown that transcription of Myc target genes is regulated by PPP and Myc recruits P-TEFb to promote transcription of target genes (Rahl *et al.,* 2010). Similarly, it has been shown that a sub set of ER target genes are regulated by PPP and loss of

NELF components leads to increase in transcription of that subset (Aiyar *et al.*, 2004). Both, ER and Myc are promoters of tumorigenesis and implication of PPP in regulation of transcription by these two tumor drivers is particularly reaffirm our findings of function of PPP in limiting tumorigenic potential of Yki.

Studies of Myc and ER mentioned above, give insights into molecular level regulation of target genes of potent transcription regulators. Our observations provide results of a combination of a potent regulator of transcription and perturbation of PPP revealing a novel physiological function of PPP. Particular emphasis from our observations is on the specificity of context besed role of PPP, wherein we showed that only Yki driven growth is limited by PPP while output of other promoters of growth such as EGFR and Notch is not affected. Depletion of 7Sk snRNP complex alone or NELF complex alone does not have any effect on wing disc tissue highlighting importance of specific context of Yki. This is consistent with studies in *Drosophila* that showed that even potent enhancers such *Vp16* cannot alleviate pause of Pol II (Krumm, Hickey and Groudine, 1995). Additionally, our observations add novel insight into physiological role of PPP factors 7SK snRNP and NELF. Although, the general phenomenon pf PPP is appreciated in developing *Drosophila* embryo for synchronous vs asynchronous expression of patterning genes (Lagha *et al.*, 2013) evidence of specific function of 7SK snRNP and NELF is uniquely provided by our observation in wing disc tissue.

In context of disease, PPP has been implicated in cancer (Rahl *et al.*, 2010; Galli *et al.*, 2015). Here we provide a direct evidence of PPP in tumorigenesis. Yki driven hyperplastic overgrowth transforms into a neoplasia when PPP is perturbed by depletion of NELF and 7SK snRNP complex. It indicates that transcriptionally, PPP is limiting Yki driven growth program and the potential of neoplastic transformation. The tumor suppressor function of 7SK snRNP is consistent with previously reported role of Larp7, human orthologue of Larp, where it has been implicated in tumorigenesis and specifically in gastric cancer via P-TEFb mediated deregulation of transcription(He *et al.*, 2008; Cheng *et al.*, 2012). However, we report tumor suppressor function of NELF complex and other 7SK snRNP components for the first time.

Consistent with observations in *Drosophila* tumor model system, we observed that depletion of NELF in mammalian cells leads to enhancement of YAP target genes such as *CTGF, Cyr61, ANKRD1*. This increase along with increase in YAP luciferase activity upon NELF depletion although preliminary findings, substantiate the potential of PPP as a regulator mechanism in

context of Hpo pathway in mammalian system as well. Intriguingly, depletion of 7SK snRNP components failed to alter expression of YAP target genes and luciferase reporter activity in mammalian system. Albeit surprising nature of these observations, we can speculate that this could be due to the fact that, in mammalian cells, YAP has been shown to recruit P-TEFb actively to a sub-set of its targets (Galli *et al.*, 2015), thus rendering function of 7SK snRNP relatively redundant as far as restricting P-TEFb from releasing PPP is concerned. However, these findings are preliminary and verification of these effects using a non-transformed cell line will shed light upon validity of such regulation of YAP activity output by PPP machinery.

Examination of TCGA cancer patient derived RNA sequencing data showed a few cancer types with correlations between low expression of NELF complex and 7SK snRNP complex and poorer survival of these patients. Breast cancer, Bladder cancer and Pancreatic adenocarcinoma are the three types that showed such correlations. Consistent with previously known role of YAP1 as an oncogene, all these three types of cancer patients also showed poorer survival when YAP1 expression is higher. Interestingly, consistent with *Drosophila* tumor model, we observed such correlations in carcinomas which are derived from epithelial tissue.

Examination of a context similar to the *in vivo* model, that is a combination of high Yki with depleted NELF or 7SK snRNP, led to an increase in the difference between survival probabilities. Particularly in case of breast cancer, we the filtering for combination of higher YAP1 and lower NELF complex components led to highly significant differences in survival, which were not seen in lower expression of individual NELF components. Thus, raising a possibility that, alike tumor model, PPP could be functioning to limit oncogenic potential of YAP. However, it has to be noted that, the analysis is limited relatively smaller number of patients, reducing statistical testing power of the analysis. Nevertheless, the trends observed here should pave the way for experimental design on patient samples.

Together, our observations give insights into context specific physiological function of 7SK snRNP and NELF complexes. These function as tumor suppressors, limiting tumorigenic potential of Yki. The specificity of Yki context also highlights cooperative mechanisms of tumorigenesis and prompts a deeper analysis of patient databases and mammalian systems for context specific function of conserved 7SK snRNP and NELF complexes in context of YAP which is also an

effector of conserved Hippo pathway. Further it will be interesting to test PPP based regulation of tumorigenesis using assays such as soft agar colony formation assay for anchorage independent growth, migrations assays for metastatic potential and xenograft assays to examine the extent of regulation PPP exerts on YAP driven tumorigenesis. From clinical perspective, it will be interesting to probe possibility of correlations between clinical sub types of breast cancer with these genes involved in PPP.

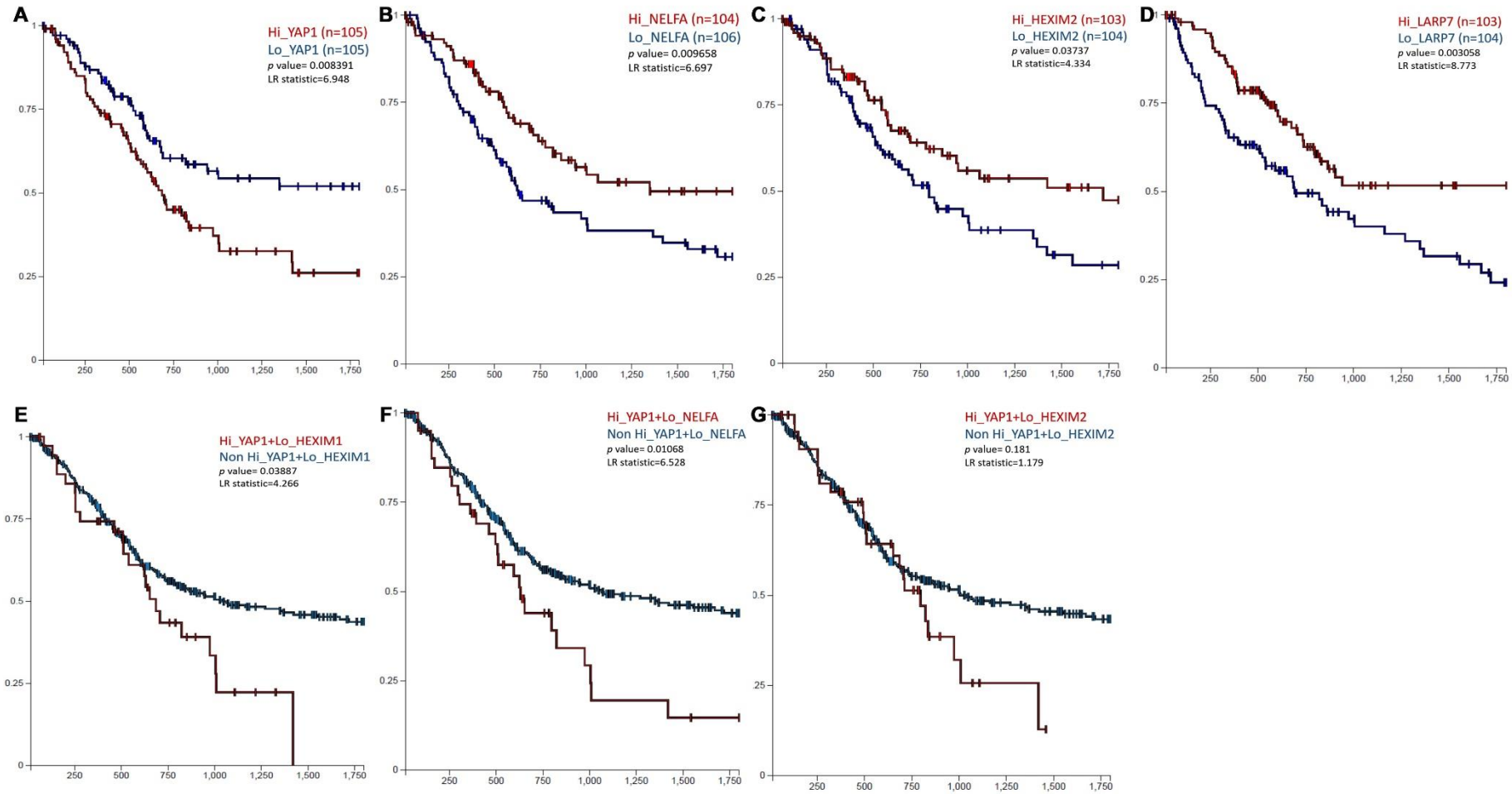


Figure 13.: Overall survival in Bladder cancer patients. (A)-(D) KM plots of bladder cancer patients divided based expression of individual genes YAP1, NELFA, HEXIM2 and LARP7 (A through D respectively) based on quartiles such that high expression represents the fourth quartile (Red) and low expression represents the first quartile (Blue). (E)-(G) KM plots of bladder cancer patients categorized based on a combination of high YAP1 expression with HEXIM1, NELFA and HEXIM2 (E through G respectively) represented by Red color.

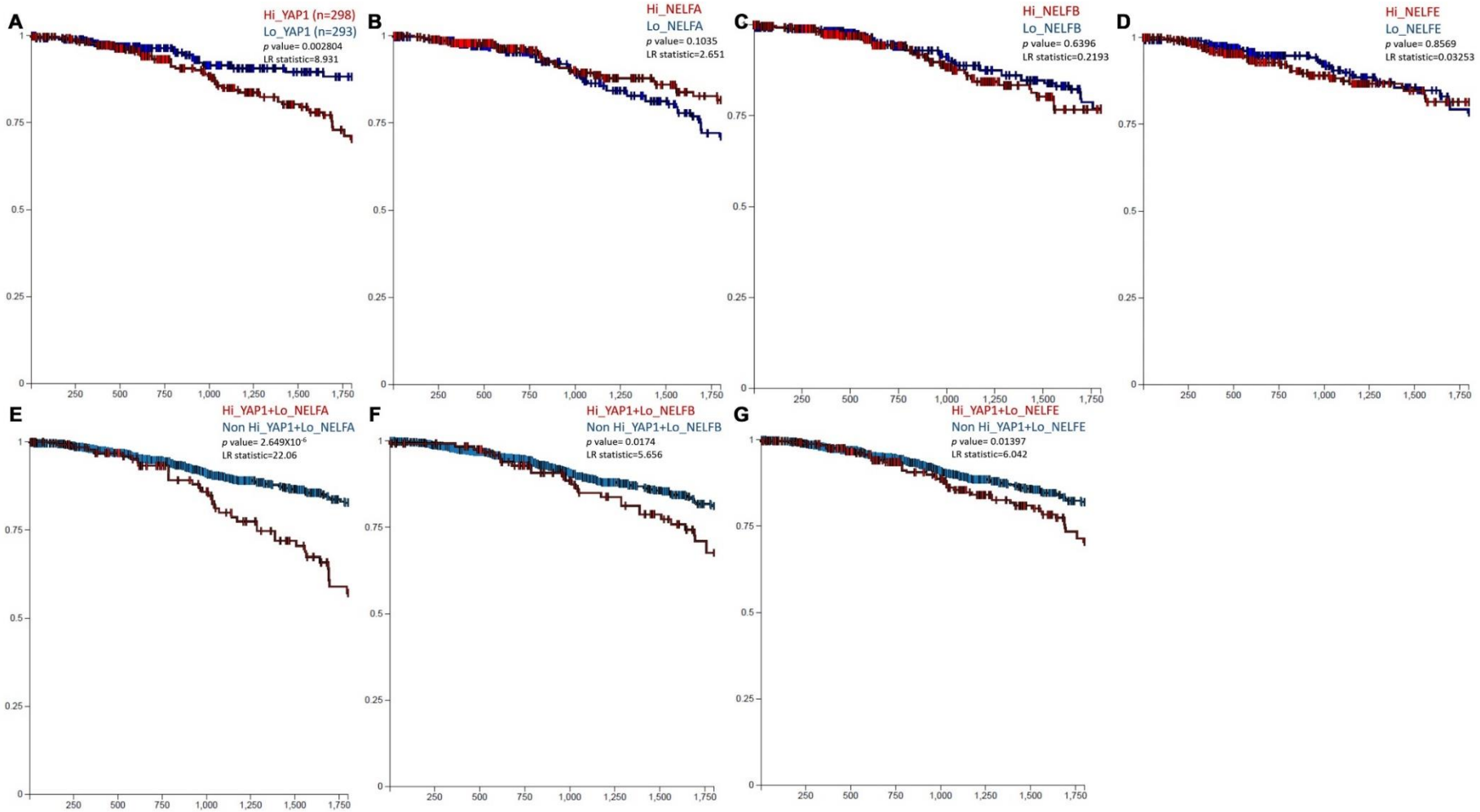


Figure 14.: Overall survival in Breast cancer patients. (A)-(D) KM plots of Breast cancer patients divided based expression of individual genes YAP1, and NELF complex NELFA, B, E (A through D respectively) based on quartiles such that high expression represents the fourth quartile (Red) and low expression represents the first quartile (Blue). Higher YAP1 expression correlates with poorer survival while individual NELF complex component gene expression is not correlated (E)-(G) KM plots of Breast cancer patients categorized based on a combination of high YAP1 expression with NELF complex components NELFA, B, E (E through G respectively) represented by Red color. All of these show poorer survival in cohort of High YAP1 combined with individual NELF complex component gene expression.

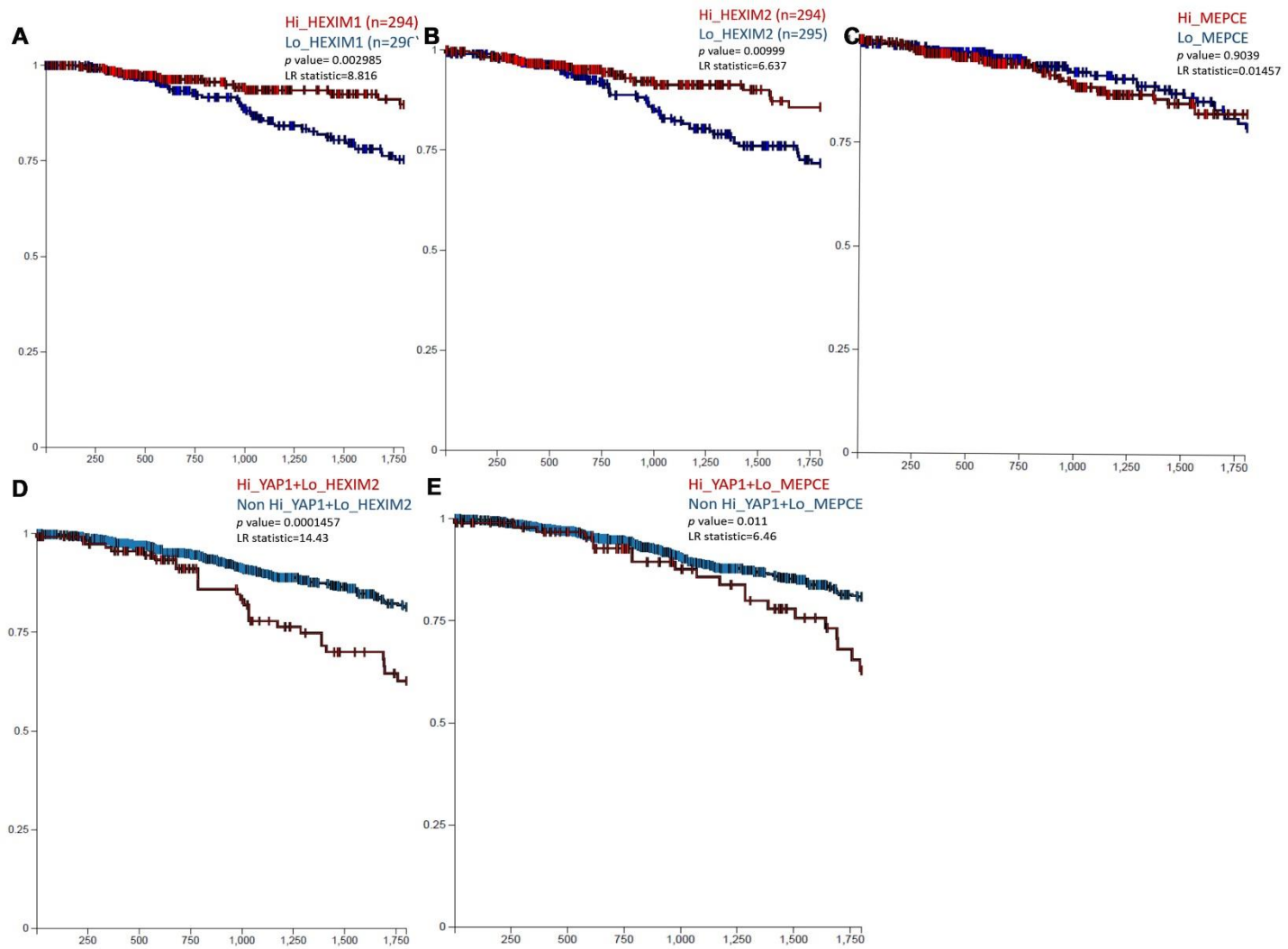
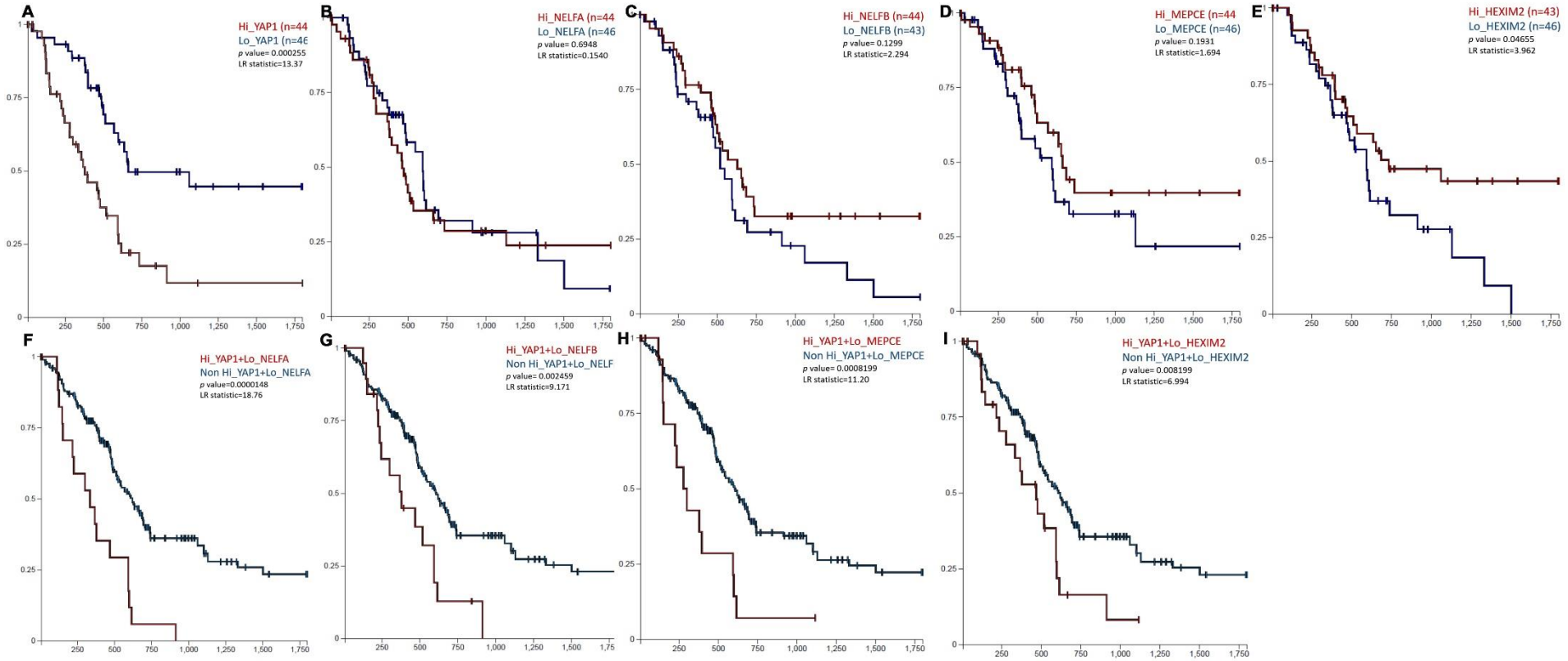


Figure 15.: Overall survival in Breast cancer patients. (A)-(C) KM plots of Breast cancer patients divided based expression of individual genes of 7SK snRNP complex- HEXIM1 and 2, MEPCE (A through C respectively) based on quartiles such that high expression represents the fourth quartile (Red) and low expression represents the first quartile (Blue). Lower HEXIM1 and 2 expressing patients show significantly poorer survival compared to high expressing cohort (D)-(E) KM plots of Breast cancer patients categorized based on a combination of high YAP1 expression with 7SK snRNP complex components HEXIM2 and MEPCE (D and E respectively) represented by Red color. All of these shows poorer survival in cohort of High YAP1 combined with individual 7SK snNRP complex component gene expression



.

Figure 16.: Overall survival in Pancreatic cancer patients. (A)-(E) KM plots of Breast cancer patients divided based expression of individual genes YAP1, and NELF complex NELFA, B, 7SK snRNP complex components MEPCE, HEXIM2 (A through E respectively) based on quartiles such that high expression represents the fourth quartile (Red) and low expression represents the first quartile (Blue). Higher YAP1 expression correlates with poorer survival while individual NELF and 7SK snRNP complex component gene expression is not correlated except HEXIM2. (F)-(I) KM plots of Breast cancer patients categorized based on a combination of high YAP1 expression with NELF and 7SK snRNP complex components (F through I respectively) represented by Red color. All of these shows poorer survival in cohort of High YAP1 combined with individual NELF and 7SK snRNP complex component gene expression.

Chapter 3: Function of CDK9 in cooperative tumorigenesis with Yki

Regulation of transcription post initiation is a critical rate limiting step and is subjected to regulation as described in previous chapter. The 7SK snRNP and NELF complexes are critical for maintaining and establishing the paused state of Pol II in promoter proximal region. In terms of regulation of gene expression as critical as the establishment and maintenance of pause is the release of paused Poll II. And release of paused Pol II is brought about by positive transcription elongation factor (P-TEFb) complex. This is perhaps the most studied regulators of PPP. The constituents of P-TEFb are cylinT1 or T2 (cyclin T in *Drosophila*) and a cyclin dependent kinase, CDK9 (Adelman and Lis, 2012).

P-TEFb brings about release of Pol II primarily through kinase activity of CDK9. A promoter proximally paused state of Polymerase II has protein complexes such as NELF, DSIF associated with it that do not allow for release of pause. Additionally, it has been shown that Pol II C-terminal domain (CTD) also needs to undergo post translational modifications, phosphorylation of Ser-2 of heptad repeats being an important elongation mark (Egloff, Dienstbier and Murphy, 2012). Thus, to release the paused Pol II CDK9 must cause pausing factors, NELF and DSIF to evict from template and Pol II, and bring about Ser-2 phosphorylation. CDK9 does that by phosphorylating NELF complex which leads to its eviction, while DSIF phosphorylation has been documented to turn the pausing factor into an assisting factor in elongation phase of transcription (Brès, Yoh and Jones, 2008; Jonkers and Lis, 2015).

PPP has been critical to trigger spontaneous and synchronous transcription in response to signaling pathways. Considering positive role of P-TEFb in regulating PPP, it has been of interest of several studies. PPP has also been observed to be widespread phenomenon. Interesting findings show that many signaling pathways tend to regulate expression of target genes by recruiting P-TEFb. c-Myc regulates target gene expression by recruiting CDK9 to highly expressed target genes as evidence by reduced level of expression on treatment with CDK9 inhibitor flavopiridol (Rahl *et al.*, 2010). Another bromodomain TF, Brd4 regulates transcription elongation by recruiting CDK9 (Jang *et al.*, 2005; Yang *et al.*, 2005). Brd4 is an activator of transcription and is known to recruit P-TEFb to increase transcription of target genes. Similar to Brd4, P-TEFB is a part of another large complex that promotes release of Pol II leading to transcription elongation called as super elongation

complex (SEC) (Luo, Lin and Shilatifard, 2012). Although P-TEFb effect on transcription is relatively global, observations that in certain context, higher impact on transcription of a sub-set of genes is observed, raising question that, there could be transcription modules that are more sensitive to P-TEFb action than their genome wide counterparts. However, evidence of such phenomenon is scarcely present.

On the other hand, P-TEFb has emerged as a strong candidate for targeted therapy, since it is at hub of transcription regulation and specific kinase inhibitors such as Flavoiridol, specifically target CDK9, inhibiting P-TEFb action. This is an important development since, observations have reported that specific inhibition of CDK9 activity reduce growth of patient derived xenografts in cases of adenocarcinoma (Allaway *et al.,* 2016). Curiously another recent study reported that, inhibition of CDK9 activity can also reduce expression of a sub-set of YAP targets (Galli *et al.*, 2015). These findings combined with identification of two tumor suppressor complexes 7Sk snRNP, which maintains PPP by sequestering P-TEFb and NELF complex which is stalling Pol II in promoter proximal regions by inhibiting its progress in context of Yki, *Drosophila* orthologue of YAP, raised a question can P-TEFb component CDK9 cooperate with Yki.

Results

CDK9 cooperates with Yki in tumorigenesis

A large portion of cellular the P-TEFb complex is unavailable due to sequestration by 7SK snRNP complex (Brès, Yoh and Jones, 2008), we hypothesized that overexpressing CDK9 might bypass regulation of pausing, exerted by the NELF complex and cause RNA Pol II release. Consistent with this hypothesis, we indeed observed massive tissue overgrowth when Yki was co-expressed with CDK9, while over-expression of CDK9 alone did not cause any such phenotype (Fig.17A). Quantitation of GFP expressing area of the wing discs was also in concordance with these observations as CDK9 expression alone did not show significantly different area measurement compared to WT wing discs in contrast, coexrepossion CDK9 and Yki led to overgrown wing discs with area greater than Yki expression alone as well (Fig.17B). Wing discs expressing UAS-CDK9 together with UAS-Yki also showed loss of apically localized E-Cad (Fig.18A,B) as well as elevated MMP1 expression (Fig. 18C,D) compared to tissue expressing UAS-CDK9 alone or UAS-Yki alone (Fig. 18E) as supported by MMP1 intensity measurement. This indicates neoplastic transformation in wing discs co-expressing Yki and CDK9, similar to the transformation

caused by depletion of 7SK snRNP and NELF complex components in combination with over-expressed Yki (Fig.18).

CDK9 is necessary for tumorigenesis induced by combination of depletion of 7SK snRNP and Yki expressoin

As a further test of if PPP deregulation is in cooperation with Yki expression causing neoplasia, we asked whether CDK9 is essential for tumorigenic outcome of 7SK snRNP depletion and Yki overexpression combination. Depletion of *cdk9* effectively suppressed the tissue overgrowth caused by depleting *bin3* or *Hexim* in Yki expressing tissue (Fig.19A). Those wing discs also showed normal apical localization of E-Cad and wildtype levels of MMP1 expression, suggesting complete suppression of tumorous growth (Fig.19B,C).

Depletion of CDK9 rescues tumor phenotype caused by cooperation between NELF depletion and Yki expression

Given that CDK9 is known to act directly on both NELF proteins and RNA Pol II, we wondered whether CDK9 activity would be required in the absence of the NELF complex. As shown above in the case of removing the 7SK snRNP complex, depletion of *cdk9* suppressed overgrowth caused by RNAi-mediated depletion of *NelfA* and over-expression of Yki (Fig.19A). This was accompanied by restoration of apico-basal polarity and MMP1 expression to wild type levels (Fig.19B,C). This finding provides evidence that alleviation of pausing by removal of NELF complex is not sufficient without CDK9 activity. This presumably reflects an importance of activation of RNA Pol II by CDK9-mediated phosphorylation.

Deregulation of PPP is not sufficient to cause overgrowth in wing disc tissue

Following the observations above, we wondered if combination of depletion of the complexes associated with PPP and increased CDK9 levels sufficient to cause over-growth phenotype or such a growth phenotype is tightly coupled to the presence of a growth driver such as Yki. Depletion of components of 7SK snRNP or NELF complexes in the background of over-expressed CDK9 did not cause any growth phenotype or morphological alteration in wing disc epithelium (Fig.20). This suggests that deregulation of RNA Pol II pausing is not sufficient on its own to produce an over-growth or neoplastic phenotype; yet it does so in the context of Yki over-expression. In the context of elevated Yki activity, there appear to be two brakes, each of which must be removed by CDK9 activity to allow excess Yki to produce tumors in *Drosophila* wing disc tissue.

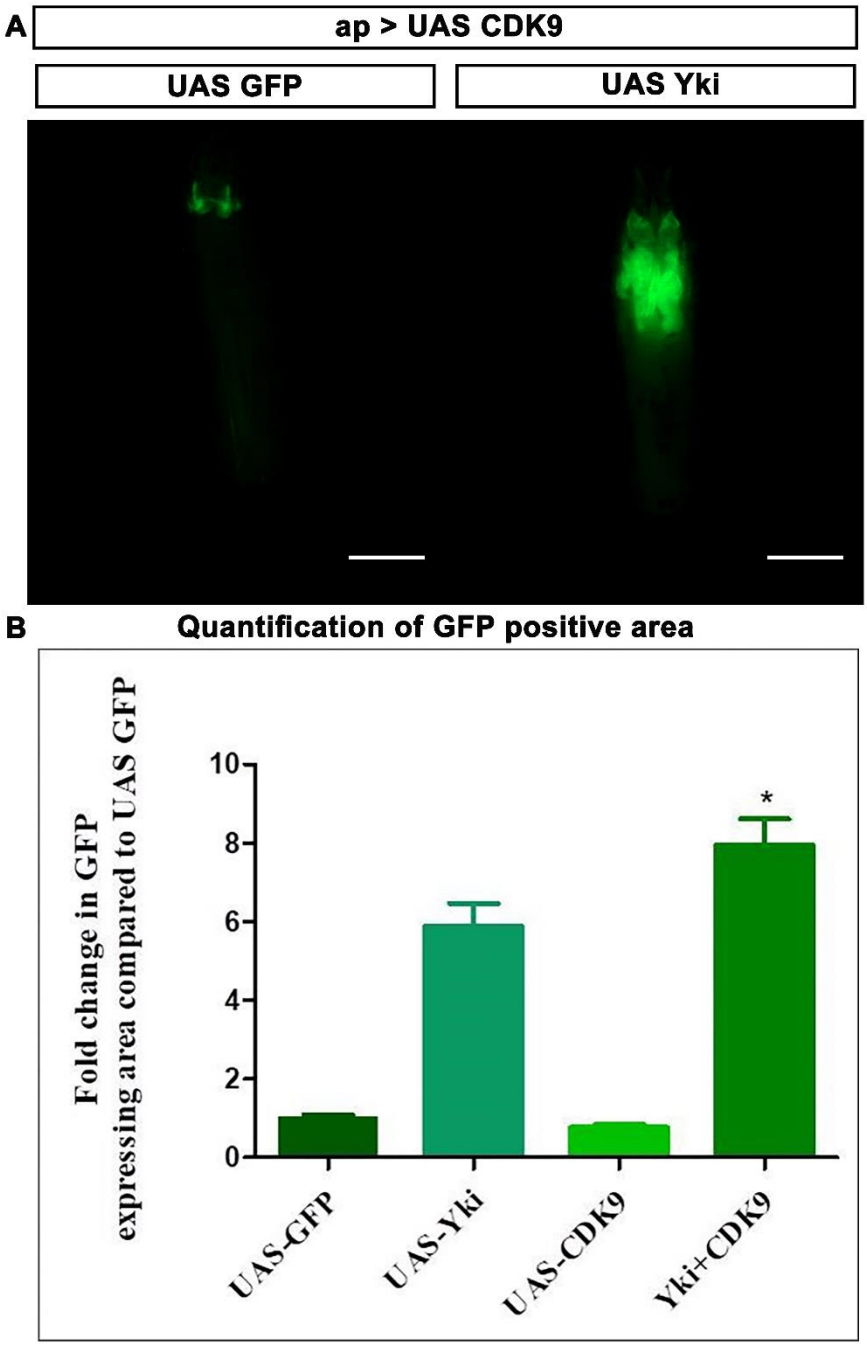


Figure 17.: Co-expression of Yki and CDK9 leads to giant larval phenotype. A) Larval images showing phenotypes of overexpression of CDK9 using *ap*-Gal4 (Left panel) and overexpression of CDK9 in combination with Yki overexpression (scale bar=0.77mm). **B)** Quantitation of GFP expressing wing disc area in respective genotypes. Yki+CDK9 coexpression showed significantly larger wing disc area compared to Yki expression alone. (*$p < 0.05$; One-way ANNOVA). N=5 wing discs of each genotype.

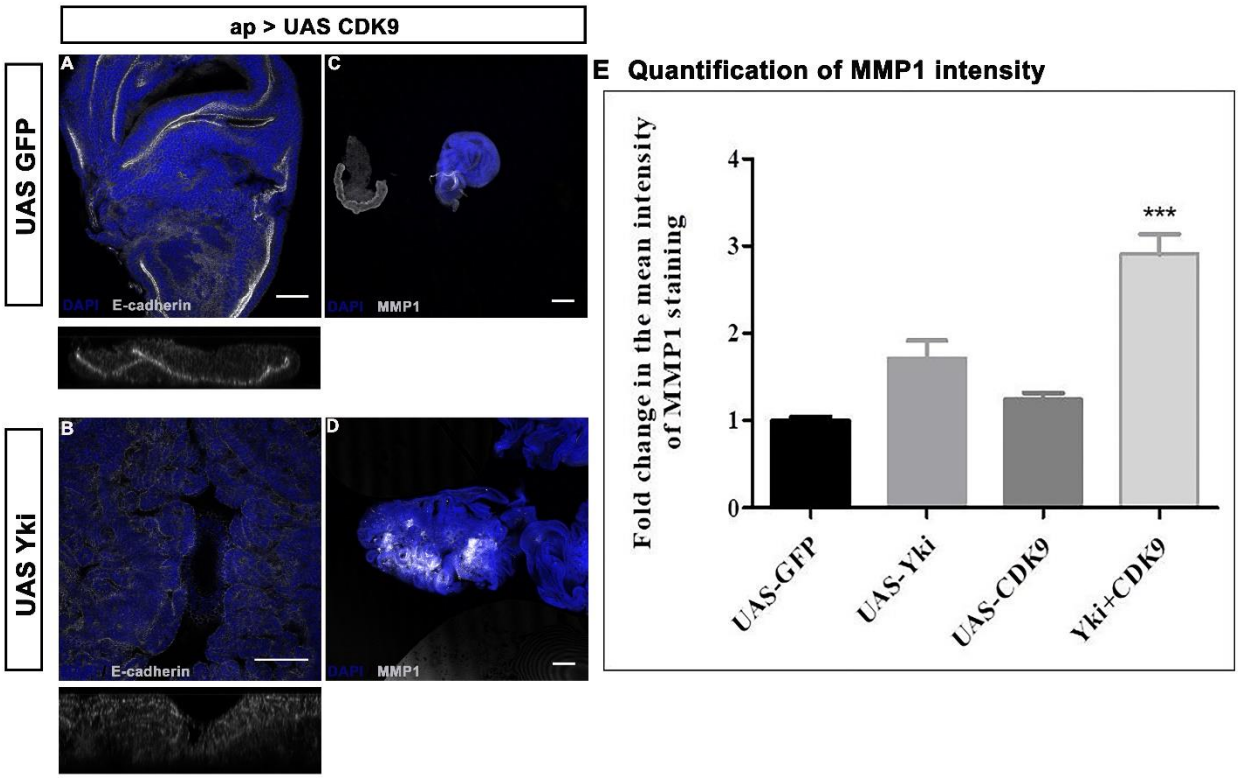


Figure 18.: Characterization of tumors induced by coexpression of CDK9 and Yki. Characterization of tumor tissue caused by combined over-expression of CDK9 and Yki using GAL80^{TS}; *ap*-GAL4. A) Shows wing disc tissue over-expressing CDK9 alone, stained for E-cadherin. Optical ortho-section shows apical localization of E-cad (scale bar = 50µm). B) Shows combined over-expression of CDK9 and Yki, stained for E-cadherin. Optical ortho-section shows perturbed localization of E-cad (scale bar = 50µm). C) Shows wing disc tissue over-expressing CDK9 alone, stained for MMP1 (white) expression (scale bar = 100µm). D) Shows combined over-expression of CDK9 and Yki, stained for MMP1 (white) expression showing increased MMP1 expression (scale bar = 100µm)., suggesting their neoplastic tumor state. All discs are also stained with DAPI (blue) to visualize nuclei. E) Quantification of MMP1 intensity in respective genotypes. Coexpression of Yki and CDK9 showed significantly higher MMP1 intensity compared to Yki expression alone. (*** $p < 0.001$; One-way ANNOVA). N=5 wing discs of each genotype

Discussion

CDK9 is the kinase component of the P-TEFb that mechanistically brings about effects of P-TEFb on the paused Pol II. We have observed that addition of CDK9 with overexpression Yki led to enhancement of growth in wing disc tissue, while Yki overexpression alone caused milder hyperplastic overgrowth. It needs to be appreciated that a large portion of cellular CDK9/P-TEFb is sequestered by7Sk snRNP and thus is unavailable. Thus, it was anticipated that overexpression of CDK9 might lead to a milder phenotype. However, CDK9 overexpression cooperated significantly with Yki leading to neoplastic transformation. These observations are consistent with a recent study sowing YAP recruits CDK9 via mediator complex to a sub set of target genes releasing the pause (Galli *et al.*, 2015).

It is intriguing that very little evidence of physiological function of P-TEFb is available. Despite the fact that we know that since P-TEFb assumes a critical position in determining the synchronicity and rapidity of transcription in response to signaling stimuli, it is critical for multiple physiological function. Here we present a direct evidence in *in vivo* model of tumorigenic growth regulated by CDK9, a core component of P-TEFb in combination with Yki. Our findings are also consistent with effect of CDK9 inhibitors reducing growth of patient derived xenografts (Allaway *et al.*, 2016) wherein it was observed that inhibition of CDK9 using two inhibitors dinaciclib and SNS-032, leads to reduced Pol II CTD phosphorylation, apoptosis and reduced growth of xenograft cells. In context of cancer, it also has to be appreciated that a very few of Hpo pathway components show somatic mutations however YAP activity has been clearly associated with poor prognosis. Thus, understanding the function of a complex that can regulate tumorigenic outcome of increased YAP activity will prove to be instrumental. The inhibitors mentioned above are in clinical trial. Considering evidence emerging from our work and Galli *et al.* CDK9 based therapeutics could be beneficial for patients with deregulated Hpo pathway. However, further analysis of CDK9 function in context of YAP in mammalian system would be crucial to determine how beneficial the therapeutic would prove to be.

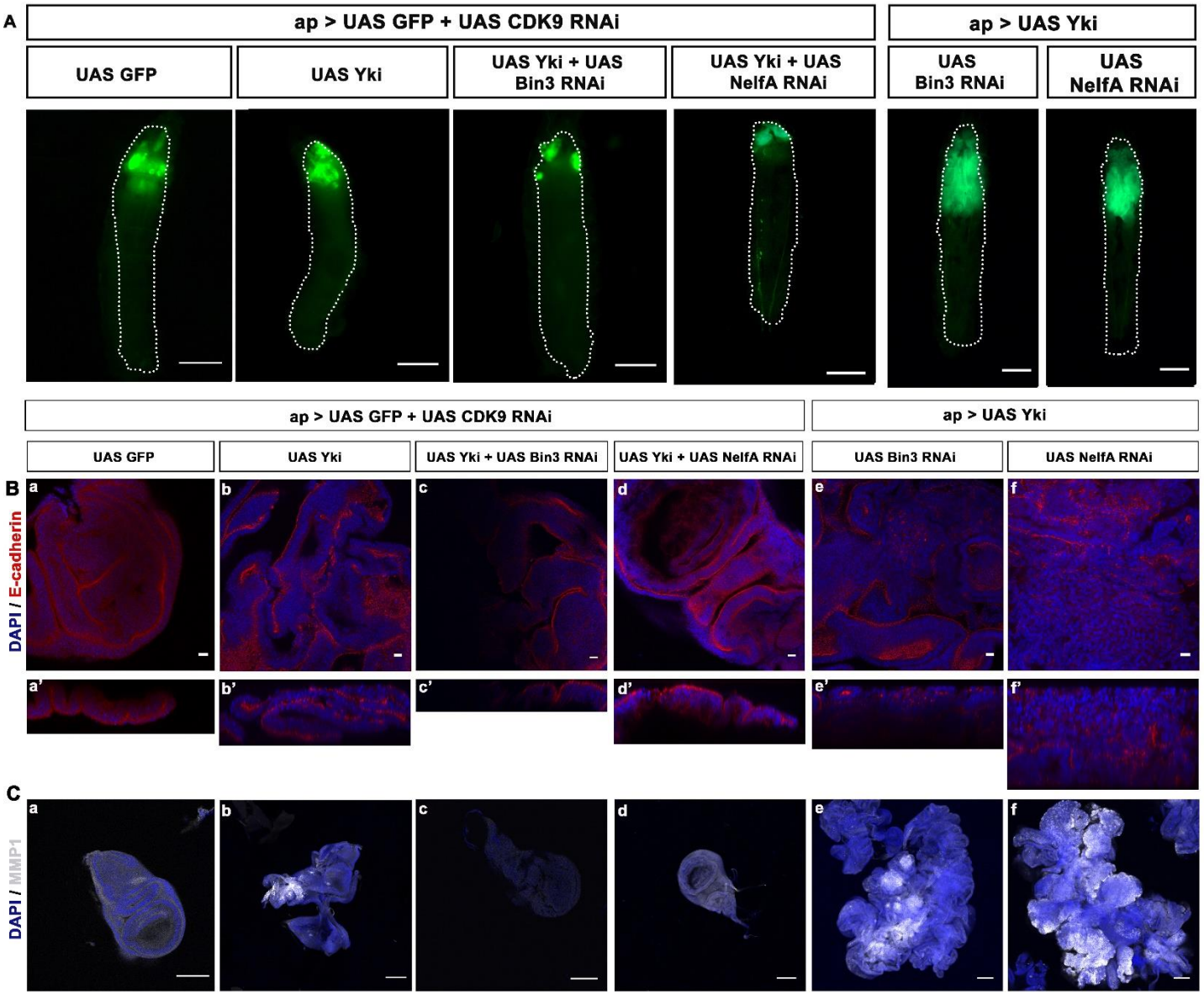


Figure 19.: CDK9 depletion prevents tumor formation in context of Yki. (A) Loss of CDK9 rescues tumor phenotype. The images show GFP-expressing wing discs of various genotypes as indicated. Size of the wing discs may be discerned by the amount of larval space occupied by GFP-expressing tissue. RNAi-mediated depletion of *cdk9* inhibited tumor formation caused by a combination of over-expression of Yki and depletion of a component of the PPP. The GFP-marked wing tissue is of the same size as in controls. All crosses were using GAL80^{TS}; *ap*-GAL4; UAS-GFP. (B) Restoration of apico-basal polarity in wing disc tissue. The images show wing discs of various genotypes as indicated stained for E-Cad (white). RNAi-mediated depletion of *cdk9* restored normal apical localization of E-cad in wing discs that over-express Yki and also depleted for a component of the PPP (a-d and a'-d') compared to tumor tissue (e-f and e'-f'). All discs are also stained with DAPI (blue) to visualize nuclei (scale bar = 10µm). (C) Restoration of MMP1 levels. The images show wing discs of various genotypes as indicated stained MMP1 (white). RNAi-mediated depletion of *cdk9* restored normal levels of MMP1 in wing discs that over-express Yki and also depleted for a component of the PPP (a-d) compared to tumor tissue (e-f). All discs are also stained with DAPI (blue) to visualize nuclei (scale bar = 100µm).

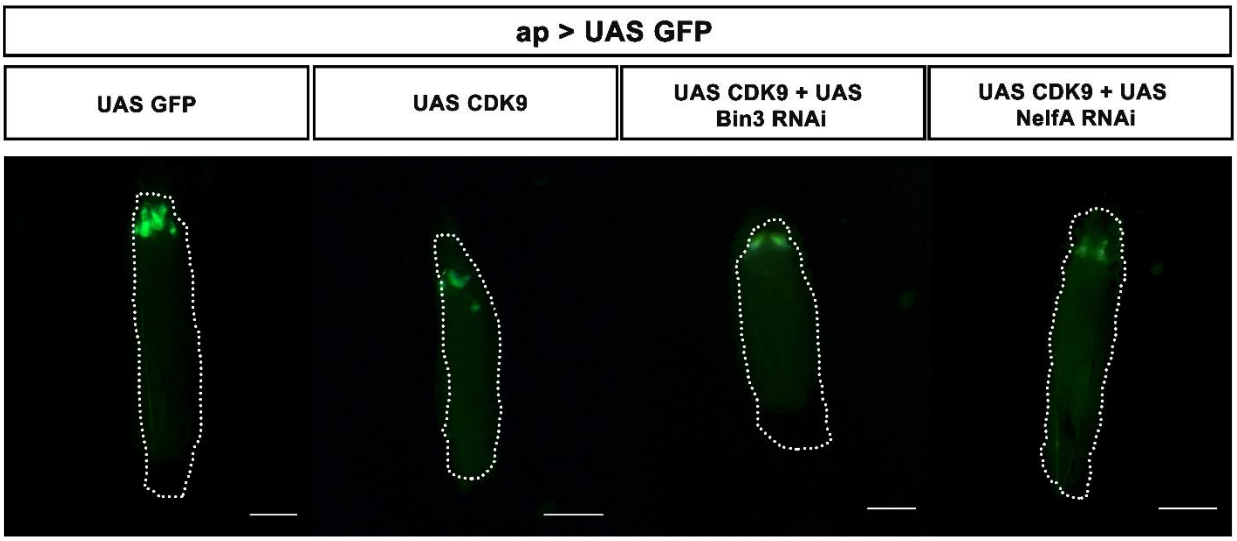


Figure 20.: Yki is the driver of growth. Larval images showing phenotype of UAS GFP in combination with (left to right) UAS-GFP, UAS-CDK9 followed by UAS-CDK9 and UAS-*bin3*^{RNAi}, UAS-CDK9 and UAS-*NelfA*^{RNAi}. None of them show over-growth phenotype as observed when Yki is over-expressed suggesting CDK9 may induce tumorous growth only in the context of over-expressed Yki. All crosses were using GAL80^{TS}; *ap*-GAL4; UAS-GFP.

The genetic experiments testing dependency of tumor on CDK9 showed that depletion of 7SK snRNP and NELF in combination with Yki does not lead to overgrowth in absence of CDK9. Additionally, neoplastic characteristics developed in the wing disc tissue by depletion of 7SK snRNP and NELF complex in combination with Yki are also restored to wild type epithelial tissue characteristics upon depletion of CDK9. These findings emphasize that 7SK snRNP depletion caused perturbation of PPP leading to tumorigenesis in context of Yki expression and not via a non-canonical mechanism. A rather interesting result was when tumors induced by depletion of NELF in combination with Yki showed dependency on CDK9. The NELF complex is bound to DSIF and Pol II in promoter proximal region blocking the active site of RNA Pol II and hindering its advance on the template DNA by restricting mobility of Pol II sub units (Vos *et al.*, 2018). Thus, in case of depletion of NELF complex it is possible that CDK9/P-TEFb action might not be necessary. However, our observations suggest that, CDK9 activity is important to have an effect on growth driven by Yki and neoplastic transformation by Yki which seems to be limited by promoter proximal pausing. It will be intriguing to examine the extent to which paused Pol II is released by loss of NELF and how the pause release is affected by presence or absence of CDK9.

Such experiments will give important insights into impact of NELF and CDK9 individually have on PPP and particularly in context of cancer. Other oncogenes such as c-Myc also trigger expression of target genes by recruiting P-TEFb (Rahl *et al.*, 2010). We also prompt an important aspect of contextual function P-TEFb and PPP have which needs to be investigated in detail for its potential in unraveling critical mechanisms that decide the tipping point which determines a normal vs disease state such as cancer.

Chapter 4: Transcriptome profiling of tumors induced by Yki and deregulated PPP

Promoter proximal pausing is widespread across genome. For eg. *Drosophila* cells showed greater than ~30% of genes with paused Pol II (Nechaev *et al.*, 2010; Core *et al.*, 2012), mammalian cells also exhibited Pol II pausing at similar extent of genes (Core and Lis, 2008). Deregulation of PPP results in release of Pol II from the pause site consequently altering the Pol II binding profile on DNA in *Drosophila,* mammalian cells and mouse embryonic stem cells (Guenther *et al.*, 2007; Muse *et al.*, 2007; Zeitlinger *et al.*, 2007). In case of increase in P-TEFb recruitment through TFs and signaling cues, highly expressed genes show that Pol II occupancy changes from promoter proximal regions to spread through the gene body (Core and Adelman, 2019). Additional evidence from global run on sequencing, a technique that identifies short reads from genes that exhibit paused Pol II, suggests that, PPP often fine tunes expression of genes rather than switching expression on or off. Similar inferences could be drawn from the studies that indicate PPP rather potentiates transcription and affects its fidelity, perhaps by regulating rates of transcription or by determining efficiency across cell and tissue types as it does in case of developing *Drosophila* embryos (Wang *et al.*, 2007, 2010; Li *et al.*, 2013; Saunders *et al.*, 2013).

Although it has been showed in multiple studies that Pol II is paused in promoter proximal regions and particularly at genes that are responsive to environmental and developmental cues, surprisingly disruption of PPP by depletion of NELF complex components has been shown to not change alteration of gene expression profiles substantially (Muse *et al.*, 2007). This is consistent with Pol II pausing being a point of regulation where dynamic stimuli can be integrated, irrespective of the model system (Guenther *et al.*, 2007; Wang *et al.*, 2007, 2010). However, these two points provide an interesting perspective, wherein broad scale stimuli are possibly an important backdrop against which regulatory fine-tuning role of PPP comes to fore. This implies that, in a specific context, deregulation of PPP can alter transcriptomic profile in a significant manner, determining the balance of a system toward regulated outcome or a deregulated diseased state. Evidence supporting this notion has been reported with targets of glucocorticoid receptor signaling (Luo *et al.*, 2013), with targets of ER signaling (Aiyar *et al.*, 2004).

The neoplastic growth in wing disc tissue we observed due to depletion of the 7SK snRNP complex or the NELF complex components or with expression of CDK9 was also dependent on specific context of Yki expression. We were curious about how PPP limits neoplastic potential of Yki. Transcriptional coactivator role played by Yki indicated possibility that PPP attenuates transcriptional output of Yki keeping the neoplastic potential in check. To explore this possibility, we decided to obtain transcriptome of tumors along with Yki expression alone and compare it with transcriptome of control genotypes- NELFA depletion and GFP expressing wing discs. We reasoned that, since Yki overexpression caused hyperplastic growth in wing disc tissue, NELF and 7SK snRNP complex limit the potential of Yki to cause neoplasia by regulating overall transcriptional output. Thus, examination of transcriptome of tumors in comparison with Yki overexpressing wing disc tissue would lead to insights about candidate pathways and genes that mediate neoplastic transformation driven by Yki.

RNA-seq methodology

To isolate total RNA, we induced expression of transgenes as per screening protocol discussed previously. Control discs expressing GFP and discs expressing NELFA RNAi were dissected along with tumor discs obtained by depletion of NELFA and overexpression of Yki and Yki overexpressing discs alone. We standardized total RNA yield from each of the sample set to determine the number of discs of each genotype to be used for RAN-seq. The numbers were, 150 wing discs each of *ap*-UAS GFP and *ap*-UAS NELFA RNAi while 50-70 wing discs were required of *ap*-UAS Yki and 20-25 tumor discs of *ap*-UAS NELFA RNAi, UAS Yki combination to yield >1µg of total RNA.

We collected each of the respective genotypes in triplicates and subjected them to quality checks followed by sequencing using Illumina platform. The sequencing results were processed using validated tools and pipeline as described in Pertea *et al* (Pertea *et al.*, 2016). Alignments and transcript levels obtained using aforementioned pipeline yielded the genes that were expressed and in all the respective genotypes. To identify the transcripts that are differentially expressed we used Bioconductor based packages and pipeline for count based comparison across genotypes (Anders and Huber, 2010; Anders *et al.*, 2013).

Results

For identifying differentially expressed genes we used *ap*-UAS GFP expressing WT wing discs as a baseline for genes expression. We compared each of the remaining three genetic conditions with it and obtained the genes that were upregulated or downregulated in each of the individual conditions.

Differentially expressed genes in *ap*-UAS NELFA RNAi

Depletion of NELFA has been reported to change Pol II occupancy from promoter proximal regions to spread across the gene body (Muse *et al.*, 2007). This is consistent with *in vitro* findings that disruption of one of the NELF complex components renders NELF complex non-functional (Yamaguchi *et al.*, 2002). Taking these findings in consideration we establishing the gene transcripts affected by depletion of NELFA are critical for delineating PPP regulated genes affect Yki driven oncogenesis.

Sequencing data showed that there were total of 894 transcripts that were significantly differentially expressed. Out of these, 483 were upregulated and 411 were downregulated in NELFA depleted wing disc tissue (Fig.21). The lesser number of genes that showed significant divergence from WT gene expression has been reported to be low in cases where PPP is perturbed alone (Muse *et al.*, 2007; Zeitlinger *et al.*, 2007). Our data is consistent with these findings. However, we must note here that since expression of UAS RNAi was driven using *apterous* driver, the knock down of NELFA is limited to the dorsal compartment of the wing disc tissue and does not cause any growth phenotype. This could lead to net underestimation of differentially expressed transcripts, as the transgene expressing tissue and wild type ventral compartment represented in almost equal proportions. Another important point to note here is that NELFA gene expression was reduced by greater than 2-fold in these wing disc tissue.

Differentially expressed genes in *ap*-UAS Yki

Yki overexpression led to change in expression of 1750 transcripts. Significantly upregulated genes were 965 genes while 785 genes were significantly downregulated (Fig.21). Out of the upregulated genes, 65 genes were common with genes upregulated in NELFA depleted wing discs (Fig.21A). On the other hand, 16 genes were common and downregulated between NELFA depleted wing discs and Yki overexpressing wing discs (Fig.21B).

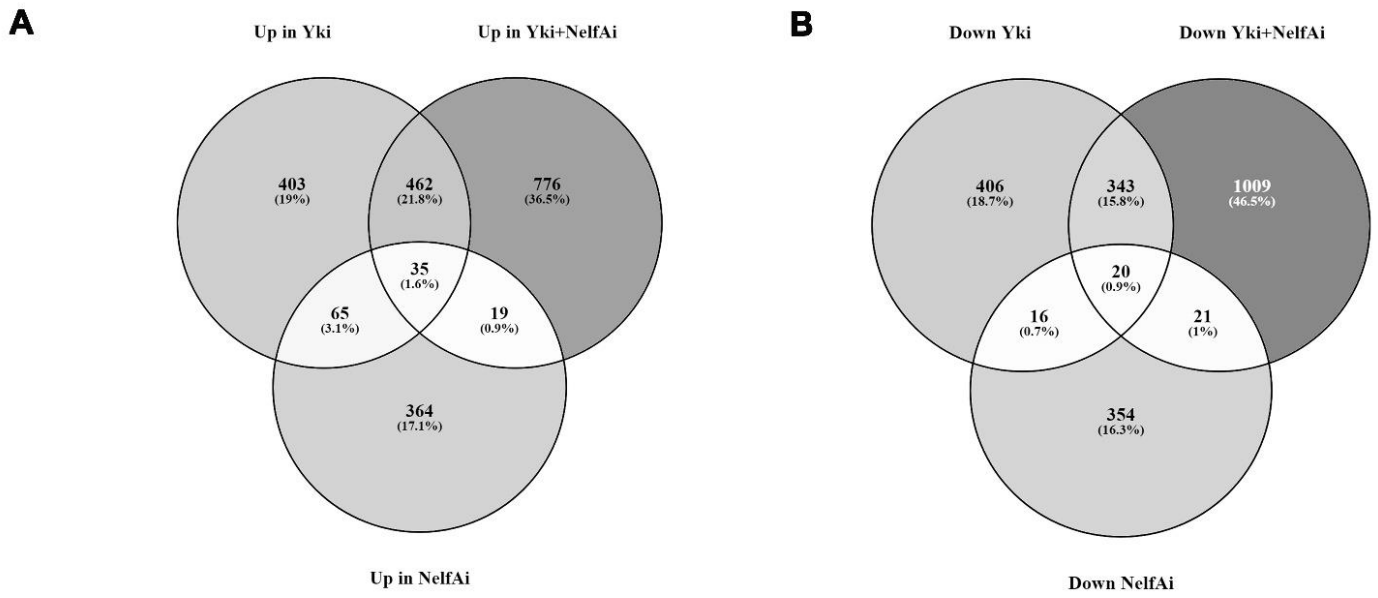


Figure 21.: Comparison of genes differentially expressed in wing disc tissue. (A) Venn diagram showing upregulated genes in Yki expression alone, Yki and NELFA RNAi combination, and NELFA RNAi alone. Number of transcripts common between different genotypes are depicted in respective mutual areas of the Venn diagram. (B) Venn diagram showing downregulated genes in Yki expression alone, Yki and NELFA RNAi combination, and NELFA RNAi alone. Number of transcripts common between different genotypes are depicted in respective mutual areas of the Venn diagram.

Differentially expressed genes in tumors

RNA-seq of tumor tissue compared to Yki expression and NEFA RNAi alone showed that 2685 transcripts were differentially expressed (Fig21). 1292 of these were upregulated and 1393 were downregulated in tumors. Among the uniquely deregulated transcripts in tumors, 776 were upregulated and 1009 were downregulated transcripts, while 462 upregulated and 343 downregulated genes were shared between Yki expression and Yki expression plus NELFA RNAi

and 35 upregulated, 20 downregulated were common between all three genotypes. NELFA RNAi alone and tumors shared 19 and 21 upregulated and downregulated genes respectively.

PPP has been showed to attenuate the expression levels of genes. Thus, the genes shared between Yki expression and Yki and NELFA RNAi combination were of particular interest, considering the possibility that PPP deregulation could enhance effect of Yki on target gene transcription. To verify this possibility, we assessed the transcripts common between Yki and Yki and NELFA RNAi combination. Out of 462 upregulated genes we observed that 150 genes showed further enhancement in expression in Yki expression and NELFA RNAi combination tumors. On the other hand, 160 of 343 commonly downregulated genes were expressed at significantly lower levels in tumors compared to Yki expression alone. These two sets of genes respectively taken together with uniquely up and downregulated genes in tumors constitute the cohort of genes at which the combination of Yki overexpression and deregulation of PPP by depletion of NELFA exert effect.

Analysis of differentially expressed genes

The next step post identifying the sets of genes that are differentially expressed, was to examine the enrichment of any particular sub set of genes. To this end we subjected individual gene sets to network based tool STRING, which categorizes protein coding genes based on Gene Ontology, KEGG and Reactome pathways as well as based on protein databses such as UniProt (Szklarczyk *et al.*, 2017).

STRING map of differentially expressed genes in NELFA RNAi alone

STRING analysis of uniquely differentially expressed genes when NELFA is depleted did not enrich for a specific category of GO terms or KEGG pathways. These observations are consistent with absence of growth phenotype in the wing disc tissue of NELFA RNAi.

STRING map of differentially expressed genes in Yki overexpression alone

Overexpression of Yki in wing disc tissue led to mild hyperplastic growth as showed previously. Since, Yki is a driver of tumorigenesis in our assay as evidenced by lack of any growth or tumor phenotype in deregulated pausing in absence of Yki, enrichment of physiological changes as reflected in transcriptome of Yki overexpressing wing discs is important. Analysis of the genes upregulated by Yki overexpression through STRING tool revealed an enriched hub of genes of

GO term for Biological process of translation and Reactome pathway of translation (Fig.23) with a very low false discovery rate. Additionally, protein components of Ribosome also formed a hub marking enrichment of these genes with a low false discovery rate (Fig.23). On the other hand, genes downregulated in Yki did not reveal a striking hub marking a specific biological activity among the enriched genes. However, it must be noted that, these were observations when all the genes upregulated or downregulated were considered for STRING analysis. When unique genes, up or downregulated in Yki overexpression were considered, the resulting STRING map showed significant PPI enrichment score but it was much higher than the PPI score in case of all genes. This will be an important consideration when appreciating differentially expressed genes in tumors.

STRING map of differentially expressed genes in tumors induced by combination of Yki and NELFA RNAi

STRING analysis of transcriptomes of NELFA depleted wing disc tissue and Yki overexpressing tissue established a platform to assess the physiological pathways and processes that are enriched and contributing to neoplastic transformation.

To do that we used two cohorts of genes, differentially expressed and unique to tumor tissue while, the other cohort was an extension of the previous with addition of common gene between Yki expressing tissue and tumor tissue, but selected for genes that showed significant up or downregulation over and above that observed in Yki alone. In case of STRING network of upregulated genes in tumor tissue, we observed that KEGG pathway for amino acyl t-RNA synthetases is enriched (Red colour in Fig.24) along with Ribosome biogenesis KEGG pathway. These terms were enriched in unique gene set as well as gene set that included additional genes that were common with Yki overexpression. Interestingly however, the FDR of amino acyl t-RNA synthetases KEGG pathway enrichment in cohort of genes in tumors and commonly upregulated in Yki was lower compared to unique and upregulated genes in tumors.

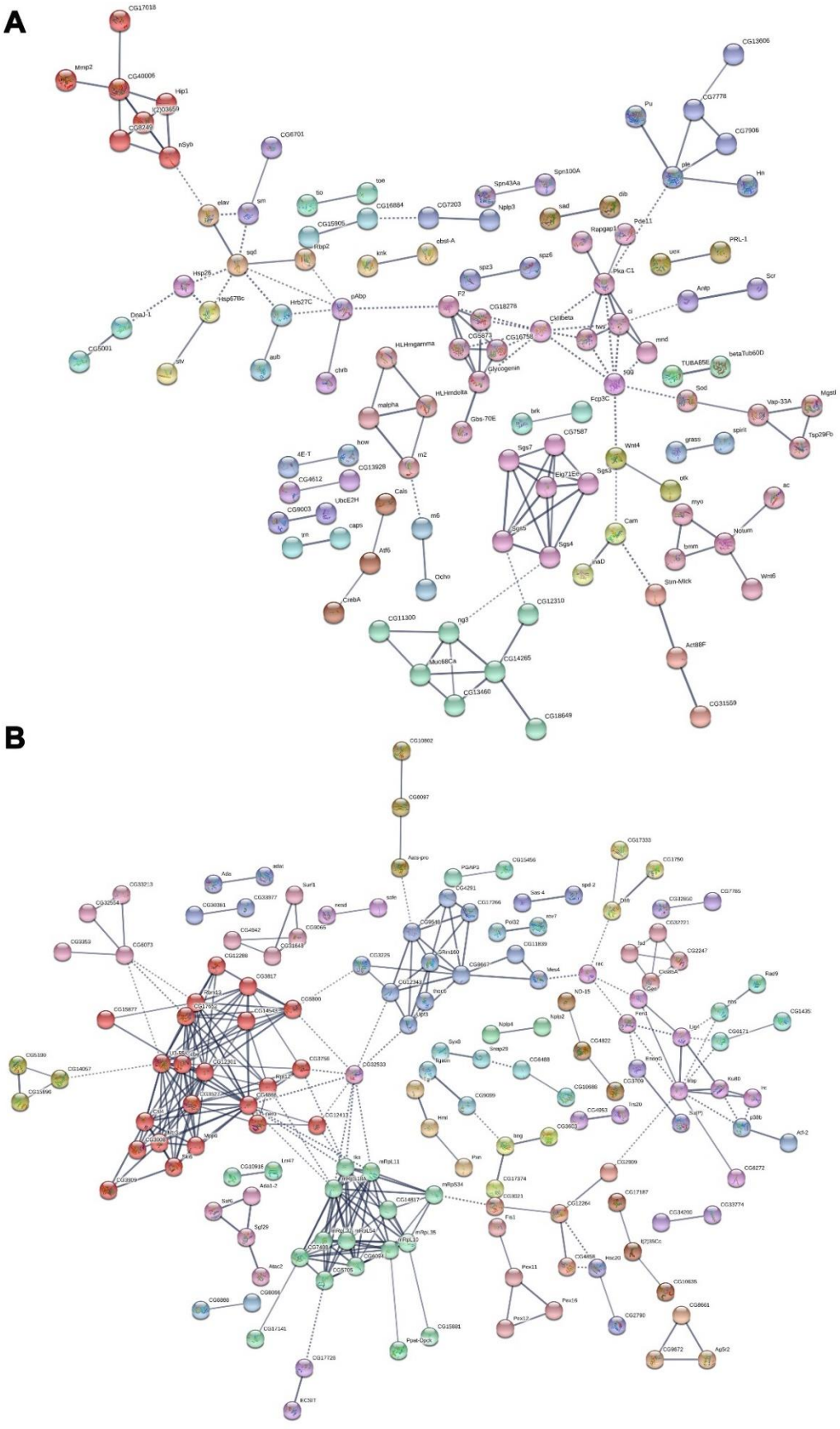


Figure 22.: STRING map of unique and differentially expressed genes in NELFA RNAi alone. (A) Represents the uniquely upregulated genes (PPI enrichment score $1e^{-16}$). (B) Represents the uniquely downregulated genes (PPI enrichment score $1e^{-16}$).

Lower FDR value indicates a much more reliable significant enrichment due to combination of NELFA RNAi with Yki overexpression. On the other hand, there is minimal difference in FDR of ribosome biogenesis KEGG pathway. STRING output of the aforementioned gene sets also showed enrichment of GO terms Pre-ribosome in Cellular components category and Translation in biological process category (Green and Yellow respectively in Fig.24). Interestingly, GO term Translation was also enriched in upregulated genes in Yki overexpressing genes alone as well (Yellow in Fig.23A and Yellow in Fig.24).

From perspective of FDR we observed that both upregulated in tumors gene sets show a huge difference in FDR of this GO term, such that FDR is much lower in case of upregulated genes in tumors compared to Yki overexpression alone. Considered together, enrichment of KEGG pathways and GO terms indicate increase in overall translation related processes selectively in tumors.

Observations of STRING map of downregulated and unique genes in tumors had enriched hub of KEGG pathway Proteasome and Protein processing in endoplasmic reticulum (ER) (Red and Blue in Fig.25A). These two KEGG pathways also constituted an enriched hub in gene cohort that included genes that were common with Yki but were significantly downregulated in tumors compared to Yki alone although with an FDR that was higher than unique downregulated genes (Red and Blue in Fig.25B). Of note was that the KEGG pathway for protein processing in ER was enriched with a much lower FDR when we analyzed all the genes that were downregulated in tumors without excluding genes that were common with Yki alone or NELFA RNAi alone, while Proteasome KEGG pathway was still enriched in this cohort with minimal difference in FDR compared to other two cohorts of downregulated genes mentioned earlier. These observations indicated that protein processing as a whole was highly deregulated in tumors. Consistent with enrichment of these enrichments, GO term for cellular component Proteasome complex was also enriched in downregulated genes in all the downregulated cohorts we analyzed using STRING (Fig.25).

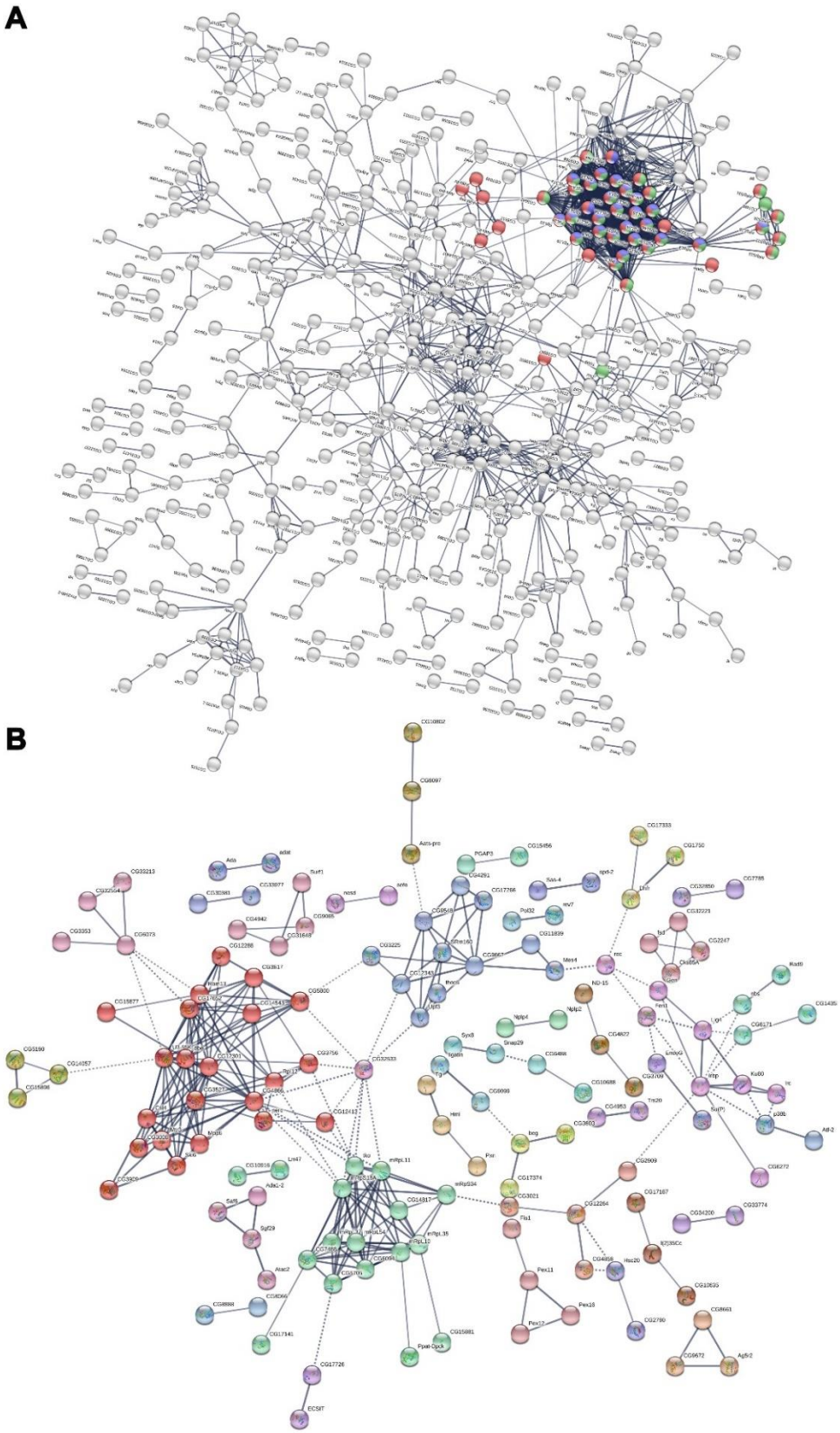


Figure 23.: STRING map of differentially expressed genes in Yki overexpressing wing disc tissue. (A) Shows the significantly upregulated genes with enriched hubs of GO term biological process- Translation (Red, FDR-0.00059), KEGG pathway- Ribosome (Blue, FDR-0.00033), Reactome pathway- Translation (Green, FDR-7.25X10^{-5}). (B) Significantly downregulated genes represented as STRING map.

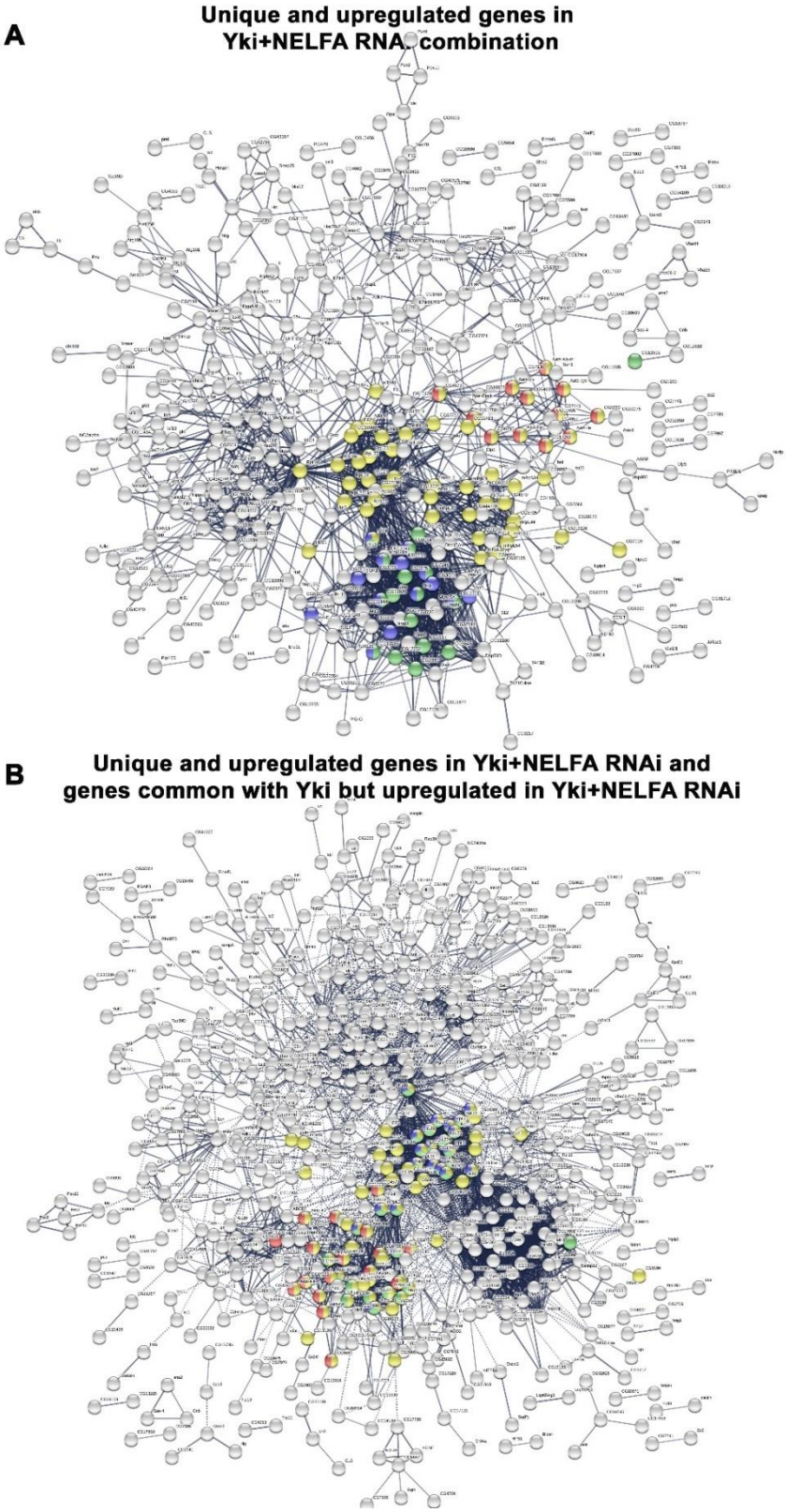


Figure 24.: STRING map of upregulated genes in Yki+NELFA RNAi wing disc tissue. (A) STRING significantly upregulated unique genes in enriched hubs for: Red- KEGG pathway Amino acyl t-RNA synthatases (FDR-0.00015), Blue- KEGG pathway Ribosome biogenesis (FDR-0.0197), Green- GO term cellular component pre-ribosome (FDR-$8.43X10^{-6}$), Yellow- GO term Biological process Translation (FDR-$1.33X10^{-10}$) (B) STRING map of unique upregulated and common and significantly upregulated genes with Yki alone: Colours represent the same terms as (A) Red (FDR-$1.69X10^{-5}$), Blue (FDR-0.0251), Green (FDR-0.0013),Yellow (FDR-$7.71X10^{-14}$).

In addition to enrichment of protein processing components, we observed an enriched hub of KEGG pathway for oxidative phosphorylation and cell component GO term of mitochondrial proton transporting ATP synthase complex (Yellow and Teal in Fig.25).

Table 2: Table enlisting upregulated genes in tumor tissue categorized under respective terms obtained from STRING analysis with corresponding fold change in gene expression levels.

List of genes whose expression is upregulated in the wing discs of *ap*-GAL4;UAS-Yki;UAS-*NelfA*^{RNAi}					
Aminoacyl-tRNA biosynthesis		**Ribosome**		**Ribosome biogenesis in eukaryotes**	
Gene name	**logFC**	**Gene name**	**logFC**	**Gene name**	**logFC**
Slimp	2.124626	RpL24-like	1.256082	Non1	2.292457
Aats-leu	1.413973	RpL5	1.202779	Ns2	0.900983
Aats-thr	0.912237	RpL15	1.130382	RIOK1	1.354352
Aats-cys	0.813931	mRpL28	0.969423	CG12301	0.997310
Aats-tyr-m	1.047768	mRpL9	0.799329	Bka	0.876333
Aats-pro	1.07342	RpS17	0.82042	eIF6	0.745744
Aats-ile	0.73741	mRpL35	0.997681	l(3)72Dn	0.800876
CG4573	1.138148	RpS23	0.782542	CG8064	0.778604
CG1750	1.487797	RpS4	0.775449	Nmd3	0.716537
CG6796	0.925494	RpL27A	0.716573	Mat89Ba	0.713426
CG7441	0.884721	RpL32	0.681588	CG11920	0.750235
CG17259	0.726080	RpL40	0.674508	CG3071	0.713535
Aats-trp	0.732224	RpS29	0.743389	CG33158	0.595732
Aats-asp	0.747097	RpL26	0.620538	CG13185	0.823244
Aats-gly	0.613889	mRpL10	0.692489	CG7246	0.798345
CG5463	1.037030	mRpL3	0.671734	CG8549	0.593618
Aats-ala-m	0.602770	RpL35	0.631348		
CG5660	0.663614	RpL27	0.594275		
		RpL28	0.629209		
		RpL21	0.600412		
		RpL22-like	1.081389		
		RpS3A	0.587288		
		RpL37A	0.3662		
		mRpL11	0.624297		

List of the genes involved in proteostasis among the upregulated and downregulated genes is provided in tables 2 and 3.

Table 3: Table enlisting downregulated genes in tumor tissue categorized under respective terms obtained from STRING analysis with corresponding fold change in gene expression levels

List of genes whose expression is down regulated in the wing discs of *ap*-GAL4;UAS-Yki;UAS-*NelfA*^{RNAi}			
Proteasome		**Protein processing in endoplasmic reticulum**	
Gene name	**logFC**	**Gene name**	**logFC**
Rpn7	-1.084698		
Rpn13	-0.949757	prtp	-1.46289
Rpn2	-0.900781	Sec61gamma	-1.65593
Prosalpha3	-0.910536	Sec61beta	-1.31964
Rpn3	-0.850895	CG5885	-1.28449
Rpn1	-0.879125	Sec61alpha	-1.25184
Pomp	-0.795409	TRAM	-1.3381
Prosalpha5	-0.834712	Pdi	-1.145
Prosbeta4	-0.76685	SsRbeta	-1.23571
Prosbeta7	-0.717883	Sec13	-1.01315
Prosbeta2	-0.706638	Sec63	-0.97489
Prosbeta5	-0.687177	CG14476	-1.03296
Prosalpha4	-0.694159	Sec24CD	-0.87522
Prosbeta6	-0.631819	Ostgamma	-0.97374
Rpn10	-0.592037	Ost48	-0.88256
Rpn12	-0.597364	CG4164	-1.21065
		ergic53	-0.86625
		Plap	-0.86843
		OstStt3	-0.90274
		Gp93	-0.89197
		l(1)G0320	-0.88951
		CG33303	-0.8065
		Hsc70-3	-0.9493
		CG5510	-0.81474
		p47	-0.78553
		Crc	-0.86323
		CG6453	-0.81869
		Sec23	-0.73903
		ERp60	-0.76882
		Der-1	-0.80252
		Csp	-0.64369
		CaBP1	-0.61193
		CG1597	-0.67306

Comparison between differentially expressed genes and Yki direct targets

Yki is a transcriptional co-activator and taking into consideration the specificity of Yki context observed for tumor suppressor function of the NELF and 7SK snRNP complexes, we were curious to know how many of these genes are direct targets of Yki.

To this end, we utilized previously published Yki ChIP-seq data from wing disc tissue (Oh *et al.*, 2013) to identify direct targets of Yki. Then we compared the direct targets of Yki with differentially expressed genes in tumors developed by combination of Yki and *NELFA* RNAi (Table 4). As the table shows a total of 69 upregulated genes were direct targets of Yki and 119 of downregulated genes were direct targets.

Table 4.: Comparison of Yki direct targets and differentially expressed genes

	Unique in Yki+NelfAi (Direct targets%)	Common with Yki_changed in same direction (Direct targets%)	Total direct targets
Upregulated genes	4.9% (38/776)	20% (31/155)	69
Downregulated genes	8.3% (84/1009)	21.9% (35/160)	119

Interestingly, the percentage of direct targets among the downregulated genes was much higher than the upregulated genes. Which was surprising considering Yki is a transcriptional activator. The list of genes, which are direct target and differentially expressed is shown in Table 5.

Discussion

The characterization of wing disc tissue showed that a combination of Yki overexpression and depletion of NELF or 7SK snRNP complex robustly leads to overgrowth accompanied with neoplastic transformation. The comparative analysis of transcriptome obtained from NELFA depletion, Yki overexpression and combination of thereof was performed to gain insight into expression changes that might be involved in driving neoplasia in the genetic the later combination. Primary results showed a much larger fraction of genes was deregulated in tumor tissue, which prompted a detailed analysis of the genes that showed such changes in expression in neoplastic tissue. Additionally, we observed that 150 of upregulated genes and 160 of the downregulated genes in tumors which were also common with genes up- and downregulated in Yki

overexpressing tissue, showed significant difference in up- and downregulation in tumors. This was an important finding that is consistent with attenuating role of PPP in regulation of transcription (Muse *et al.*, 2007; Core and Adelman, 2019). STRING tool allowed to visualize enrichment in differentially expressed genes as hubs in a network. Enrichment of Amino acyl tRNA synthetases (ARSs) among the upregulated genes was an intriguing find as it has been showed that some of the amino acyl tRNA synthetases are observed to be overexpressed in some types of cancers. One such example is elevated methionine tRNA synthetase activity in colon cancer (Kushner *et al.*, 1976). Classically ARSs have been deemed as housekeeping type of genes, however, there role in disease have come to fore. Furthermore, some of the ARSs have been shown to have non-canonical functions such as cytokine activity, promoting TF activity, promoting rRNA transcription (Razin *et al.*, 1999; Wakasugi and Schimmel, 1999; Ko *et al.*, 2000, 2001). From a clinical perspective, a systematic analysis of ARSs in cancer samples has showed a significant association with cancer associated genes and ARS expression deregulation in these samples. Many of these deregulations result in many fold increase in ARS expression in cancer tissue (Cheng and Deming, 2011; Kim, You and Hwang, 2011). Since these many ARSs are upregulated in uniquely in Yki+NELFA RNAi combination induced tumors, it indicates a possibility that ARSs could be a contributing factor to neoplastic transformation. In what capacity these contribute in process of transformation however, needs further exploration.

The next enriched KEGG pathway we observed was ribosome biogenesis. Increased transcription of rRNA by RNA Pol I, which causes increased ribosome biogenesis, has been linked to different types of cancers (Pelletier *et al.*, 2017) and additional evidence showed that increase in rRNA production by oncogenic drivers Ect1 and Netrin-1 confer malignant characteristics (Serafini *et al.*, 1996; Justilien *et al.*, 2017). RNA-seq data from our tumor takes into account transcripts made by Pol II however and the enrichment is in upregulated ribosomal protein genes. Interestingly, Myc, a well-known oncogenic driver has been shown to depend on increased expression of genes encoding ribosomal proteins. Loss of ribosomal proteins in Myc driven pre-cancer cells halts their growth and malignant transformation (Barna *et al.*, 2008).

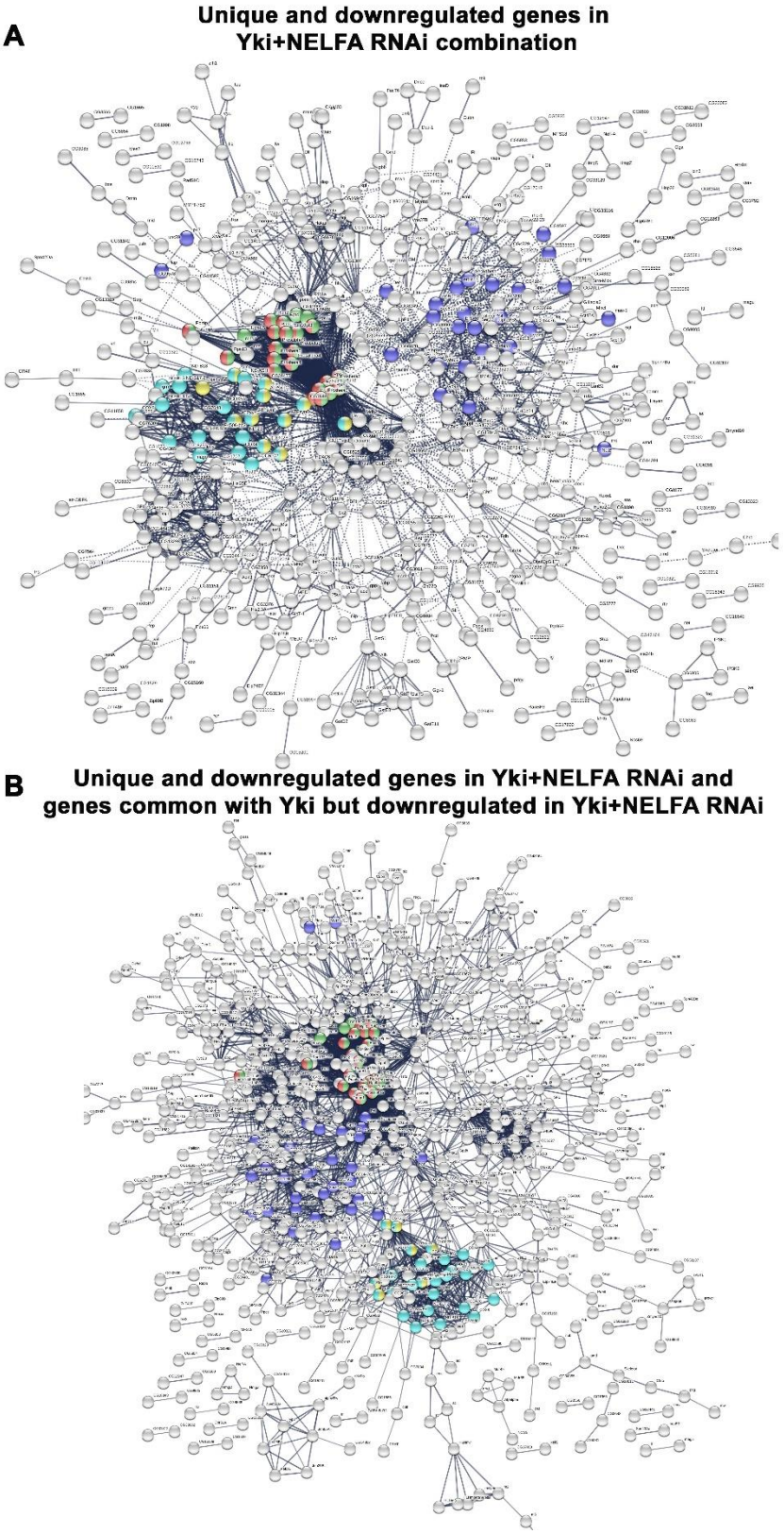


Figure 25.: STRING map of upregulated genes in Yki+NELFA RNAi wing disc tissue. (A) STRING significantly upregulated unique genes in enriched hubs for: Red- KEGG pathway Amino acyl t-RNA synthatases (FDR-0.00015), Blue- KEGG pathway Ribosome biogenesis (FDR-0.0197), Green- GO term cellular component pre-ribosome (FDR-$8.43X10^{-6}$), Yellow- GO term Biological process Translation (FDR-$1.33X10^{-10}$) (B) STRING map of unique upregulated and common and significantly upregulated genes with Yki alone: Colours represent the same terms as (A) Red (FDR-$1.69X10^{-5}$), Blue (FDR-0.0251), Green (FDR-0.0013),Yellow (FDR-$7.71X10^{-14}$).

Our observations, are following a similar trend where we see that genes encoding ribosomal proteins are significantly enriched in the upregulated genes that are unique to tumors. Taken together, the results from our study and evidence reported in literature implicate increased ribosomal biogenesis is a potential mechanism via which, Yki and NELFA depletion bring about neoplastic transformation. This is the first time that an epithelial *in vivo* model has implicated such a mechanism for neoplastic transformation and also a first study where ribosomal biogenesis is linked to Hpo pathway driven tumorigenesis. It will be interesting to test, if loss of critical proteins in the ribosome biogenesis pathway affect tumor growth in our model of tumorigenesis. Additionally, a high degree of conservation in ribosome biogenesis pathway make it a possible therapeutic target, as variety of drugs that target ribosome biogenesis have been shown to efficiently kill cancer cells selectively (Drygin *et al.*, 2009; Bywater *et al.*, 2012; Bruno *et al.*, 2017). Another biological process which was enriched in tumors was translation, which continues to highlight that overall protein production has been increased in the cells of the tumor tissue, especially when considered with enriched KEGG pathways of Ribosome biogenesis and amino acyl tRNA synthetases. Overall increase in translational output has been known to be a feature of cancer cells (Wu *et al.,* 2003; Truitt and Ruggero, 2016), however, such increase has been shown to be driven largely by deregulating rRNA translation by oncogenic drivers Myc, ERK pathway mTOR pathway or by loss of tumor suppressors p53, Rb (White *et al.*, 1996; Cairns and White, 1998; Kim *et al.*, 2000; Hannan *et al.*, 2003; Arabi *et al.*, 2005; Grandori *et al.*, 2005; Stumpf and Ruggero, 2011). Evidence from our study uniquely indicated that translational regulation is brought about via a combination of oncogene Yki and novel tumor suppressor NELFA, and is potentially involved in tipping the growth scales towards a neoplasia in wing epithelial tissue. These observations although demand a detailed investigation, the potential of such tumorigenic mechanism downstream of conserved Hpo pathway is remarkable.

In the cohort of downregulated genes, proteasomal machinery encoding genes and protein processing in ER were enriched indicating downregulation of proteasome dependent processes and rise in deregulated protein homeostasis. Taken together with increased translation as showed by upregulated genes, the scenario indicates a proteostatic stress in tumor cells. Cancer cells have been showed to have proteostatic stress, which leads to elicitation of unfolded protein stress response, which in turn invokes cell death (Buszczak, Signer and Morrison, 2014). However, in case of the tumors in this study, Yki overexpression strongly inhibits cell death (Pan, 2010). Thus,

it will be interesting to examine further to understand how these tumor cells deal with proteostatic stress. It has been reported that, cancer cells activate factors such as HSF1 to guard against the detrimental consequences of proteostatic stress (Dai and Sampson, 2016). Additionally, the components of proteosomal machinery are also critical for signaling pathways such as mTOR and hedgehog signaling (Liu, 2019). Thus, downregulation of proteasomal machinery could also result in increased signaling via these pathways which inturn augments growth caused by Yki overexpression.

Interestingly, comparison of Yki target genes identified by Irvine's group (Oh *et al.*, 2013) with differentially expressed genes in tumors indicates a certain degree of overlap. It is tempting to speculate that these genes are possible candidates that could be driving larger proportion of neoplastic transformation as they are direct targets and are perturbed in neoplastic tumors. However, a detailed molecular analysis of paused Pol II at this candidate gene loci will be required to definitively answer this question. Intriguingly, not many genes from the proteosomal machinery fall in the category of direct target genes of Yki. Thus, it remains to be determined whether the perturbed proteostasis is a direct effect of Yki or an indirect effect. However, it is important ot note here that, Yki and its mammalian counterparts are known to function via enhancers to a large extent and thus it is not possible to rule out proteostatic machinery being directly downstream of Yki and PPP based regulation.

Table 5.: List of direct targets of Yki whose transcripts are up/down-regulated in the wing discs of *ap*-GAL4/UAS-*NELFA* RNAi; UAS-Yki

Upregulated in *ap*-GAL4/UAS-*NelfA*^{RNAi}; UAS-Yki		**Down-regulated in *ap*-GAL4/UAS- *NelfA*^{RNAi}; UAS-Yki**			
aru	***fru***	*Actn*	*CG6770*	*kek5*	*simj*
Bsg	*ft*	*Akap200*	*CG7065*	***klu***	*siz*
cbt	***ftz-f1***	***alt***	*CG7272*	***knrl***	*Sk2*
CG10075	***Gclc***	*Amun*	*CG7914*	*ko*	***smt3***
CG10462	***GlyP***	***Argk***	*CG8243*	*l(2)03659*	*sn*
CG10628	*GlyS*	*Atpalpha*	***CG8498***	***lama***	*sns*
CG10914	*l(3)02640*	***bchs***	***CG9331***	*LanB1*	*SppL*
CG11360	***Lasp***	*bowl*	***CG9650***	***LpR2***	***svp***
CG11658	***MFS17***	*brat*	*CG9663*	*Lrt*	*tio*
CG12065	*mTerf3*	***caup***	***chic***	***Lsp1alpha***	***tsh***
CG13185	*mthl1*	*CG10237*	***chm***	*mam*	*tup*
CG13398	*Mys45A*	***CG10960***	***CrebA***	*Mhcl*	*tutl*
CG13893	***neb***	*CG11050*	*cv-d*	*modSP*	*vri*
CG14322	*Nrg*	*CG12769*	*CycE*	*msi*	***wit***
CG1648	***nrv2***	***CG13875***	*Dh31-R*	*mtd*	
CG2247	***NTPase***	***CG15628***	*DOR*	*mth*	
CG2540	*Phb2*	*CG17684*	*dpr16*	*Mvl*	
CG2875	*Ptp10D*	*CG31475*	***drm***	*MYPT-75D*	
CG30069	*Ptp61F*	***CG32447***	***E(spl)m2-BFM***	*nrv1*	
CG31635	*qm*	*CG3270*	*Eip74EF*	*nub*	
CG32095	*rau*	*CG33129*	*Eip75B*	***opa***	
CG32365	*S6k*	*CG33229*	*emc*	*path*	
CG32369	***SNF4Agamma***	*CG3529*	***fax***	*Pep*	
CG33158	*Socs36E*	*CG4020*	*Fmr1*	*pk*	
CG3838	***spi***	*CG42272*	*Gale*	*Prosbeta7*	
CG5059	***Spt***	***CG42340***	*Gpdh*	*Rab5*	
CG6175	*Su(Tpl)*	***CG4374***	*grp*	***Reph***	
CG7841	*Tgt*	*CG4562*	*Gug*	*RnrS*	
CG8360	***Traf4***	*CG4615*	***hbs***	*Rtnl1*	
CG9932	*Tsp39D*	*CG5001*	*Hnf4*	*rump*	
dbe	***uif***	*CG5756*	***Hr39***	***Sb***	
dpp	*Vha26*	***CG5758***	*Hs6st*	*sbb*	
ex	***wg***	***CG5885***	*Hsp26*	*SC35*	
fng	***zormin***	*CG6163*	***Hsp83***	*Sec63*	
for		***CG6287***	*kek1*	*side*	

The genes shown in bold letters are those direct targets of Yki that are upregulated in both non-tumorous *ap*>UAS-Yki discs and tumorous *ap*>UAS-Yki; UAS- *NELFA* RNAi discs, but degree of enhancement was higher in tumorous tissue.

Summary and Future perspective

Summary

Many aspects of growth control during development of an organism have commonly been observed to be disrupted in case of diseases, more evidently so in case of cancer. Such parallels between development and disease also prompt animal models for exploring disease progression and tumorigenesis is no exception. *Drosophila* is a known model in which elucidation of myriad of developmental pathways has occurred and over the past one-two decades it has also emerged as a model of testing tumorigenesis *in vivo*, particularly in deciphering complex network of various signaling pathways regulating growth. Classical genetics and more modern molecular biology methods depend on alternative versions (mutations) of a gene to decipher its normal function. Cancer too has provided enormous opportunity to better understand normal developmental and growth control pathways, which in turn provide cues for controlling the disease.

Our work was based on an *in vivo* model in *Drosophila* for tumorigenesis in epithelial tissue of developing wing imaginal disc tissue. We utilized visual screening to identify growth regulator in wing imaginal disc tissue in specific context of growth promoting EGFR and Yki expression. This screen identified several genes, depletion of which led to massive overgrowth in wing imaginal disc tissue in background of EGFR and Yki overexpression. These included a few of the known tumor suppressor genes such as core kinases of Hpo pathway- *hpo, wts* in addition to several novel genes with previously unknown growth regulatory function. These signify a context dependent role and potential of these candidates for studying growth regulation in further detail. Additionally, we also performed a screen wherein, a known neoplasia causing genetic background of EGFR+*SOCS36E*^{RNAi} (SOCS screen) was combined with depletion of genes with goal of identifying genes that can rescue the neoplastic transformation. These novel genes have the potential to be prognostic markers and/or therapeutic targets for tumors depending on EGFR and SOCS36E.

Hpo pathway effector Yki is a growth regulator that promotes cell growth and survival. Amongst the tumor suppressor genes identified in the background of Yki, were genes that encoded two complexes- the 7SK snRNP complex and the NELF complex. Interestingly, these two complexes regulate elongation of transcription via a process called as promoter proximal pausing (PPP). Yki being a transcriptional coactivator, identification of PPP as a process that limits oncogenic potential was an interesting finding. Further characterization of tumors obtained by combination

of depletion of individual 7SK snRNP or NELF complex components with Yki overexpression, showed that these are neoplastic in nature. These observations indicated that deregulation of PPP which is established and maintained through the NELF and 7SK snRNP complex unleashed the neoplastic output by Yki in wing imaginal disc tissue. In support of the PPP being critical for regulating Yki driven tumorigenesis, we found that CDK9, the kinase component of pause releasing P-TEFb complex, cooperated with Yki. We observed that coexpression of Yki and CDK9 led to massive overgrowth accompanied by development of neoplastic characteristics in wing imaginal disc tissue. On the other hand, co-depletion of CDK9 and individual components of the 7SK snRNP complex in Yki expression background repressed tumorigenesis observed in combination of 7SK snRNP depletion and Yki. Consistent with canonical function of 7SK snRNP, necessity of CDK9 shows tumor suppressor function of 7SK snRNP through PPP. Genetic experiments of co-depletion of CDK9 and NELF complex components in combination with Yki surprisingly showed repression of tumor phenotype in wing imaginal disc tissue. In this case, the necessity of CDK9 is indicative of basal function of CDK9 required for release of pause in spite of absence/depleted NELF complex.

Next step to understand the process of neoplastic transformation, transcriptome analysis was performed. These data presented with a novel link between proteostasis in tumor cells and higher activity of Yki as genes involved in protein synthesis were highly enriched among upregulated genes in tumor tissue while genes encoding proteasomal machinery were enriched in downregulated gene set. We propose that, YKi driven tumorigenesis occurs via deregulated protein synthesis in tumor tissue.

To examine relevance of our findings from mammalian context, we tested expression of YAP target genes in 293T cells and found that depletion of NELF complex enhanced effects of YAP overexpression on a sub set of YAP target genes. Additionally, we also examined TCGA patient data and found that in case of Breast, Bladder and Pancreatic cancer, patients with lower expression of components of the NELF and 7SK snRNP complexes survived poorly over a period of 5 years compared to those with higher expression. These results from mammalian systems of study are suggestive of tumor suppressor function of the NELF complex and the 7SK snRNP complex. Interestingly, when we categorized patients in aforementioned three cancer types in a way that combined higher YAP expression and lower expression of individual components of the NELF and 7Sk snRNP complex, poorer survival of such patient cohort was much more significant in some of the combinations.

Taken together this study presents the evidence of novel function of two complexes- NELF and 7Sk snRNP as tumor suppressors via regulation of a critical rate limiting step of promoter proximal pausing in specific context of Yki and potentially in context of YAP as well. From the context of Yki or YAP these findings have a unique implication of probing in depth the transcriptional program downstream of Yki or YAP that contributes to process of tumorigenesis.

Future perspective

PPP in growth control is an under investigated phenomenon with limited studies highlighting role of PPP in synchronous expression of patterning genes in *Drosophila* embryos (Lagha *et al.*, 2013), while very less propensity of PPP in embryonic stem cell over differentiated cells (Min *et al.*, 2011). Our work brings a novel perspective to regulation of growth and development by considering a potent regulator of organ and tissue growth together with PPP regulators. In-depth testing of how PPP specifically regulates Yki driven growth will provide great insights into importance of PPP from developmental as well as disease point of view. Taking a candidate gene approach using RNA-seq data we have provided in this study can serve as foundation for future studies. Additionally, it has been showed that super elongation complex, and BRD4 recruit P-TEFb depending on cellular context and gene promoter (Core and Adelman, 2019). Interestingly, components of SEC and BRD4 are negative in our screens, which is consistent with their known function, making these interesting candidates to test in context of Yki for their necessity for recruiting P-TEFb. These genetic experiments involving overexpression with Yki and depletion in Yki+CDK9 expression combination, will give an essential insight into molecular mechanism by which PPP is specific for Yki expression background. Since PPP provides an attenuator type of regulation via multinodal regulation depending on the context, addition of upstream mediators of context specificity is important to expand understanding of regulatory landscape over which PPP is influential.

Depletion of NELF complex using RNA interference has been shown to alter RNA pol II occupancy across the genome in *Drosophila* (Muse *et al.*, 2007), similarly occupancy of Pol II at a sub set of YAP target genes has been shown to change in CDK9 dependent manner in mammalian cells (Galli *et al.*, 2015). Wing disc model reported here suffers from a limitation that tissue with RNAi alone and Yki overexpression alone retain a significant part of ventral compartment which does not express the transgene. This renders the tissue not very suitable for probing Pol II occupancy in genome wide manner. Analysis of Pol II occupancy in response to Yki expression alone or NELF RNAi alone compared to Yki+NELFA RNAi will be crucial to pin point the direct target genes as well as regulatory regions susceptible to Yki expression in absence of PPP. This in

turn will help in shedding light upon regulatory networks at play and in combination with the RNA-seq data we have we will have much better understanding of the transcriptional unit involved in neoplastic transformation.

Alternatively, similar analysis can be performed in mammalian cells using siRNA and retroviral transductions. Since we have already reported that depletion of NELFA in 293T cells leads to changes in YAP target gene expression, a global genome wide analysis in mammalian system will be insightful especially in terms of elucidating how mammalian cells respond to these genetic alterations and perhaps will be helpful when extrapolating this mechanism to patient samples.

Analysis of survival data in TCGA patient cohort is particularly interesting because we see a direct correlation of patient survival in BRCA, PAAD and BLCA. Examination of mechanistic parallels of what we have observed in *Drosophila* using bioinformatics approach as a primary method followed by verification of these in tissue sample cohorts will be useful to advancing these findings in a clinically beneficial direction. For that we can primarily look at proteostasis related genes using orthologues of deregulated genes we have identified in *Drosophila* tumor model. In case of BRCA an additional analysis based on sub types of BRCA and correlation of expression of NELF complex, 7SK snRNP complex and YAP with survival of patients of each individual PAM50 sub types will be crucial for clinical approach.

Materials and Methods

Drosophila strains

Following *Drosophila* strains are used in this study. *ap-Gal4* (Cohen *et al.*, 1992), *UAS-Yki* (Huang *et al.*, 2005b). Following RNAi stocks were obtained from Vienna *Drosophila* RNAi Center and Bloomington *Drosophila* stock Center: *UAS-NelfA*^{RNAi} (KK106245, TRiP #32897), *UAS-NelfB*^{RNAi} (KK108441, TRiP #42547), *UAS-NelfE*^{RNAi} (TRiP # 32835), *UAS-NelfD*^{RNAi} (KK100009, TRiP # 38934, #42931), *UAS-bin3*^{RNAi} (KK101090, TRiP #41527), *UAS-Hexim*^{RNAi} (KK100500, #32898). *UAS-CDK9* was obtained from FlyORF (#F001571). KK RNAi stocks screened are listed elsewhere.

Cross schedule

Mon	Tues	Wed	Thurs	Fri
Collect virgins	Collect virgins	Collect virgins	Collect virgins	Set up cross → VIAL-1. Keep cross at 18°C.
	Flip cross VIAL-1 → VIAL-2. Keep VIAL-2 at 18°C Discard VIAL-1.			Flip cross VIAL-2 → VIAL-3. Keep VIAL-2 & VIAL-3 at 18°C.
Put VIAL-2 at 29°C Discard flies VIAL-3 and keep vial at 18°C		Put VIAL-3 at 29°C		
	Look for the presence of giant larvae in VIAL-3 (day 8). Take pics of the positive ones (GFP-microscope).		Look for the presence of giant larvae in VIAL-3 (day 8). Take pics of the positive ones (GFP-microscope).	

Whole Larval imaging

Larval images were taken in bright field and in GFP channel with Leica stereomicroscope. Image processing was done using Adobe Photoshop 6 and ImegeJ.

Antibodies used

DE-Cadherin (DCAD2 from DSHB) at 1:100, MMP1 (DSHB 3A6B4, 3B8D12 and 5H7B11) each antibody at 1:50, Rhodamine Phalloidin (Invitrogen) at 1:200.

Immunostaining Protocol

1. Fix the tissue using formaldehyde (4% in PBS). Incubate at RT on the rotator for 20 minutes. Alternatively, 100µl formalin can be added to 900µl PBS instead of 4% PFA.
2. Remove the solution and wash with 500µl PBS + TritonX100 (0.2%) (PBST) and at RT on rotator for 10 minutes.
3. Repeat the washes two more times.
4. Remove the wash solution and add blocking solution (0.2% PBST + o.5%BSA) and incubate the samples at RT on rotator for 2 hours.
5. Remove blocking solution and add primary antibody (diluted to desired concentration in blocking solution) and incubate overnight at 4°C on rotator.
6. Remove the primary antibody (use the aliquot two more times for immunostaining) and wash with PBST at RT for 10 minutes on rotator.
7. Repeat the washes two more times.
8. After washes, add secondary antibody (diluted to desired concentration in blocking solution) and incubate at RT for 2 hours at room temperature. Cover the tubes with aluminum foil before addition of secondary antibody or incubate in a dark place
9. Remove secondary antibody and wash with PBST for 10 minutes at RT.
10. Repeat the washes two more times.
11. After washes, remove PBST and add mounting media (along with DAPI) and store at 4°C in a dark place.
12. The discs were then mounted on slides in the same mounting medium. While mounting tumorous wing disc tissue, use larval carcass as spacers between coverslip and slide. This is a critical step especially for confocal imaging of epithelial polarity markers in order to avoid squashing the tissue which causes ambiguous perception of tissue depth and polarity.

Quantitation

Quantitation was done using ImageJ. For area measurements GFP channel was used and custome selection was done for whole area that shows GFP expression. MMP1 mean intensity measurement was done across the whole wing disc. Custom selection tool in ImageJ was used to select whole wing disc and mean intensity across wing disc was measure. Normalization to calculate fold change in area and mean intensity was done with respect to area of WT GFP expressing wing disc

and MMP1 intensity in WT GFP expressing wing disc, respectively. Statistical analysis was performed in Prism GraphPad (Version 5.0) using One-way ANNOVA.

Wing disc collection for RNA-seq

1. Larvae were taken out from vials in PBS made in RNase free water (GIBCO) and cleaned and water replaced with fresh clean PBS. GFP positive larvae were selected from these under epifluorescence microscope.
2. Dissection of wing discs was carried out in chilled PBS made in RNase free water using cleaned forceps and insulin needles. Followed by cleaning of imaginal disc tissue to avoid picking up minimal possible fat and tracheal tissue.
3. The wing imaginal tissue was transferred to TRIzol reagent (Invitrogen) using a fresh dropper. The dissections were carried out such that, *ap-GAL4* GFP, *ap-GAL4* NELFA RNAi would pool 150 wing discs each, *ap-GAL4* UAS Yki would pool 50-60 wing discs and *ap-GAL4* UAS Yki, UAS NELFA RNAi would pool 20-25 wing discs. Each genotype was collected in triplicates.
4. RNA-seq of these samples was carried out on Illumina platform.

RNA-seq data analysis

RNA-seq data was subjected to pipeline as described by Pertea and colleagues (Pertea *et al.*, 2016). This pipeline utilized HISAT, StringTie and Ballgown tools on Bioconductor on Linux operating system. In combination of HISAT and StringTie we utilized DE-seq package to obtain count data from aligned sequences (Anders *et al.*, 2013). These data were Log transformed and we considered fold change of >0.58 as significant which roughly means approximately 1.5-fold change in expression. Using these criteria, we obtained differentially expressed genes in UAS Yki, UAS NELFA RNAi and UAS Yki+UAS NELFA RNAi in comparison to genes expressed in UAS GFP. Additionally, we also compared expression of differentially expressed genes between UAS Yki and UAS Yki+UAS NELFA RNAi using same criterion.

Venn diagram was obtained using Venny tool. This also was used to obtained unique and common genes between different genotypes.

STRING analysis

STRING tool can be accessed at https://string-db.org/. List of genes in each of the depicted genotypes were uploaded as files in 'multiple proteins' tab on STRING page. Next it displays list of valid genes. Since it considers protein encoding genes, it returned most of the long non coding

RNA transcripts from differentially expressed genes as invalid entries for further analysis and our analysis of differentially expressed genes thus excludes such transcripts.
For stringent analysis, we apply high confidence (0.007) and MCL clustering as criterion for STRING network output.

Survival analysis in patient samples

UCSC Xena browser was accessed at https://xenabrowser.net/. We followed protocol for survival analysis as follows.

1. At UCSC Xena interface select data sets from GDC-TCGA studies.
2. From these studies select individual cancer study. For eg, Breast cancer. Then go to visualize data.
3. On the next screen selected the gene of interest and the type of data (phenotype and genetic). We selected genetic for the analysis presented here. This showed the data of gene expression for the selected gene across the cancer patient samples.
4. From here select Kaplan-Meyer (KM) plot and in KM plot use X-axis upto 1800 days (approximately 5 years) and also quartile-based division of data which compared the top and the lower quartile samples.
5. For combining expression level filters of two genes, we used Filter feature of the browser. The filters were applied based on lower quartile value of TSGs identified in our study while for YAP we used the upper quartile as filtering criterion. We used 'AND' command here to combine the two filters and created a new column that divided the whole dataset in 'TRUE' for the condition and 'FALSE' for the condition. This was then compared using KM plot as described in earlier step.

Transfection of mammalian cells

1. Split the cells day before or in morning if they can attach fast
2. In afternoon: take 2xHBS buffer from fridge and let warm to RT (pH 7.05, filter sterilized through 0.22 um filter)
3. Take CaCl2 solution from fridge (DO NOT FREEZE) (0.25M in MQ, filter sterilized through 0.22um filter)
4. For a 10 cm dish: Put DNA (15-20 ug total) in a 15 ml tube
5. add 500 ul CaCl2 solution (do six tubes or less at once)
6. place tube on vortex, add 600 ul of 2XHBS while vortexing. Let it stand for a few minutes (the time that it takes to get and label the dishes with cells)

7. Add the mix to the cells, pipet it into the medium
8. For a well in 6 well plate: put DNA (3 ug total) in a 2 ml tube
9. Add 100 ul CaCl2 solution.
10. Place tube on vortex, add 110 ul of 2XHBS while vortexing. Let stand for a few minutes time to label the dishes with cells)
11. Add the mix to the cells, pipet/pour it into the medium.
12. Incubate overnight (at 37^{o}C).
13. Wash cells next day with PBS + 4mM EDTA (for a few minutes, so CaPO4 can dissolve), then put on fresh medium

2X HBS

HEPES for Calcium Phosphate Coprecipitation Transfection

* 1. Make a stock solution of Na2HPO4 dibasic (5.25 g in 500 ml of water)

* 2. Make 2 x HBS: 8.0 g NaCl 6.5 g HEPES (sodium salt) 10 ml Na2HPO4 stock solution for 500 ml.

* 3. pH to 7.0 using NaOH or HCl. Bring volume up to 500 mls. Check pH again. The pH is very important, it must be exactly 7.05

Constructs and siRNA

pMSCV (Blasticidin resistance) this the empty vector in which remaining genes used in cell culture assays were cloned, Flag-YAP S127A/S397A cloned in pMSCV (Blast) (Nguyen *et al.*, 2014). siRNAs for NELFA, HEXIM1, HEXIM2 and MEPCE were obtained from siRNA SMARTpool, Sigma-Aldrich.

RNA extraction using TRIzol reagent

1. Collected cells or wing disc tissue in TRIzol (1ml).
2. Wing disc tissue was crushed in TRIzol using autoclaved piston while cells were scraped off the surface of tissue culture plate.
3. These steps were followed by standard protocol for TRIzol based RNA extraction.
4. Following RNA extraction, we checked the concentration and purity of RNA using nanodrop.

Luciferase assay

8xGTIIC-luciferase was used as a reporter construct (Addgene plasmid #34615) (Dupont *et al.*, 2011). pRL-CMV (Renilla #E2261) purchased from Promega. Dual luciferase kit (E1960, Promega).

Luciferase reporter assay protocol

1. Remove growth medium from cultured cells carefully.
2. Rinse cells in 1X PBS (Do not dislodge cells). Remove as much of the final wash as possible.
3. Dispense a minimal volume of 1X lysis reagent (CCLR, RLB or PLB) into each culture vessel (e.g.,400µl/well in 24 well plate).
4. For culture dishes, scrape attached cells from the dish, and transfer the cells and solution to a microcentrifuge tube. Pellet debris by brief centrifugation (@5000 RPM for 2 min), and transfer the supernatant to a new tube.
5. Mix 20µl of cell lysate with 100µl of Luciferase Assay Reagent and measure the light produced.

cDNA synthesis

We used Superscript III kit for cDNA synthesis. Total RNA amount used for all cDNA synthesis reactions was between 250-500ng.

qRT-PCR

Quantitative PCR was carried out using Sybr green kit. Primers used for qRT-PCR are as following

Name	Primer Sequence
Yap1_F	ACGTTCATCTGGGACAGCAT
Yap1_R	GTTGGGAGATGGCAAAGACA
NELFA_F	TGGATGATCTCCATTAGGGC
NELFA_R	TCATCGACAACATCCGTCTC
COBRA1_F	AACTGCAGCACCATGTCGTA
COBRA1_R	ACTTTTTCAGTCCTTCCCCC
TH1L_F	CAGGCATTCTTGCTGAACCT
TH1L_R	TGGACGAGGACTACTACGGG
RDBP_F	GATATGCTCCTCTGGAACGG
RDBP_R	ACTCAGGCTTCAAGCGTTCT
MEPCE_F	AGGCAGAGCACCACATCATA
MEPCE_R	GGAGCGGACACATCAGTCTT
Hexim1_F	TTACGAAACCAACCAAAGCC
Hexim1_R	GGGCAAAGGGGACTTTTTAC
Hexim2_F	CAGGGAACCACCAGAGTCAT
Hexim2_R	ACCGCCTGTAATGCAGAGTC
GAPDH_F	AATGAAGGGGTCATTGATGG
GAPDH_R	AAGGTGAAGGTCGGAGTCAA

References

Adelman, K. and Lis, J. T. (2012) 'Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans', *Nature Reviews Genetics*, 13(10), pp. 720–731. doi: 10.1038/nrg3293.

Aiyar, S. E. *et al.* (2004) 'Attenuation of estrogen receptor ??-mediated transcription through estrogen-stimulated recruitment of a negative elongation factor', *Genes and Development*, 18(17), pp. 2134–2146. doi: 10.1101/gad.1214104.

Allaway, R. J. *et al.* (2016) 'Genomic characterization of patient-derived xenograft models established from fine needle aspirate biopsies of a primary pancreatic ductal adenocarcinoma and from patient-matched metastatic sites', *Oncotarget*, 7(13), pp. 17087–17102. doi: 10.18632/oncotarget.7718.

Amleh, A. *et al.* (2009) 'Mouse cofactor of BRCA1 (Cobra1) is required for early embryogenesis', *PLoS ONE*, 4(4), pp. 2–9. doi: 10.1371/journal.pone.0005034.

Anders, S. *et al.* (2013) 'Count-based differential expression analysis of RNA sequencing data using R and Bioconductor', *Nature Protocols*, 8(9), pp. 1765–1786. doi: 10.1038/nprot.2013.099.

Anders, S. and Huber, W. (2010) 'Differential expression analysis for sequence count data', *Genome Biology*, 11(10), p. R106. doi: 10.1186/gb-2010-11-10-r106.

Arabi, A. *et al.* (2005) 'c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription', *Nature Cell Biology*, 7(3), pp. 303–310. doi: 10.1038/ncb1225.

Avirneni-Vadlamudi, U. *et al.* (2012) 'Drosophila and mammalian models uncover a role for the myoblast fusion gene TANC1 in rhabdomyosarcoma', *Journal of Clinical Investigation*, 122(1), pp. 403–407. doi: 10.1172/JCI59877.

Barna, M. *et al.* (2008) 'Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency', *Nature*. Nature Publishing Group, 456(7224), pp. 971–975. doi: 10.1038/nature07449.

Barolo, S. and Posakony, J. W. (2002) 'Enhanced fisheries announcement - FINAL', (1), pp. 1167–1181. doi: 10.1101/gad.976502.naling.

Bartman, C. R. *et al.* (2019) 'Transcriptional Burst Initiation and Polymerase Pause Release Are Key Control Points of Transcriptional Regulation', *Molecular Cell*. Elsevier Inc., 73(3), pp. 519-532.e4. doi: 10.1016/j.molcel.2018.11.004.

Biswas, D. K. and Iglehart, J. D. (2006) 'Linkage between EGFR family receptors and nuclear factor kappaB (NF-κB) signaling in breast cancer', *Journal of Cellular Physiology*, 209(3), pp. 645–652. doi: 10.1002/jcp.20785.

Bradner, J. E., Hnisz, D. and Young, R. A. (2017) 'Transcriptional Addiction in Cancer', *Cell*. Elsevier Inc., 168(4), pp. 629–643. doi: 10.1016/j.cell.2016.12.013.

Brès, V., Yoh, S. M. and Jones, K. A. (2008) 'The multi-tasking P-TEFb complex', *Current Opinion in Cell Biology*, 20(3), pp. 334–340. doi: 10.1016/j.ceb.2008.04.008.

Brumby, A. M. and Richardson, H. E. (2003) 'scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila', *EMBO Journal*, 22(21), pp. 5769–5779. doi: 10.1093/emboj/cdg548.

Bruno, P. M. *et al.* (2017) 'A subset of platinum-containing chemotherapeutic agents kills cells by inducing ribosome biogenesis stress', *Nature medicine*. 2017/02/27, 23(4), pp. 461–471. doi: 10.1038/nm.4291.

Buszczak, M., Signer, R. A. J. and Morrison, S. J. (2014) 'Cellular differences in protein synthesis regulate tissue homeostasis', *Cell*. Elsevier Inc., 159(2), pp. 242–251. doi: 10.1016/j.cell.2014.09.016.

Bywater, M. J. *et al.* (2012) 'Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53', *Cancer cell*, 22(1), pp. 51–65. doi: 10.1016/j.ccr.2012.05.019.

Bywater, M. J. *et al.* (2013) 'Dysregulation of the basal RNA polymerase transcription apparatus in cancer', *Nature Reviews Cancer*. Nature Publishing Group, 13(5), pp. 299–314. doi: 10.1038/nrc3496.

Cai, J. *et al.* (2015) 'β-catenin destruction complex-independent regulation of Hippo-YAP signaling by APC in intestinal tumorigenesis', *Genes and Development*, 29(14), pp. 1493–1506. doi: 10.1101/gad.264515.115.

Cairns, C. A. and White, R. J. (1998) 'p53 is a general repressor of RNA polymerase III transcription', *EMBO Journal*, 17(11), pp. 3112–3123. doi: 10.1093/emboj/17.11.3112.

Camargo, F. D. *et al.* (2007) 'YAP1 Increases Organ Size and Expands Undifferentiated Progenitor Cells', *Current Biology*, 17(23), pp. 2054–2060. doi: 10.1016/j.cub.2007.10.039.

Campbell, K. *et al.* (2018) 'Differential roles of the Drosophila EMT-inducing transcription factors Snail and Serpent in driving primary tumour growth', *PLoS Genetics*, 14(2), pp. 1–17.

doi: 10.1371/journal.pgen.1007167.

Cheng, J. and Deming, T. J. (2011) 'synthesis of polypeptides by ROP of NCAs', *Peptide-Based Materials*, 310(July 2013), pp. 1–26. doi: 10.1007/128.

Cheng, Y. *et al.* (2012) 'LARP7 is a potential tumor suppressor gene in gastric cancer', *Laboratory Investigation*, 92(7), pp. 1013–1019. doi: 10.1038/labinvest.2012.59.

Cheon, D.-J. and Orsulic, S. (2011) 'Mouse Models of Cancer', *Annual Review of Pathology: Mechanisms of Disease*, 6(1), pp. 95–119. doi: 10.1146/annurev.pathol.3.121806.154244.

Cherrier, T. *et al.* (2013) 'CTIP2 is a negative regulator of P-TEFb', *Proceedings of the National Academy of Sciences of the United States of America*, 110(31), pp. 12655–12660. doi: 10.1073/pnas.1220136110.

Citri, A. and Yarden, Y. (2006) 'EGF-ERBB signalling: Towards the systems level', *Nature Reviews Molecular Cell Biology*, 7(7), pp. 505–516. doi: 10.1038/nrm1962.

Cohen, B. *et al.* (1992) 'apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins', *Genes and Development*, 6(5), pp. 715–729. doi: 10.1101/gad.6.5.715.

Collier, L. S. *et al.* (2005) 'Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse', *Nature*, 436(7048), pp. 272–276. doi: 10.1038/nature03681.

Copeland, N. G. and Jenkins, N. A. (2010) 'Harnessing transposons for cancer gene discovery', *Nature Reviews Cancer*. Nature Publishing Group, 10(10), pp. 696–706. doi: 10.1038/nrc2916.

Core, L. and Adelman, K. (2019) 'polymerase II : a nexus of gene regulation', pp. 1–24. doi: 10.1101/gad.325142.119.and.

Core, L. J. *et al.* (2012) 'Defining the Status of RNA Polymerase at Promoters', *Cell Reports*. The Authors, 2(4), pp. 1025–1035. doi: 10.1016/j.celrep.2012.08.034.

Core, L. J. and Lis, J. T. (2008) 'Transcription regulation through promoter-proximal pausing of RNA polymerase II', *Science (New York, N.Y.)*, 319(5871), pp. 1791–1792. doi: 10.1126/science.1150843.

D'Orso, I. (2016) '7SKiing on chromatin: Move globally, act locally', *RNA Biology*. Taylor & Francis, 13(6), pp. 545–553. doi: 10.1080/15476286.2016.1181254.

Dai, C. and Sampson, S. B. (2016) 'HSF1: Guardian of Proteostasis in Cancer', *Trends in Cell*

Biology. Elsevier Ltd, 26(1), pp. 17–28. doi: 10.1016/j.tcb.2015.10.011.

Dow, L. E. *et al.* (2008) 'Loss of human Scribble cooperates with H-Ras to promote cell invasion through deregulation of MAPK signalling', *Oncogene*, 27(46), pp. 5988–6001. doi: 10.1038/onc.2008.219.

Drygin, D. *et al.* (2009) 'Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis', *Cancer research*. 2009/09/08. United States, 69(19), pp. 7653–7661. doi: 10.1158/0008-5472.CAN-09-1304.

Dupont, S. *et al.* (2011) 'Role of YAP/TAZ in mechanotransduction.', *Nature*. Nature Publishing Group, 474(7350), pp. 179–183. doi: 10.1038/nature10137.

Dupuy, A. J. *et al.* (2009) 'A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice', *Cancer Research*, 69(20), pp. 8150–8156. doi: 10.1158/0008-5472.CAN-09-1135.

Egloff, S., Dienstbier, M. and Murphy, S. (2012) 'Updating the RNA polymerase CTD code: Adding gene-specific layers', *Trends in Genetics*. Elsevier Ltd, 28(7), pp. 333–341. doi: 10.1016/j.tig.2012.03.007.

Eichenlaub, T. *et al.* (2018) 'Warburg Effect Metabolism Drives Neoplasia in a Drosophila Genetic Model of Epithelial Cancer', *Current Biology*, 28(20), pp. 3220-3228.e6. doi: 10.1016/j.cub.2018.08.035.

Ellenbroek, S. I. J., Iden, S. and Collard, J. G. (2012) 'Seminars in Cancer Biology Cell polarity proteins and cancer', *Seminars in Cancer Biology*. Elsevier Ltd, 22(3), pp. 208–215. doi: 10.1016/j.semcancer.2012.02.012.

Fernandez-L, a *et al.* (2012) 'Oncogenic YAP promotes radioresistance and genomic instability in medulloblastoma through IGF2-mediated Akt activation', *Oncogene*, 31(15), pp. 1923–1937. doi: 10.1038/onc.2011.379.

Galli, G. G. *et al.* (2015) 'YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers', *Molecular Cell*. Elsevier Inc., 60(2), pp. 328–337. doi: 10.1016/j.molcel.2015.09.001.

Garraway, L. A. and Lander, E. S. (2013) 'Lessons from the cancer genome', *Cell*. Elsevier Inc., 153(1), pp. 17–37. doi: 10.1016/j.cell.2013.03.002.

Gateff, E. (1978) 'Malignant neoplasms of genetic origin in Drosophila melanogaster', *Science*,

200(4349), pp. 1448–1459. doi: 10.1126/science.96525.

Gateff, E., Löffler, T. and Wismar, J. (1993) ‘A temperature-sensitive brain tumor suppressor mutation of Drosophila melanogaster: Developmental studies and molecular localization of the gene’, *Mechanisms of Development*, 41(1), pp. 15–31. doi: 10.1016/0925-4773(93)90052-Y.

Gilmour, D. S. and Lis, J. T. (1986) ‘RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells’, *Molecular and cellular biology*, 6(11), pp. 3984–3989. doi: 10.1128/mcb.6.11.3984.

Goldman, M. *et al.* (2019) ‘The UCSC Xena platform for public and private cancer genomics data visualization and interpretation’, *bioRxiv*, p. 326470. doi: 10.1101/326470.

Gonzalez, C. (2013) ‘Drosophila melanogaster: a model and a tool to investigate malignancy and identify new therapeutics’, *Nature Reviews Cancer*. Nature Publishing Group, 13(3), pp. 172–183. doi: 10.1038/nrc3461.

Grandori, C. *et al.* (2005) ‘c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I’, *Nature Cell Biology*, 7(3), pp. 311–318. doi: 10.1038/ncb1224.

Groth, C. *et al.* (2020) ‘Genome Wide Screen for Context-Dependent Tumor Suppressors Identified Using in Vivo Models for Neoplasia in Drosophila’;, *G3: Genes|Genomes|Genetics*, p. g3.401545.2020. doi: 10.1534/g3.120.401545.

Gschwind, A., Fischer, O. M. and Ullrich, A. (2004) ‘The discovery of receptor tyrosine kinases: Targets for cancer therapy’, *Nature Reviews Cancer*, 4(5), pp. 361–370. doi: 10.1038/nrc1360.

Guenther, M. G. *et al.* (2007) ‘A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells’, *Cell*, 130(1), pp. 77–88. doi: 10.1016/j.cell.2007.05.042.

Hanahan, D. and Weinberg, R. a. (2011) ‘Hallmarks of cancer: The next generation’, *Cell*. Elsevier Inc., 144(5), pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Hannan, K. M. *et al.* (2003) ‘mTOR-Dependent Regulation of Ribosomal Gene Transcription Requires S6K1 and Is Mediated by Phosphorylation of the Carboxy-Terminal Activation Domain of the Nucleolar Transcription Factor UBF ’, *Molecular and Cellular Biology*, 23(23), pp. 8862–8877. doi: 10.1128/mcb.23.23.8862-8877.2003.

Hansemann, D. (1890) ‘Ueber asymmetrische Zelltheilung in Epithelkrebsen und deren

biologische Bedeutung', *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin*, 119(2), pp. 299–326. doi: 10.1007/BF01882039.

Hansen, C. G., Moroishi, T. and Guan, K.-L. (2015) 'YAP and TAZ: a nexus for Hippo signaling and beyond', *Trends in Cell Biology*. Elsevier Ltd, pp. 1–15. doi: 10.1016/j.tcb.2015.05.002.

Harden, N. (2017) 'New Insights from Drosophila into the Regulation of EGFR Signaling', in Wang, Z. (ed.) *ErbB Receptor Signaling: Methods and Protocols*. New York, NY: Springer New York, pp. 37–42. doi: 10.1007/978-1-4939-7219-7_2.

Hariharan, I. K. (2015) 'Organ Size Control: Lessons from Drosophila', *Developmental Cell*. Elsevier Inc., 34(3), pp. 255–265. doi: 10.1016/j.devcel.2015.07.012.

Harvey, K. F., Zhang, X. and Thomas, D. M. (2013) 'The Hippo pathway and human cancer.', *Nature reviews. Cancer*. Nature Publishing Group, 13(April), pp. 246–257. doi: 10.1038/nrc3458.

Harvey, K. and Tapon, N. (2007) 'The Salvador-Warts-Hippo pathway - An emerging tumour-suppressor network', *Nature Reviews Cancer*, 7(3), pp. 182–191. doi: 10.1038/nrc2070.

He, N. *et al.* (2008) 'A La-Related Protein Modulates 7SK snRNP Integrity to Suppress P-TEFb-Dependent Transcriptional Elongation and Tumorigenesis', *Molecular Cell*, 29(5), pp. 588–599. doi: 10.1016/j.molcel.2008.01.003.

Herranz, H. *et al.* (2012) 'Oncogenic cooperation between SOCS family proteins and EGFR identified using a Drosophila epithelial transformation model', *Genes and Development*, 26, pp. 1602–1611. doi: 10.1101/gad.192021.112.

Herranz, H., Eichenlaub, T. and Cohen, S. M. (2016) 'Cancer in Drosophila. Imaginal Discs as a Model for Epithelial Tumor Formation', *Current Topics in Developmental Biology*, 116, pp. 181–199. doi: 10.1016/bs.ctdb.2015.11.037.

Herranz, H., Weng, R. and Cohen, S. M. (2014) 'Crosstalk between epithelial and mesenchymal tissues in tumorigenesis and imaginal disc development', *Current Biology*, 24(13), pp. 1476–1484. doi: 10.1016/j.cub.2014.05.043.

Hipfner, D. R. and Cohen, S. M. (2004) 'Connecting proliferation and apoptosis in development and disease', *Nature Reviews Molecular Cell Biology*, 5(10), pp. 805–815. doi: 10.1038/nrm1491.

Huang, J. *et al.* (2005a) 'The Hippo signaling pathway coordinately regulates cell proliferation

and apoptosis by inactivating Yorkie, the Drosophila homolog of YAP', *Cell*, 122, pp. 421–434. doi: 10.1016/j.cell.2005.06.007.

Huang, J. *et al.* (2005b) 'The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila homolog of YAP', *Cell*, 122(3), pp. 421–434. doi: 10.1016/j.cell.2005.06.007.

Hynes, N. E. and Lane, H. a (2006) 'Reviews', *Nature Reviews Cancer*, 5(May), pp. 341–354. doi: 10.1038/nrc1609.

Jang, M. K. *et al.* (2005) 'The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription', *Molecular Cell*. Elsevier, 19(4), pp. 523–534. doi: 10.1016/j.molcel.2005.06.027.

Jezowska, B. *et al.* (2011) 'A dual function of Drosophila capping protein on DE-cadherin maintains epithelial integrity and prevents JNK-mediated apoptosis', *Developmental Biology*. Elsevier Inc., 360(1), pp. 143–159. doi: 10.1016/j.ydbio.2011.09.016.

Jia, J. *et al.* (2003) 'The Drosophila Ste20 family kinase dMST functions as a tumor suppressor by restricting cell proliferation and promoting apoptosis', *Genes and Development*, 17(20), pp. 2514–2519. doi: 10.1101/gad.1134003.

Jiang, T. *et al.* (2011) 'CARMA3 is crucial for EGFR-induced activation of NF-κB and tumor progression', *Cancer Research*, 71(6), pp. 2183–2192. doi: 10.1158/0008-5472.CAN-10-3626.

Jones, S. *et al.* (2008) 'Comparative lesion sequencing provides insights into tumor evolution.', *Proceedings of the National Academy of Sciences of the United States of America*, 105(11), pp. 4283–8. doi: 10.1073/pnas.0712345105.

Jonkers, I. and Lis, J. T. (2015) 'Getting up to speed with transcription elongation by RNA polymerase II', *Nat Rev Mol Cell Biol*. Nature Publishing Group, 16(3), pp. 167–177. doi: 10.1038/nrm3953.

Justice, R. W. *et al.* (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), pp. 534–546. doi: 10.1101/gad.9.5.534.

Justilien, V. *et al.* (2017) 'Ect2-Dependent rRNA Synthesis Is Required for KRAS-TRP53-Driven Lung Adenocarcinoma', *Cancer Cell*. Elsevier Inc., 31(2), pp. 256–269. doi: 10.1016/j.ccell.2016.12.010.

Kandoth, C. *et al.* (2013) 'Mutational landscape and significance across 12 major cancer types', *Nature*. Nature Publishing Group, 502(7471), pp. 333–339. doi: 10.1038/nature12634.

Keng, V. W. *et al.* (2009) 'A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma', *Nature Biotechnology*, 27(3), pp. 264–274. doi: 10.1038/nbt.1526.

Kim, S. *et al.* (2000) 'Induction of ribosomal genes and hepatocyte hypertrophy by adenovirus-mediated expression of c-Myc in vivo', *Proceedings of the National Academy of Sciences of the United States of America*, 97(21), pp. 11198–11202. doi: 10.1073/pnas.200372597.

Kim, S., You, S. and Hwang, D. (2011) 'Aminoacyl-tRNA synthetases and tumorigenesis: More than housekeeping', *Nature Reviews Cancer*, 11(10), pp. 708–718. doi: 10.1038/nrc3124.

Knudson, A. G. (1971) 'Mutation and Cancer: Statistical Study of Retinoblastoma', *Proceedings of the National Academy of Sciences*, 68(4), pp. 820 LP – 823. doi: 10.1073/pnas.68.4.820.

Ko, Y. G. *et al.* (2000) 'Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis', *The Journal of cell biology*. The Rockefeller University Press, 149(3), pp. 567–574. doi: 10.1083/jcb.149.3.567.

Ko, Y. G. *et al.* (2001) 'Glutamine-dependent antiapoptotic interaction of human glutaminyl-tRNA synthetase with apoptosis signal-regulating kinase 1', *The Journal of biological chemistry*. 2000/11/28. United States, 276(8), pp. 6030–6036. doi: 10.1074/jbc.M006189200.

Krueger, B. J. *et al.* (2008) 'LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated', *Nucleic Acids Research*, 36(7), pp. 2219–2229. doi: 10.1093/nar/gkn061.

Krumm, A., Hickey, L. B. and Groudine, M. (1995) 'Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation', *Genes and Development*, 9(5), pp. 559–572. doi: 10.1101/gad.9.5.559.

Kushner, J. P. *et al.* (1976) 'Elevated methionine-tRNA synthetase activity in human colon cancer', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*. United States, 153(2), pp. 273–276. doi: 10.3181/00379727-153-39526.

Kwak, H. (2013) 'Initiation and Pausing', *Science*, 339, p. 950. doi: 10.1126/science.1229386.

Kwak, H. and Lis, J. T. (2013) 'Control of Transcriptional Elongation', *Annual Review of*

Genetics, 47(1), pp. 483–508. doi: 10.1146/annurev-genet-110711-155440.

Lagha, M. *et al.* (2013) 'XPaused Pol II coordinates tissue morphogenesis in the drosophila embryo', *Cell*. Elsevier Inc., 153(5), p. 976. doi: 10.1016/j.cell.2013.04.045.

Lai, Z. C. *et al.* (2005) 'Control of cell proliferation and apoptosis by mob as tumor suppressor, mats', *Cell*, 120(5), pp. 675–685. doi: 10.1016/j.cell.2004.12.036.

Lee, Y.-R., Chen, M. and Pandolfi, P. P. (2018) 'The functions and regulation of the PTEN tumour suppressor: new modes and prospects', *Nature Reviews Molecular Cell Biology*, 19(9), pp. 547–562. doi: 10.1038/s41580-018-0015-0.

Levine, M. (2011) 'Paused RNA polymerase II as a developmental checkpoint', *Cell*. Elsevier Inc., 145(4), pp. 502–511. doi: 10.1016/j.cell.2011.04.021.

Levine, M., Cattoglio, C. and Tjian, R. (2014) 'Looping back to leap forward: Transcription enters a new era', *Cell*. Elsevier Inc., 157(1), pp. 13–25. doi: 10.1016/j.cell.2014.02.009.

Li, J. *et al.* (2013) 'Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing', *Molecular Cell*. Elsevier Inc., 50(5), pp. 711–722. doi: 10.1016/j.molcel.2013.05.016.

Lin, K. C., Park, H. W. and Guan, K.-L. (2018) 'Deregulation and Therapeutic Potential of the Hippo Pathway in Cancer', *Annual Review of Cancer Biology*, 2(1), pp. 59–79. doi: 10.1146/annurev-cancerbio-030617-050202.

Liu, A. (2019) 'Proteostasis in the Hedgehog signaling pathway', *Seminars in Cell and Developmental Biology*. Elsevier, 93(July 2018), pp. 153–163. doi: 10.1016/j.semcdb.2018.10.009.

Luo, M. *et al.* (2013) 'A conserved protein motif is required for full modulatory activity of negative elongation factor subunits NELF-A and NELF-B in modifying glucocorticoid receptor-regulated gene induction properties', *Journal of Biological Chemistry*, 288(47), pp. 34055–34072. doi: 10.1074/jbc.M113.512426.

Luo, Z., Lin, C. and Shilatifard, A. (2012) 'The super elongation complex (SEC) family in transcriptional control', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 13(9), pp. 543–547. doi: 10.1038/nrm3417.

Marquardt, J. U. *et al.* (2014) 'Sequential transcriptome analysis of human liver cancer indicates late stage acquisition of malignant traits', *Journal of Hepatology*. European Association for the

Study of the Liver, 60(2), pp. 346–353. doi: 10.1016/j.jhep.2013.10.014.

Marquardt, J. U., Andersen, J. B. and Thorgeirsson, S. S. (2015) 'Functional and genetic deconstruction of the cellular origin in liver cancer', *Nature Reviews Cancer*. Nature Publishing Group, 15(11), pp. 653–667. doi: 10.1038/nrc4017.

Martin-belmonte, F. and Perez-moreno, M. (2012) 'Epithelial cell polarity , stem cells and cancer', *Nature Reviews Cancer*. Nature Publishing Group, 12(January). doi: 10.1038/nrc3169.

McGranahan, N. and Swanton, C. (2017) 'Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future', *Cell*. Elsevier Inc., 168(4), pp. 613–628. doi: 10.1016/j.cell.2017.01.018.

McNamara, R. P. *et al.* (2013) 'Transcription factors mediate the enzymatic disassembly of promoter-bound 7sk snrnp to locally recruit p-tefb for transcription elongation', *Cell Reports*, 5(5), pp. 1256–1268. doi: 10.1016/j.celrep.2013.11.003.

McNamara, R. P., Bacon, C. W. and D'Orso, I. (2016) 'Transcription elongation control by the 7SK snRNP complex: Releasing the pause', *Cell Cycle*. Taylor & Francis, 15(16), pp. 2115–2123. doi: 10.1080/15384101.2016.1181241.

Miles, W. O., Dyson, N. J. and Walker, J. a. (2011) 'Modeling tumor invasion and metastasis in Drosophila', *Disease Models & Mechanisms*, 4, pp. 753–761. doi: 10.1242/dmm.006908.

Min, I. M. *et al.* (2011) 'Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells', *Genes and Development*, 25(7), pp. 742–754. doi: 10.1101/gad.2005511.

Muse, G. W. *et al.* (2007) 'RNA polymerase is poised for activation across the genome.', *Nature genetics*, 39(12), pp. 1507–11. doi: 10.1038/ng.2007.21.

Narita, T. *et al.* (2003) 'Human Transcription Elongation Factor NELF : Identification of Novel Subunits and Reconstitution of the Functionally Active Complex Human Transcription Elongation Factor NELF : Identification of Novel Subunits and Reconstitution of the Functionally Active', *Molecular and Cellular Biology*, 23(6), pp. 1863–1873. doi: 10.1128/MCB.23.6.1863.

Nechaev, S. *et al.* (2010) 'Global Analysis of Short RNAs Supplementary', *Science*, 327(5963), pp. 335–338. doi: 10.1126/science.1181421.

Nguyen, D. *et al.* (2012) 'The Drosophila 7SK snRNP and the essential role of dHEXIM in development', *Nucleic Acids Research*, 40(12), pp. 5283–5297. doi: 10.1093/nar/gks191.

Nguyen, H. T. *et al.* (2014) 'Viral small T oncoproteins transform cells by alleviating Hippo-pathway-mediated inhibition of the YAP proto-oncogene', *Cell Reports*. The Authors, 8(3), pp. 707–713. doi: 10.1016/j.celrep.2014.06.062.

Nguyen, V. T. *et al.* (2001) '7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes', *Nature*, 414(6861), pp. 322–325. doi: 10.1038/35104581.

Normanno, N. *et al.* (2006) 'Epidermal growth factor receptor (EGFR) signaling in cancer', *Gene*, 366(1), pp. 2–16. doi: 10.1016/j.gene.2005.10.018.

Oh, H. *et al.* (2013) 'Genome-wide Association of Yorkie with Chromatin and Chromatin-Remodeling Complexes', *Cell Reports*. The Authors, 3(2), pp. 309–318. doi: 10.1016/j.celrep.2013.01.008.

Overholtzer, M. *et al.* (2006) 'Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(33), pp. 12405–10. doi: 10.1073/pnas.0605579103.

Pagliarini, R. a and Xu, T. (2003) 'A genetic screen in Drosophila for metastatic behavior.', *Science (New York, N.Y.)*, 302(5648), pp. 1227–1231. doi: 10.1126/science.1088474.

Pallavi, S. K. and Shashidhara, L. S. (2005) 'Signaling interactions between squamous and columnar epithelia of the Drosophila wing disc.', *Journal of cell science*, 118, pp. 3363–3370. doi: 10.1242/jcs.02464.

Pan, D. (2010) 'The hippo signaling pathway in development and cancer', *Developmental Cell*. Elsevier Inc., 19(4), pp. 491–505. doi: 10.1016/j.devcel.2010.09.011.

Parsa, A. T. and Holland, E. C. (2004) 'Cooperative translational control of gene expression by Ras and Akt in cancer', *Trends in Molecular Medicine*, 10(12), pp. 607–613. doi: 10.1016/j.molmed.2004.10.009.

Pelletier, J. *et al.* (2017) 'Nrc.2017.104', *Nature*, 18, p. 51. Available at: https://doi.org/10.1038/nrc.2017.104%0Ahttp://10.0.4.14/nrc.2017.104.

Peng, X. *et al.* (2018) 'Molecular Characterization and Clinical Relevance of Metabolic Expression Subtypes in Human Cancers', *Cell Reports*, 23(1), pp. 255-269.e4. doi: 10.1016/j.celrep.2018.03.077.

Pertea, M. *et al.* (2016) 'RNA-seq experiments with HISAT , StringTie and Ballgown', *Nature Protocols*. Nature Publishing Group, 11(9), pp. 1650–1667. doi: 10.1038/nprot.2016-095.

Peterlin, B. M. and Price, D. H. (2006) ‘Controlling the Elongation Phase of Transcription with P-TEFb’, *Molecular Cell*, 23(3), pp. 297–305. doi: 10.1016/j.molcel.2006.06.014.

Quaresma, A. J. C., Bugai, A. and Barboric, M. (2016) ‘Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb’, *Nucleic Acids Research*, 44(16), pp. 7527–7539. doi: 10.1093/nar/gkw585.

Rahl, P. B. *et al.* (2010) ‘C-Myc regulates transcriptional pause release’, *Cell*, 141(3), pp. 432–445. doi: 10.1016/j.cell.2010.03.030.

Rajasekhar, V. K. *et al.* (2003) ‘Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes’, *Molecular Cell*, 12(4), pp. 889–901. doi: 10.1016/S1097-2765(03)00395-2.

Rasmussen, E. B. and Lis, J. T. (1993) ‘In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes’, *Proceedings of the National Academy of Sciences of the United States of America*, 90(17), pp. 7923–7927. doi: 10.1073/pnas.90.17.7923.

Razin, E. *et al.* (1999) ‘Suppression of microphthalmia transcriptional activity by its association with protein kinase C-interacting protein 1 in mast cells’, *Journal of Biological Chemistry*, 274(48), pp. 34272–34276. doi: 10.1074/jbc.274.48.34272.

Read, R. D. *et al.* (2009) ‘A Drosophila model for EGFR-Ras and PI3K-dependent human glioma’, *PLoS Genetics*, 5(2). doi: 10.1371/journal.pgen.1000374.

Rosenbluh, J. *et al.* (2012) ‘β-Catenin-driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis’, *Cell*. Elsevier Inc., 151(7), pp. 1457–1473. doi: 10.1016/j.cell.2012.11.026.

Rougvie, A. E. and Lis, J. T. (1988) ‘The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged’, *Cell*, 54(6), pp. 795–804. doi: 10.1016/S0092-8674(88)91087-2.

Sanchez-Vega, F. *et al.* (2018) ‘Oncogenic Signaling Pathways in The Cancer Genome Atlas’, *Cell*, 173(2), pp. 321-337.e10. doi: 10.1016/j.cell.2018.03.035.

Saunders, A. *et al.* (2013) ‘Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription’, *Genes and Development*, 27(10), pp. 1146–1158. doi: 10.1101/gad.215459.113.

Schulze, K. *et al.* (2015) ‘Exome sequencing of hepatocellular carcinomas identifies new

mutational signatures and potential therapeutic targets', *Nature Genetics*. Nature Publishing Group, 47(5), pp. 505–511. doi: 10.1038/ng.3252.

Serafini, T. *et al.* (1996) 'Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system', *Cell*. United States, 87(6), pp. 1001–1014. doi: 10.1016/s0092-8674(00)81795-x.

Shostak, K. and Chariot, A. (2015) 'EGFR and NF-κB: Partners in cancer', *Trends in Molecular Medicine*. Elsevier Ltd, 21(6), pp. 385–393. doi: 10.1016/j.molmed.2015.04.001.

Song, S., Herranz, H. and Cohen, S. M. (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), pp. 1201–1209. doi: 10.1242/dmm.030122.

Starr, T. K. *et al.* (2009) 'A Transposon-Based Genetic Screen in Mice Identifies Genes Altered in Colorectal Cancer', *Science*, 323(5922), pp. 1747 LP – 1750. doi: 10.1126/science.1163040.

Stratton, M. R. (2011) 'Exploring the genomes of cancer cells: progress and promise.', *Science (New York, N.Y.)*, 331(March), pp. 1553–1558. doi: 10.1126/science.1204040.

Stumpf, C. R. and Ruggero, D. (2011) 'The cancerous translation apparatus', *Current Opinion in Genetics and Development*. Elsevier Ltd, 21(4), pp. 474–483. doi: 10.1016/j.gde.2011.03.007.

Sur, I. and Taipale, J. (2016) 'The role of enhancers in cancer', *Nature Reviews Cancer*. Nature Publishing Group, 16(8), pp. 483–493. doi: 10.1038/nrc.2016.62.

Szklarczyk, D. *et al.* (2017) 'The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible', *Nucleic Acids Research*, 45(D1), pp. D362–D368. doi: 10.1093/nar/gkw937.

Tabassum, D. P. and Polyak, K. (2015) 'Tumorigenesis: it takes a village', *Nature Reviews Cancer*. Nature Publishing Group, 15(8), pp. 473–483. doi: 10.1038/nrc3971.

Takeda, H. *et al.* (2015) 'Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression', *Nature Genetics*, 47(2), pp. 142–150. doi: 10.1038/ng.3175.

Tanos, B. and Rodriguez-Boulan, E. (2008) 'The epithelial polarity program: machineries involved and their hijacking by cancer.', *Oncogene*, 27(55), pp. 6939–6957. doi: 10.1038/onc.2008.345.

Tapon, N. *et al.* (2002) ‘salvador promotes both cell cycle exit and apoptosis in Drosophila and is mutated in human cancer cell lines’, *Cell*, 110(4), pp. 467–478. doi: 10.1016/S0092-8674(02)00824-3.

Truitt, M. L. and Ruggero, D. (2016) ‘New frontiers in translational control of the cancer genome’, *Nature Reviews Cancer*, 16(5), pp. 288–304. doi: 10.1038/nrc.2016.27.

Udan, R. S. *et al.* (2003) ‘Hippo promotes proliferation arrest and apoptosis in the Salvador/Warts pathway’, *Nature Cell Biology*, 5(10), pp. 914–920. doi: 10.1038/ncb1050.

Uhlirova, M. and Bohmann, D. (2006) ‘JNK- and Fos-regulated Mmp1 expression cooperates with Ras to induce invasive tumors in Drosophila’, *EMBO Journal*, 25(22), pp. 5294–5304. doi: 10.1038/sj.emboj.7601401.

Vissers, J. H. A. *et al.* (2016) ‘A Drosophila RNAi library modulates Hippo pathway-dependent tissue growth’, *Nature Communications*, 7(1), p. 10368. doi: 10.1038/ncomms10368.

Vogelstein, B. *et al.* (2013) ‘Cancer Genome Landscapes’, *Science*, 339(6127), pp. 1546–1558. doi: 10.1126/science.1235122.

Vos, S. M. *et al.* (2018) ‘Structure of paused transcription complex Pol II–DSIF–NELF’, *Nature*. Springer US, 560(7720), pp. 601–606. doi: 10.1038/s41586-018-0442-2.

Wakasugi, K. and Schimmel, P. (1999) ‘Two Distinct Cytokines Released from a Human Aminoacyl-tRNA Synthetase’, *Science*. American Association for the Advancement of Science, 284(5411), pp. 147–151. doi: 10.1126/science.284.5411.147.

Wang, X. *et al.* (2007) ‘Transcription elongation controls cell fate specification in the Drosophila embryo’, *Genes and Development*, 21(9), pp. 1031–1036. doi: 10.1101/gad.1521207.

Wang, X. *et al.* (2010) ‘NELF potentiates gene transcription in the drosophila embryo’, *PLoS ONE*, 5(7), pp. 1–8. doi: 10.1371/journal.pone.0011498.

Watson, K. L., Justice, R. W. and Bryant, P. J. (1994) ‘Drosophila in cancer research: The first fifty tumor suppressor genes’, *Journal of Cell Science*, 107(SUPPL. 18), pp. 19–33. doi: 10.1242/jcs.1994.supplement_18.4.

White, R. J. *et al.* (1996) ‘Retinoblastoma Protein’, 382(July), pp. 88–90.

Wu, C. H. *et al.* (2003) ‘NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila’, *Genes and Development*, 17(11), pp. 1402–1414. doi: 10.1101/gad.1091403.

Wu, C. H. *et al.* (2005) 'Molecular characterization of Drosophila NELF', *Nucleic Acids Research*, 33(4), pp. 1269–1279. doi: 10.1093/nar/gki274.

Wu, M., Pastor-Pareja, J. C. and Xu, T. (2010) 'Interaction between Ras(V12) and scribbled clones induces tumour growth and invasion.', *Nature*, 463(January), pp. 545–548. doi: 10.1038/nature08702.

Wu, X. and Pandolfi, P. P. (2001) 'Mouse models for multistep tumorigenesis', *Trends in Cell Biology*, 11(11), pp. 2–9. doi: 10.1016/S0962-8924(01)02127-4.

Xu, T. *et al.* (1995) 'Identifying tumor suppressors in genetic mosaics: The Drosophila lats gene encodes a putative protein kinase', *Development*, 121(4), pp. 1053–1063.

Xue, Y. *et al.* (2009) 'A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP', *Nucleic Acids Research*, 38(2), pp. 360–369. doi: 10.1093/nar/gkp977.

Yachida, S. *et al.* (2010) 'Distant metastasis occurs late during the genetic evolution of pancreatic cancer', *Nature*, 467(7319), pp. 1114–1117. doi: 10.1038/nature09515.

Yamaguchi, Y. Y. *et al.* (1999) 'NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation.', *Cell*, 97(1), p. 11. doi: 10.1016/S0092-8674(00)80713-8.

Yamaguchi, Y. Y. *et al.* (2002) 'Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA.', *Molecular and cellular biology*, 22(9), pp. 2918–2927. doi: 10.1128/MCB.22.9.2918.

Yang, W. *et al.* (2012) 'EGFR-Induced and PKCε Monoubiquitylation-Dependent NF-κB Activation Upregulates PKM2 Expression and Promotes Tumorigenesis', *Molecular Cell*, 48(5), pp. 771–784. doi: 10.1016/j.molcel.2012.09.028.

Yang, Z. *et al.* (2005) 'Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4', *Molecular Cell*, 19(4), pp. 535–545. doi: 10.1016/j.molcel.2005.06.029.

Yik, J. H. N. *et al.* (2003) 'Inhibition of P-TEFb (CDK9/cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA', *Molecular Cell*, 12(4), pp. 971–982. doi: 10.1016/S1097-2765(03)00388-5.

Zanconato, F. *et al.* (2015) 'Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth', *Nature Cell Biology*, 17(9). doi: 10.1038/ncb3216.

Zeitlinger, J. *et al.* (2007) 'RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo.', *Nature genetics*, 39(12), pp. 1512–6. doi: 10.1038/ng.2007.26.

Zender, L. *et al.* (2006) 'Identification and Validation of Oncogenes in Liver Cancer Using an Integrative Oncogenomic Approach', *Cell*, 125(7), pp. 1253–1267. doi: 10.1016/j.cell.2006.05.030.

Appendix I

Following table represents the screened RNAi lines with three backgrounds. Each entry has been depicted with initials of the screen background for ease of visualization. A blank cell under any background header shows that particular line was not subjected to that screen.

CG Number	Trans-ID	Yki	EGFR	SOCS
CG42783	105624	y	e	s
CG5279	101930	y	e	s
CG6964	107413	y	e	s
CG12663	108171	y	e	s
CG15112	106484	y	e	s
CG12333	103639	y	e	s
CG5336	107590	y	e	s
CG1775	106767	y	e	s
CG2762	104102	y	e	s
CG4254	110599	y	e	s
CG10061	106051	y	e	s
CG12919	108814	y	e	s
CG33336	103001	y	e	s
CG17712	105119	y	e	s
CG4337	107332	y	e	s
CG6167	104486	y	e	s
CG3249	105107	y	e	s
CG17603	106119	y	e	s
CG11354	104468	y	e	s
CG6007	106175	y	e	s
CG8862	105749	y	e	s
CG5045	103423	y	e	s
CG4866	105768	y	e	s
CG5546	103926	y	e	s
CG11085	103232	y	e	s
CG6570	103231	y	e	s
CG1107	103426	y	e	s
CG9353	105729	y	e	s
CG3780	107304	y	e	s
CG34345	110487	y	e	s
CG9450	105108	y	e	s
CG14882	103984	y	e	s
CG14032	104515	y	e	s
CG10699	105746	y	e	s
CG10184	104488	y	e	s
CG10094	110504	y	e	s
CG5247	110409	y	e	s
CG34412	105732	y	e	s
CG4205	104499	y	e	s

CG13410	103388	y	e	s
CG3845	105121	y	e	s
CG9143	105825	y	e	s
CG14838	105707	y	e	s
CG8302	104458	y	e	s
CG8696	106220	y	e	s
CG9323	110410	y	e	s
CG2135	103338	y	e	s
CG17737	105763	y	e	s
CG14879	103284	y	e	s
CG32405	104058	y	e	s
CG3466	103975	y	e	s
CG13551	104540	y	e	s
CG14221	107841	y	e	s
CG12096	104492	y	e	s
CG7583	107313	y	e	s
CG3360	104130	y	e	s
CG17642	107308	y	e	s
CG12264	105106	y	e	s
CG17870	104496	y	e	s
CG17928	103969	y	e	s
CG8789	103410	y	e	s
CG4233	106120	y	e	s
CG8767	110435	y	e	s
CG2028	110768	y	e	s
CG15128	100238	y	e	s
CG4465	100202	y	e	s
CG2899	110621	y	e	s
CG10391	100318	y	e	s
CG7028	107042	y	e	s
CG10267	100204	y	e	s
CG32181	100186	y	e	s
CG6272	101871	y	e	s
CG3324	103513	y	e	s
CG15591	100240	y	e	s
CG2102	100305	y	e	s
CG15820	100203	y	e	s
CG2244	110632	y	e	s
CG3853	100253	y	e	s
CG16800	100207	y	e	s
CG9649	100194	y	e	s
CG8801	100270	y	e	s
CG10447	103655	y	e	s
CG14227	100192	y	e	s
CG34143	100181	y	e	s
CG10750	100239	y	e	s
CG5925	103666	y	e	s

CG14031	103489	y	e	s
CG2023	100264	y	e	s
CG10345	100252	y	e	s
CG8902	100235	y	e	s
CG11196	100255	y	e	s
CG3724	100269	y	e	s
CG1979	100191	y	e	s
CG18522	100330	y	e	s
CG33522	100199	y	e	s
CG8673	100185	y	e	s
CG1650	100301	y	e	s
CG33213	103517	y	e	s
CG3078	100258	y	e	s
CG10952	100260	y	e	s
CG3679	100249	y	e	s
CG5206	101737	y	e	s
CG8635	100278	y	e	s
CG31679	100187	y	e	s
CG5231	100267	y	e	s
CG1978	100201	y	e	s
CG30428	100243	y	e	s
CG34389	100247	y	e	s
CG16910	100257	y	e	s
CG7184	101694	y	e	s
CG1644	101723	y	e	s
CG31077	100188	y	e	s
CG9045	110672	y	e	s
CG18599	110312	y	e	s
CG17724	102068	y	e	s
CG8545	108849	y	e	s
CG14740	100676	y	e	s
CG11859	109296	y	e	s
CG40486	105948	y	e	s
CG3567	107735	y	e	s
CG18823	110023	y	e	s
CG1621	108055	y	e	s
CG14981	109502	y	e	s
CG2397	104735	y	e	s
CG4193	105924	y	e	s
CG7693	106919	y	e	s
CG12196	101677	y	e	s
CG10842	106354	y	e	s
CG9307	109467	y	e	s
CG17221	106931	y	e	s
CG10627	105398	y	e	s
CG14938	104313	y	e	s
CG31371	110252	y	e	s

CG9222	104259	y	e	s
CG14200	107468	y	e	s
CG30489	109256	y	e	s
CG3315	102348	y	e	s
CG11828	101020	y	e	s
CG30360	104905	y	e	s
CG1973	110214	y	e	s
CG11033	109295	y	e	s
CG3719	102366	y	e	s
CG18231	101201	y	e	s
CG12299	102146	y	e	s
CG7069	101116	y	e	s
CG10093	100982	y	e	s
CG1333	110454	y	e	s
CG12229	101097	y	e	s
CG5154	100977	y	e	s
CG5444	109640	y	e	s
CG31368	110348	y	e	s
CG3086	110317	y	e	s
CG2889	102340	y	e	s
CG8567	104881	y	e	s
CG17110	101120	y	e	s
CG15738	103029	y	e	s
CG4170	109655	y	e	s
CG2543	110444	y	e	s
CG6428	110692	y	e	s
CG7659	110705	y		
CG14030	101096	y	e	s
CG31414	101212	y	e	s
CG4839	100999	y	e	s
CG10895	110342	y	e	s
CG2054	102249	y	e	s
CG3008	103828	y	e	s
CG9726	101283	y	e	s
CG13098	101173	y	e	s
CG10638	102914	y	e	s
CG5064	104867	y	e	s
CG8863	104880	y	e	s
CG4360	105000	y	e	s
CG9088	103830	y	e	s
CG12236	100972	y	e	s
CG33957	100969	y	e	s
CG12287	102126	y	e	s
CG15845	102176	y	e	s
CG10197	103837	y	e	s
CG12630	101085	y	e	s
CG4936	101256	y	e	s

CG5008	106591	y	e	s
CG2254	101149	y	e	s
CG8888	110467	y		
CG8878	100985	y	e	s
CG11012	100027	y	e	s
CG5872	100035	y	e	s
CG4827	100050	y	e	s
CG4389	100021	y	e	s
CG14438	100109	y	e	s
CG32971	100048	y	e	s
CG9984	100009	y	e	s
CG15435	100074	y	e	s
CG16852	100065	y		s
CG2380	100018	y		s
CG9773	100013	y	e	s
CG11475	100087	y	e	s
CG9373	100001	y	e	s
CG8147	100052	y	e	s
CG31530	100059	y	e	s
CG16710	100054	y	e	s
CG33119	100043	y	e	s
CG2453	100031	y	e	s
CG32553	100058	y	e	s
CG9064	100064	y	e	s
CG11906	100083	y	e	s
CG17929	100093	y	e	s
CG31862	100086	y	e	s
CG31991	100003	y	e	s
CG32407	100022	y	e	s
CG10137	100007	y	e	s
CG10232	100033	y	e	s
CG2968	100621	y	e	s
CG13155	100041	y	e	s
CG15550	100063	y	e	s
CG31121	100046	y	e	s
CG14651	100026	y	e	s
CG11619	100091	y	e	s
CG42280	100077	y	e	s
CG5414	100000	y	e	s
CG4120	100049	y	e	s
CG3788	100076	y	e	s
CG7780	100014	y	e	s
CG10541	100094	y	e	s
CG14925	100032	y	e	s
CG11767	100057	y	e	s
CG5961	100023	y	e	s
CG5718	100071	y	e	s

CG7357	100097	y	e	s
CG10217	100081	y	e	s
CG11486	106497	y	e	s
CG6724	100008	y	e	s
CG17348	100039	y	e	s
CG10827	100073	y	e	s
CG5958	100038	y	e	s
CG14685	100047	y	e	s
CG10449	100095	y	e	s
CG42345	100070	y	e	s
CG5440	100099	y	e	s
CG4616	100082	y	e	s
CG8947	100102	y	e	s
CG6193	100104	y	e	s
CG17028	100100	y	e	s
CG4168	100080	y	e	s
CG8328	100056	y	e	s
CG13305	101892	y		s
CG12021	101877	y		s
CG13924	100380	y	e	s
CG17932	101859	y	e	s
CG10483	100151	y	e	s
CG11888	106457	y	e	s
CG8641	110183	y		s
CG13847	105931	y	e	s
CG7630	107667	y	e	s
CG1887	100219	y		s
CG10203	100226	y	e	s
CG30471	100166	y	e	s
CG6380	100121	y	e	s
CG8211	105887	y	e	s
CG4090	100128	y		s
CG14034	100114	y	e	s
CG31509	106548	y		s
CG7259	100320	y	e	s
CG11967	100233	y	e	s
CG42266	100144	y	e	s
CG12439	100230	y	e	s
CG32813	101839	y	e	s
CG5429	110197	y		s
CG8191	100139	y	e	s
CG1605	100150	y	e	s
CG10506	101761	y		s
CG3651	100115	y	e	s
CG7417	100326	y	e	s
CG6757	105886	y	e	s
CG7164	100152	y	e	s

CG7557	100132	y	e	s
CG13062	101832	y	e	s
CG31811	100123	y	e	s
CG7742	100125	y	e	s
CG8029	101726	y	e	s
CG5056	105929	y	e	s
CG4147	101766	y	e	s
CG8654	100112	y	e	s
CG7228	100391	y	e	s
CG12126	100365	y	e	s
CG15220	101833	y	e	s
CG7726	101731	y	e	s
CG6142	100131	y	e	s
CG33558	100164	y	e	s
CG15878	100116	y	e	s
CG5372	100120	y	e	s
CG42271	100176	y	e	s
CG30099	100117	y	e	s
CG17224	101772	y	e	s
CG16723	100092	y	e	s
CG15461	103441	y	e	s
CG1817	110443	y	e	s
CG4204	101542	y	e	s
CG2574	105725	y	e	s
CG13202	103440	y	e	s
CG34114	103456	y	e	s
CG1977	110417	y	e	s
CG7275	103948	y	e	s
CG2467	105189	y	e	s
CG1821	104467	y	e	s
CG8395	103922	y	e	s
CG4694	103967	y	e	s
CG4396	101508	y	e	s
CG32625	103216	y	e	s
CG4609	103929	y	e	s
CG7181	104047	y	e	s
CG6292	103387	y	e	s
CG3193	103393	y	e	s
CG30429	103389	y	e	s
CG17027	103270	y	e	s
CG16993	103407	y	e	s
CG13608	101549	y	e	s
CG8237	103932	y	e	s
CG3999	105165	y	e	s
CG15125	105180	y	e	s
CG4759	103401	y	e	s
CG9079	103282	y	e	s

CG11858	106769	y	e	s
CG12986	105857	y	e	s
CG6712	103907	y	e	s
CG13121/44152	105199	y	e	s
CG11201	104449	y	e	s
CG3388	107940	y	e	s
CG7762	103939	y	e	s
CG10406	105182	y	e	s
CG31473	104039	y	e	s
CG6443	103428	y	e	s
CG8781	107385	y	e	s
CG32812	104081	y	e	s
CG11249	108319	y	e	s
CG13298	105704	y		s
CG31206	108771	y	e	s
CG14174	108724	y	e	s
CG2577	105471	y	e	s
CG30415	106403	y	e	s
CG32402	103959	y	e	s
CG5874	106245	y	e	s
CG6815	110208	y	e	s
CG1341	108834	y	e	s
CG3661	108402	y	e	s
CG2960	108730	y	e	s
CG6384	108694	y	e	s
CG7669	108918	y		s
CG2095	105653	y	e	s
CG6227	110778	y	e	s
CG14127	106258	y	e	s
CG10280	101800	y	e	s
CG32428	104945	y	e	s
CG14605	105310	y	e	s
CG13096	108860	y	e	s
CG10480	110321	y	e	s
CG1515	105648	y	e	s
CG14491	100196	y		s
CG4975	107558	y	e	s
CG11994	105631	y		s
CG17707	108786	y		s
CG14788	108840	y		s
CG1109	108830	y	e	s
CG4501	105635	y	e	s
CG6220	108909	y	e	s
CG14787	101857	y	e	s
CG4651	109503	y	e	s
CG9975	106300	y		s
CG8383	106336	y		s

CG5341	105836	y		s
CG3975/ CG15267	106371	y		s
CG13951	108881	y		s
CG6095	108650	y	e	s
CG4663	108829	y	e	s
CG18094	100138	y	e	s
CG2925	105354	y	e	s
CG10851	101740	y	e	s
CG7516	105260	y	e	s
CG7398	105181	y	e	s
CG15347	107479	y	e	s
CG14275	107513	y	e	s
CG7766	110184	y		s
CG31612	108700	y	e	s
CG4008	110459	y	e	s
CG10309	108173	y	e	s
CG1703	105998	y	e	s
CG5242	101442	y	e	s
CG3174	101452	y	e	s
CG14434	107085	y	e	s
CG4494	105980	y		s
CG13742	107138	y		s
CG10604	106634	y	e	s
CG8954	107160	y	e	s
CG5904	109662	y	e	s
CG4164	108576	y	e	s
CG8205	107575	y	e	s
CG5524	107055	y	e	s
CG6791	106627	y	e	s
CG12357	107112	y	e	s
CG6751	107563	y	e	s
CG7985	107448	y	e	s
CG9177	105992	y	e	s
CG30409	107093	y	e	s
CG6686	107167	y	e	s
CG1546	107425	y	e	s
CG42673	108571	y	e	s
CG6370	107068	y	e	s
CG3267	105961	y	e	s
CG31692	108554	y	e	s
CG9005	106589	y	e	s
CG14906	107782	y	e	s
CG3959	108606	y	e	s
CG32350	107420	y	e	s
CG30046	106940	y	e	s
CG3361	106951	y	e	s
CG9072	110249	y	e	s

CG1970	108068	y		s
CG17896	107006	y		s
CG8233	108002	y	e	s
CG32743	108450	y	e	s
CG3262	105489	y	e	s
CG1673	110229	y	e	s
CG6593	105525	y	e	s
CG11253	105377	y	e	s
CG8207	104395	y	e	s
CG8987	106955	y	e	s
CG1438	106954	y		s
CG31973	104771	y	e	s
CG3402	110191	y	e	s
CG30488	100513	y	e	s
CG15715	108387	y	e	s
CG12389	104362	y	e	s
CG8332	104439	y	e	s
CG18783/45074	107935	y		s
CG9384	109285	y	e	s
CG11761	108456	y	e	s
CG32446	104437	y	e	s
CG11403	105364	y	e	s
CG10801	100397	y	e	s
CG4207	104412	y	e	s
CG7392	106971	y	e	s
CG5809	108439	y	e	s
CG33103	108005	y	e	s
CG3740	105455	y		s
CG7055	104361	y	e	s
CG1404	104417	y	e	s
CG11121	104386	y	e	s
CG11325	109300	y	e	s
CG11586	104359	y	e	s
CG6279	105378	y		s
CG3180	110216	y		s
CG3258	108511	y		s
CG2964	109509	y		s
CG3948	104405	y		s
CG2655	104381	y	e	s
CG32721	108441	y	e	s
CG4673	109309	y	e	s
CG5941	104349	y	e	s
CG6413	108013	y	e	s
CG10881	106972	y	e	s
CG6027	109409	y	e	s
CG7959	106492	y		s
CG10932	107027	y	e	s

CG10672	109311	y		s
CG6764	105488	y	e	s
CG13022	108033	y	e	s
CG1354	104415	y	e	s
CG15732	100422	y	e	s
CG18104	104407	y	e	s
CG8908	100472	y	e	s
CG6289	105515	y		s
CG9078	106665	y		s
CG33197	105486	y	e	s
CG9954	109303	y		s
CG6851	106996	y	e	s
CG1135	108017	y	e	s
CG3509	100451	y	e	s
CG3699	101959	y	e	s
CG11086	100413	y		s
CG1488	107664	y	e	s
CG18250	107029	y		s
CG1945	107716	y	e	s
CG6488	104383	y	e	s
CG31137	104442	y	e	s
CG17018	106964	y	e	s
CG9327	104373	y	e	s
CG6667	105491	y	e	s
CG17560	104756	y	e	s
CG10281	110225	y	e	s
CG9331	107680	y	e	s
CG6657/44239	104411	y	e	s
CG2060	108025	y	e	s
CG15475	100509	y	e	s
CG4920	102357	y	e	s
CG8493	100591	y	e	s
CG2194	100560	y	e	s
CG31102	100543	y	e	s
CG13379	100581	y	e	s
CG14323	101160	y	e	s
CG7144	109650	y	e	s
CG6794	100537	y	e	s
CG1089	100526	y	e	s
CG32177	104434	y	e	s
CG10090	101139	y	e	s
CG10043	100583	y	e	s
CG13232	100571	y	e	s
CG13898	100929	y	e	s
CG8750	100551	y	e	s
CG14883	100589	y	e	s
CG1049	100575	y	e	s

CG10075	100623	y	e	s
CG11313	101266	y	e	s
CG4545	100584	y	e	s
CG4405	100555	y	e	s
CG6762	100627	y	e	s
CG2278	100549	y	e	s
CG6868	100930	y	e	s
CG5559	100957	y	e	s
CG7293	101058	y		s
CG4899	104428	y		s
CG8677	110463	y		s
CG14569	101158	y		s
CG12290	100939	y		s
CG11971	100553	y	e	s
CG18540	100538	y	e	s
CG7954	100582	y	e	s
CG7422	101136	y		s
CG12664	110068	y		s
CG9198	110343	y		s
CG32246	101262	y		s
CG14131	101269	y	e	s
CG13310	101276	y	e	s
CG2291	101157	y	e	s
CG2691	110718	y		s
CG5720	110710	y		s
CG3897	101083	y	e	s
CG3621	109816	y	e	s
CG9706	100567	y	e	s
CG5442	104978	y	e	s
CG32239	105252	y	e	s
CG2674	103143	y	e	s
CG18467	100539	y	e	s
CG9918	101115	y	e	s
CG32392	100592	y	e	s
CG31650	100590	y	e	s
CG31247	101175	y	e	s
CG6637	110350	y	e	s
CG6936	102303	y	e	s
CG8913	101098	y	e	s
CG1591	110156	y	e	s
CG8678	100536	y	e	s
CG43366	101186	y	e	s
CG14102	101155	y	e	s
CG4711	101017	y	e	s
CG5052	100534	y	e	s
CG17931	100894	y	e	s
CG1736/CG none	100528	y	e	s

CG5656	110733	y	e	s
CG10862	101113	y	e	s
CG8100	101026	y	e	s
CG12582	110464	y	e	s
CG6251	100588	y	e	s
CG9125	101165	y	e	s
CG4213	100535	y	e	s
CG9196	100897	y	e	s
CG13409	100569	y	e	s
CG42726	101163	y	e	s
CG31369/CG45263	102322	y	e	s
CG31740	101156	y	e	s
CG1389	101154	y	e	s
CG12919	108814	y	e	s
CG13317	105024	y	e	s
CG42244	101189	y	e	s
CG1796	101441	y	e	s
CG1744	108053	y	e	s
CG7051	101248	y	e	s
CG5458	109403	y	e	s
CG9734	101830	y	e	s
CG2790	101619	y	e	s
CG6801	101038	y	e	s
CG18103/ CG44122	102440	y	e	s
CG10682	107720	y	e	s
CG1358	101453	y	e	s
CG31453	110161	y	e	s
CG42797	103302	y	e	s
CG9210	110750	y	e	s
CG12945	100936	y	e	s
CG9854	105978	y	e	s
CG17734	102605	y	e	s
CG5671	101475	y	e	s
CG12529	101507	y	e	s
CG12203	101489	y	e	s
CG13601	110752	y	e	s
CG8640	103252	y	e	s
CG12045	103041	y	e	s
CG15101	103249	y	e	s
CG8585	110274	y	e	s
CG10583	106237	y	e	s
CG17816	103210	y		s
CG3675	109925	y		s
CG34398	103861	y		s

CG14741	102648	y		s
CG33483	102589	y		s
CG4379	101524	y	e	s
CG2750	101287	y		s
CG13045	102537	y		s
CG10631	101159	y		s
CG2781	102543	y		s
CG42699	102511	y		s
CG8422	110708	y	e	s
CG8505	109968	y	e	s
CG32810	108816	y		s
CG9629	101544	y		s
CG12025	104336			s
CG12242	102234			s
CG12602	106611	y	e	s
CG7542	108973			s
CG5905	108660	y	e	s
CG5427	105560			s
CG7874	102395			s
CG6839	105598			s
CG8808	106641			s
CG11199	106588	y	e	s
CG10559	106623			s
CG1705	100638	y	e	s
CG2974	110202			s
CG1737	106613			s
CG15623	106603			s
CG18480	106622	y	e	s
CG32580	105591			s
CG10584	106620			s
CG34166	102074	y		s
CG31676	100658			s
CG30401	105571			s
CG9375	106642			s
CG12964	105593	y		s
CG43225	106616	y		s
CG16716	106602			s
CG5857	101264	y		s
CG17330	103958	y		s
CG1444	110678			s
CG5133	104927			s
CG10992	108315			s
CG10794	102607	y	e	s
CG6554	110391			s
CG17104	107804	y	e	s
CG3287/44154	108236	y	e	s
CG30203	108235	y	e	s

CG2022	103027			s
CG30272	105475	y		s
CG11340	108337			s
CG17531	104977	y	e	s
CG17856	102408	y	e	s
CG13630	107078	y	e	s
CG4363	102579	y	e	s
CG9662	100769	y	e	s
CG4421	106022	y	e	s
CG4005	104523			s
CG12359	106625			s
CG7629	104918			s
CG30187	106612	y	e	s
CG3669	106607			s
CG6113	106614	y	e	s
CG9159	109910	y	e	s
CG32473	106586	y	e	s
CG3620	105676			s
CG12507	106624			s
CG3700	108237			s
CG5322	106609	y	e	s
CG1789	103722			s
CG1956	110757	y	e	s
CG7298	100861	y	e	s
CG32843	101995			s
CG7830	105649	y	e	s
CG8869	105042	y	e	s
CG3143	106097			s
CG33156	103902			s
CG11128	110425	y	e	s
CG8846	100739			s
CG18594	101957			s
CG13061	105584			s
CG14995	109283	y	e	s
CG4317	108018			s
CG11382	105782			s
CG1615	104883	y		s
CG6341	106636	y	e	s
CG5638	109688	y	e	s
CG15896	106643			s
CG10264	106590			s
CG30440	108605	y	e	s
CG2701	106635			s
CG6948	106632	y	e	s
CG17352	100646	y	e	s
CG6151	104993			s
CG14168	103225			s

CG16885	102569			s
CG8586	106597	y	e	s
CG34100	108478	y	e	s
CG42327	106630			s
CG12947	105561			s
CG15395	106617			s
CG11081	107004	y	e	s
CG10346	100830	y	e	s
CG11413	106939			s
CG42458	106608			s
CG8284	106600	y		s
CG11144	103736	y		s
CG12763	104287			s
CG30265	106618	y		s
CG14584	109324	y		s
CG31199	106621	y		s
CG18278	109944	y		s
CG43313	106610	y		s
CG11907	109885	y	e	s
CG34442	109877	y		s
CG14045/CG43947	109769	y	e	s
CG14300	105793	y		s
CG18773	109243	y		s
CG5537	110072	y		s
CG9914	106649	y		s
CG40440	109257			s
CG40801	109253			s
CG3127	110081			s
CG1725	109274			s
CG7287	109231	y		s
CG16712	109238	y		s
CG32474	110381	y		s
CG6755	110387	y		s
CG42671	109186	y		s
CG8629	109198	y		s
CG9807	109220			s
CG41343	109249			s
CG32164	109183			s
CG32172	109270		e	s
CG12338	108398			s
CG11242	103943			s
CG1826	110685		e	s
CG32279	109207			s
CG6016	110371		e	s
CG11369	109778		e	s
CG13314	107670		e	s

CG34251	109218		e	s
CG10174	109227		e	s
CG3889	110477			s
CG14265	109262		e	s
CG18064	109244		e	s
CG18350/CG43770	109221	y	e	s
CG15112	106484	y		s
CG4216	109195	y		s
CG10675	102615	y		s
CG1179	109233	y	e	s
CG13270	102613	y		s
CG32451	110379	y	e	s
CG32706	109212			s
CG1527	109250	y	e	s
CG10654	105179			s
CG7740	102612	y		s
CG12559	109200	y	e	s
CG8811	107043			s
CG34267	109192			s
CG2412	102617	y	e	s
CG11741	109223	y		s
CG12592	109213	y		s
CG10495	109883	y		s
CG8441	110365			s
CG13810	109780	y	e	s
CG1986	109032	y	e	s
CG7565	110362	y	e	s
CG8486/CG44122	105132	y	e	s
CG1034	109235	y	e	s
CG31960	109247	y	e	s
CG40733	109267	y	e	s
CG13041	109222			s
CG14013	110065	y	e	s
CG12090	110386	y		s
CG7717	110339	y	e	s
CG34450	109884	y	e	s
CG6342	110637	y		s
CG42329	110085	y		s
CG34293	109163	y	e	s
CG31362	109209	y	e	s
CG34268	109208	y	e	s
CG42323	109182	y	e	s
CG7220	104478		e	s
CG10041	110063		e	s
CG12559	109193		e	s

CG10473	102407		e	s
CG17610	101701		e	s
CG6754	110366		e	s
CG33508	110084		e	s
CG42280	110361		e	s
CG31465	110061		e	s
CG1675	110351		e	s
CG41527	109276	y	e	s
CG32671	109230			s
CG34112	109272	y		s
CG7111	104470	y		s
CG10039/CG43774	103237	y		s
CG7377	109269	y		s
CG5475	102484	y	e	s
CG9066	110480	y	e	s
CG17946	109196	y	e	s
CG1252	109191	y	e	s
CG6673	109255	y	e	s
CG40002	109239	y	e	s
CG9028	105723			s
CG31522	106652	y	e	s
CG3021	107045	y	e	s
CG3622	104737	y	e	s
CG32616	109155			s
CG42668	110074			s
CG40733	109264	y	e	s
CG31230	104060	y	e	s
CG3157	107572	y	e	s
CG31196	108129	y	e	s
CG3068	108446	y	e	s
CG5586	110397	y	e	s
CG33957	100969	y	e	s
CG7643	110572	y	e	s
CG8887	108832	y	e	s
CG6103	101512	y	e	
CG8418	104782	y	e	
CG17743	110466	y	e	
CG4760	101435	y	e	
CG4202	103352	y	e	
CG17031	105585	y	e	
CG6375	106078	y	e	
CG15440	106754	y	e	
CG6224	105407		e	s
CG3936	100002		e	s
CG7186	105102		e	s
CG7003	108076		e	s

CG34403	108679		e	s
CG13758	106381		e	s
CG5373	100296		e	s
CG4173	110652		e	s
CG6127	108348		e	s
CG32858	105747		e	s
CG32191	101578		e	s
CG32484	101018		e	s
CG9067	102015		e	s
CG4843	107970		e	s
CG33094	104580		e	s
CG18361	101525		e	s
CG11183	105638		e	s
CG4313	107434		e	s
CG4252	103624		e	
CG1007	100587		e	
CG12759	104183		e	
CG6438	110788		e	
CG9985	101624		e	
CG17212	107760		e	
CG7659	110705	y		
CG8908	100472	y		
CG6289	105515	y		
CG6502	107072			s
CG4859	101505			s
CG16858	106812			s
CG33950	110494			s
CG33261	106433			s
CG6805	107728			s
CG10182	105599			s
CG8445	107757			s
CG5047	110446			s
CG1597	108675	y		s
CG4676	108669	y	e	s
CG4919	108737	y	e	s
CG12387	108756	y	e	s
CG15427	108746	y	e	s
CG16836	103059	y	e	s
CG3346	110738	y	e	s
CG1599	108733	y	e	s
CG9339	108736	y	e	s
CG31072	108740	y	e	s
CG4422	108693	y	e	s
CG31002	108768	y	e	s
CG33062	108751	y	e	s
CG6440	108760	y	e	s
CG34346/45477	108769	y	e	s

CG15141	108728	y	e	s
CG6972	108757	y	e	s
CG12403	108701	y	e	s
CG32280	108754	y	e	s
CG30472	108777		e	s
CG15279	108759		e	s
CG3178	108661		e	s
CG8548	108741	y	e	s
CG1502	108750	y	e	s
CG12773	108667	y	e	s
CG11881	108799	y	e	s
CG8648	108738	y	e	s
CG7458	108770	y	e	s
CG6477	108743	y	e	s
CG9962	108721	y	e	s
CG31675	108765	y	e	s
CG10254	108657	y	e	s
CG17321	108762	y	e	s
CG33639	108753	y	e	s
CG14007	108776	y	e	s
CG13331	108752	y	e	s
CG10089	108744	y	e	s
CG14088	108687	y	e	s
CG6993	108732	y	e	s
CG18041	108668	y	e	s
CG6569	108702	y	e	s
CG11811	110740	y	e	s
CG17977	108774	y	e	s
CG5220	108672	y	e	s
CG8085	108670	y	e	s
CG30035	103045	y	e	s
CG15177	103053	y	e	s
CG42820	108684			s
CG14982	108704	y	e	s
CG9796	103038	y	e	s
CG9631	108696	y	e	s
CG17064	110670	y	e	s
CG43109	103080	y	e	s
CG14445	103056	y	e	s
CG13054	103057		e	s
CG1844	103046	y	e	s
CG4552	110700	y	e	s
CG6470	108699	y	e	s
CG4766	109483	y	e	s
CG40198	109980	y	e	s
CG18155	109492	y	e	s
CG33872	109540	y	e	s

CG40188	109995	y	e	s
CG3560/CG32576	109542	y	e	s
CG31618	109541	y	e	s
CG5510	109494	y	e	s
CG6106	109517	y	e	s
CG34448	109548	y	e	s
CG14877	109320	y	e	s
CG6961	109951	y	e	s
CG15699/CG44162	109937	y	e	s
CG32132/CG43749	109928	y	e	s
CG42676	109330	y		s
CG33083	104734	y	e	s
CG3345	109477	y	e	s
CG10880	109476	y	e	s
CG12487	109357			s
CG12680	109484	y	e	s
CG32425	109975	y	e	s
CG34249/CG45057	109957	y	e	s
CG15145	109489	y	e	s
CG13031	109474	y	e	s
CG10812	109350	y	e	s
CG17515	109991	y	e	s
CG4286	106482	y	e	s
CG2789	106470	y	e	s
CG5721	106501	y	e	s
CG13519	105908	y	e	s
CG18635	104764	y	e	s
CG16790	104333	y	e	s
CG16733	105889	y	e	s
CG11836	109470	y	e	s
CG7544	106472	y	e	s
CG16964	109926		e	s
CG31662	109329	y	e	s
CG4822	105922	y	e	s
CG42665	105885	y	e	s
CG31861	109480	y	e	s
CG1467	109504	y	e	s
CG7313	105867			s
CG11262	105920		e	s
CG31472	105941		e	s
CG13576	105895	y		s

CG5116	106458	y		s
CG1014	104759		e	s
CG15593	105915	y		s
CG33461	105904	y		
CG13760	106483	y	e	s
CG12768	105917	y	e	s
CG12393	104339			s
CG12309	105894	y	e	s
CG1319	106473			s
CG15446	104760	y	e	s
CG3517	104742	y	e	s
CG14814	106454	y	e	s
CG32986	105891	y	e	s
CG6966	105934	y	e	s
CG7204	105892	y	e	s
CG7280	105942	y	e	s
CG31839	105925		e	
CG31773	104738	y	e	s
CG43366	109932	y	e	s
CG9945	105944	y	e	s
CG6450	106476	y	e	s
CG6977	105901	y	e	s
CG14506	105907	y	e	s
CG1587	106498	y	e	s
CG34398	105909	y	e	s
CG9492	105898	y	e	s
CG1927	105926	y	e	s
CG9186	105945	y	e	s
CG11261	105927	y	e	s
CG12797	105939	y	e	s
CG6607	110240	y	e	s
CG1163	105937	y	e	s
CG1516	105936	y	e	s
CG18578	105923	y	e	s
CG31547	105911	y	e	s
CG6788	101786	y	e	s
CG9090	101848	y	e	s
CG30150	101783	y	e	s
CG13004	101281	y	e	s
CG6518	101719	y	e	s
CG10005	101799	y	e	s
CG15599	108268	y	e	s
CG33673	101780	y	e	s
CG10067	102129	y	e	s
CG10536	101755	y	e	s
CG16813	101794	y		s
CG30445	101858	y		s

CG33958	101861	y		s
CG17525	100986	y		s
CG5174	101767	y		s
CG15520	101705	y		s
CG15399	101711	y		s
CG13643	101773	y		s
CG6017	101736	y		s
CG9559	101125	y		s
CG5594	101742	y		s
CG42747	101775	y		s
CG4563	101849	y		s
CG5065	101744	y		s
CG12147	101875	y		s
CG14818	101863	y		s
CG42564	101703	y		s
CG7131	105507			s
CG31922	101747			s
CG17147	101722			s
CG4741	101770			s
CG3429	102368			s
CG43066	101768			s
CG42254	101795	y	e	s
CG17738	101840	y	e	s
CG31099	101788	y	e	s
CG17272	101784	y	e	s
CG42352	101725	y	e	s
CG13224	101854	y	e	s
CG6171	101741	y	e	s
CG7540	101757	y	e	s
CG8172	101724	y	e	s
CG7810	101850	y	e	s
CG7786	102305	y	e	s
CG42701	101745	y	e	s
CG15877	101698	y	e	s
CG13744	101735	y	e	s
CG32655	101720	y	e	s
CG18769	110781	y	e	s
CG7753	101798	y	e	s
CG1236	110779	y	e	s
CG2827	106308	y	e	s
CG12511	101785	y	e	s
CG3964	102197	y	e	s
CG8167	102158	y	e	s
CG11159	101748	y	e	s
CG33214	108284	y	e	s
CG2174/43657	101729	y	e	s
CG7333	101790	y	e	s

CG1867	101693	y	e	s
CG32016	101047	y	e	s
CG15595	101778	y	e	s
CG15862	101763	y	e	s
CG32243	101716	y	e	s
CG17799	101710	y	e	s
CG12123	101797	y	e	s
CG7895	101825	y	e	s
CG2944	101738	y	e	s
CG1584	101728	y	e	s
CG3056	101781	y	e	s
CG6217	106302	y	e	s
CG13788	101727	y	e	s
CG8997	101706	y	e	s
CG11656	108277	y	e	s
CG17237	100917	y	e	s
CG1391	110770	y	e	s
CG17364	101844	y	e	s
CG10521	100840	y	e	s
CG8948	110812	y	e	s
CG13640	101707	y	e	s
CG17048	101890	y	e	s
CG33927	101889	y	e	s
CG4688	101884	y	e	s

Appendix II

Following is the list of positive genes:

CG Number	Trans-ID				Gene name
CG6751	107563	**YKI**			No Child Left Behind (NCLB)
CG5874	106563	**YKI**			Negative elongation Factor A (NELFA)
CG4008	110459	**YKI**			MAP2/Uninitiated (Und)
CG8948	110812	**YKI**			Graf
CG10521	100840	**YKI**			Netrin-B
CG1391	110770	**YKI**			CalpD, small optic Lobes (SOL)
CG10638	102914	**YKI**			Aldo-keto reductase/AKR
CG11967	100233	**YKI**			
CG8878	100985	**YKI**			
CG13062	101832	**YKI**			
CG2621	101538	**YKI**			GKS-beta
CG14882	103984	**YKI**			
CG12072	106174	**YKI**	**EGFR**		warts
CG6964	107413	**YKI**	**EGFR**		Atrophin
CG4204	101542		**EGFR**		elongin B
CG8654	100112		**EGFR**		CG8654
CG5859	100004			**SOCS**	INTS8
CG5546	103926			**SOCS**	Med19
CG3780	107304			**SOCS**	Spx
CG10079	107130			**SOCS**	EGFR

Appendix III: Publications

Promoter Proximal Pausing Limits Tumorous Growth Induced by the Yki Transcription Factor in Drosophila

Sanket Nagarkar,* Ruchi Wasnik,* Pravallika Govada,* Stephen Cohen,† and L. S. Shashidhara*,‡

*Indian Institute of Science Education and Research (IISER), Pashan, Pune 411008, †Department of Cellular and Molecular Medicine, University of Copenhagen, 2200N, Denmark, and ‡Ashoka University, Sonepat, Haryana 131029, India

ORCID ID: 0000-0001-9306-1432 (L.S.S.)

ABSTRACT Promoter proximal pausing (PPP) of RNA polymerase II has emerged as a crucial rate-limiting step in the regulation of gene expression. Regulation of PPP is brought about by complexes 7SK snRNP, P-TEFb (Cdk9/cycT), and the negative elongation factor (NELF), which are highly conserved from Drosophila to humans. Here, we show that RNAi-mediated depletion of bin3 or Hexim of the 7SK snRNP complex or depletion of individual components of the NELF complex enhances Yki-driven growth, leading to neoplastic transformation of Drosophila wing imaginal discs. We also show that increased CDK9 expression cooperates with Yki in driving neoplastic growth. Interestingly, overexpression of CDK9 on its own or in the background of depletion of one of the components of 7SK snRNP or the NELF complex necessarily, and specifically, needed Yki overexpression to cause tumorous growth. Genome-wide gene expression analyses suggested that deregulation of protein homeostasis is associated with tumorous growth of wing imaginal discs. As both Fat/Hippo/Yki pathway and PPP are highly conserved, our observations may provide insights into mechanisms of oncogenic function of YAP—the ortholog of Yki in humans.

KEYWORDS tumorigenesis; Drosophila; Hippo pathway; promoter proximal pausing; transcription regulation in growth and cancer

REGULATION of growth is arguably the most critical phenomenon that establishes size and shape of all tissues, organs,andoverallbodysizeinmetazoananimals.Itisalsoan important homeostatic process, failure of which is linked to diseases and disorders, particularly cancer in humans. Regulated growth is achieved by an intricate interplay between factors promoting growth(oncogenes) andthosesuppressing it (tumor suppressors).

Yorkie (Yki),theDrosophila orthologoftheYes-Associated Protein 1 (YAP1), acts as a transcriptional cofactor that mediates the effects of the Hippo tumor suppressor pathway. The Hippo pathway is highly conserved from Drosophila to

doi: https://doi.org/10.1534/genetics.120.303419
Manuscript received June 8, 2020; accepted for publication July 27, 2020; published Early Online July 30, 2020.
Available freely online through the author-supported open access option.

Supplemental material available at figshare: https://doi.org/10.25386/genetics.12689318. 1 Corresponding author: Department of Biology, Main Bldg., IISER Pune, Dr Homi Bhabha Road, Pashan, Pune 411008. E-mail: ls.shashidhara@iiserpune.ac.in

humans (Pan 2010). The Hippo (Hpo)/MST kinases and the Warts (Wts)/LATS kinases and their cofactors form kinase cassettes that directly phosphorylate Yki (YAP/TAZ) to regulate protein stability and activity (Zhao et al. 2011). Members of this pathway were initially found to limit tissue growth in Drosophila by limiting Yki activity (Huang et al. 2005; Dong et al. 2007). Consistent with this, YAP overexpression has been reported as a driver of tissue growth and cancer in a mouse model (Dong et al. 2007; Zanconato et al. 2015). In humans, the YAP1 locus was found to be amplified in different types of cancer (Overholtzer et al. 2006; Zender et al. 2006). These findings have sparked a great deal of interest in understanding of regulation of Yki/YAP function.

In Drosophila, Yki regulates expression of regulators of cell growth and survival such as Diap1, dMyc, bantam, etc. Targets of YAP in humans include the EGFR-ligand AREG as well as CTGF, Cyr61 (Johnson and Halder 2014). While these target genes are necessary for growth induced by Yki/YAP activity, they are not sufficient to phenocopy effects of Yki/YAP. This indicates possibility of more regulators that are involved in Yki/YAP induced growth.

We have reported an in vivo screen in Drosophila (Groth et al. 2020), wherein we have identified a large number of

genes, which, when depleted, enhanced growth induced by Yki and EGFR. More importantly, these genes function like classical tumor suppressors as, when downregulated in the background of overexpressed Yki or EGFR, we observed neoplastic growth. Amongthese, weidentified anumber ofgenes involved in the control of promoter proximal transcriptional pausing.

Promoter proximal pausing (PPP) of RNA polymerase (Pol) II has been identified as a key step in transcriptional regulation for many genes, during development and in stem cells (Guenther et al. 2007; Muse et al. 2007; Zeitlinger et al. 2007). At paused loci, after initiation, RNA Pol II first passes through the promoter but then stops at 30–60 bp from the transcription start site (Kwak and Lis 2013). Productive transcription then requires release from the paused state. PPP is brought about by the negative transcription elongation factor (NELF) and 5,6-dichloro-1-b-D-ribofuranosylbenzimidazole (DRB)-sensitivity inducing factor (DSIF) protein complexes, which were identified as factors responsible for DRB sensitivity of transcription elongation (Wada et al. 1998; Yamaguchi et al. 1999). These complexes bind RNA Pol II and halt its progress downstream of the

promoter. This pause is alleviatedby a positive transcription elongation factor complex (P-TEFb) (Figure 1A), which consists of cyclin T and cyclin dependent kinase-CDK9 (Marshall and Price 1995). Once recruited to the paused complex, CDK9 phosphorylates NELF and DSIF leading to ejection of NELF from the paused complex while DSIF assists Ser-5 phosphorylated RNA Pol II in productive elongation (Jonkers and Lis 2015). The PTEFb complex is, in turn, regulated through sequestration by 7SK snRNP complex. P-TEFb is required for release paused RNA polymerase II into productive elongation (Kwak and Lis 2013). Sequestration of P-TEFb by 7SK snRNP leads to its unavailability for mediating pause release, which, in turn, regulates transcription elongation via sustained pause of RNA Pol II. Mammalian 7sk-snRNP complex consists of 7sk RNA, Hexim1/2, Larp7, and MePCE. Drosophila homologs of components of mammalian 7sk-snRNP complex were identified and characterized recently (Nguyen et al. 2012). These include Bin3 (MePCE ortholog), Larp (Larp7 ortholog), Hexim (HEXIM1/2 ortholog), and d7SK RNA. All are highly conserved at functional levels with their mammalian counterparts. Signaling events of pathways such as ERK, TCR, etc., trigger liberation of P-TEFb. Thus, making sequestration and liberation of P-TEFb a context dependent process that is critical for regulating expression of gene regulation depending on the context.

Interestingly, CDK9 has been shown to be important for transcription of target genes of oncogenes such as Myc (Kanazawa et al. 2003) and YAP (Galli et al. 2015). Here, we present evidence of tumor suppressor function of various components involved in PPP, specifically in the context of elevated Yki activity. Our findings show that factors involved in PPP and its regulation are important to restrict Yki driven growth and to prevent neoplastic transformation in vivo.

Material and Methods

Drosophila strains

The following Drosophila strains are used in this study: ap-Gal4 (Cohen et al. 1992) and UAS-Yki (Huang et al. 2005). The following RNAi stocks were obtained from the Vienna Drosophila RNAi Center and Bloomington Drosophila stock Center: UAS-NelfA^{RNAi} (KK106245, TRiP #32897), UAS-NelfB^{RNAi} (KK108441, TRiP #42547), UAS-NelfE^{RNAi} (TRiP # 32835), UAS-NelfD^{RNAi} (KK100009, TRiP # 38934, #42931), UAS-bin3^{RNAi} (KK101090, TRiP #41527), UAS-Hexim^{RNAi} (KK100500). UAS-CDK9 was obtained from FlyORF (#F001571).

Spatio-temporal regulation of transgene expression in wing imaginal disc

The apterous enhancer was used to drive expression of Gal4 conditionally in dorsal compartment of wing imaginal discs. Gal4 activity was regulated using Gal80^{TS}, which allows expression of transgenes at permissive temperature of 29 as against restrictive 18 temperature. In all experiments, tubulin-Gal80^{TS} wasused.Drosophilacrosseswereallowedtolayeggsfor3 days at 18, and were then flipped or discarded. Larvae were then allowed to grow for additional 5 days before switching to temperature of 29. At 29 they were maintained for 4–14 days. All crosses were using tubulin-GAL80^{TS}; ap-GAL4; UAS-GFP. Thus, all experimental crosses had one copy of GFP, while control crosseshadtwocopiesofGFP.Detailedmethodologyisprovided in Groth et al. (2020). Larval images were taken in bright field and in GFP channel with a Leica stereomicroscope. Image processing was done using Adobe Photoshop 6 and ImageJ.

Immunohistochemistry

Thefollowingprimaryantibodieswereused:ratanti-Ecadherin, mouse anti-MMP1 (Developmental Studies Hybridoma Bank). Rhodamine-phalloidin (ThermoFisher Scientific, Cat no R415) was used to stain actin in tissue.

Third instar larvae were dissected in PBS. Samples were fixed in 4% PFA for 20 min, followed by three 10-min washes in PBT (PBS-Tween20) at room temperature. Then, 5% BSA in PBS was used for blocking followed by overnight incubation in primary antibody at 4. Next day, the samples were washed with PBT, three times for 10 min each followed by incubation with secondary antibody for 2 hr at room temperature. Samples were

then washed with PBT and stained for DNA using 49,6-diamidino-2-phenylindole (DAPI; Sigma Aldrich) for 5 min. Wing disc tissue was then mounted on slides in Antifade Gold mountant (ThermoFisher Scientific). Imaging was done on a Leica SP8 confocal laser-scanning microscope. Image processing was done using ImageJ and Adobe Photoshop 6. Measurement MMP1 intensities and comparison between different genotypes was carried out using ImageJ, statistical analysis(one-wayANOVA)wasdoneusingPrism-Graphpad5.

RNA-seq

Induction procedure for transgenes was followed as mentioned earlier. Wing imaginal disc tissue was dissected on

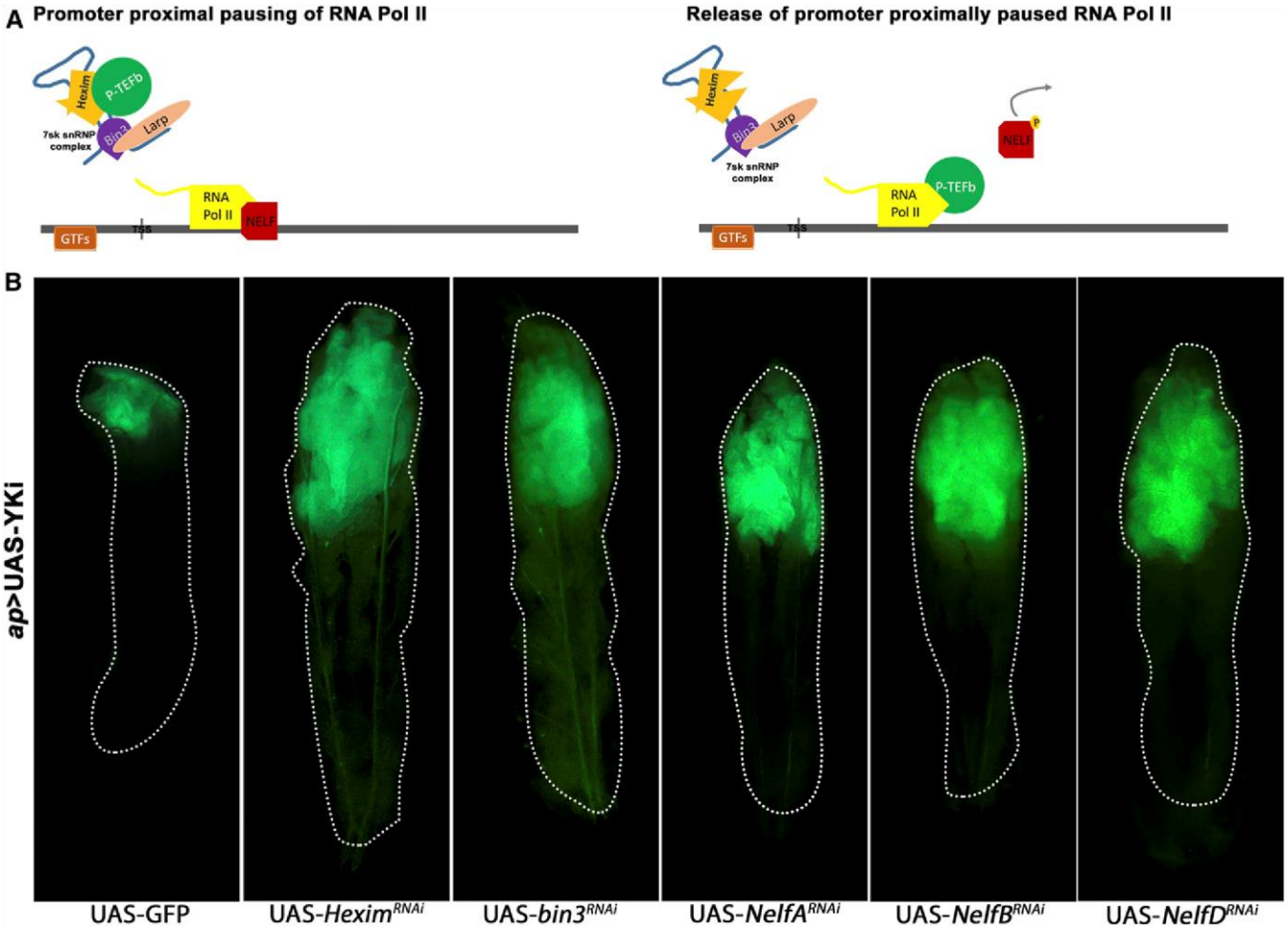


Figure 1 Identification of complexes involved in promoter proximal pausing as tumor suppressors. (A) A schematic representing known function of two complexes we identified as candidate tumor suppressors. The 7SK snRNP complex regulates promoter proximal pausing by sequestring the P-TEFb complex, while the NELF complex is involved in the formation of a stall of RNA Pol II at the promoter proximal region. As dictated by surrounding cues, P-TEFb is released by the 7SK snRNP complex. Thus, freed P-TEFb is recruited to stalled RNA Pol II, where it brings about release of RNA Pol II from the paused state. (B) Larval images showing wing imaginal discs expressing GFP at low magnification. Dimensions of GFP-expressing tissue is indication of growth in imaginal discs. Top row: Larvae overexpressing only Yki (crossed to UAS-GFP as control) and those in combination with RNAi-mediated knockdown of 7sksnRNP components: Hexim or bin3 using GAL80^{TS}; ap-GAL4; UAS-GFP. Bottom row: wing discs overexpressing Yki in combination with RNAi-mediated knockdown of NELF components (from left to right) NelfA, NelfB, and NelfD (also known as TH1) using GAL80^{TS}; ap-GAL4; UASGFP. Note significantly larger GFP expressing-wing discs (green) in larvae that are overexpressing Yki and also depleted for a component of PPP.

4th–5th day after induction for ap . GFP, ap . UAS-Yki, ap . Nelf-A RNAi (KK106245), ap . UAS-Yki, UAS Nelf-A RNAi. Larvae were washed in RNase-, DNase-free ultrapure water (GiBCO), and then dissections were done in RNase-, DNasefree PBS (GiBCo). Number of wing imaginal discs collected was 150, 70, 150, 25, respectively for ap . GFP, ap . UASYki, ap . Nelf-A RNAi (KK106245), ap . UAS-Yki, and UAS Nelf-A RNAi. Collection was done in TRiZOL reagent (ThermoFisher Scientific). Each genotype was collected in three biological replicates. RNA sequencing was done on an Illumina platform.

RNA-seq data analysis

RNA-seqanalysiswasperformedusingtheHISAT2.0package protocol as explained in Pertea et al. (2016). To identify significantly differentially expressing genes in different combinations of comparisons, DEseq package and EdgeR were used (Anders and Huber 2010). The entire RNA-seq data set is available on GEO database (https://www.ncbi.nlm. nih.gov/geo/query/acc.cgi?acc=GSE151935).

The list of genes obtained was then used as input for the web-based tool venny (http://bioinfogp.cnb.csic.es/tools/ venny/index.html) to obtain a list of genes that are unique to each genotype, overlapping between all three or combination of any two genotypes.

Gene ontology analysis

Forgeneontology(GO)andpathwayenrichmentanalysis,we utilized STRING10 (Szklarczyk et al. 2017). We used gene lists that are significantly differentially expressed in single

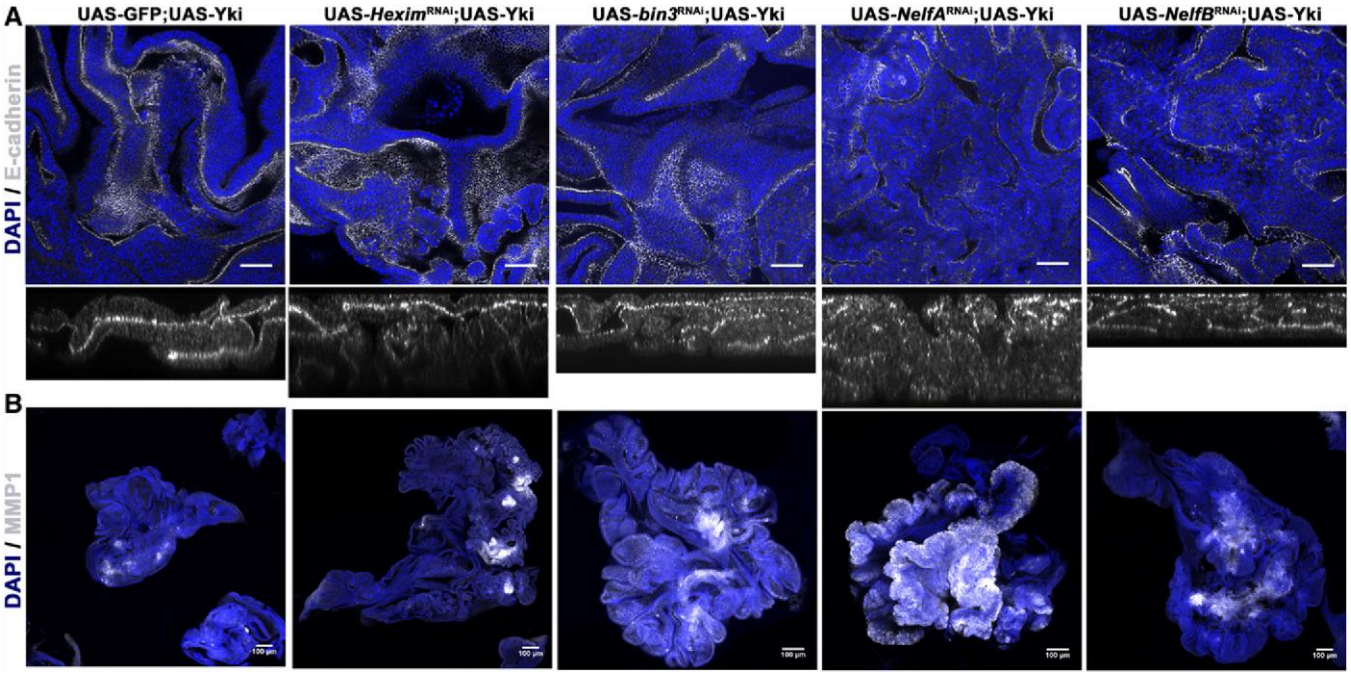


Figure 2 Characterization of tumors induced in the wing disc. (A) Disruption of characteristic epithelial apico-basal polarity in tumor discs. Images of wing discs overexpressing Yki alone (crossed to UAS-GFP as control) or in combination with RNAi-mediated knockdown of Hexim, bin3, NelfA or NelfB using GAL80^{TS}; ap-GAL4; UAS-GFP (Bar, 10 mm). Discs are stained for E-Cadherin (white) expression and localization. Bottom panel of each image shows orthogonal optical section of respective genotype. Note delocalization of E-Cad in tumorous tissues caused by the depletion of a component of PPP and Yki overexpression (higher magnification images are shown for few genotypes). All discs are also stained with DAPI (blue) to visualize nuclei. (B) Increased expression of MMP1 is observed in tumor discs. Images of wing discs overexpressing Yki alone (crossed to UAS-GFP as control) or in combination with RNAi-mediated knockdown of Hexim, bin3, NelfA, or NelfB using GAL80^{TS}; ap-GAL4; UAS-GFP (Bar, 100 mm). Wing discs are stained for MMP1 (white). Note increased MMP1 staining in tumorous tissues caused by the depletion of a component of PPP and Yki overexpression.
All discs are also stained with DAPI (blue) to visualize nuclei.

genotype or a combination of genotypes as mentioned in the results section, as input to the STRING. The output files were downloaded as interaction network and list of genes from input that are enriched in different GO categories or as KEGG pathways.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. All Drosophila stocks are available upon request. RNA-seq data are available at GEO with the accession number: GSE151935. Supplemental material available at figshare: https://doi.org/10.25386/genetics.12689318.

Results

Depletion 7SK snRNP complex components cooperates with Yki in causing tumorous growth

Studies using Drosophila tumor models have found that larvae containing proliferating tumors are unable to enter pupariation and continue to grow (Gateff et al. 1993). The resulting giantlarva phenotype canbe usedingenetic screens to identify tumor-causing genotypes. We made use of this property to identify candidate genes in a genetic screen for tumor suppressors cooperating with Yki [the entire screen is published elsewhere (Groth et al. 2019)]. We found that RNAi-mediated depletion of bin3 or Hexim, components of the 7SK snRNA complex in combination with Yki overexpression led to massive overgrowth in wing disc tissue and giant larval phenotype (Figure 1B). Wing discs expressing Yki alone show only moderate overgrowth phenotype, and larvae eventually pupate (Figure 1B). Depletion of 7SK snRNP components did not produce overgrowth on their own (Supplemental Material, Figure S1), but only did so when coupled with Yki overexpression. We also did not observe wing disc overgrowth when depletion of 7SK snRNP components in combination with overexpression of other well-known oncoproteins such as epidermal growth factor receptor (EGFR) or notch intracellular domain (NICD) (Figure S2). Thus, our observations suggest that, Drosophila 7SK snRNP complex may function, specifically, to repress tumorigenic potential of Yki in vivo in an epithelial tumor model.

Components of the NELF complex may function as tumor suppressors

The NELF complex is composed of four subunits: NELF-A, -B, -C/-Dand-E.DepletionofeachoftheNELFcomponentsusing RNAi in combination with Yki also produced a giant larval phenotype (Figure 1B) and massively overgrown wing disc tissue compared to the larvae overexpressing only Yki (Figure 1B). Depletion of the NELF components on their own did not cause such giant larval phenotype or overgrowth of the wing disc tissue (Figure S1). These components too did not show any tumor phenotype in the context of overexpressed EGFR or NICD (Figure S2).

It was intriguing to find multiple components of the two spatio-temporally separated protein complexes, involved in the regulation of transcription elongation, among the tumor suppressors identified in a genome-wide screen for factors cooperatingwithYkiingrowthregulation(Grothetal.2019).

Neoplastic transformation induced by Yki combined with depletion of 7SK snRNP or NELF complexes

Yki is known to promote cell proliferation and cell survival. Thus, it is possible that larger size of the wing disc tissue observed upon loss of either 7SK snRNP or NELF complex is a resultofenhancementofgrowthandsurvivaleffectofYki,and not a neoplastic transformation. To distinguish between the two possibilities, we analyzed the tumor tissue using markers that indicate neoplastic transformation.

First,weexaminedepithelialcellpolarity.Neoplastictransformation of an epithelial tissue is accompanied by the loss of their characteristic apico-basal cell polarity. E-cadherin (E-Cad) is a subapically localized protein that provides a convenient marker for epithelial polarization (Tanos and Rodriguez-Boulan 2008). Wing discs overexpressing Yki alone showed localization of E-Cad, in a pattern similar to the wild-type wing discs, although the former discs are much larger (Figure 2A). This indicated that Yki overexpression caused overgrowth of the epithelium without perturbation of epithelial cell polarity. In contrast, when Yki overexpression was combined with depletion of a component of the 7SK snRNP complex or the NELF complex, subapical localization of E-cad was lost or perturbed (Figure 2A). Additionally, we analyzed F-Actin, which localizes near the apical junctions of the wing disc epithelial cells, using rhodamine-labeled phalloidin. As with E-Cad, we observed loss of apical localization of F-actin in the Yki expressing tissue depleted of a component of the 7SK snRNP or the NELF complex, but not in wing disc tissue expressing Yki alone (Figure S3).

We did not observe any change in cell polarity, as indicated by E-Cad or F-Actin localization in wing discs with depletion of components of 7SK snRNP and NELF complexes alone (Figure S4A; data not shown for F-Actin).

The matrix metallo-protease MMP1 has been used as a marker of epithelial to mesenchymal transition (EMT) and neoplastic transformation in Drosophila tumor models. MMP1expression is elevated in tumor models and its depletion by RNAi has been reported to block metastasis (Uhlirova and Bohmann 2006; Beaucher et al. 2007; Miles et al. 2011). We examined the effects of depleting components of 7SK snRNP and NELF complexes in Yki-expressing tissue on the levels of MMP1 by immunohistochemistry. We observed significantly elevated levels of MMP1 in wing discs

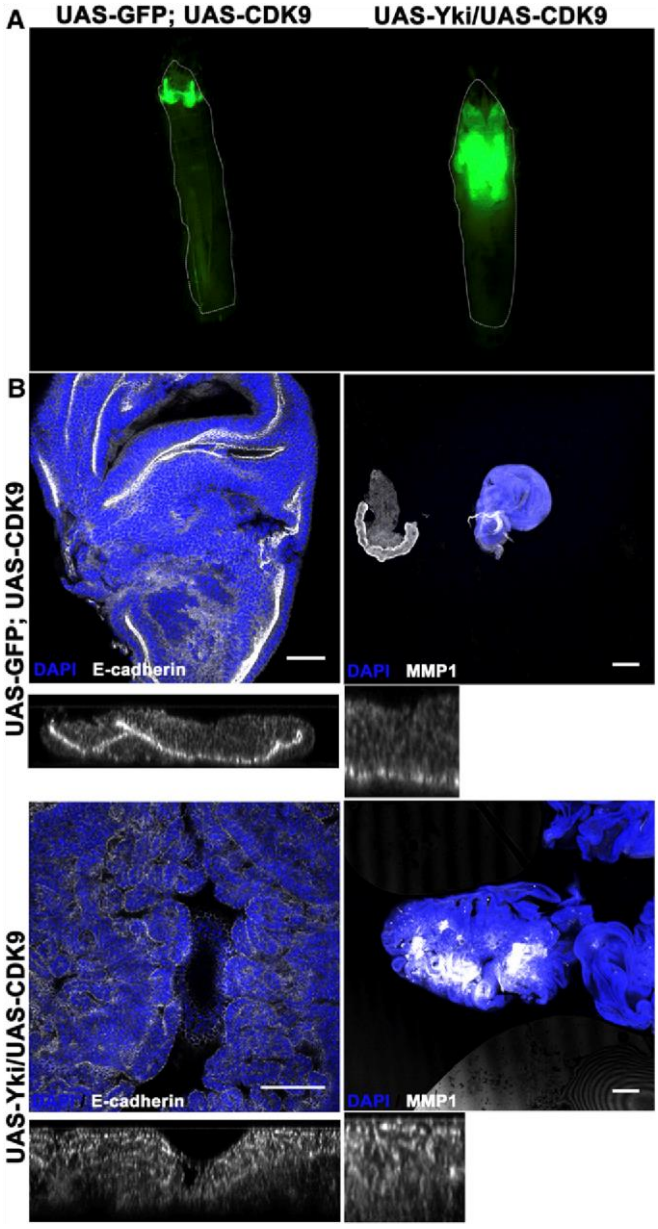


Figure 3 CDK9 cooperates with Yki in tumorigenesis. (A) Larval images showing growth observed in combination of UAS-CDK9 and UAS-Yki as compared to UAS-CDK9 alone (crossed to UAS-GFP as control) using GAL80^{TS}; ap-GAL4; UAS-GFP. The combined overexpression phenocopies the phenotypes observed in Figure 2B. (B) Characterization of tumor tissue caused by combined overexpression of CDK9 and Yki using GAL80^{TS}; ap-GAL4. Top row of images shows wing disc tissue overexpressing CDK9 alone, while the bottom row shows combined overexpression of CDK9 and Yki. Discs in the left column are stained for E-Cadherin (white) (Bar, 10 mm) and those in the right column are stained for MMP1 (white) expression (Bar, 100 mm). Please note deregulated E-cad localization (optical z-sections and two different magnification levels are shown below the discs) and increased MMP1 expression in tissues that overexpress both CDK9 and Yki, suggesting their neoplastic tumor state. All discs are also stained with DAPI (blue) to visualize nuclei. Both the discs stained for MMP1 are imaged at lower magnification (10X) for better comparison, as tumorous disc is too large to show at higher magnification.

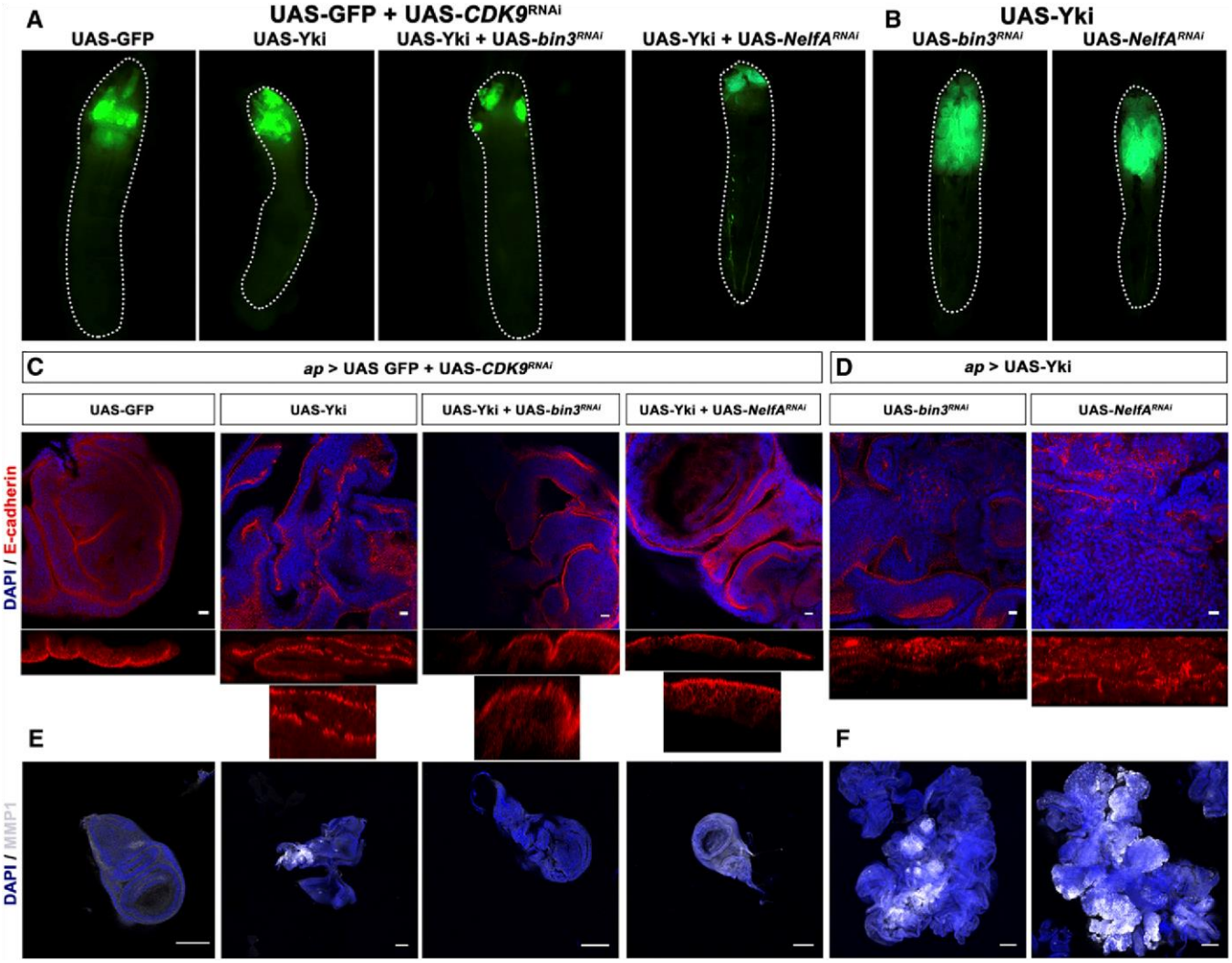


Figure 4 CDK9 is necessary for Yki-mediated tumorigenesis. (A) Loss of CDK9 rescues tumor phenotype. The images show GFP-expressing wing discs of various genotypes as indicated. Size of the wing discs may be discerned by the amount of larval space occupied by GFP-expressing tissue. RNAimediated depletion of cdk9 inhibited tumor formation caused by a combination of overexpression of Yki and depletion of a component of the PPP. The GFP-marked wing tissue is of the same size as in controls. All crosses were using GAL80^{TS}; ap-GAL4; UAS-GFP. (B) Tumorous wing disc phenotypes caused by the overexpression of Yki in the background of depletion of bin3 or NelfA shown here again as a control to (A). (C) Restoration of apico-basal polarity in wing disc tissue. The images show wing discs of various genotypes as indicated stained for E-Cad (red). RNAi-mediated depletion of cdk9 restored normal apical localization of E-Cad (optical z-sections are shown below the discs) in wing discs that overexpress Yki and are also depleted for a component of the PPP. All discs are also stained with DAPI (blue) to visualize nuclei (Bar, 10 mm). (D) Tumorous wing discs (stained for E-Cad) of larvae overexpressing Yki in the background of depletion of bin3 or NelfA shown here again as a control to (C). (E) Restoration of MMP1 levels. The images show wing discs of various genotypes as indicated stained MMP1 (white). RNAi-mediated depletion of cdk9 restored normal levels of MMP1 in wing discs that overexpress Yki and also depleted for a component of the PPP. All discs are also stained with DAPI (blue) to visualize nuclei (Bar, 100 mm). (F) Tumorous wing discs (stained for MMP1) of larvae overexpressing Yki in the background of depletion of bin3 or NelfA shown here again as a control to (E).

overexpressing Yki and depleted for Bin3, Hexim, or the NELF complex (Figure 2B, Figure S5A). We observed only marginal increase(statisticallyinsignificant)inMMP1levelsinwingdiscs expressing Yki alone (Figure 2B, Figure S5A). We did not observe any detectable change in the intensity of MMP1 levels in the wing discs depleted for the components of 7SK snRNP and NELF complexes alone (Figure S4B and Figure S5B).

Taken together, tumors formed upon depletion of 7SK snRNP or NELF complex components in combination with Yki exhibit neoplastic characters. As neither genetic change alone produced these results, it appears that they act in combination to promote neoplasia, a classical mechanism of cooperative tumorigenesis as known in mammals. These observations provide evidence that the activity of 7SK snRNP andNELFcomplexes mayhave atumor-suppressing function, but only in the context of elevated Yki activity.

CDK9 is required for Yki-mediated tumorigenesis

The 7SK snRNP and NELF complexes help in maintaining the paused state of RNA Pol II. Our findings raised the question of whether pausing of RNA Pol II per se served to limit the tumor promoting potential of Yki activity. If this is the case, we reasoned that using an alternative means to release RNA Pol II should also lead to tumorigenesis in the context of Yki overexpression. The P-TEFb complex, comprising cycT/CDK9, is required for release of paused RNA Pol II and effective elongation of mRNA. CDK9 phosphorylates the NELF complex, leading to eviction of NELF from the pause site. This in turn facilitates release of paused RNA Pol II, aiding in productive elongation. CDK9 also acts directly on RNA Pol II, phosphorylating it on S5 in the C-terminal domain, a known mark of elongating RNA Pol II (Jennings 2013). As the P-TEFb complex is normally rendered inactive through sequestration by 7SK snRNP complex, we hypothesized that overexpressing CDK9 might bypass normal regulation of pausing, leading to inactivation of NELF complex and RNA Pol II release. Consistent with this hypothesis, we indeed observed massive tissue overgrowth when Yki was co-expressed with CDK9, while overexpression of CDK9 alone did not cause any such phenotype (Figure 3A). Such overgrowth phenotype was not observed when CDK9 was overexpressed in the background of elevated activities of EGFRor Notch (Figure S2). This suggests that PPP-mediated regulation of growth is Yki-specific.

Wing discs expressing UAS-CDK9 together with UAS-Yki alsoshowedlossofapicallylocalizedE-Cadaswellaselevated MMP1 expression (Figure 3B), compared to tissue expressing UAS-CDK9 alone or UAS-Yki alone. This indicates neoplastic transformation in wing discs co-expressing Yki and CDK9, similar to the transformation caused by depletion of 7SK snRNP and NELF complex components in combination with overexpressed Yki.

As further test of this model, we asked whether CDK9 is essential for tumorigenic cooperation between depletion of 7SK snRNP complex components and Yki. Depletion of cdk9 effectively suppressed the tissue overgrowth caused by depleting bin3 or Hexim in Yki expressing tissue (Figure 4A). Those wingdiscs also showed normal apical localization of E-Cad and wildtype levels of MMP1 expression, suggesting complete suppression of tumorous growth (Figure 4, B and C).

Given that CDK9 is known to act directly on both NELF proteins and RNA Pol II, we wondered whether CDK9 activity would be required in the absence of the NELF complex. As shown above in the case of removing the 7SK snRNP complex, depletion of cdk9 suppressed overgrowth caused by RNAimediated depletion of NelfA and overexpression of Yki (Figure 4A). This was accompanied by restoration of apico-basal polarity and MMP1expression to wild-type levels(Figure 4, B and C). This finding provides evidence that alleviation of pausing by removal of NELF complex is not sufficient without CDK9 activity. This presumably reflects an importance of activation of RNA Pol II by CDK9-mediated phosphorylation.

Wethenexaminedifdepletionofthecomplexesassociated with PPP and increased CDK9 levels are sufficient to cause overgrowth phenotype, or whether the growth is tightly coupled to the presence of a growth driver such as Yki. Depletion of components of 7SK snRNP or NELF complexes in the background of overexpressed CDK9 did not cause any growth phenotype or morphological alteration in wing disc

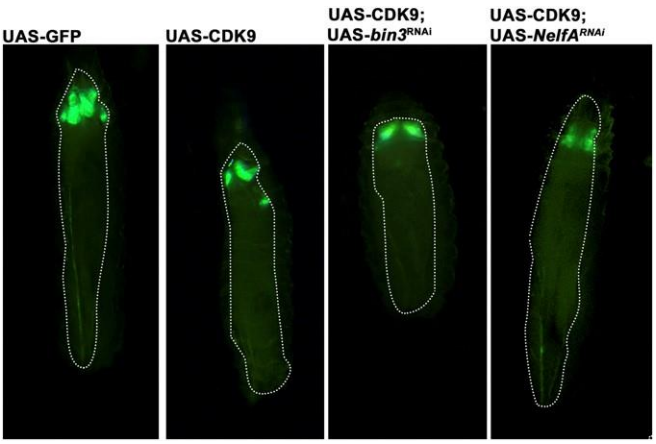


Figure 5 Yki is the driver of tumorigenesis. Larval images showing phenotype of UAS GFP in combination with (left to right) UAS-GFP, UASCDK9 followed by UAS-CDK9 and UAS-bin3^{RNAi}, UAS-CDK9, and UAS-NelfA^{RNAi}. None of them show overgrowth phenotype as observed when Yki is overexpressed, suggesting CDK9 may induce tumorous growth only in the context of overexpressed Yki. All crosses were using GAL80^{TS}; ap-GAL4; UAS-GFP.

epithelium (Figure 5). This suggests that deregulation of RNA Pol II pausing is not sufficient on its own to produce an overgrowth or neoplastic phenotype; yet it does so in the context of Yki overexpression. In the context of elevated Yki activity, there appear to be two brakes, each of which must be removed by CDK9 activity to allow excess Yki to produce tumors in Drosophila wing disc tissue.

Tumorigenesis induced by alleviation of pausing is associated with deregulated proteostasis

AsoverexpressionofYkiwasessential,althoughnotsufficient, to cause neoplastic tumors, genetic experiments above provided an opportunity to distinguish between Yki-activated genes that cause simple hyperplastic growth of the discs (when Yki is overexpressed in a wild-type background) vs. causing neoplastic growth (when Yki is overexpressed along with depletion of bin3, Hexim, or NELFs).

We carried out RNA-seq to identify differentially expressed genesindiscsdepletedforNelfAandoverexpressingYkiaswell as both individual treatments. We also carried out RNA-seq for GFP expressing wild-type wing discs as a control. We find that transcripts corresponding to 776 genes were uniquely upregulated (Figure 6A) and 1009 genes were uniquely downregulated (Figure 6B) in the tumorous wing discs (ap . UAS-Yki; UAS-NelfA^{RNAi}), compared to all other genotypes including wild-type discs (noncoding transcripts are not included in this estimation). When compared to the list of direct targets of Yki (reportedbybasedonChIP-seqdata),wefind38(4.9%)ofthe upregulated genes and 84 (8.3%) of the downregulated genes are presumptive direct targets of Yki (Table S1).

We also observed an enhancement of effect of Yki (compared to wildtype discs) in a subset of transcripts that were common to nontumorous tissue overexpressing Yki alone (ap . UAS-Yki) and tumorous ap . UAS-Yki; UAS-NelfA^{RNAi}

Table 1 List of genes whose expression is upregulated in the wing discs of ap-GAL4/UAS-NelfA^{RNAi}; UAS-Yki

Aminoacyl-tRNA biosynthesis		Ribosome		Ribosome biogenesis in eukaryotes	
Gene name	logFC	Gene name	logFC	Gene name	logFC
Slimp	2.124626	RpL24-like	1.256082	Non1	2.292457
Aats-leu	1.413973	RpL5	1.202779	Ns2	0.900983
Aats-thr	0.912237	RpL15	1.130382	RIOK1	1.354352
Aats-cys	0.813931	mRpL28	0.969423	CG12301	0.997310
Aats-tyr-m	1.047768	mRpL9	0.799329	Bka	0.876333
Aats-pro	1.07342	RpS17	0.82042	eIF6	0.745744
Aats-ile	0.73741	mRpL35	0.997681	l(3)72Dn	0.800876

CG4573	1.138148	RpS23	0.782542	CG8064	0.778604
CG1750	1.487797	RpS4	0.775449	Nmd3	0.716537
CG6796	0.925494	RpL27A	0.716573	Mat89Ba	0.713426
CG7441	0.884721	RpL32	0.681588	CG11920	0.750235
CG17259	0.726080	RpL40	0.674508	CG3071	0.713535
Aats-trp	0.732224	RpS29	0.743389	CG33158	0.595732
Aats-asp	0.747097	RpL26	0.620538	CG13185	0.823244
Aats-gly	0.613889	mRpL10	0.692489	CG7246	0.798345
CG5463	1.037030	mRpL3	0.671734	CG8549	0.593618
Aats-ala-m	0.602770	RpL35	0.631348		
CG5660	0.663614	RpL27	0.594275		
		RpL28	0.629209		
		RpL21	0.600412		
		RpL22-like	1.081389		
		RpS3A	0.587288		
		RpL37A	0.3662		
		mRpL11	0.624297		

tissue. We reasoned that since PPP functions to attenuate expression of genes, identifying transcripts whose expression is further up- or downregulated in ap . UAS-Yki; UAS$\mathrm{NelfA}^{\mathrm{RNAi}}$ tissue (compared to ap. UAS-Yki) may give a better indication of the role of PPP in Yki-mediated growth. We find that transcripts corresponding to 155 genes that are upregulated in both nontumorous ap . UAS-Yki discs and tumorous ap . UAS-Yki; UAS-$\mathrm{NelfA}^{\mathrm{RNAi}}$ discs, but degree of enhancement was higher in tumorous tissue. Likewise, these transcripts corresponding to 160 genes, whose expression was downregulated compared to wildtype discs, were common to both nontumorous and tumorous tissue, but degree of downregulation was higher in tumorous tissue. Interestingly,

31 (20%) of these upregulated genes (n = 155) and

35 (21.9%) of downregulated genes were presumptive direct targets of Yki, suggesting that we indeed have captured many targets of Yki that are regulated by PPP and misregulated due to RNAi medicated knockdown of many components of the pausing machinery.

We used genes corresponding to these transcripts to perform GO analysis in order to explore gene sets that show enrichmentandmightindicatepathwaysorprocessesthatare involved in tumorigenesis. For this purpose, the STRING tool was utilized (Szklarczyk et al. 2017). STRING output is based on statistical enrichment score of interactions obtained from the input compared to a random set of genes from the genome of the organism, in this case D. melanogaster. STRING also collates data from manually curated databases of interactions such as Kyoto Encyclopedia of Genes and Genomes (KEGG) and GO terms.

We observed enrichment for pathways involved in ribosome and its biogenesis in the upregulated set (Table 1 and Figure 6). Interestingly, protein processing in endoplasmic reticulum, regulators of proteasome function, and different components of proteasome were enriched among genes downregulated in tumorous tissues (Table 2). These observations indicate overall deregulation of protein homeostasis (proteostasis) in tumors caused by depletion of NelfA in combination with Yki overexpression, consistent with recent data on human cancers (Ruggero 2013; Pelletier et al. 2017).

Discussion

PPP has emerged as a critical regulatory step in gene expression (Core and Adelman 2019). It involves stalling of RNA Pol II 20–60 nucleotides downstream of the transcription start site, and controlled release of RNA Pol II when triggered by signals from the surroundings. Many studies in recent years have elucidated mechanisms by which RNA Pol II is stalled and the factors that bring about pausing as well as release of the paused RNA Pol II. Our in vivo model for tumorigenesis has allowed us to elucidate the functions of the NELF, 7SKsnRNP, and P-TEFb complexes in the context of growth control in vivo. Previous studies have implicated NELF in regulating the response of embryonic stem cells to signaling cues such as fibroblast growth factor (FGF; Williams et al. 2015). Furthermore, PPP has been shown to be important for coordination of expression genes involved in morphogenesis of Drosophila embryo (Lagha et al. 2013). Our findings provide direct evidence that PPP can limit tumor

Table 2 List of genes whose expression is downregulated in the wing discs of ap-GAL4/UAS-NelfA^{RNAi}; UAS-Yki

Proteasome		Protein processing in endoplasmic reticulum	
Gene name	logFC	Gene name	logFC
Rpn7	21.084698	prtp	21.46289
Rpn13	20.949757	Sec61gamma	21.65593
Rpn2	20.900781	Sec61beta	21.31964
Prosalpha3	20.910536	CG5885	21.28449
Rpn3	20.850895	Sec61alpha	21.25184
Rpn1	20.879125	TRAM	21.3381
Pomp	20.795409	Pdi	21.145
Prosalpha5	20.834712	SsRbeta	21.23571
Prosbeta4	20.76685	Sec13	21.01315
Prosbeta7	20.717883	Sec63	20.97489
Prosbeta2	20.706638	CG14476	21.03296
Prosbeta5	20.687177	Sec24CD	20.87522
Prosalpha4	20.694159	Ostgamma	20.97374
Prosbeta6	20.631819	Ost48	20.88256
Rpn10	20.592037	CG4164	21.21065
Rpn12	20.597364	ergic53	20.86625
		Plap	20.86843
		OstStt3	20.90274
		Gp93	20.89197
		l(1)G0320	20.88951
		CG33303	20.8065
		Hsc70-3	20.9493
		CG5510	20.81474
		p47	20.78553
		Crc	20.86323
		CG6453	20.81869
		Sec23	20.73903

ERp60	20.76882
Der-1	20.80252
Csp	20.64369
CaBP1	20.61193
CG1597	20.67306

formation in the context of the Hippo tumor suppressor pathway. Depletion of these factors alone, or even in combination with overexpression of CDK9, was not sufficient to induce tumorous growth but did so when combined with overexpression of Yki. This cooperation appears to be specific to Yki-induced tumors as there was no cooperation with other oncogenic drivers such as EGFR or activated Notch in wing

disc tumor models. This suggests that pausing plays a previously unappreciated role regulating the output of Hippo pathway in growth control, thereby limiting its tumorigenic potential.

We were intrigued by the finding that CDK9 activity is required for Yki-driven tumor formation, even when the upstreamanddownstreampausingcomplexfactorshavebeen removed. These observations suggest that CDK9 activity is

required not only to remove the “brake” exerted by NELF pausing complex, but also required to increase RNA Pol II activity through direct phosphorylation. Neither alone is sufficient. This suggests an overlapping “belt and suspenders” regulation to ensure that expression of Yki targets is maintained at appropriate levels for

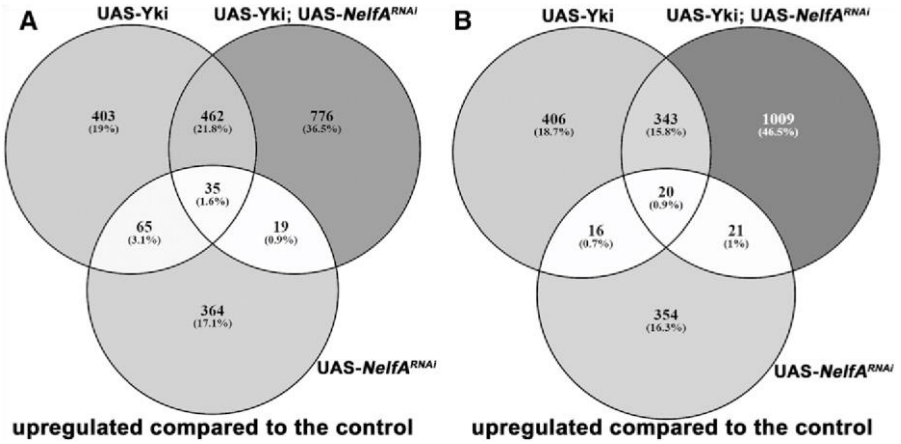


Figure 6 Identification of genes potentially involved in Yki-mediated tumorigenesis. (A) Venn diagram showing number of common and unique genes, who expression is upregulated in comparison with ap . GFP from different genotypes as indicated in the figure. (B) Venn diagram depicting number of common and unique genes downregulated in comparison with ap . GFP from different genotypes.

normal growth control, while preventing overexpression, which may lead to tumorigenesis. A mechanism of this sort allows for the possibility that other growth regulatory or metabolic homeostasis pathways might impact on the outcome of Yki activity via regulation of the CDK9. Indeed, evidence of a role for CDK9 in YAP/TAZmediated cell growth via regulation of a subset of YAP/TAZ target genes in mammalian liver cells has been demonstrated (Galli et al. 2015). Inhibition of CDK9 activity using flavopiridol nullified the effect of YAP S127A mutant form (the constitutively active form of YAP) on the expression of YAP target genes studied (Galli et al. 2015). Although this observation is not validated in fly tissues, perhaps PPP (including 7skRNP-, CDK9-, and NELFs)-dependent regulation of Yki is independent of the phosphorylation status of Yki, which implies a parallel function for PPP rather than it being upstream of Yki.

Our genetic model is also useful to study the importance of PPP in attenuating transcriptional output at genome wide scale. Preliminary observations of data generated by RNA-seq suggestthatmostgenesthatarediferentiallyexpressedwhen Yki is overexpressed show further changes in the same direction (up or down regulation) in combination of Yki overexpression with depletion of Nelf-A. Furthermore, we also report deregulation of proteostasis uniquely in tumor tissue. This is consistent with

recent reports that deregulation of translation and deregulation of protein processing are important factors in progression of cancers and might be target for therapy (Ruggero 2013; Pelletier et al. 2017).

To conclude, our study has highlighted additional regulatory module on Yki driven tumorigenic activity, which impingesdirectlyontranscription.Itwillbeinterestingtoseethe role of the PPP machinery, which has been reported to be highly conserved from Drosophila to humans (Peterlin and Price 2006), in the context of highly conserved Hippo pathway effectors YAP/TAZ. Considering the reported function of CDK9 in YAP-driven transcription, and the therapeutic accessibility of CDK9activity (Galli etal. 2015;Blakeetal. 2019), it is critical to understand the function of 7SK snRNP and NELF complexes in this context.

Acknowledgments

We thank G. Deshpande and members of the LSS and SMC laboratories for critical input. This work was supported primarily by an Indo-Danish research grant from the Department of Biotechnology, Government of India to L.S.S. and from the Innovation fund Denmark, Novo Nordisk Foundation NNF12OC0000552 and Neye Foundation to SMC; a JC Bose Fellowship and grant from the Department of Science and Technology, Government of India to LSS; and a University Grants Commission (UGC) Research Fellowship to SN.

Author contributions: S.N., P.G. and R.W. carried out all fly experiments. S.N. did RNA-seq and its analysis, all image analyses, and wrote the MS. L.S.S. and SM conceived the project and wrote the MS. We declare “no-conflict-of-interest”.

Note added in proof: See Groth et al. 2020 (pp. 2999–3008) in G3 10:9 for a related work.

Literature Cited

Anders, S., and W. Huber, 2010 Differential expression analysis for sequence count data. Genome Biol. 11: R106. https://doi.org/10.1186/gb-2010-11-10-r106

Beaucher, M., E. Hersperger, A. Page-McCaw, and A. Shearn, 2007 Metastatic ability of Drosophila tumors depends on MMP activity. Dev. Biol. 303: 625–634. https://doi.org/10.1016/j.ydbio.2006.12.001

Blake D. R., A. V. Vaseva, R. G. Hodge, M. P. Kline, T. S. K. Gilbert et al., 2019 Application of a MYC degradation screen identifies sensitivity to CDK9 inhibitors in KRAS-mutant pancreatic cancer. Sci. Signal. 12: eaav7259. https://doi.org/10.1126/scisignal.aav7259

Cohen, B., M. E. McGuffin, C. Pfeifle, D. Segal, and S. M. Cohen, 1992 apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. Genes Dev. 6: 715–729. https://doi.org/10.1101/gad.6.5.715

Core L., and K. Adelman, 2019 Promoter-proximal pausing of RNA polymerase II : a nexus of gene regulation. Genes Dev. 33: 960–982. https://doi.org/10.1101/gad.325142.119

Dong, J., G. Feldmann, J. Huang, S. Wu, N. Zhangetal., 2007 Elucidation of a universal size-control mechanism in Drosophila and mammals. Cell 130: 1120–1133. https://doi.org/10.1016/j.cell.2007.07.019

Galli, G. G., M. Carrara, W. C. Yuan, C. Valdes-Quezada, B. Gurung et al., 2015 YAP drives growth by controlling transcriptional pause release from dynamic enhancers. Mol. Cell 60: 328–337. https://doi.org/10.1016/j.molcel.2015.09.001

Gateff, E., T. Löffler, and J. Wismar, 1993 A temperature-sensitive brain tumor suppressor mutation of Drosophila melanogaster: developmental studies and molecular localization of the gene. Mech. Dev. 41: 15–31. https://doi.org/10.1016/0925-4773(93)90052-Y

Groth, C., P. Vaid, A. Khatpe, N. Prashali, A. Ahiya et al., 2020 Genome-wide screen for context-dependent tumor suppressors identified using in vivo models for neoplasia in Drosophila. G3 (Bethesda) DOI: 10.1534/g3.120.401545.

Guenther, M. G., S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young, 2007 A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130: 77–88. https:// doi.org/10.1016/j.cell.2007.05.042

Huang, J., S. Wu, J. Barrera, K. Matthews, and D. Pan, 2005 The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila homolog of YAP. Cell 122: 421–434. https://doi.org/10.1016/j.cell.2005.06.007

Jennings, B. H., 2013 Pausing for thought: disrupting the early transcription elongation checkpoint leads to developmental defects and tumourigenesis. BioEssays 35: 553–560. https:// doi.org/10.1002/bies.201200179

Johnson, R., and G. Halder, 2014 The two faces of Hippo: targeting the Hippo pathway for regenerative medicine and cancer treatment. Nat. Rev. Drug Discov. 13: 63–79. https://doi.org/10.1038/nrd4161

Jonkers, I., and J. T. Lis, 2015 Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. 16: 167–177. https://doi.org/10.1038/nrm3953

Kanazawa, S., L. Soucek, G. Evan, T. Okamoto, and B. M. Peterlin, 2003 c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. Oncogene 22: 5707–5711. https://doi.org/10.1038/sj.onc.1206800

Kwak, H., and J. T. Lis, 2013 Control of transcriptional elongation. Annu. Rev. Genet. 47: 483–508. https://doi.org/10.1146/ annurev-genet-110711-155440

Lagha, M., J. P. Bothma, E. Esposito, S. Ng, L. Stefanik et al., 2013 Paused Pol II coordinates tissue morphogenesis in the drosophila embryo. Cell 153: 976–987. https://doi.org/10.1016/ j.cell.2013.04.045

Marshall, N. F., and D. H. Price, 1995 Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. 270: 12335–12338. https://doi.org/ 10.1074/jbc.270.21.12335

Miles, W. O., N. J. Dyson, and J. Walker, 2011 Modeling tumor invasion and metastasis in Drosophila. Dis. Model. Mech. 4: 753–761. https://doi.org/10.1242/dmm.006908

Muse, G. W., D. A. Gilchrist, S. Nechaev, R. Shah, J. S. Parker et al., 2007 RNA polymerase is poised for activation across the genome. Nat. Genet. 39: 1507–1511. https://doi.org/10.1038/ng.2007.21

Nguyen, D., B. J. Krueger, S. C. Sedore, J. E. Brogie, J. T. Rogers et al., 2012 The Drosophila 7SK snRNP and the essential role of dHEXIM in development. Nucleic Acids Res. 40: 5283–5297. https://doi.org/10.1093/nar/gks191

Overholtzer, M., J. Zhang, G. A. Smolen, B. Muir, W. Li et al., 2006 Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon. Proc. Natl. Acad. Sci. USA 103: 12405–12410. https://doi.org/10.1073/pnas.0605579103

Pan, D., 2010 The hippo signaling pathway in development and cancer. Dev. Cell 19: 491–505. https://doi.org/10.1016/j.devcel.2010.09.011

Peterlin, B. M., and D. H. Price, 2006 Controlling the elongation phase of transcription with P-TEFb. Mol. Cell 23: 297–305. https://doi.org/10.1016/j.molcel.2006.06.014 Pelletier, J., G. Thomas, S. Volarevic, and S. Volarevi´c, 2017 Nrc. 2017.104. Nature 18: 51.

Ruggero, D., 2013 Translational control in cancer etiology. Cold Spring Harb. Perspect. Biol. 5: a012336 [Corrigenda: Cold Spring Harb. Perspect. Biol. 4 (2012)]. https://doi.org/10.1101/ cshperspect.a012336

Szklarczyk, D., J. H. Morris, H. Cook, M. Kuhn, S. Wyder et al., 2017 The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 45: D362–D368. https://doi.org/10.1093/ nar/gkw937

Tanos, B., and E. Rodriguez-Boulan, 2008 The epithelial polarity program: machineries involved and their hijacking by cancer. Oncogene 27: 6939–6957. https://doi.org/10.1038/onc.2008.345

Uhlirova, M., and D. Bohmann, 2006 JNK- and Fos-regulated Mmp1 expression cooperates with Ras to induce invasive tumors in Drosophila. EMBO J. 25: 5294–5304. https://doi.org/ 10.1038/sj.emboj.7601401

Wada, T., T. Takagi, Y. Yamaguchi, A. Ferdous, T. Imai et al., 1998 DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. 12: 343–356. https://doi.org/10.1101/gad.12.3.343

Williams, L. H., G. Fromm, N. G. Gokey, T. Henriques, G. W. Muse et al., 2015 Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. Mol. Cell 58: 311–322. https://doi.org/10.1016/j.molcel.2015.02.003

Yamaguchi, Y. Y., T. T. Takagi, T. T. Wada, K. K. Yano, A. A. Furuya et al., 1999 NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell 97: 41–51. https://doi.org/10.1016/S0092-8674(00)80713-8

Zanconato, F., M. Forcato, G. Battilana, L. Azzolin, E. Quaranta et al., 2015 Genome-wide association between YAP/TAZ/ TEAD and AP-1 at enhancers drives oncogenic growth. Nat. Cell Biol. 17: 1218–1227. https://doi.org/10.1038/ncb3216

Zeitlinger, J., A. Stark, M. Kellis, J.-W. W. Hong, S. Nechaev et al., 2007 RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet. 39: 1512– 1516. https://doi.org/10.1038/ng.2007.26

Zender, L., M. S. Spector, W. Xue, P. Flemming, C. Cordon-Cardo et al., 2006 Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell 125: 1253–1267. https://doi.org/10.1016/j.cell.2006.05.030

Zhao, B., K. Tumaneng, and K. L. Guan, 2011 The Hippo pathway in organ size control, tissue regeneration and stem cell self-renewal. Nat. Cell Biol. 13: 877–883. https://doi.org/10.1038/ncb2303

Communicating editor: P. Geyer

G3 Genes | Genomes | Genetics

MUTANT SCREEN REPORT

Genome-Wide Screen for Context-Dependent Tumor Suppressors Identified Using in Vivo Models for Neoplasia in Drosophila

Casper Groth,*,† Pooja Vaid,* Aditi Khatpe,* Nelchi Prashali,* Avantika Ahiya,* Diana Andrejeva,† Madhumita Chakladar,* Sanket Nagarkar,* Rachel Paul,* Devaki Kelkar,* Teresa Eichenlaub,† Hector Herranz,† TS Sridhar,‡ Stephen M. Cohen,† and LS Shashidhara*,§,1

*Indian Institute of Science Education and Research (IISER) Pune, Dr. Homi Bhabha Road, Pashan, Pune 411008, India, †Department of Cellular and Molecular Medicine, University of Copenhagen, Blegdamsvej 3, Copenhagen 2200N, Denmark, ‡Division of Molecular Medicine, St Johns Research Institute, Bangalore, India, and §Department of Biology, Ashoka University, Sonipat, India

ORCID IDs: 0000-0001-5652-3351 (A.K.); 0000-0001-9306-1432 (L.S.)

ABSTRACT Genetic approaches in Drosophila have successfully identified many genes involved in regulation of growth control as well as genetic interactions relevant to the initiation and progression of cancer in vivo. Here, we report on large-scale RNAi-based screens to identify potential tumor suppressor genes that interact with known cancer-drivers: the Epidermal Growth Factor Receptor and the Hippo pathway transcriptional cofactor Yorkie. These screens were designed to identify genes whose depletion drove tissue expressing EGFR or Yki from a state of benign overgrowth into neoplastic transformation in vivo. We also report on an independent screen aimed to identify genes whose depletion suppressed formation of neoplastic tumors in an existing EGFR-dependent neoplasia model. Many of the positives identified here are known to be functional in growth control pathways. We also find a number of novel connections to Yki and EGFR driven tissue growth, mostly unique to one of the two. Thus, resources provided here would be useful to all researchers who study negative regulators of growth during development and cancer in the context of activated EGFR and/or Yki and positive regulators of growth in the context of activated EGFR. Resources reported here are available freely for anyone to use.

KEYWORDS
Tumorigenesis
Neoplasia
Drosophila
EGFR
Hippo pathway

Studies in genetic models of tissue growth have identified networks of signaling pathways that cooperate to control growth during animal development (reviewed in (Harvey et al. 2013; Richardson and Portela 2017). Normal tissue growth involves controlling the rates of cell proliferation and cell death, as well as cell size, cell shape, etc. Signaling

Growth regulatory pathways include both positive and negative elements to allow for feedback regulation. These feedback systems confer robustness to deal with intrinsic biological noise, and with a fluctuating external environment (Herranz and Cohen 2010). They also provide the means for different regulatory pathways to interact (Ren et al. 2010;

pathways mediate hormonal and neuroendocrine regulation of growth, which depend on nutritional status. Cell interactions also contribute to coordinating growth of cells within a tissue.

Herranz et al. 2012a; Reddy and Irvine 2013). In the context of tumor formation, this robustness is reflected in the difficulty in generating significant misregulation of growth - a twofold change in expression of many growth regulators seldom has a substantial effect on tissue size in Drosophila genetic models. More striking is the difficulty in transitioning from benign overgrowth to neoplasia: hyperplasia does not normally lead to neoplasia without additional genetic alterations (e.g., (Huang et al. 2005; Herranz et al. 2012b 2014).

Cancers typically show mis-regulation of multiple growth regulatory pathways. Mutational changes and changes in gene expression status contribute to driving cell proliferation, overcoming cell death and cellular senescence, as well as to allowing cells to evade the

doi: https://doi.org/10.1534/g3.120.401545
Manuscript received July 6, 2020; accepted for publication July 27, 2020; published Early Online July 31, 2020.

Supplemental material available at figshare: https://doi.org/10.25387/g3.12746513.
[1] Corresponding author: Department of Biology, Main Building, IISER Pune, Dr Homi Bhabha Road, Pashan, Pune 411008. E-mail: ls.shashidhara@iiserpune.ac.in

S

checkpoints that normally serve to eliminate aberrant cells. These changes alter the normal balance of cellular regulatory mechanisms, from initial cellular transformation through disease progression (Stratton 2011; Alexandrov et al. 2013). For many tumor types, specific mutations have been identified as potent cancer drivers, with well-defined roles in disease (Kandoth et al. 2013; Zehir et al. 2017). However, most human tumors carry hundreds of mutations, whose functional relevance is unknown. The spectrum of mutation varies from patient to patient, and also within different parts of the same tumor (McGranahan and Swanton 2017). Evidence is emerging that some of these genetic variants can cooperate with known cancer drivers during cellular transformation or disease progression. The mutational landscape of an individual tumor is likely to contain conditional oncogenes or tumor suppressors that modulate important cellular regulatory networks.

Sequence-based approaches used to identify cancer genes favor those with large individual effects that stand out from the 'background noise' of the mutational landscape in individual cancers (Stratton 2011; Alexandrov et al. 2013). In vivo experimental approaches are needed to assign function to candidate cancer genes identified by tumor genome sequencing, and to identify functionally significant contributions of genes that have not attracted notice in genomics studies due to low mutational frequency, or due to changes in activity not associated with mutation. In vivo functional screens using transposon mutagenesis of the mouse genome have begun to identify mutations that cooperate with known cancer driver mutations, such as K-Ras, in specific tumor models (Copeland and Jenkins 2010; Pérez-Mancera et al. 2012; Takeda et al. 2015). Genetic approaches using Drosophila models of oncogene cooperation have also been used to identify genes that act together with known cancer drivers in tumor formation (Brumby and Richardson 2003; Pagliarini and Xu 2003; Wu et al. 2010; Brumby et al. 2011; Herranz et al. 2012b 2014; Eichenlaub et al. 2016; Richardson and Portela 2017; Song et al. 2017). The simplicity of the Drosophila genome, coupled with the ease of large-scale genetic screens and the high degree of conservation of major signaling pathways with humans, make Drosophila an interesting model to identify novel cancer genes and to study the cellular and molecular mechanisms that underlie tumor formation in vivo (reviewed in (Gonzalez 2013; Herranz et al. 2016; Sonoshita and Cagan 2017; Richardson and Portela 2018).

In Drosophila, overexpression of the Epidermal Growth Factor Receptor, EGFR, or Yorkie (Yki, the fly ortholog of the YAP oncoprotein) cause benign tissue over-growth (Huang et al. 2005; Herranz et al. 2012a 2014). Combining these with additional genetic alterations can lead to neoplastic transformation and eventually metastasis (Herranz et al. 2012b 2014; Eichenlaub et al. 2016, 2018; Song et al. 2017). Here, we report results of large-scale screens combining UAS-RNAi transgenes with EGFR or Yki expression to identify negative regulators of these growth regulatory networks that can lead to aggressive tumor formation in vivo. We also performed an independent screen to identify factors that could suppress EGFRdriven neoplasia. These screens have identified an expanded genomic repertoire of potential tumor suppressors that cooperate with EGFR or Yki. We have also identified few positive regulators of growth in the context of activated EGFR. Interestingly, there was limited overlap among the genes that cooperated with EGFR and those that cooperated with Yki. Gene intractome analysis and analyses of cancer databases for human orthologs of positives of these screens suggest that a large number of them have strong correlations to many clinical parameters. The output of this screen would, therefore, be useful to all researchers who study negative regulators of growth during development and cancer in the context of activated EGFR and/ or Yki. Resources reported here are freely available for anyone to use.

MATERIALS AND METHODS

RNAi Screens

The KK transgenic RNAi stock library was obtained from the Vienna Drosophila RNAi Center (www.vdrc.at; also listed in Table S1) carrying inducible UAS-RNAi constructs on Chromosome II. For each cross, 5 males from the KK transgenic RNAi stock were crossed separately to 10-15 virgins from each of the following three driver stocks

(see Supplemental Fig. S1A for the schematics of fly stocks): w, ap-Gal4, UAS-GFP/CyO; UAS-Yki, tub-Gal80^{ts}/TM6B (Yki driver; Song et al. 2017); w; ap-Gal4, UAS-GFP/CyO; UAS-EGFR, tubGal80^{ts}/TM6B (EGFR driver; Herranz et al. 2012b); and w; ap-Gal4, UAS-GFP/CyO; and w; ap-Gal4, UAS-GFP, Socs36E^{RNAi}/CyO; UAS-EGFR, tub-Gal80^{ts}/TM6B (EGFR driver +SOCS36E^{RNAi}). The combination of UAS-EGFR and UAS- SOCS36E^{RNAi} inducing tumorous growth is reported in Herranz et al. (2012b).

Virgin female flies were collected over 4-5 days and stored at 18 in temperature-controlled incubators on medium supplemented with dry yeast, prior to setting up crosses. Virgin females were mated to KK stock males (day 1) and the crosses were stored at 18 for 4 days to provide ample time for mating before starting the timed rearing protocol used for the screen. On day 5, the crosses were transferred into new, freshly yeasted vials for another 3 days at 18. On day 8, the adult flies were discarded, and larvae were allowed to develop until day 11, at which time the vials were moved to 29 incubators to induce Gal4 driver activity. Crosses were aged at 29 for a further 8-9 days, after which larvae were scored for size and wing disc overgrowth phenotypes for Yki and EGFR driver screen crosses. Flies were scored for suppression of the tumor phenotype for the EGFR driver +SOCS36E^{RNAi} crosses (see Supplemental Fig. S1B for the screen workflow).

In order to verify the integrity of the driver stocks during the course of the screen, we examined their expression patterns in conjunction with setting up screen crosses each week. For each driver, 2-3 of the bottles used for virgin collection were induced at 29 for 24 hr and analyzed using fluorescence microscopy for apterous-Gal4 specific expression in wandering 3-instar larvae (see Supplemental Fig. S2 for larval images of quality control). Any batch that showed tumorous growth on its own without a cross with KK-RNAi line (in case of SOCS stocks, if the batch didn't show tumorous growth) were discarded and new batches were made from the original clean stock.

Positive hits form the initial screen were retested by setting up 2 or more additional crosses. The hits were scored as verified if 2 out of 3 tests scored positive. Wandering third instar larvae of confirmed positives were imaged and documented using fluorescence microscopy.

Genomic DNA PCR 40D landing site occupancy test
Genomic DNA from a select number of Drosophila KK transgenic RNAi library stocks was isolated following a protocol available at the VDRC (www.vdrc.at). The presence or absence of the KK RNAi transgene at the 40D insertion site on the second chromosome was determined by multiplex PCR using the following primers: 40D primer (C_Genomic_F): 59-GCCCACTGTCAGCTCTCAAC-39 pKC26_R: 59-TGTAAAACGACGGCCAGT-39 pKC43_R: 59-TCGCTCGTTGCAGAATAGTCC-39 PCR amplification was performed using GoTaq G2 Hot Start Green Master Mix kit (Promega) in a 25 mL standard reaction mix and the following program: initial denaturation at 95 for 2 min, followed by 33 cycles with denaturation at 95 for 15 sec, annealing at 58 for 15 sec and extension at 72 for 90 sec. One final extension reaction was carried out at 72 for 10 min. Reactions were stored at -20 prior to gel loading. PCR using these primers generate an approximately 450 bp product in case of a transgene insertion or a 1050 bp product in case of no transgene insertion site at 40D.

Screen database
Results from the three screening projects were added to a screen management database, http://www.iiserpune.ac.in/rnai/, including images of positive hits and background information such as RNAi line ID, corresponding gene information from the Flybase etc. The database was developed by Livetek Software Consultant Services (Pune, Maharashtra, INDIA).

Pathway and gene set enrichment analysis
Gene set enrichment analysis was performed using genes that upon down regulation induced tumor formation (EGFR, YKI background) or suppressed tumor formation (EGFR+SOCS background). For D. melanogaster enrichment analysis all D. melanogaster protein coding genes were used as the "gene universe" together with organism specific datasets. For human ortholog enrichment analysis all human protein coding genes were used as the "gene universe" together with organism specific datasets. The algorithm packages and databases used in analysis are listed in Supplemental

Tables S2 and S3. Unless otherwise specified, pathway databases included in these packages were used. The KEGG database was downloaded directly from source on 10.10.2018. Organ system specific and disease related pathway maps were excluded from this analysis. Minimum and maximum number of genes per pathway or gene set, significant criteria, minimum enriched gene count and annotated gene counts for each test and database are indicated in Supplemental Tables S2 and S3. GO results were filtered for level .2, to eliminate broad highlevel categories and ,10 to minimize duplication among subcategories. A representative term was selected in the cases were identical set of genes mapped to multiple terms within the same database. After filtering, the top 10 terms from each database were used for clustering analysis.

Pathway and gene set enrichment analysis results were visualized as enrichment map with appropriate layout based on gene overlap ration using igraph. Gene overlap ratio was set as edge width. Edges with low overlap were deleted, filtering threshold was based on a number of "terms" in the results table – from 0 to 50 by 10; increasing filtering thresholds from 0.16 to 0.26 by 0.2. Clusters were detected using "Edge betweenness community" algorithm. Similar biological processes were color-coded.

R packages

clusterProfiler (3.8.1) - (Yu et al. 2012).

ReactomePA (1.24.0) - (Yu and He 2016).

http://pubs.rsc.org/en/Content/ArticleLanding/2015/MB/ C5MB00663E.

graphite (1.26.1) - Sales G, Calura E, Romualdi C (2018). graphite: GRAPH Interaction from pathway Topological Environment. R package version 1.26.1.

igraph (1.2.2) - Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.org

Database references

KEGG – (Kanehisa et al. 2016, 2017).

REACTOME – (Fabregat et al. 2018) Panther – (Thomas et al. 2003) GO – (Ashburner et al. 2000).

STRING interaction maps

STRING v10 is a computational tool for protein interaction network and pathway analysis (Szklarczyk et al. 2017)), to identify significant functional clustering among the candidate genes. STRING builds interaction maps by combining experimental data (including protein interaction data) with information about functional associations from text mining. STRING interactome maps were used to search for statistically significant enrichment of KEGG pathways.

Data availability

All stocks are available on request. Supplement Table S1 provides details of all RNAi lines used and link to the corresponding genes in the Flybase. Complete screen information along with larval images of the positives is also accessible from: http://www.iiserpune.ac.in/rnai/. Supplemental material available at figshare: https://doi.org/10.25387/ g3.12746513.

RESULTS

Overexpression of EGFR or Yki proteins in the Drosophila wing imaginal disc produces tissue overgrowth. Under these conditions the imaginal discs retain normal epithelial organization, but grow considerably larger than normal. However, in combination with additional genetic or environmental changes, the tissue can become neoplastic and form malignant tumors (Herranz et al. 2012b 2014; Song et al. 2017; Eichenlaub et al. 2018). In this context, we carried out large-scale screens using UAS-RNAi lines from the Vienna Drosophila RNAi KK library to identify genes which would drive hyperplastic growth to neoplastic transformation when down-regulated. To facilitate screening for tumorous growth, we expressed UAS-GFP with UAS-EGFR or UAS-Yki to allow imaginal disc size to be scored in the intact 3rd instar larva (Figure 1A; screen design, examples and quality controls are shown in Supplemental Figures S1 and S2).

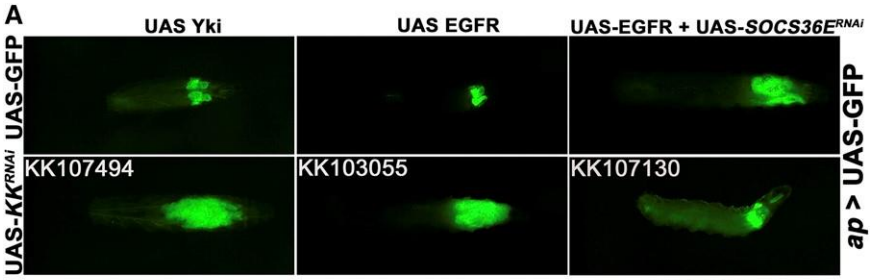
A large panel of independent UAS-RNAi lines were tested for their effects on tissue growth in the EGFR and

Yki expression backgrounds (Figure 1B). Of 8800 lines tested (Table S1), 74 interacted with EGFR to produce tumors (1%), whereas 904 interacted with Yki (10%) (Table S2). There was limited overlap, with only 21 RNAi lines producing tumors in both screens (Figure 1B), but we note that some loci that would be expected to score as hits in both screens, such as dlg, scrib and l(2)gl, were not targeted by RNAi lines in the KK collection, and so were not tested. In a parallel screen, we started with neoplastic tumors produced by co-expression of UASEGFR and UAS-SOCS36E^{RNAi} [Herranz et al. 2012] and asked whether including expression of another RNAi transgene could suppress neoplasia (Figure 1A, right panels). SOCS36E depletion has been reported to potentiate EGFR driven tumor formation by alleviating repression of JAK Stat activity [8]. Of 8900 lines tested (listed in Supplemental Table S1), 32 suppressed tumor formation in this assay (Figure 1B). Supplemental Table S2 (A) lists the genes identified in these three screens. In previous studies, massive disc overgrowth as in Figure 1(A) was often associated with loss of apically localized Actin and E-Cadherin: features indicative of Epithelial Mesenchymal Transition (EMT); and with formation of malignant transplantable tumors [Herranz et al. 2012b, 2014; Song et al. 2017). Apico-basal polarity and Matrix Metalloprotease 1 (MMP1) expression were assessed for a randomly selected subset of lines from the EGFR and Yki screens to assess neoplastic transformation (Figure S3).

To identify the processes and pathways responsible for the interaction with the screen drivers, we looked for over-representation of biological functions among the screen positives using gene set enrichment analysis and the KEGG, REACTOME, GO and PANTHER databases. Figure 2 presents the results of the enrichment analysis as graphical interaction maps, with similar biological processes color-coded. Edge length represents similarity between genes associated with significantly enriched terms. Thus, similar terms are closer together and form a community of biological process. The genes in each cluster are shown in Figure 2 and listed in Supplemental Table S3.

Genes that potentially modulate EGFR function during growth control

For discs overexpressing EGFR, we observed enrichment of RNAi lines targeting the Hippo pathway, growth signaling, and apoptosis (Figure 2A, B). Many of the genes in the Hippo pathway act as negative regulators of tissue growth, so their depletion by RNAi is expected to promote growth. The Hippo pathway is known to interact with the EGFR pathway to regulate normal developmental growth (Ren et al. 2010; Herranz et al. 2012a; Reddy and Irvine 2013). The Hippo pathway hits included core elements of the pathway, hpo, wts and mats, which serve as negative growth regulators; the upstream pathway regulators fat (an atypical cadherin) and expanded; as well as the transcriptional corepressor grunge, which is linked to Hippo pathway activity (Table S3). Several of these loci also contributed to the enrichment of terms linked to apoptosis, along with pten,



B

		UAS Yki	UAS EGFR	UAS-EGFR + UAS-*SOCS36E*^{RNAi}
Number of RNAi lines screened		8798	8795	8948
Confirmed positives		904	74	32
	Overlapped Yki/EGFR →	21		
Confirmed positives with human orthologues		582	48	31
	Overlapped Yki/EGFR →	12		

Figure 1 tumor formation/suppression visualized in intact larvae (A) Larvae co-expressed UAS-GFP with the indicated transgenes to permit visualization of the imaginal discs in the intact animal. All samples carried the ap-Gal4 driver and UASGFP. In addition, they carried either a second copy of UAS-GFP or one of the following: UASYki, UAS-EGFR or UAS-EGFR+UAS-SOCS36E$^{RNAi.}$ (B) Table summarizing the number of RNAi lines screened and identified in the three largescale screens (represents those many number of interacting genes).

a phospholipase that serves as a negative regulator of PI3K/AKT signaling, protein kinase A-C1, Src42A, the insulin-like peptide, ilp4, which are also linked to growth control (Table S3).

For suppression of tumors in discs overexpressing EGFR together with SOCS36E RNAi, we observed enrichment of RNAi lines targeting signaling pathways related to growth, including elements of the AKT/PI3K pathway (Figure 2E, F, Table S3). These pathways may be required for neoplasia in this EGFR driven tumor model. Interestingly, this pathway was also identified in a screen for synthetic lethals interacting with RasV12 (Willecke et al. 2011). As would be expected, depletion of Egfr limited tumor growth. Also enriched was a set of genes involved in protein synthesis (Table S3). This may reflect a need for active cellular growth machinery to support tumor growth. The significance of genes involved in RNA splicing merits further investigation.

Genes that potentially modulate Yki function during growth control

For discs overexpressing Yki, RNAi lines targeting the Hippo pathway and associated growth regulators led to tumor production (Figure 2C, D, Table S3). These include hpo, sav, wts, mats, ft and Grunge (Gug). Although wts null mutants show some loss of neuronal differentiation and impairment of polarity (Menut et al. 2007) tumor formation solely due to elevated Yki activity has not been observed previously in Drosophila. It is worth noting that overexpression of YAP has been shown to lead to neoplasia in mouse liver and intestinal epithelial models (Dong et al. 2007; Cai et al. 2010). While most cancers appear to result from activation/inactivation of multiple genes and pathways, sufficient activation of the Yki or Yap can result in neoplasia.

The Hippo tumor suppressor pathway is regulated by cell polarity, cell contact, and mechanical forces (Wada et al. 2011; Halder et al. 2012; Aragona et al. 2013) as well as by other growth signaling pathways. The atypical Cadherin Fat mediates cell interactions and acts upstream of the Hippo pathway. Gug is the fly ortholog of the mammalian Atrophin/RERE proteins, and has been reported to interact physically and genetically with Fat (Fanto et al. 2003). Growth signaling pathways involving the sgg, pten, PKA-C1, TSC1 genes among others, were also identified. Additionally, a number of genes linked to membrane-cytoskeleton interaction and transmembrane transport were found to interact, including Arf and Rab family members. We also noted the enrichment of terms related to lipid and general metabolism. Regulation of lipid metabolism might affect the properties of cellular membranes. An intriguing subgroup contain genes related to glutamatergic signaling, including the vesicular glutamate transporter VGlut and the Eaat plasma membrane glutamate transporters. This finding is of interest in light of the results of an in vivo chemical screen which showed that that scribble mutant RasV12 tumors are glutamine-dependent (Willoughby et al. 2013). These tumors upregulate Yki and require Yki for tumor growth (Doggett et al. 2011).

Another major finding from this screen is the fact that many components of the machinery causing Promoter proximal pausing of

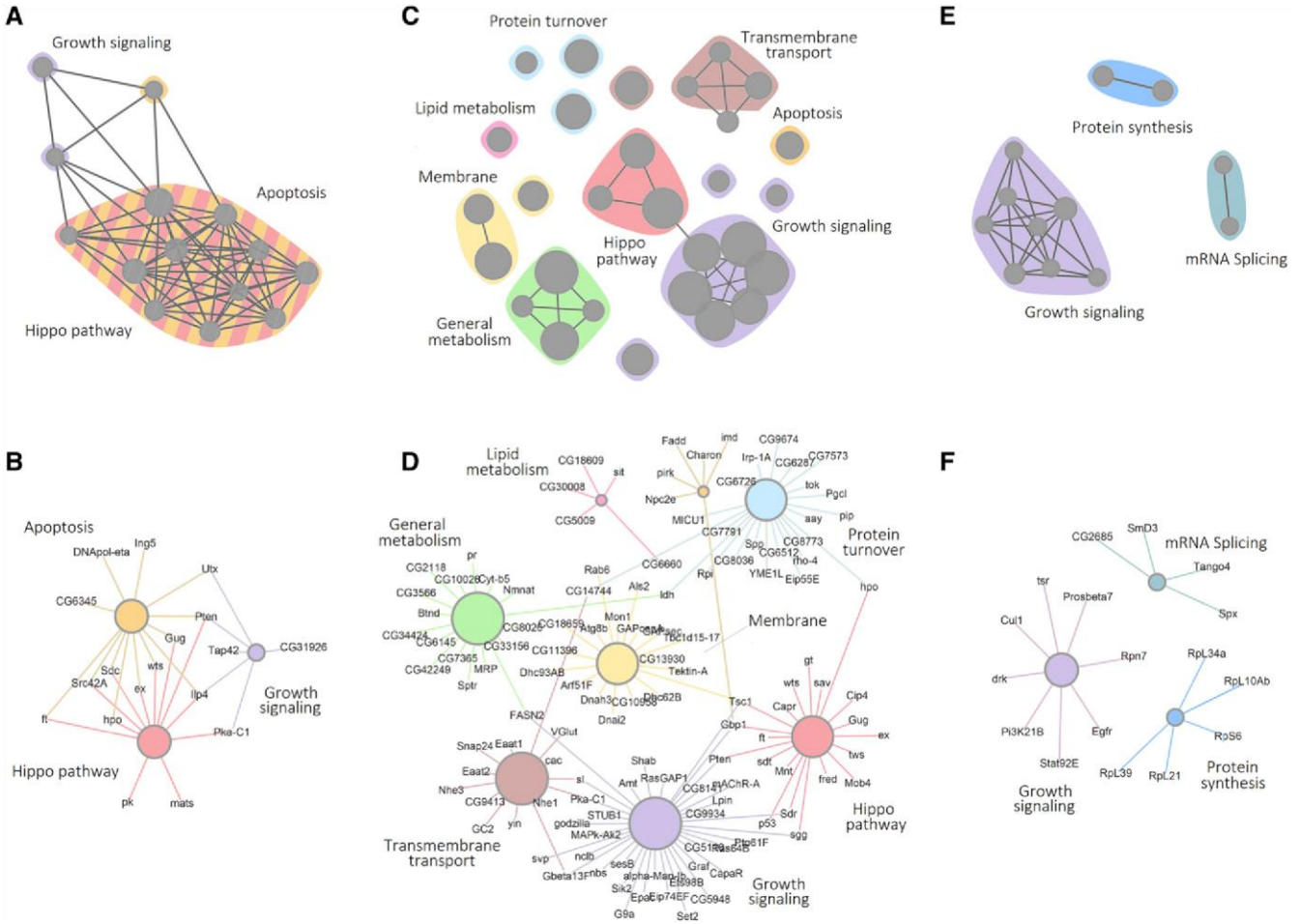


Figure 2 Summary of pathway enrichment analysis of fly genes identify in the in vivo screens reported here. (A, C, E) The results of the pathway and gene set enrichment analysis are shown as graphical interaction maps. Each node represents a significantly enriched term or pathway from the GO, KEGG, Reactome and Panther databases (Table S3). Color-coding indicates functionally related groups of terms. Lines indicate genes shared among different terms. (B, D, F) show the individual genes associated with functionally enriched cluster. (A, B) UAS-EGFR screen (C, D) UAS-Yki screen (E, F) UAS-EGFR+UAS-SOCS36E^{RNAi} screen

RNA Polymerase II (such as components of the 7SK snRNAP and NELF complexes) are when depleted, enhanced Yki-driven growth leading to neoplastic transformation of Drosophila wing imaginal discs (Nagarkar et al. 2020). Additional work suggested that this phenomenon is dependent on CDK9 function and also specific to Yki-induced growth context (Nagarkar et al. 2020).

The large number of Yki interactors could reflect greater sensitivity of the screen. Alternatively, it might indicate a high false positive rate. While this screen was in progress, Vissers et al. (Vissers et al. 2016), reported that some of the RNAi lines from the Vienna Drosophila RNAi KK library have the potential to produce false positives in screens based on sensitized Hippo pathway phenotypes. This proved to be due to the presence of a second transgene landing site at 40D that was found in a subset of KK lines, in addition to the 30B landing site (Green et al. 2014; Vissers et al. 2016). We tested the 40D landing site strain (Vissers et al. 2016) and found that it did not cause a tumor phenotype under the conditions used for the screen. Nonetheless, we sampled the 40D status for a large subset of our Yki interactors (Table S2, 734/904) and found that 45% of them had insertions at 40D. A small survey comparing KK lines with Trip and GD lines showed that 65% of genes for which the KK line had a 40D site retested positive for interaction with Yki using an independent (non-KK) transgene (15/23). The Yki-interaction screen should therefore be viewed as a more sensitized sampling of potential interactors, compared to the EGFR-interaction screen.

STRING interactome analyses

To view all genes identified in the three screens as one functional unit (for the fact that they were all growth regulators in one or the other contexts), we made use of STRING v10 (Szklarczyk et al. 2017) to produce protein interaction maps. STRING v10 builds interaction maps by combining experimental data (including protein interaction data) with information about functional associations from text mining. STRING v10 also uses information of co-occurrence, co-expression, gene neighborhood, gene fusion, and does sequence similarity search to predict functional interaction between proteins. An interaction pair supported by multiple lines of evidence has higher confidence score than other pairs.

Figure 3A shows the STRING interaction map for the genes identified as interactors of EGFR. As noted above, Hippo pathway (red) components were prominent among the genes identified as cooperating with EGFR to drive tumor formation. Figure 3(B) shows the interaction map for the genes identified as interactors of Yki. The larger number of hits in this screen results in a more complex

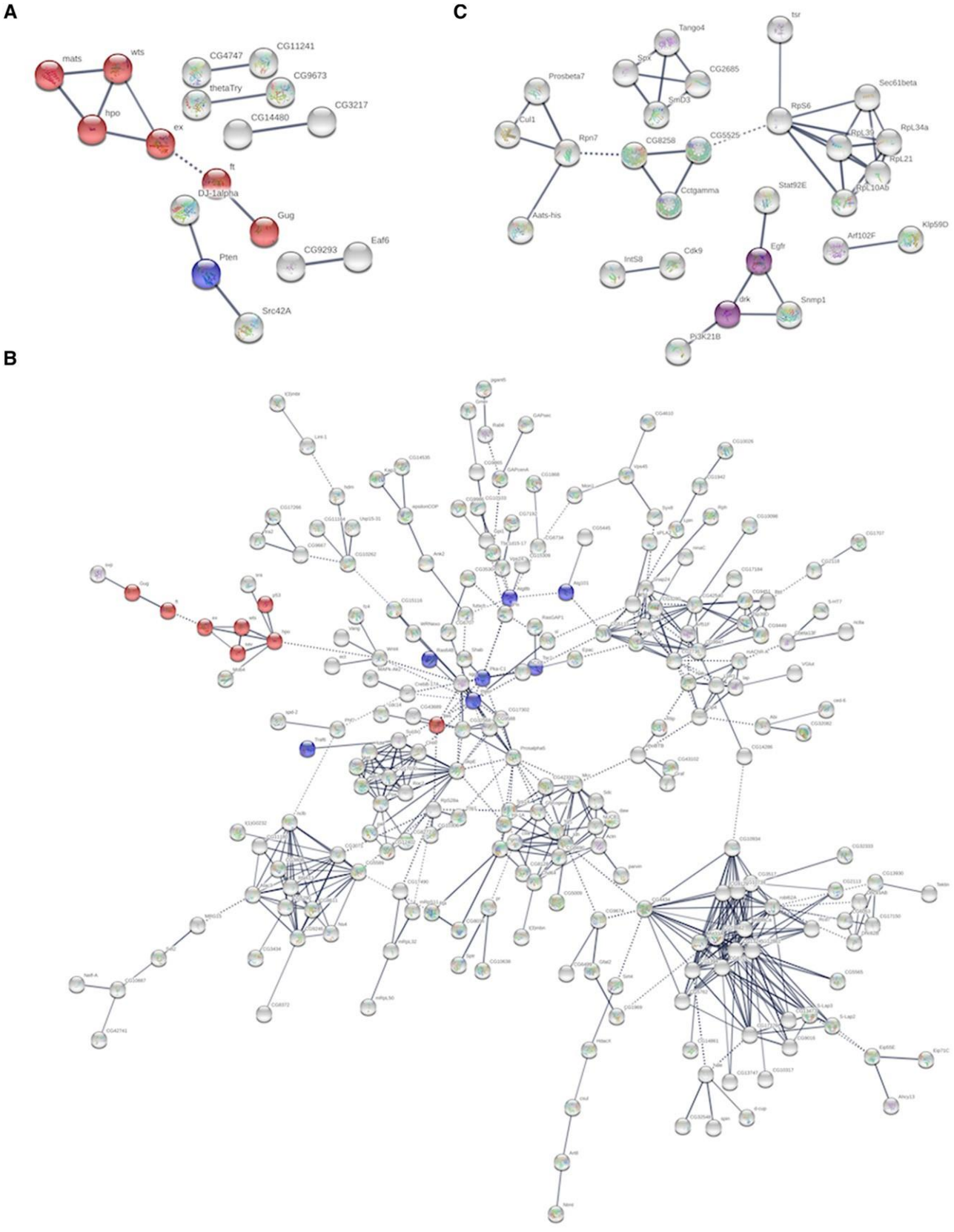


Figure 3 STRING interactome analysis of potential interactors of EGFR and YKi in Drosophila. STRING analysis was performed with confidence score of 0.7 and MCL clustering value of 2. (A) STRING Interactome of 73 fly genes identified as potential negative regulators in the context of over expression of EGFR.

17 out of those formed molecular clusters (with PPI enrichment value of 0.000482), largest being a cluster of 6 genes, all of which are constitutes of Fat/Hippo pathway (shown in red; FDR-1.39E^{-5}). (B) STRING Interactome of 888 genes of identified as potential negative

interaction map, with multiple interconnected clusters. The Hippo pathway (red) was again prominent in the fly screen. We also noted clusters containing elements of the ubiquitin mediated proteolysis pathway (green) and the PI3K/TOR (blue). As noted above, the higher sensitivity of this screen leads to the inclusion of weaker interactors, which may add to the complexity of these interaction maps. A focus on the stronger clusters and the interaction between them should guide future studies. Figure 3(C) shows interaction map for the genes identified as interactors of EGFR in the suppressor screen (in discs overexpressing EGFR together with SOCS36E RNAi). Among fly genes, as expected, we observed suppression of the tumor phenotype when components of EGFR pathway are down regulated.

Human orthologs of the fly genes identified in the three screens

To identify human orthologs for the candidate genes, we used the DRSC Integrative Ortholog Prediction Tool, DIOPT (Version 7.1, March 2018; www.flybase.org). DIOPT scores reflect the number of independent prediction tools that identify an ortholog for a given Drosophila gene. Orthology relationships are usually unambiguous when found by most of the 12 independent prediction tools in DIOPT. Table S2 lists the primary human orthologs (highest weighted DIOPT score), as well as the other orthologs with a weighted DIOPT score .2 for each of the hits in the fly screen. The primary human ortholog was used for subsequent analysis. In cases where multiple human orthologs had the same score, all orthologs with highest weighted DIOPT score were used. Out of 73 EGFR positive hits, 46 genes had one or more human orthologs, in total mapping to 50 human genes. Out of 32 SOCS positive hits 30 genes had one or more human orthologs, in total mapping to 31 human genes. Out of 904 YAP positive hits 570 genes had one or more human orthologs, in total mapping to 611 human genes.

To view the human orthologs in a functional context, we performed a gene set enrichment analysis and the KEGG, REACTOME, GO, PANTHER, NCI, MsigDB, BIOCARTA databases. Figure 4 presents the results of the enrichment analyses as graphical interaction maps, with similar biological processes color-coded. Edge length represents similarity between genes associated with significantly enriched terms. Thus, similar terms are closer together and form a community of biological processes. The genes in each cluster are shown in Figure 4 and listed in Supplemental Table S4. Because the enrichment analysis is highly sensitive to the number of orthologs for each of the fly genes, we used the minimal set consisting of only the primary human orthologs.

Hippo pathway components were enriched among the orthologs cooperating with EGFR to drive tumor formation (Figure 4A, B; Table S3). Two of these, LATS1 and STK3, also contributed to enrichment for a term linked to protein turnover. Regulation of protein turnover is an important mechanism for controlling the activity of a number of Hippo pathway components. For the screen for suppression of tumors in discs overexpressing EGFR together with SOCS36E RNAi, we observed enrichment of orthologs targeting growth signaling pathways, protein synthesis and mRNA splicing (Figure 4E, F, Table S4), similar to what was seen for the fly gene set analysis. We also observed enrichment of pathways related to protein folding and molecular chaperones, in the human gene set. For the Yki screen, the human ortholog set was enriched for terms related to general metabolism, and membrane transport, as well as growth signaling, and other signaling pathways, including genes involved in protein turnover (Figure 4C, D).

METABRIC Analysis

We also studied gene expression levels in cancer patients by systematically querying METABRIC (Pereira et al. 2016) a large database on breast cancer. We chose this as breast cancer is an epithelial cancer and the distribution of treatment-naïve samples from very early to late stages are well characterized. More importantly, gene expression patterns have been well studied at genomic level for all stages of the cancer. For each of the human orthologs of the genes identified in the Yki screen, we examined how their expression levels (low levels, median levels and high levels) are correlated to

clinical parameters/ attributes such as months of disease-free survival, early vs. old age of the patients at diagnosis, Lymph node status at diagnosis, tumor grade III or above at diagnosis, early vs. late stages of cancer at diagnosis and small vs. large tumors at diagnosis. Total 365 human orthologs showed significant correlation to disease-free survival. Among them 186 were associated with their low levels of expression and 179 with high levels of expression (see Supplement Table S4 and Supplemental_Information_METABRIC analysis). The fact that higher levels of expression correlate to aggressive tumors suggest that they are potential growth promoters, while their fly homologs were identified as potential tumor suppressors in our screen. This discrepancy could be due to more complex nature of growth control in human, wherein a conserved pathway may have different outcomes in different contexts. Expression levels of 76 genes also showed strong correlations to the three clinical parameters as listed above (see Supplement Table S4 and Supplemental_Information_METABRIC analysis) indicating their critical role in growth control and impairment in their expression causing tumorous growth. Taken together, the positive hits in these screens would be useful for studies on growth control in development model organisms and in the context of cancer in human.

DISCUSSION

The Hippo pathway has emerged from this study as the single most important pathway limiting tumor formation in Drosophila. Increasing Yki activity by depletion of upstream negative regulators promoted tumor formation in both the EGFR and Yki hyperplasia models. Yki controls tissue growth by promoting cell proliferation and by concurrently inhibiting cell death through targets including Diap1, cycE and bantam miRNA (Tapon et al. 2002; Huang et al. 2005; Nolo et al. 2006; Thompson and Cohen 2006; Wu et al. 2008). The central role of the Hippo pathway as an integrator of other growth-related signals may also contribute to the abundance of tumor suppressors associated with Yki-driven growth (Harvey et al. 2013; Richardson and Portela 2017, 2018). Mis-regulation of this pathway also contributes to tumor formation in mouse models (Yu et al. 2015).

The potential of Yki/YAP expression to drive cellular transformation has been highlighted by studies of primary human cells in

regulators in the context of over expression of Yki. 228 of those formed a single cluster with PPI enrichment value 1.4E-06. Components of Fat/ Hippo pathway (red: FDR-0.00076) and Autophagy genes (blue: FDR-0.0241) are enriched in this cluster. (C) STRING Interactome of 32 fly genes identified as potential oncogenes in the context of SOCS suppression. 27 out of those formed molecular clusters (with PPI enrichment value of 0.0122), largest being a cluster of 14 genes. A smaller cluster comprising of EGFR and DrK were enriched in Dorso-ventral axis formation (shown in purple: FDR-0.0089).

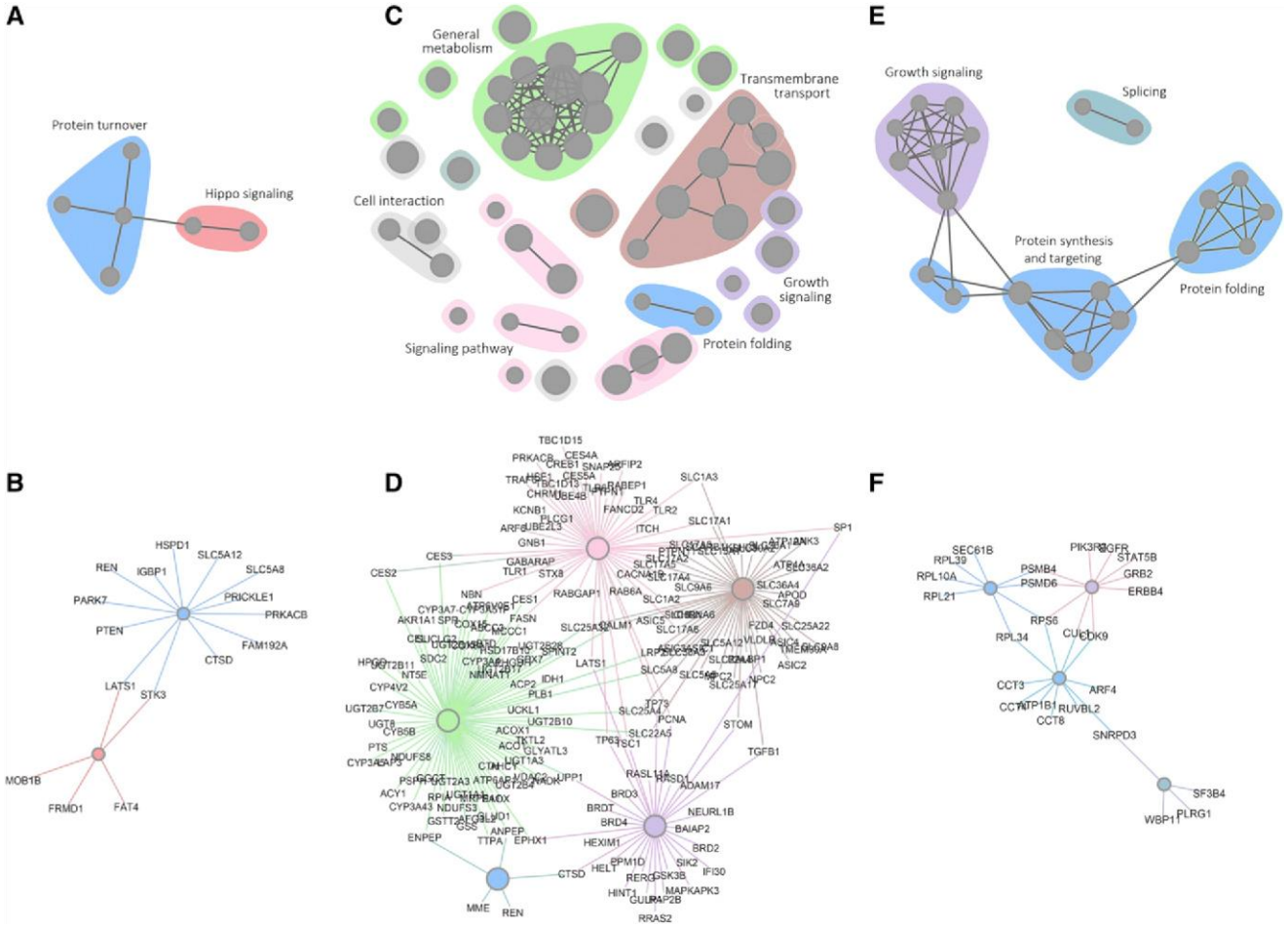


Figure 4 Summary of pathway enrichment analysis of human orthologs (A, C, E) The results of the pathway and gene set enrichment analysis are shown as enrichment maps. Each node represents a significantly enriched term or pathway from the GO, KEGG, Reactome and PANTHER, NCI, MsigDB, BIOCARTA databases (Table S3). Color-coding indicates functionally related groups of terms. Lines indicate genes shared among different terms. (B, D, F) show the individual genes associated with functionally enriched cluster. (A, B) UAS-EGFR screen (C, D) UAS-Yki screen (E, F) UAS-EGFR+UAS-SOCS36E^{RNAi} screen.

culture, which have shown that YAP expression is both necessary and sufficient to confer a transformed phenotype involving anchorage independent growth and the ability to form tumors in xenograft models (Hong et al. 2014; Nguyen et al. 2014). We therefore consider it likely that the consequence of Yki overexpression predispose the tissue to transformation, allowing identification of a richer repertoire of cooperating factors. Indeed, YAP overexpression has been causally linked to formation of specific human tumors (Kapoor et al. 2014; Shao et al. 2014). The Hippo pathway has also been implicated in tumor formation resulting from cytokinesis failure (Ganem et al. 2014) and this has recently been linked to Yki-mediated regulation of string (CDC25) expression (Gerlach et al. 2018). The sensitivity of Yki-expressing tissue to tumor formation might be explained by the finding that Yki promotes cell cycle progression at both the G1-S transition (through regulation of cycE (Huang et al. 2005) and at the G2-M transition through regulation of string. In contrast, mitogens and growth factors such as EGFR typically induce growth by promoting G1-S, and therefore remain somewhat constrained by the G2-M checkpoint.

We have analyzed in more detail one group of genes, all related to regulating promoter proximal pausing of RNA Poly II, identified in this screen to validate the importance of the repertoire of genes provided here. We have observed that Yki-driven growth is limited by the pausing of RNA Pol II, release of which is controlled by potential tumor suppressor genes (Nagarkar et al. 2020).

While our manuscript was in preparation, another group reported an RNAi screen to identify loci cooperating in tumorigenesis driven by expression in eye discs of the oncogenic activated mutant form of Ras (Zoranovic et al. 2018). We note that the activated Ras RNAi screen produced over 900 hits, compared with 74 for our EGFR screen, suggesting that the Ras screen was considerably more sensitized. We were surprised to note that there was almost no overlap between the two screens with only 3 hits in common: Elongin B, CG7966 and CG7313. This suggests that the genetic interactions required to promote tumorigenesis in the context of expression of an activated mutant form of RAS are distinct from those required to promote tumorigenesis in the context of native EGRF overexpression. And perhaps, the differences between the tissue contexts (eye discs in (Zoranovic et al. 2018) vs. wing discs in our screen). It will be of interest, in future, to learn whether this distinction holds true for factors promoting tumor formation in human cancers that depend on EGFR overexpression vs. those dependent on Ras mutants.

To conclude, the results reported here provide an extensive assessment of the genes that can serve as negative regulators of growth that can contribute to the formation of neoplastic tumors in vivo in Drosophila. In addition to finding genes linked to known growth control pathways, a number of novel connections to Yki and EGFR driven tissue growth have been identified, which merit further investigation in the Drosophila genetic model. Exploring the potential relevance of genes identified in this manner to human cancer will involve assessing the correlation of candidate gene expression with clinical outcome across a broad range of cancers (e.g., (Andrejeva et al. 2018; Eichenlaub et al. 2018)), as a starting point to identify biomarkers as well as novel candidate drug targets.

ACKNOWLEDGMENTS

The professionalism, tireless support and goodwill provided by Snehal Patil, Yashwant Pawar and Bhargavi Naik from the IISER Pune fly facility contributed greatly to the success of the screening projects reported here. We thank other members of the laboratories of LSS, SMC and TSS for critical input. This work was supported by an Indo-Danish research grant from Department of Biotechnology, Govt. of India to TSS and LSS and from Innovation fund Denmark, Novo Nordisk Foundation NNF12OC0000552 and Neye Foundation to SMC. JC Bose Fellowship and grant from Department of Science & Technology, Govt of India to LSS. HH, CG, TE, SMC and LSS designed the screen. CG, PV, AK, NP, AA, SN and RP carried out the genetic screen and participated in data analysis. DA and MC carried out computational analysis. SMC, TLS and LSS conceived of, designed and coordinated the study. SMC and LSS analyzed the data and drafted the manuscript.

Note added in proof: See Nagarkar et al. 2020 (pp. 67–77) in Genetics 216:1 for a related work.

LITERATURE CITED

Alexandrov, L. B., S. Nik-Zainal, D. C. Wedge, S. A. J. R. Aparicio, S. Behjati et al., 2013 Signatures of mutational processes in human cancer. Nature 500: 415–421. https://doi.org/10.1038/nature12477

Andrejeva, D., J. M. Kugler, H. T. Nguyen, A. Malmendal, M. L. Holm et al., 2018 Metabolic control of PPAR activity by aldehyde dehydrogenase regulates invasive cell behavior and predicts survival in hepatocellular and renal clear cell carcinoma 11 Medical and Health Sciences 1112 Oncology and Carcinogenesis. BMC Cancer 18: 1–15.

Aragona, M., T. Panciera, A. Manfrin, S. Giulitti, F. Michielin et al., 2013 A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. Cell 154: 1047–1059. https://doi.org/10.1016/j.cell.2013.07.042

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler et al., 2000 Gene Ontology: tool for the unification of biology. Nat. Genet. 25: 25–29.

Brumby, A. M., K. R. Goulding, T. Schlosser, S. Loi, R. Galea et al., 2011 Identification of novel Ras-cooperating oncogenes in Drosophila melanogaster: A RhoGEF/Rho-family/JNK pathway is a central driver of tumorigenesis. Genetics 188: 105–125. https://doi.org/10.1534/genetics.111.127910

Brumby, A. M., and H. E. Richardson, 2003 scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila. EMBO J. 22: 5769–5779. https://doi.org/10.1093/emboj/cdg548

Cai, J., N. Zhang, Y. Zheng, R. F. De Wilde, A. Maitra et al., 2010 The Hippo signaling pathway restricts the oncogenic potential of an intestinal regeneration program. Genes Dev. 24: 2383–2388. https://doi.org/10.1101/ gad.1978810

Copeland, N. G., and N. A. Jenkins, 2010 Harnessing transposons for cancer gene discovery. Nat. Rev. Cancer 10: 696–706. https://doi.org/10.1038/nrc2916

Doggett, K., F. A. Grusche, H. E. Richardson, and A. 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 Dev. Biol. 11: 57. https://doi.org/ 10.1186/1471-213X-11-57

Dong, J., G. Feldmann, J. Huang, S. Wu, N. Zhang et al., 2007 Elucidation of a Universal Size-Control Mechanism in Drosophila and Mammals. Cell 130: 1120–1133. https://doi.org/10.1016/j.cell.2007.07.019

Eichenlaub, T., S. M. Cohen, and H. Herranz, 2016 Cell competition drives the formation of metastatic tumors in a drosophila model of epithelial tumor formation. Curr. Biol. 26: 419–427. https://doi.org/10.1016/j.cub.2015.12.042

Eichenlaub, T., R. Villadsen, F. C. P. Freitas, D. Andrejeva, B. I. Aldana et al., 2018 Warburg Effect Metabolism Drives Neoplasia in a Drosophila Genetic Model of Epithelial Cancer. Curr. Biol. 28: 3220–3228.e6. https:// doi.org/10.1016/j.cub.2018.08.035

Fabregat, A., S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie et al., 2018 The Reactome Pathway Knowledgebase. Nucleic Acids Res. 46: D649–D655. https://doi.org/10.1093/nar/gkx1132

Fanto, M., L. Clayton, J. Meredith, K. Hardiman, B. Charroux et al., 2003 The tumor-suppressor and cell adhesion molecule fat controls planar polarity via physical interactions with Atrophin, a transcriptional co-repressor. Development 130: 763–774. https://doi.org/10.1242/dev.00304

Ganem, N. J., H. Cornils, S. Y. Chiu, K. P. O'Rourke, J. Arnaud et al., 2014 Cytokinesis failure triggers hippo tumor suppressor pathway activation. Cell 158: 833–848. https://doi.org/10.1016/j.cell.2014.06.029

Gerlach, S. U., T. Eichenlaub, and H. Herranz, 2018 Yorkie and JNK Control Tumorigenesis in Drosophila Cells with Cytokinesis Failure. Cell Rep. 23: 1491–1503. https://doi.org/10.1016/j.celrep.2018.04.006

Gonzalez, C., 2013 Drosophila melanogaster: a model and a tool to investigate malignancy and identify new therapeutics. Nat. Rev. Cancer 13: 172–183. https://doi.org/10.1038/nrc3461

Green, E. W., G. Fedele, F. Giorgini, and C. P. Kyriacou, 2014 A Drosophila RNAi collection is subject to dominant phenotypic effects. Nat. Methods 11: 222–223. https://doi.org/10.1038/nmeth.2856

Halder, G., S. Dupont, and S. Piccolo, 2012 Transduction of mechanical and cytoskeletal cues by YAP and TAZ. Nat. Rev. Mol. Cell Biol. 13: 591–600. https://doi.org/10.1038/nrm3416

Harvey, K. F., X. Zhang, and D. M. Thomas, 2013 The Hippo pathway and human cancer. Nat. Rev. Cancer 13: 246–257. https://doi.org/10.1038/nrc3458

Herranz, H., and S. M. Cohen, 2010 MicroRNAs and gene regulatory networks: Managing the impact of noise in biological systems. Genes Dev. 24: 1339–1344. https://doi.org/10.1101/gad.1937010

Herranz, H., T. Eichenlaub, and S. M. Cohen, 2016 Cancer in Drosophila: Imaginal Discs as a Model for Epithelial Tumor Formation. Curr. Top. Dev. Biol. 116: 181–199. https://doi.org/10.1016/bs.ctdb.2015.11.037

Herranz, H., X. Hong, and S. M. Cohen, 2012a Mutual repression by bantam miRNA and capicua links the EGFR/MAPK and hippo pathways in growth control. Curr. Biol. 22: 651–657. https://doi.org/10.1016/j.cub.2012.02.050

Herranz, H., X. Hong, N. T. Hung, P. Mathijs Voorhoeve, and S. M. Cohen, 2012b Oncogenic cooperation between SOCS family proteins and EGFR identified using a Drosophila epithelial transformation model. Genes Dev. 26: 1602–1611. https://doi.org/10.1101/gad.192021.112

Herranz, H., R. Weng, and S. M. Cohen, 2014 Crosstalk between epithelial and mesenchymal tissues in tumorigenesis and imaginal disc development. Curr. Biol. 24: 1476–1484. https://doi.org/10.1016/j.cub.2014.05.043

Hong, X., H. T. Nguyen, Q. Chen, R. Zhang, Z. Hagman et al., 2014 Opposing activities of the R as and H ippo pathways converge on regulation of YAP protein turnover. EMBO J. 33: 2447–2457. https://doi.org/10.15252/embj.201489385

Huang, J., S. Wu, J. Barrera, K. Matthews, and D. Pan, 2005 The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the Drosophila homolog of YAP. Cell 122: 421–434. https://doi.org/10.1016/j.cell.2005.06.007

Kandoth, C., M. D. McLellan, F. Vandin, K. Ye, B. Niu et al., 2013 Mutational landscape and significance across 12 major cancer types. Nature 502: 333–339. https://doi.org/10.1038/nature12634 Kanehisa, M., M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, 2017 KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45: D353–D361. https://doi.org/10.1093/nar/gkw1092

Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, 2016 KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44: D457–D462. https://doi.org/10.1093/nar/gkv1070

Kapoor, A., W. Yao, H. Ying, S. Hua, A. Liewen et al., 2014 Yap1 activation enables bypass of oncogenic KRAS addiction in pancreatic cancer. Cell 158: 185–197. https://doi.org/10.1016/j.cell.2014.06.003

McGranahan, N., and C. Swanton, 2017 Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 168: 613–628. https://doi.org/ 10.1016/j.cell.2017.01.018

Menut, L., T. Vaccari, H. Dionne, J. Hill, G. Wu et al., 2007 A mosaic genetic screen for Drosophila neoplastic tumor suppressor genes based on defective pupation. Genetics 177: 1667–1677. https://doi.org/10.1534/genetics.107.078360

Nagarkar, S., R. Wasnik, P. Govada, S. M. Cohen, and L. S. Shashidhara, 2020 Promoter proximal pausing limits Yki-induced tumorous growth in Drosophila. Genetics. https://doi.org/10.1534/genetics.120.303419.

Nguyen, H. T., X. Hong, S. Tan, Q. Chen, L. Chan et al., 2014 Viral small T oncoproteins transform cells by alleviating Hippo-pathway-mediated inhibition of the YAP proto-oncogene. Cell Rep. 8: 707–713. https://doi.org/10.1016/j.celrep.2014.06.062

Nolo, R., C. M. Morrison, C. Tao, X. Zhang, and G. Halder, 2006 The bantam MicroRNA Is a Target of the Hippo Tumor-Suppressor Pathway. Curr. Biol. 16: 1895–1904. https://doi.org/10.1016/j.cub.2006.08.057

Pagliarini, R. A., and T. Xu, 2003 A genetic screen in Drosophila for metastatic behavior. Science 302: 1227–1231. https://doi.org/10.1126/science.1088474

Pereira, B., S.-F. Chin, O. M. Rueda, H.-K. M. Vollan, E. Provenzano et al., 2016 The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nat. Commun. 7: 11479.

Pérez-Mancera, P. A., A. G. Rust, L. Van Der Weyden, G. Kristiansen, A. Li et al., 2012 The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. Nature 486: 266–270. https://doi.org/10.1038/nature11114

Reddy, B. V. V. G., and K. D. Irvine, 2013 Regulation of Hippo Signaling by EGFR-MAPK Signaling through Ajuba Family Proteins. Dev. Cell 24: 459–471. https://doi.org/10.1016/j.devcel.2013.01.020

Ren, F., L. Zhang, and J. Jiang, 2010 Hippo signaling regulates Yorkie nuclear localization and activity through 14–3-3 dependent and independent mechanisms. Dev. Biol. 337: 303–312. https://doi.org/10.1016/j.ydbio.2009.10.046

Richardson, H. E., and M. Portela, 2018 Modelling Cooperative Tumorigenesis in Drosophila. BioMed Res. Int. 2018: 4258387. https://doi.org/10.1155/2018/4258387

Richardson, H. E., and M. Portela, 2017 Tissue growth and tumorigenesis in Drosophila: cell polarity and the Hippo pathway. Curr. Opin. Cell Biol. 48: 1–9. https://doi.org/10.1016/j.ceb.2017.03.006

Shao, D. D., W. Xue, E. B. Krall, A. Bhutkar, F. Piccioni et al., 2014 KRAS and YAP1 converge to regulate EMT and tumor survival. Cell 158: 171–184. https://doi.org/10.1016/j.cell.2014.06.004

Song, S., H. Herranz, and S. 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 Dis. Model. Mech. 10: 1201–1209. https://doi.org/10.1242/dmm.030122

Sonoshita, M., and R. L. Cagan, 2017 Modeling Human Cancers in Drosophila, Elsevier Inc., Amsterdam. https://doi.org/10.1016/bs.ctdb.2016.07.008

Stratton, M. R., 2011 Exploring the genomes of cancer cells: progress and promise. Science 331: 1553–1558. https://doi.org/10.1126/science.1204040

Szklarczyk, D., J. H. Morris, H. Cook, M. Kuhn, S. Wyder et al., 2017 The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 45: D362–D368. https://doi.org/10.1093/nar/gkw937

Takeda, H., Z. Wei, H. Koso, A. G. Rust, C. C. K. Yew et al., 2015 Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. Nat. Genet. 47: 142–150. https://doi.org/ 10.1038/ng.3175

Tapon, N., K. F. Harvey, D. W. Bell, D. C. R. Wahrer, T. A. Schiripo et al., 2002 salvador promotes both cell cycle exit and apoptosis in Drosophila and is mutated in human cancer cell lines. Cell 110: 467–478. https://doi.org/10.1016/S0092-8674(02)00824-3

Thomas, P. D., M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak et al., 2003 PANTHER: A library of protein families and subfamilies indexed by function. Genome Res. 13: 2129–2141. https://doi.org/10.1101/gr.772403

Thompson, B. J., and S. M. Cohen, 2006 The Hippo Pathway Regulates the bantam microRNA to Control Cell Proliferation and Apoptosis in Drosophila. Cell 126: 767–774. https://doi.org/10.1016/j.cell.2006.07.013

Vissers, J. H. A., S. A. Manning, A. Kulkarni, and K. F. Harvey, 2016 A Drosophila RNAi library modulates Hippo pathway-dependent tissue growth. Nat. Commun. 7: 10368. https://doi.org/10.1038/ncomms10368

Wada, K. I., K. Itoga, T. Okano, S. Yonemura, and H. Sasaki, 2011 Hippo pathway regulation by cell morphology and stress fibers. Development 138: 3907–3914. https://doi.org/10.1242/dev.070987

Willecke, M., J. Toggweiler, and K. Basler, 2011 Loss of PI3K blocks cellcycle progression in a Drosophila tumor model. Oncogene 30: 4067–4074. https://doi.org/10.1038/onc.2011.125

Willoughby, L. F., T. Schlosser, S. A. Manning, J. P. Parisot, I. P. Street et al., 2013 An in vivo large-scale chemical screening platform using Drosophila for anti-cancer drug discovery. DMM Dis. Model. Mech. 6: 521–529. https://doi.org/10.1242/dmm.009985

Wu, S., Y. Liu, Y. Zheng, J. Dong, and D. Pan, 2008 The TEAD/TEF family protein Scalloped mediates transcriptional output of the Hippo growthregulatory pathway. Dev. Cell 14: 388–398. https://doi.org/10.1016/ j.devcel.2008.01.007

Wu, M., J. C. Pastor-Pareja, and T. Xu, 2010 Interaction between Ras(V12) and scribbled clones induces tumour growth and invasion. Nature 463: 545–548. https://doi.org/10.1038/nature08702

Yu, G., and Q. Y. He, 2016 ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. Mol. Biosyst. 12: 477–479. https://doi.org/10.1039/C5MB00663E

Yu, G., L. G. Wang, Y. Han, and Q. Y. He, 2012 ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi. A J. Integr. Biol. 16: 284–287.

Yu, F. X., B. Zhao, and K. L. Guan, 2015 Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer. Cell 163: 811–828. https://doi.org/10.1016/j.cell.2015.10.044

Zehir, A., R. Benayed, R. H. Shah, A. Syed, S. Middha et al., 2017 Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat. Med. 23: 703–713. https://doi.org/10.1038/nm.4333

Zoranovic, T., J. Manent, L. Willoughby, R. Matos de Simoes, J. E. La Marca et al., 2018 A genome-wide Drosophila epithelial tumorigenesis screen identifies Tetraspanin 29Fb as an evolutionarily conserved suppressor of Ras-driven cancer. PLoS Genet. 14: e1007688. https://doi.org/10.1371/ journal.pgen.1007688

Communicating editor: B. Andrews