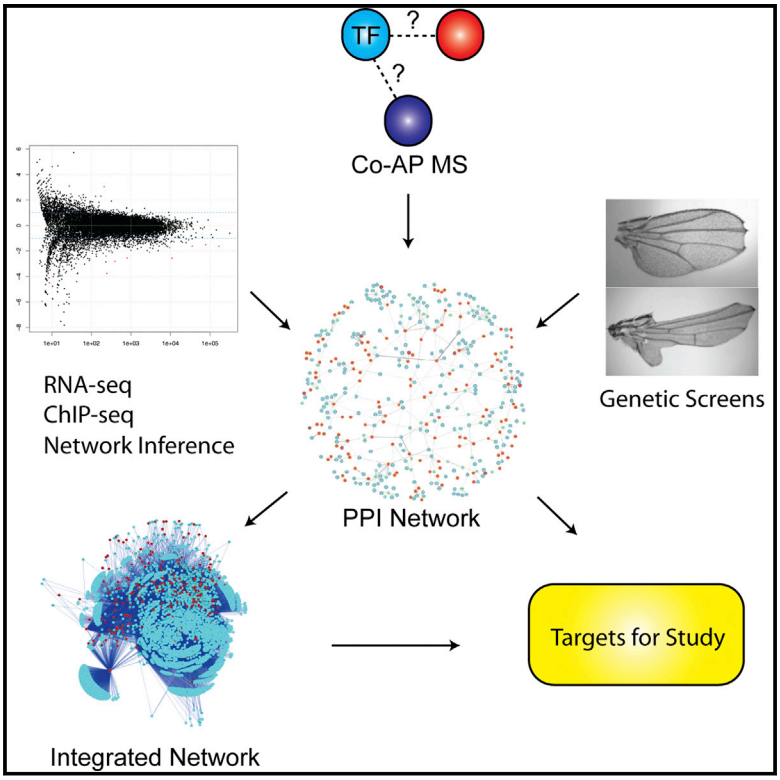
Resource

Cell Reports

Transcription Factor Networks in *Drosophila melanogaster*

Graphical Abstract



Highlights

Direct physical interactions for the majority of *Drosophila* TFs are analyzed

TF protein interactions are integrated with existing genomic data sets

A resource for studying the biology of transcription factors is provided

TF interactions are used to functionally interrogate the *mastermind* signaling network

Authors

David Y. Rhee, Dong-Yeon Cho, ..., Robert A. Obar, Spyros Artavanis-Tsakonas

Correspondence

sartavanis@cb.med.harvard.edu

In Brief

In this study, Rhee et al. systematically identify protein interactions for *Drosophila* transcription factors using a co-affinity purification/mass spectrometry approach. These data are integrated with existing genomic data sets to probe aspects of the Notch signaling pathway, and they provide a rich resource for forming biological hypotheses.

Rhee et al., 2014, Cell Reports 8, 2031–2043
September 25, 2014 ©2014 The Authors
http://dx.doi.org/10.1016/j.celrep.2014.08.038

Cell Reports
Resource

Transcription Factor Networks in *Drosophila melanogaster*

David Y. Rhee,[1] Dong-Yeon Cho,[2] Bo Zhai,[1] Matthew Slattery,[3] Lijia Ma,[3] Julian Mintseris,[1] Christina Y. Wong,[1] Kevin P. White,[3] Susan E. Celniker,[4] Teresa M. Przytycka,[2] Steven P. Gygi,[1] Robert A. Obar,[1] and Spyros Artavanis-Tsakonas[1,5,*]

[1]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA
[2]Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA
[3]Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL 60637, USA
[4]Berkeley *Drosophila* Genome Project, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[5]Biogen Idec, Inc., Cambridge, MA 02142, USA
*Correspondence: sartavanis@cb.med.harvard.edu
http://dx.doi.org/10.1016/j.celrep.2014.08.038

SUMMARY

Specific cellular fates and functions depend on differential gene expression, which occurs primarily at the transcriptional level and is controlled by complex regulatory networks of transcription factors (TFs). TFs act through combinatorial interactions with other TFs, cofactors, and chromatin-remodeling proteins. Here, we define protein-protein interactions using a coaffinity purification/mass spectrometry method and study 459 *Drosophila melanogaster* transcription-related factors, representing approximately half of the established catalog of TFs. We probe this network in vivo, demonstrating functional interactions for many interacting proteins, and test the predictive value of our data set. Building on these analyses, we combine regulatory network inference models with physical interactions to define an integrated network that connects combinatorial TF protein interactions to the transcriptional regulatory network of the cell. We use this integrated network as a tool to connect the functional network of genetic modifiers related to *mastermind*, a transcriptional cofactor of the Notch pathway.

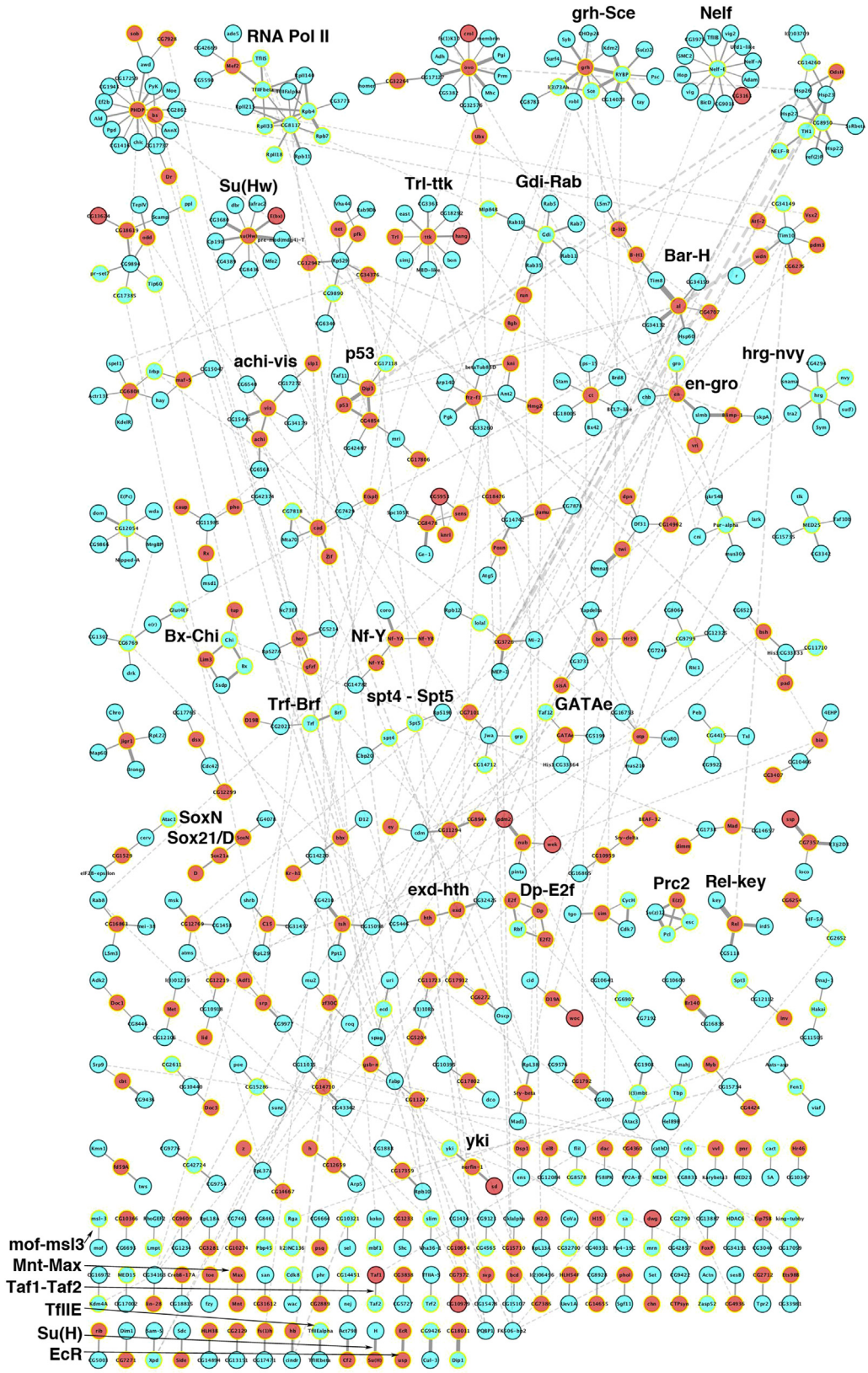
INTRODUCTION

As the targets of signaling pathways and the focal point of gene regulatory networks, transcription factors (TFs) represent a crucial point of regulation relating to the vast majority of cellular processes. As a rule, TFs function through interactions with a wide range of proteins, including other TFs, cofactors, and chromatin modifiers (D'Alessio et al., 2009; Grove and Walhout 2008; Näär et al., 2001; Spitz and Furlong 2012). The biological activity of each TF depends upon these protein interactions, which ultimately govern DNA-binding affinity, activation of chromatin remodeling, and DNA-binding-sequence specificity (Siggers et al., 2011; Slattery et al., 2011). Given the importance of the cooperative action of TFs, defining their protein interaction profile is essential for understanding the regulation of cellular gene expression.

Several studies over the last decade have made inroads into defining the *Drosophila* protein "interactome" (Giot et al., 2003; Guruharsha et al., 2011; Stanyon et al., 2004). Although these studies defined large networks of protein-protein interactions (PPIs), TFs are largely underrepresented in the resulting data sets, likely because the protein preparations (whole-cell extracts) used in these studies were biased against nuclear proteins. Other studies have taken advantage of two-hybrid screening strategies to focus on TF-TF interactions in both *Caenorhabditis elegans* and mammals (Grove et al., 2009; Ravasi et al., 2010) and identified a large number of novel connections between TFs. However, these studies examined only a small portion of the entire TF interactome and, by experimental design, revealed only TF pairs in isolation without taking into account the large repertoire of protein interactions between TFs and other non-TF proteins.

Alternative approaches for exploring the TF interactome include interaction predictions based on coexpression (Adryan and Teichmann, 2010; Suzuki et al., 2009; Tomancak et al., 2007) and combined multiple-TF-occupancy studies (Cole et al., 2008; Lee et al., 2006; Mathur et al., 2008; Roy et al., 2010). In each case, direct interactions must still be confirmed through additional experimental means. Furthermore, TF-occupancy studies treat each TF in isolation, and it has been estimated that only 10%–25% of bound DNA sites in higher eukaryotes result in expression changes in the cognate targets (Spitz and Furlong, 2012). Given the combinatorial nature of TFs and the absence of general rules for their incorporation into protein complexes, systematically defining their interactions would help explain the disconnect between physical binding and functional output, and would contribute substantially to our understanding of gene regulatory networks in the cell.

Toward this goal, we interrogated the protein interaction network of *Drosophila* TFs using a coaffinity purification/mass spectrometry (co-AP/MS) platform. The vast majority of edges



(legend on next page)

in our network are novel, representing new avenues for investigation. As a proof of principle, we used this PPI framework to predict and validate proteins that function in vivo in the Notch signaling network. Building on large-scale expression data sets from modENCODE, we defined tissue-specific PPI networks, addressing the importance of TFs in tissue specification. We also integrated our PPI network with learned regulatory network inference models to create an integrated regulatory network that is linked directly to TF protein complexes. The resulting network enabled us to bridge the gap between our physical PPI data and functional data sets, which we demonstrated by connecting genetic modifiers identified in a genome-wide screen for *mastermind*, a Notch transcriptional coactivator. As regulatory programs are often conserved across species (Erwin and Davidson, 2009), these analyses provide a universal framework for interrogating the biology of TFs and their targets.

RESULTS

Drosophila TF Protein Interaction Networks

TFs are defined as proteins that bind specific sequences of DNA and either activate or repress transcription. They are customarily defined by the presence of one or more sequence-specific DNA-binding domains and are classified into several families depending on the type of DNA-binding domain that is present (Adryan and Teichmann, 2006; Babu et al., 2004; Reece-Hoyes et al., 2005). Of the 14,000 protein-coding genes in the *Drosophila* genome, approximately 708 are TFs with characterized DNA-binding domains (Hammonds et al., 2013). We surveyed the literature and compiled a list of 996 genes, including TFs with characterized binding domains, computationally predicted (putative) TFs, chromatin-related proteins, and transcriptional machinery components (Adryan and Teichmann, 2006; Pfreundt et al., 2010; Table S1). We obtained FLAG-HA-tagged clones encoding 668 of these proteins from the Universal Proteomics Resource (Yu et al., 2011; http://www.fruitfly.org/EST/proteomics.shtml), a part of the Berkeley *Drosophila* Genome Project.

We transiently transfected these clones into *Drosophila* S2R+ cells and generated nuclear extracts, which allowed us to address TF interactions specifically in the context of the nucleus. Protein complexes were isolated by single-step affinity purification, fragmented with trypsin, and analyzed by high-pressure liquid chromatography followed by tandem MS (LC/MS/MS). Approximately 80% of the transfected clones were expressed successfully, as their unique cognate peptides were detected by LC/MS/MS. Across all experiments, we recovered 2,065 proteins with a 2.27% false discovery rate (FDR) from 468 individual affinity purifications (Table S2). This represents approximately one-third of the expressed S2R+ proteome, based on transcriptome and whole-proteome analyses (Cherbas et al., 2011; Guruharsha et al., 2011). From these raw data, we identified 3,407 binary TF-TF interactions, as well as interaction data for 72 chromatin-related proteins and 327 TFs with characterized DNA-binding domains (Table S2).

We subsequently filtered our data using the HyperGeometric Spectral Count Score (HGSCore) method (Guruharsha et al., 2011), taking into account only bait-prey interactions to focus the network specifically on TF interactions and to decrease network noise. In total, 174,561 interactions between the 2,065 identified proteins were analyzed and scored. These scored interactions were filtered to an FDR of 2% based on the use of random data sets, leading to a high-confidence network containing 647 proteins, 229 (35%) of which are characterized TFs (Figure 1; Data S1; Table S3). This interaction network shows a group of 406 proteins (63%) as the giant component of the network and a second group of 241 proteins in smaller, independent protein complexes. Of particular interest, 39% (253) of the proteins in the high-confidence network had no previous functional annotation or were annotated only in silico (by inferred electronic annotation); thus, our map provides direct physical evidence for the functions of these previously uncharacterized proteins. We next analyzed these high-confidence interactions using the Markov clustering algorithm (MCL) (Enright et al., 2002), defining 171 putative protein complexes (Table S3).

TF Network Quality Assessment

In large-scale protein interaction studies, it is difficult to define a reference set of positive interactions due to the small degree of overlap between existing data sets and the lack of a high-quality, manually curated set of interactions for *Drosophila*, such as those established for yeast (Yu et al., 2008). We utilized the *Drosophila* Interactions Database (DroID) (Murali et al., 2011), which contains protein interaction data from nine discrete sources, including recently published large-scale data sets (Friedman et al., 2011; Guruharsha et al., 2011). We found that 18.4% of our high-confidence network overlapped with interactions in DroID with a p value of < 0.001 (Table S4).

We recovered a number of well-characterized complexes, such as RNA polymerase II, Polycomb Repressive Complex 2, and the extradenticle-homothorax TF complex (exd-hth) (Figure 1), and examined several of these complexes in the light of existing interaction data (Figure 2). First, we examined achintya (*achi*) and vismay (*vis*), two homeobox-containing TFs that have been shown to interact with one another to play an essential role in *Drosophila* spermatogenesis, forming a complex with two meiotic arrest genes, always early (*aly*) and cookie monster (*comr*) (Wang and Mann, 2003). Despite their role in spermatogenesis, both genes are expressed across multiple stages in development in both sexes, suggesting additional roles for these

Figure 1. *Drosophila* Transcription Factor Interaction Network

High-confidence interaction network map representing interactions involving 229 site-specific TFs (red nodes). The network contains 647 proteins connected by 624 edges. A total of 117 putative protein complexes were defined using MCL clustering (Enright et al., 2002; Table S3); 9.46% of interactions are binary TF-TF interactions, 21.79% are nonTF-nonTF interactions, and 68.75% are interactions between a TF and a non-TF protein. Protein interactions are shown as gray lines, with line thickness proportional to the HGSCore for the interaction and intercomplex interactions shown in light gray. A number of previously characterized protein complexes are labeled.

See also Tables S1 and S2 and Data S1.

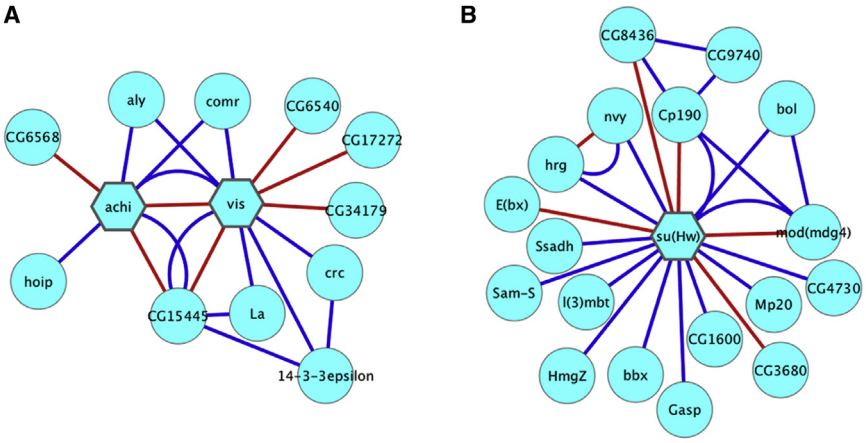


Figure 2. TF Protein Complexes

Interactions identified in our TF study are marked in red, and blue edges represent interactions from the DroID database.

(A) The achintya and vismay protein complex. Previously identified interactions between achi, vis, and CG15445 are recovered. Novel interactions with CG34179, CG6568, CG6540, and CG17272 represent targets for functional studies.

(B) The Su(Hw) protein complex. Known interactions with Cp190 and mod(mdg)4 are recovered. An interaction between CG8436 and Cp190 connects a novel interactor to the known Su(Hw) protein complex.

See also Table S4.

TFs. Since we used an embryonically derived, macrophage-like cell line for our analysis, it may not be surprising that we did not recover interactions with aly and comr. Nevertheless, we recovered a previously identified interaction between achi-vis and CG15445, a protein with unknown function (Figure 2A). In addition, our analysis identified novel interactions with four additional unstudied proteins, providing targets for future work.

We also examined protein interactions related to suppressor of Hairy wing (Su(Hw)), a zinc-finger-containing C2H2 protein that is essential for *gypsy* insulator function and the development of the female germline (Figure 2B). *Su(Hw)* functions through the recruitment of Centrosomal Protein 190 kD (*Cp190*) and Modifier of mdg4 (*mod(mdg)4*) (Georgiev and Kozycina, 1996; Pai et al., 2004). We recovered these interactions in our experiments, confirming the interactions described previously. Furthermore, we recovered an interaction with CG8436, a protein of unknown function that was previously shown to interact with Cp190 (Guruharsha et al., 2011), suggesting a Su(Hw)-related role and providing additional evidence for the validity of this particular interaction.

As an additional means of quality control (outlined in the next section), we performed a genetic screen looking specifically at *Drosophila* proteins that physically interact, according to our data, with known genetic modifiers of the Notch transcription coactivator, mastermind (*mam*). In vivo tests showed that functional genetic interactions were recovered more frequently with the genetic screen compared with random screening. This not only validates a number of our physical interactions functionally but also demonstrates the predictive value of our PPI data with regard to biological function.

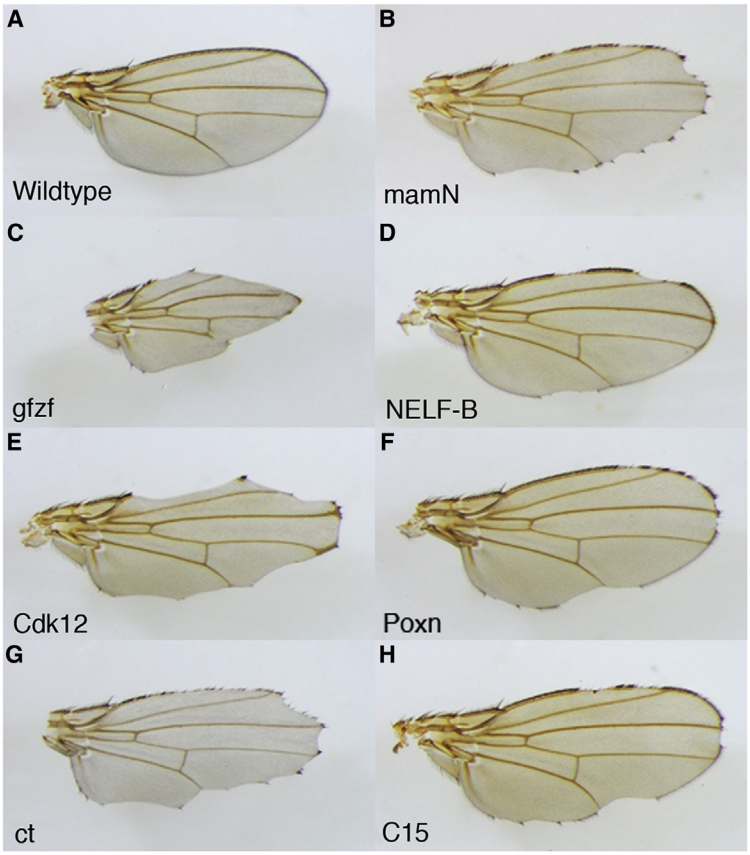
It is important to emphasize that demonstrating the high quality of our data presents a unique challenge due to the lack of a "gold standard" reference set of PPI interactions in *Drosophila* to compare with our data, and the fact that 39% of the proteins in our network are otherwise unstudied. As such, we have used rigorous, established statistical methods to define interactions, leaning heavily on strict statistical cutoffs to limit the number of false-positive interactions in our high-confidence interaction network. The recovery of well-characterized protein complexes and, as outlined below, our ability to functionally validate in vivo relationships predicted by our proteomic data indicate that the network we generated is reliable.

Functional Validation of the TF Interaction Network

An essential aspect of PPI networks is their utility for predicting biological functions and generating hypotheses. We tested predictions from our interactions in vivo, specifically focusing our efforts on the Notch pathway, a conserved fundamental signaling mechanism that broadly controls cell fate during development in metazoans (Artavanis-Tsakonas et al., 1999). In previous genome-wide genetic-modifier studies of a dominant-negative allele of *mam*, a Notch transcriptional coactivator, (Kankel et al., 2007) identified 408 genes that genetically interact with *mam* and recovered genetic modifiers in 4% of the genes screened. This particular screen utilized the Exelixis collection, a transposon-induced mutant collection with insertions in just over half of all genes in the *Drosophila* genome (Thibault et al., 2004; Parks et al., 2004).

Operating under a simple guilt-by-association hypothesis that proteins that interact often share function, we mapped these previously identified genetic modifiers onto our interaction data and identified 88 proteins that physically interact with *mam* modifiers that had not been identified as Notch signal modifiers before (Table S5). To interrogate these 88 genes functionally, we obtained transgenic RNAi alleles under UAS control and crossed them to a dominant-negative C-terminal *mam* truncation driven in the developing wing, *1/2C96-GAL4, UAS-MamN* (*C96-MamN*) (Helms et al., 1999; Kankel et al., 2007; Kitagawa et al., 2001; Wu et al., 2000). We recovered genetic interactions in 35% of our crosses (Figures 3A–3H), representing a 7-fold increase when compared with a random screening and demonstrating clear predictive power for our PPI data.

One of the biggest challenges in interpreting genetic screens is to understand how disparate genes that modify the same pathway are related to one another at a mechanistic level. In one instance, we found that five previously characterized modifiers of the *mam* phenotype—*simj*, *Lim1*, *CG11334*, *fd68A*, and *CG34417*—though previously unlinked to one another, physically interact with cut (*ct*), a transcriptional target of the Notch pathway. ct itself is a TF that was also shown to interact with *mam* in our genetic screen (Figure 3I). As three of the interacting proteins are TFs (the other two are unstudied), this strongly suggests that their functional connection to the Notch signaling pathway may be mediated through TF-TF interactions with ct.



Symbol	c96-MamN x RNAi	%	Total
Arp66B	Suppressor	36.9	111
bin	Suppressor	21.4	126
C15	Suppressor	28.3	46
Cdk12	Enhancer	95.0	100
CG10321	Suppressor	100.0	14
CG1218	Enhancer	22.8	123
CG13183	Suppressor	66.1	62
CG3226	Enhancer	18.8	101
CG6364	Suppressor	26.5	68
ct	Enhancer	100.0	147
Cyp1	Suppressor	40.4	109
EloA	Suppressor	100.0	103
geminin	Suppressor	54.5	88
gfzf	Enhancer	24.2	66
HLH3B	Suppressor	100.0	35
Kdm4A	Suppressor	58.2	55
Ku80	Suppressor	66.0	47
lolal	Suppressor	97.5	40
maf-S	Suppressor	70.6	17
NELF-B	Suppressor	96.3	134
nerfin-1	Suppressor	44.2	120
nub	Suppressor	64.5	31
Poxn	Suppressor	55.5	128
RpS9	Suppressor	63.8	58
SCAR	Enhancer	23.9	113
Sgf29	Suppressor	38.3	133
TfIIS	Suppressor	98.1	104
wdn	Suppressor	64.9	77

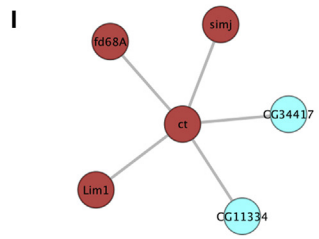


Figure 3. *mastermind* Genetic Screen

(A) Wild-type fly wing.

(B) Dominant-negative *mam* (*c96-mamN*) phenotype.

(C, E, and G) Enhancer phenotypes seen with loss of *gfzf*, *Cdk12*, and *ct*.

(D, F, and H) Suppressor phenotypes seen with loss of *NELF-B*, *Poxn*, and *C15*. Note the presence of patches of wild-type wing margin.

(I) Interactions between previously identified *mam* modifiers and the Notch target gene, *ct*. Red nodes represent TFs.

See also Table S5.

Tissue-Specific Interaction Networks

As a general rule, we expect that proteins that interact are expressed in the same place at the same time. To examine the coexpression and tissue specificity of proteins in our interaction network, we integrated our PPI network with RNA sequencing (RNA-seq) data from the modENCODE project, spanning 29 tissues and developmental time points (Graveley et al., 2011; Smibert et al., 2012). This type of integration between PPI and expression data sets has been used previously to infer network dynamics and identify functional modules within PPI networks (Lin et al., 2010; Przytycka et al., 2010; Tang et al., 2011). Although TFs are often discussed in the context of conferring tissue specificity, a significant proportion of *Drosophila* TFs are expressed ubiquitously at some point during embryonic development (Hammonds et al., 2013; Tomancak et al., 2007) and most exhibit a broad pattern of expression in the adult animal (Adryan and Teichmann, 2010). TFs that show tissue specificity embryonically are usually not limited to a single tissue, but rather exhibit a narrow range of expression in several tissues. These findings suggest that it is not only the presence of a specific TF that defines a particular tissue but also the interactions of these TFs that establish tissue identity.

We scored all proteins in our network using the tissue specificity score (TSPS) (Ravasi et al., 2010). The distribution of TSPS-scored proteins revealed three categories of expression: broad or "general" expression across tissues, expression across several tissues, and high or "specific" tissue specificity (Figure 4A). Low-TSPS proteins, representing broad expression, were assembled into a "core" network of 128 interactions that, based on their ubiquitous expression, are likely to be present across many tissues. We then focused on the group of high-scoring TSPS proteins, utilizing an outlier method (Kadota et al., 2003) to assign each protein to specific tissues (Table S6). We combined these high-specificity proteins with our "core" network to build 24 different tissue-specific interaction networks (Data S2).

Two very different protein complexes, one specific to the testis and the other specific to the larval CNS, illustrate the value of this tissue-specificity analysis (Figures 4B and 4C). The first complex is centered on an unnamed protein, CG8117, which according to

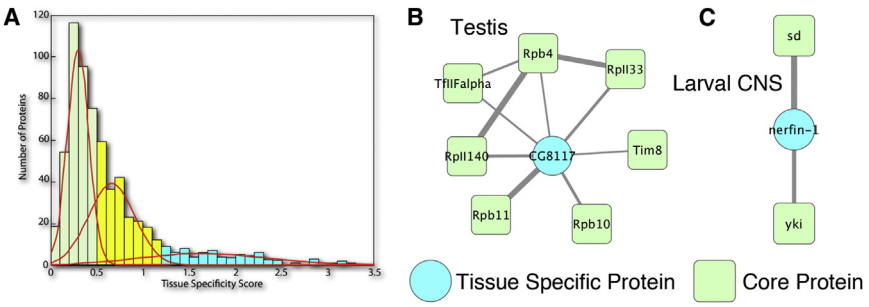


Figure 4. Tissue-Specific Protein Complexes

(A) Tissue-specificity distribution for all proteins in the high-confidence interaction network scored using the TSPS algorithm. Low-specificity proteins are labeled in green, moderate-specificity proteins are labeled in yellow, and high-specificity proteins are labeled in blue. Distribution was fit to a trimodal distribution and bins were defined with cutoffs of 0.4781 and 1.1741.

(B) Testis-specific protein complex. Rounded squares represent "core" network proteins and blue circles represent "specific" proteins. *CG8117* is an ortholog of a human testis-specific transcription elongation factor, also expressed specifically in the *Drosophila* testis. The other polymerase II components are expressed broadly.

(C) Larval CNS-specific protein complex. *Nerfin-1* is highly specific to the larval CNS. It interacts with two low-specificity proteins: the TF sd and the transcription coactivator yki.

See also Table S6 and Data S2.

our results is a part of the RNA polymerase II complex, and is connected in our map through eight physical edges (Figure 4B). CG8117 is electronically inferred to have transcription regulatory activity and to bind both zinc ions and nucleic acids. It is expressed at high levels in the adult testis, but is largely absent from other tissues (Chintapalli et al., 2007). Outside of large-scale screens, CG8117 has not been independently studied in *Drosophila*. However, the human ortholog of this protein, TCEA2, has been characterized to be a testis-specific TF (Weaver and Kane, 1997), suggesting that this gene could play a similar tissue-specific role in *Drosophila*.

The second protein complex we wish to highlight links two TFs, nervous fingers 1 (*nerfin-1*) and scalloped (*sd*), to the transcriptional coactivator yorkie (*yki*) (Figure 4C). sd is expressed in the developing nervous system, where it is essential for development of the sensory organs (Campbell et al., 1992). nerfin-1 has been shown to be important for axon guidance during early CNS development (Kuzin et al., 2005). yki, the *Drosophila* ortholog of the human protein YAP, is a transcriptional coactivator that functions in the hippo-yap pathway and was previously shown to interact with *sd* (Goulev et al., 2008). It has also been suggested that nerfin-1 is a binding partner of sd. Both nerfin-1 and sd are expressed in a highly specific manner in the larval CNS, and given their established importance in CNS development and their physical interaction in our map, we can formulate the hypothesis that they work together to regulate CNS development, possibly in tandem with the coactivator yki.

Combinatorial Targets of Interacting TFs

Given the importance of combinatorial TF interactions in gene regulation, we compared the in vitro PPI data with in vivo DNA-binding data for all protein pairs for which genome-wide chromatin immunoprecipitation (ChIP) data were available (Table S7). We identified multiple pairs in which the PPIs and regulatory-factor cobinding were consistent with the existing literature. For example, we observed an interaction between ecdysone receptor (*EcR*) and ultraspiracle (*USP*), which are the two proteins that comprise the complete EcR complex. Upon ligand binding, EcR and USP are activated and coordinately regulate genes such as *Eip75B* and *DHR3* (Yao et al., 1993; Figure 5A). We also recovered an interaction between polycomblike (*Pcl*) and enhancer of zeste (*E(z)*), two proteins that are members of Pcl-polycomb repressive complex 2 (Pcl-PRC2; Figure 5B), as well as an interaction between the segment polarity gene engrailed (*en*) and the corepressor groucho (*gro*) (Figure 5C; Hittinger and Carroll, 2008).

Beyond these well-characterized complexes, we found several examples of PPIs that are supported by colocalization on DNA. For instance, we observed an interaction between tramtrack (*ttk*) and Trithorax-like (*Trl*) (Figure 5D). Both are BTB/POZ (Br-C, ttk, and bab/Pox virus and Zinc finger) domain-containing proteins. This interaction was previously identified in a large-scale yeast two-hybrid screen and in *Drosophila* S2 cells, providing additional evidence for this particular interaction (Pagans et al., 2002). ttk has been shown to function as both a transcriptional repressor and activator, playing a variety of roles in processes such as nervous system development, photoreceptor differentiation, and tracheal development (Araújo et al., 2007; Badenhorst, 2001; Lai and Li, 1999). Trl (also known as GAGA factor or GAF) has been suggested to play a role in transcriptional activation through chromatin changes, and in some cases is necessary for transcriptional activation driven by some TF complexes (Bayarmagnai et al., 2012; Granok et al., 1995). This raises the possibility that ttk activity is modulated through Trl interactions and likely plays a role in activating the gene expression of shared targets.

Inferred Regulatory Motifs for TF Complexes

To gain insight into the regulatory consequences of the PPI in our network, we integrated our results with existing inferred regulatory network models (Marbach et al., 2012). These inferred networks integrate a wide range of data sets, including TF binding, gene expression, and chromatin modifications, utilizing supervised and unsupervised machine learning frameworks to predict regulatory edges. These networks have been shown to be useful tools for predicting gene function (Marbach et al., 2012). It is important to note, however, that PPI data were not included in the assembly of these particular networks. By integrating our PPI data with such transcriptional regulatory networks, we were able to provide a new dimension to our analysis and gain insight into the combinatorial action of interacting TFs by linking their regulatory edges directly to TF protein complexes.

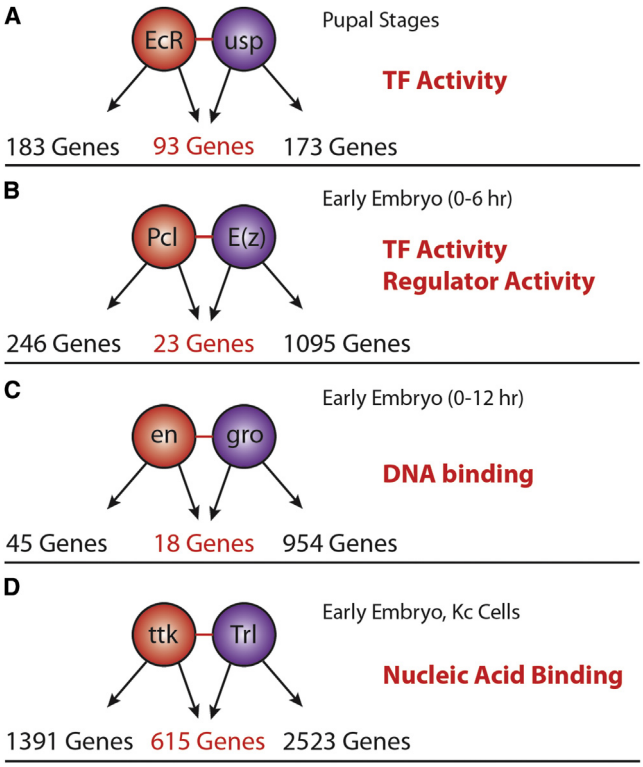


Figure 5. Combinatorial Targets of Interacting TFs

Shared physical targets of interacting TF pairs. Enriched Gene Ontology terms for shared targets are delineated in red.

(A) Ecdysone receptor (*EcR*) and ultraspiracle (*usp*) comprise the two parts of the complete EcR complex. They co-occupy 93 shared targets during pupal stages.

(B) Polycomblike (*Pcl*) and Enhancer of zeste (*E(z)*), two members of the Pcl-PRC2 complex.

(C) engrailed (*en*) and groucho (*gro*).

(D) tramtrack (*ttk*) and Trithoraxlike (*Trl*), two BTB/POZ domain-containing proteins.

See also Table S7.

To combine the PPIs with regulatory interaction and probe these large integrated networks, we defined a set of TF regulatory motifs based on physical and regulatory interactions (Figure 6A; Table S8). These three motifs represent instances in which (1) an interacting protein is regulated by its binding partner, (2) two interacting proteins regulate the same target, and (3) a single factor regulates interacting proteins. Each instance of each of these motifs essentially defines a biological hypothesis, representing an avenue for future inquiry.

By permuting the edges of both our high-confidence PPI network and the inferred regulatory networks independently, we confirmed that these motifs occur significantly more frequently than expected by chance (Figure 6A). Furthermore, as we have demonstrated the predictive power of the high-confidence interactions in our PPI network, focusing only on motifs that contain one of our PPI edges effectively filters the regulatory network based on experimental evidence. We combined these motifs to build networks containing 22,781 edges between 3,145 proteins in a supervised model, and 19,062 edges between 2,331 proteins in an unsupervised model (Data S3).

Within the supervised model, we highlighted the regulatory network related to Dp TF and E2F, members of the dREAM (RBF, dE2F2, and dMyb) complex (Figures 6B–6I). The dREAM complex is conserved in most eukaryotes and plays multiple roles, including regulation of cell fate, cell division, and apoptosis (van den Heuvel and Dyson, 2008). Dp and E2f comprise a dimeric TF that is important in the G1/S phase transition during the cell cycle, where E2f levels are rate limiting for cell proliferation (Johnson et al., 1993).

Previous work has described interactions between E2f and both Dp and Rbf, consistent with interactions revealed in our network (Frolov et al., 2001; Figure 6B). Another component of the dREAM complex, Myb, acts in a mutually exclusive manner with Dp-E2f to regulate target selection (Georlette et al., 2007). Although we did not recover Myb as a physical interactor, it is one of only three proteins that are inferred to both regulate Dp-E2f and in turn be targeted by the TF pair. The other two proteins are MTA1-like and CG17385, which have not previously been tied to dREAM functions and thus represent targets for functional analyses (Figure 6B). As expected, downstream targets of Dp-E2f in our network include genes that are important for the cell cycle (Figure 6E) and DNA replication (Figure 6F).

The dREAM complex is thought to modulate transcription through the repressive binding of Rbf to E2f, by inhibiting the basal transcription machinery, and by recruiting chromatin-modifying proteins (Georlette et al., 2007). Our regulatory network reflects all three of these possibilities, showing a physical interaction between Rbf and E2f, the targeting of a number of basal transcriptional machinery components (Figure 6D), and the regulation of chromatin-modifying proteins such as brahma and MRG15 (Figure 6H). Other downstream targets of Dp-E2f in our network include a group that is largely enriched for transcription-related proteins (Figure 6G) and 28 targets that are not annotated (Figure 6I). *Dp* and *E2f* are themselves targeted by a cohort of TFs and cofactors, including DREF, Mad, and Trl (Figure 6C). Consequently, we have identified a well-characterized protein complex, a number of its known regulatory targets, and, most interestingly, targets that have not previously been linked to dREAM complex function, thereby defining new specific functional hypotheses.

Connecting Functional Networks

Genetic screens, especially in *Drosophila*, have been used as a powerful tool to define networks of proteins that share function (Fortini and Artavanis-Tsakonas, 1994; Go and Artavanis-Tsakonas, 1998; Sen et al., 2013; St Johnston, 2002; Xu and Artavanis-Tsakonas, 1990). However, it is difficult to understand, at a mechanistic level, how these proteins are connected to one another. On the other end of the spectrum, PPI networks describe the physical relationships between proteins, but do not capture functional relationships. Although there is some overlap between these two network types, not every functional relationship is the result of a direct PPI. As such, the majority of network edges between these two data types do not typically overlap. By combining transcriptional regulatory networks with our PPI data, we developed an integrated network that can

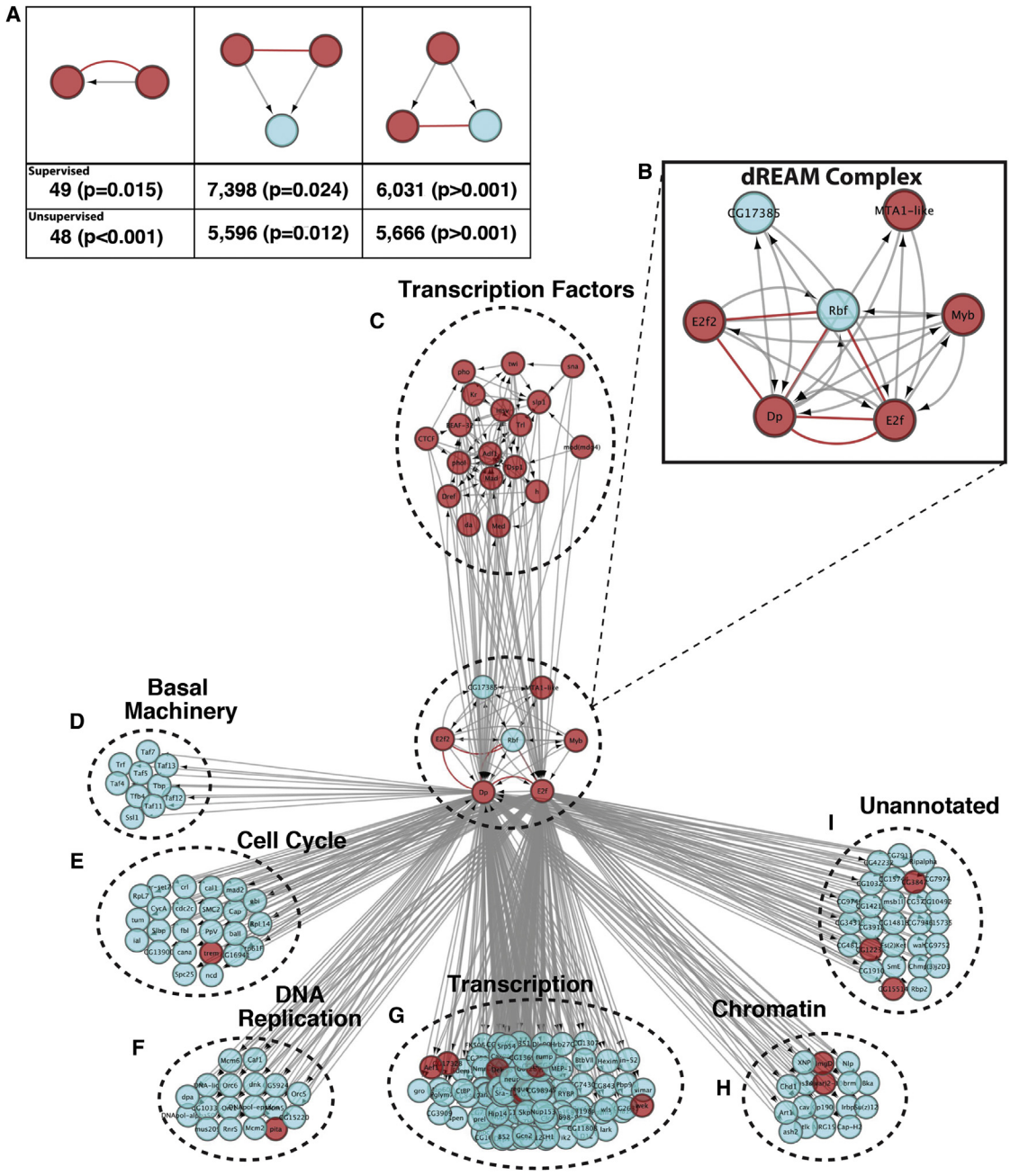


Figure 6. Inferred Regulatory Edges for Transcriptional Complexes

(A) Transcriptional regulatory motifs, representing instances in which an interacting protein regulates its binding partner (1:1), combinatorial regulation of a target by two interacting factors (2:1), and regulation of interacting proteins by a single factor (1:2). Red edges indicate PPIs and gray edges with arrows indicate directional regulatory edges.

(B) Components of the *Drosophila* dREAM complex recovered in our interaction network.

(C) Transcriptional regulators of *Dp-E2f*.

(D) Basal transcription machinery components.

(E) Cell-cycle proteins.

(F) DNA replication-related proteins.

(G) Transcription.

(H) Chromatin related.

(I) Unannotated targets of Dp/E2f.

See also Table S8 and Data S3.

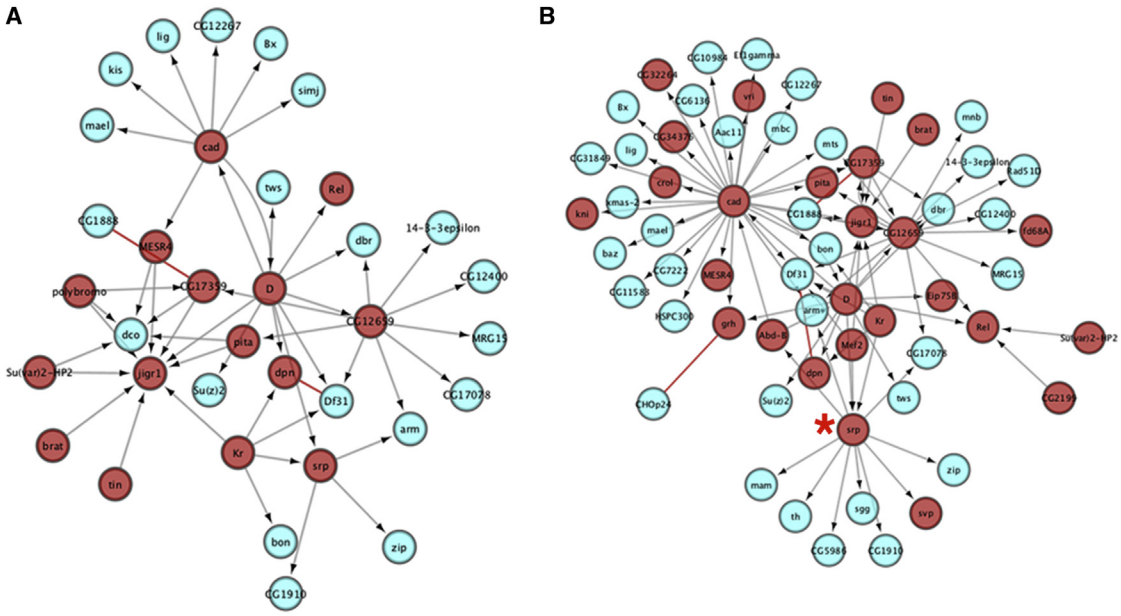


Figure 7. Connecting the *mastermind* Genetic Network

(A) Unsupervised network view of 35 *mastermind* modifiers.

(B) Supervised network view of 62 *mastermind* modifiers. All nodes in interaction network are previously identified *mastermind* modifiers. Red nodes represent TFs. Blue nodes represent non-TF proteins. Red edges represent PPIs. Gray edges with arrows represent directional regulatory edges. The red asterisk indicates interactions related to serpent (*srp*).

See also Data S3.

bridge the gap between physical and functional relationships through defined regulatory edges, and provide insight into the combinatorial regulation of targets by interacting TFs.

As an example, we once again focus on the genetic interaction network of *mam*, as defined in a genome-wide screen in *Drosophila* (Kankel et al., 2007). A total of 408 genes were shown to interact genetically with *mam* in vivo, and our supervised and unsupervised integrated networks contain 140 and 103 of these modifiers, respectively. If we examine the direct relationships between these genes in our networks, we find that 62 and 35 proteins, respectively, are directly linked to one another (Figures 7A and 7B). If we expand this view to include first-neighbor interactions, all *mastermind* modifiers in both instances are connected to one another.

The organization of these networks reveals several potential "hubs" of regulation based on the total number of edges that connect to a particular node. For instance, the TF serpent (*srp*) is connected by 12 separate network edges in our supervised network (Figure 7B). Although *srp* itself has not been demonstrated to be directly regulated by the Notch pathway, it was previously shown to function upstream of direct Notch targets during *Drosophila* larval hematopoiesis (Duvic et al., 2002). This would suggest a potential mechanism by which loss of *srp* could modulate Notch activity downstream, thus explaining the genetic interaction between *srp* and *mam*. Interestingly, our network identifies *mam* as a direct target of *srp*.

While these regulatory edges will certainly vary depending on the context, this approach provides a network of hypotheses regarding the connections between functional data points. These connections can be used as the basis for probing the mechanisms that link these functionally related proteins. We expect that as more data become available, these networks will be further refined and expanded to provide higher-resolution insight into the mechanisms that drive biological function. As things stand, our integrated networks provide a substantial foundation for exploring the mechanisms that connect functional data sets.

DISCUSSION

In this work, we performed a network analysis of TFs in *Drosophila melanogaster* by determining TF protein interactions using a co-AP/MS approach. In addition to PPIs for TFs, our integrated analysis takes into account tissue-specific interaction subnetworks and candidate combinatorial gene targets. This integrated regulatory network approach can be used to examine combinatorial transcriptional regulation and probe functional data sets.

Although a considerable number of our interaction results are novel, we demonstrated the high quality of our findings by recovering previously identified interactions, as well as by functional validation in vivo. Given the lack of a reference set of positive interactions based on a high-quality, manually curated set of interactions for *Drosophila*, we based our interaction network on strict statistical cutoffs, which minimize false positives but may also mask interactions of interest. It will therefore be important for those interested in a specific protein to examine our raw interaction data, as many well-characterized interactions fell below

our deliberately strict statistical cutoff. We acknowledge several limitations to our methods, in particular, the use of epitope-tagged fusion proteins expressed at nonphysiological levels. Although we cannot ignore the fact that in some cases epitope tags can perturb protein folding and function, the recovery of previously characterized interactions, including those identified via alternative methods (e.g., two-hybrid screening), provides additional evidence of the validity of our experimental pipeline. Furthermore, similar methods have been used successfully to identify confirmed interactions in a number of settings, including the human autophagy system, as well as a proteome-wide analysis in *Drosophila* (Behrends et al., 2010; Guruharsha et al., 2011; Sowa et al., 2009).

Our tissue-specific subnetworks emphasize the importance of context with regard to TF function. We defined groups of proteins based on their broad or specific expression, and then connected these categories, providing insight into how general and specific TFs cooperate with one another to drive transcriptional programs. As suggested in a previous study (Ravasi et al., 2010), it is likely that the presence of a particular TF protein interaction within a specific tissue, rather than the expression of a single tissue-specific TF, confers tissue identity.

Building on the recent availability of large-scale genomic data sets from the modENCODE project and others, we focused on connecting the TF PPI network with the gene-regulatory network of the cell. As previous work has shown (Spitz and Furlong, 2012), TFs do not function in isolation, nor does physical binding necessarily correlate to a change in gene expression. It is the combination of various TFs and their interacting proteins that confers a specific activity and thus defined common physical targets between interacting sets of proteins. Indeed, we find multiple examples of protein interactions that are supported by the genome-wide DNA-binding data (e.g., EcR-Usp, Pcl-E(z), and En-Gro), as well as interactions that warrant further exploration. Our data provide an extensive catalog of physical interactions that can be used to probe function at the level of the complex.

We also connected TF protein complexes to the gene-regulatory network using inferred regulatory edges, which allowed us to expand target prediction beyond direct physical targets. We established the predictive value of the physical edges in our network, which likely improves the quality of the inferred regulatory network, as we examined only edges that are directly linked to an experimentally validated physical interaction. The analysis of the *Drosophila* dREAM protein complex presented here, including the identification of both characterized and novel targets, is indicative of the utility of this integrated network.

Finally, we used our integrated networks to interrogate large-scale functional data sets. Although genetic screens have been used for decades, connecting the large number of functional modifiers identified in these screens to one another has been a significant challenge. Gene Ontology analysis certainly provides insight into the categorization of genes within these data sets; however, the complex relationships between these components can only be captured from a network perspective. Our integrated network analysis provides a considerable framework from which to build hypotheses as to how various functionally connected proteins are related to one another.

We view our data as a framework for developing specific hypotheses for future studies in *Drosophila* and other metazoans. Given the conservation of regulatory programs, it is likely that many of the regulatory connections presented here have been preserved in other species, though possibly (and interestingly) used in different biological contexts. As TFs represent fundamental points of regulation in the cell, we expect the present work to be relevant to a broad spectrum of biological processes.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

C-terminal FLAG-HA-tagged TF clones in the pMK33-CFH-BD vector were acquired from the Berkeley *Drosophila* Genome Project (Yu et al., 2011). Each clone was transiently transfected into two 54 ml cultures of *Drosophila* S2R+ cells using Effectene (QIAGEN) and subsequently cultured in Schneider's media with 10% fetal bovine serum. At 24 hr posttransfection, gene expression was induced with 0.35 mM $CuSO_4$, and cells were harvested 24 hr after induction (Veraksa et al., 2005). Nuclear extracts were prepared as previously described, with the exception that cells were lysed using an 18-gauge syringe (Dignam et al., 1983). Nuclear extracts were diluted 1:1 with dialysis buffer (20 mM HEPES pH 7.6, 20% glycerol, 100 mM KCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM dithiothreitol, 0.25 mM phenylmethanesulfonyl fluoride, and Roche cOmplete Protease Inhibitor) to reduce the overall salt concentration. Each extract was incubated with 40 μl of dimethyl pimelimidate crosslinked anti-HA immunoaffinity resin (Sigma) for 3 hr at 4°C. Following incubation, the resin was washed twice with dialysis buffer followed by two PBS washes. Bound proteins were eluted using IgG elution buffer (400 μl total; Thermo Scientific Pierce) divided into two separate 5 min incubations performed at room temperature with gentle shaking. The elution was then neutralized with 52 μl 1 M Tris, pH 8.0.

MS and Network Construction

Copurified proteins were precipitated with trichloroacetic acid (TCA), followed by a 10% TCA wash and two acetone washes. The samples were then dried, digested overnight with trypsin, cleaned with c18 Stage Tips (Thermo Scientific), and analyzed by LC-MS/MS on a linear ion trap quadrupole (Thermo Scientific) instrument. MS/MS spectra were searched with SEQUEST (Eng et al., 2008) against FlyBase release 5.41 and filtered to a 2.27% protein FDR for the entire data set with the reverse database approach (Elias and Gygi 2007). Column carry-over between experiments was corrected with a statistical approach, incorporating peptide abundance and the probability of consecutive observations. Following processing and filtering, a high-confidence TF interaction map was generated using the HGSCore method to distinguish specific interactions as described previously, but with indirect prey-prey interactions filtered out to focus the network on the TF-interacting subspace. To draw the cutoff for interaction specificity and determine the FDR, we ran HGSCore on 40 simulated data sets randomly sampled from the real data set, until convergence on a cutoff score, resulting in a 2% FDR. This high-confidence interaction network was clustered using MCL (Enright et al., 2002) with an inflation value of 1.8.

Genetic Screen

Flies were cultured on standard media and crosses were carried out at 23°C. The *C96-Gal4*, *UASMamN* (*C96-MamN*) stocks were previously described (Helms et al., 1999). UAS-RNAi fly stocks were obtained from the TRiP collection at Harvard Medical School (NIH/NIGMRS R01-GM084947). Adult fly wings were dehydrated in isopropanol and mounted in a 3:1 dilution of CMCP-10 (Masters Company) and lactic acid.

Tissue Specificity Analysis

The TSPS was executed as previously described (Ravasi et al., 2010) utilizing 24 mRNA-sequencing data sets (Smibert et al., 2012), encompassing 24 groups containing various tissues dissected from *Oregon R* wild-type flies. The distribution for all proteins based on their TSPS was fit to a trimodal

Gaussian distribution, with cutoff values of 0.4781 for low (general)-specificity proteins and 1.17406 for high (specific)-specificity proteins. High-specificity proteins, based on TSPS distribution, were assigned to specific tissues according to previously described methods (Kadota et al., 2003).

ChIP Data

ChIP data from both the modENCODE Project (Roy et al., 2010) and the Berkeley *Drosophila* Transcription Network Project (MacArthur et al., 2009) were used. For published ChIP-chip and ChIP-seq data sets, filtered peaks were taken directly from the published analyses. New ChIP-seq data sets were generated as described previously (Roy et al., 2010) and analyzed through the Irreproducible Discovery Rate data analysis pipeline (described in detail at https://sites.google.com/site/anshulkundaje/projects/idr).

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight tables and three data files and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.08.038.

ACKNOWLEDGMENTS

We thank Mark Kankel, Mat Sowa, Jean-Francois Rual, and members of S.A.-T.'s lab for helpful discussions and comments. This work was supported by a grant from the NIH (5RO1HG003616) to S.A.-T. Generation of the clone set was supported by a grant from the NHGRI (P41HG3487) to S.E.C. D.-Y.C. and T.M.P. were supported by the Intramural Research Program of the National Library of Medicine, NIH.

Received: February 25, 2014
Revised: June 9, 2014
Accepted: August 16, 2014
Published: September 18, 2014

REFERENCES

Adryan, B., and Teichmann, S.A. (2006). FlyTF: a systematic review of site-specific transcription factors in the fruit fly Drosophila melanogaster. Bioinformatics *22*, 1532–1533.

Adryan, B., and Teichmann, S.A. (2010). The developmental expression dynamics of Drosophila melanogaster transcription factors. Genome Biol. *11*, R40.

Araújo, S.J., Cela, C., and Llimargas, M. (2007). Tramtrack regulates different morphogenetic events during Drosophila tracheal development. Development *134*, 3665–3676.

Artavanis-Tsakonas, S., Rand, M.D., and Lake, R.J. (1999). Notch signaling: cell fate control and signal integration in development. Science *284*, 770–776.

Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M., and Teichmann, S.A. (2004). Structure and evolution of transcriptional regulatory networks. Curr. Opin. Struct. Biol. *14*, 283–291.

Badenhorst, P. (2001). Tramtrack controls glial number and identity in the Drosophila embryonic CNS. Development *128*, 4093–4101.

Bayarmagnai, B., Nicolay, B.N., Islam, A.B., Lopez-Bigas, N., and Frolov, M.V. (2012). Drosophila GAGA factor is required for full activation of the dE2f1-Yki/Sd transcriptional program. Cell Cycle *11*, 4191–4202.

Behrends, C., Sowa, M.E., Gygi, S.P., and Harper, J.W. (2010). Network organization of the human autophagy system. Nature *466*, 68–76.

Campbell, S., Inamdar, M., Rodrigues, V., Raghavan, V., Palazzolo, M., and Chovnick, A. (1992). The scalloped gene encodes a novel, evolutionarily conserved transcription factor required for sensory organ differentiation in Drosophila. Genes Dev. *6*, 367–379.

Cherbas, L., Willingham, A., Zhang, D., Yang, L., Zou, Y., Eads, B.D., Carlson, J.W., Landolin, J.M., Kapranov, P., Dumais, J., et al. (2011). The transcriptional diversity of 25 Drosophila cell lines. Genome Res. *21*, 301–314.

Chintapalli, V.R., Wang, J., and Dow, J.A. (2007). Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. *39*, 715–720.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev. *22*, 746–755.

D'Alessio, J.A., Wright, K.J., and Tjian, R. (2009). Shifting players and paradigms in cell-specific transcription. Mol. Cell *36*, 924–931.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

Duvic, B., Hoffmann, J.A., Meister, M., and Royet, J. (2002). Notch signaling controls lineage specification during Drosophila larval hematopoiesis. Curr. Biol. *12*, 1923–1927.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods *4*, 207–214.

Eng, J.K., Fischer, B., Grossmann, J., and Maccoss, M.J. (2008). A fast SEQUEST cross correlation algorithm. J. Proteome Res. *7*, 4598–4602.

Enright, A.J., Van Dongen, S., and Ouzounis, C.A. (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. *30*, 1575–1584.

Erwin, D.H., and Davidson, E.H. (2009). The evolution of hierarchical gene regulatory networks. Nat. Rev. Genet. *10*, 141–148.

Fortini, M.E., and Artavanis-Tsakonas, S. (1994). The suppressor of hairless protein participates in notch receptor signaling. Cell *79*, 273–282.

Friedman, A.A., Tucker, G., Singh, R., Yan, D., Vinayagam, A., Hu, Y., Binari, R., Hong, P., Sun, X., Porto, M., et al. (2011). Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal-regulated kinase signaling. Sci. Signal. *4*, rs10.

Frolov, M.V., Huen, D.S., Stevaux, O., Dimova, D., Balczarek-Strang, K., Elsdon, M., and Dyson, N.J. (2001). Functional antagonism between E2F family members. Genes Dev. *15*, 2146–2160.

Georgiev, P., and Kozycina, M. (1996). Interaction between mutations in the suppressor of Hairy wing and modifier of mdg4 genes of Drosophila melanogaster affecting the phenotype of gypsy-induced mutations. Genetics *142*, 425–436.

Georlette, D., Ahn, S., MacAlpine, D.M., Cheung, E., Lewis, P.W., Beall, E.L., Bell, S.P., Speed, T., Manak, J.R., and Botchan, M.R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. Genes Dev. *21*, 2880–2896.

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., et al. (2003). A protein interaction map of Drosophila melanogaster. Science *302*, 1727–1736.

Go, M.J., and Artavanis-Tsakonas, S. (1998). A genetic screen for novel components of the notch signaling pathway during Drosophila bristle development. Genetics *150*, 211–220.

Goulev, Y., Fauny, J.D., Gonzalez-Marti, B., Flagiello, D., Silber, J., and Zider, A. (2008). SCALLOPED interacts with YORKIE, the nuclear effector of the hippo tumor-suppressor pathway in Drosophila. Curr. Biol. *18*, 435–441.

Granok, H., Leibovitch, B.A., Shaffer, C.D., and Elgin, S.C. (1995). Chromatin. Ga-ga over GAGA factor. Curr. Biol. *5*, 238–241.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of Drosophila melanogaster. Nature *471*, 473–479.

Grove, C.A., and Walhout, A.J. (2008). Transcription factor functionality and transcription regulatory networks. Mol. Biosyst. *4*, 309–314.

Grove, C.A., De Masi, F., Barrasa, M.I., Newburger, D.E., Alkema, M.J., Bulyk, M.L., and Walhout, A.J. (2009). A multiparameter network reveals extensive divergence between C. elegans bHLH transcription factors. Cell *138*, 314–327.

Guruharsha, K.G., Rual, J.F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., et al. (2011). A protein complex network of Drosophila melanogaster. Cell *147*, 690–703.

Hammonds, A.S., Bristow, C.A., Fisher, W.W., Weiszmann, R., Wu, S., Hartenstein, V., Kellis, M., Yu, B., Frise, E., and Celniker, S.E. (2013). Spatial expression of transcription factors in Drosophila embryonic organ development. Genome Biol. *14*, R140.

Helms, W., Lee, H., Ammerman, M., Parks, A.L., Muskavitch, M.A., and Yedvobnick, B. (1999). Engineered truncations in the Drosophila mastermind protein disrupt Notch pathway function. Dev. Biol. *215*, 358–374.

Hittinger, C.T., and Carroll, S.B. (2008). Evolution of an insect-specific GROUCHO-interaction motif in the ENGRAILED selector protein. Evol. Dev. *10*, 537–545.

Johnson, D.G., Schwarz, J.K., Cress, W.D., and Nevins, J.R. (1993). Expression of transcription factor E2F1 induces quiescent cells to enter S phase. Nature *365*, 349–352.

Kadota, K., Nishimura, S., Bono, H., Nakamura, S., Hayashizaki, Y., Okazaki, Y., and Takahashi, K. (2003). Detection of genes with tissue-specific expression patterns using Akaike's information criterion procedure. Physiol. Genomics *12*, 251–259.

Kankel, M.W., Hurlbut, G.D., Upadhyay, G., Yajnik, V., Yedvobnick, B., and Artavanis-Tsakonas, S. (2007). Investigating the genetic circuitry of mastermind in Drosophila, a notch signal effector. Genetics *177*, 2493–2505.

Kitagawa, M., Oyama, T., Kawashima, T., Yedvobnick, B., Kumar, A., Matsuno, K., and Harigaya, K. (2001). A human protein with sequence similarity to Drosophila mastermind coordinates the nuclear form of notch and a CSL protein to build a transcriptional activator complex on target promoters. Mol. Cell. Biol. *21*, 4337–4346.

Kuzin, A., Brody, T., Moore, A.W., and Odenwald, W.F. (2005). Nerfin-1 is required for early axon guidance decisions in the developing Drosophila CNS. Dev. Biol. *277*, 347–365.

Lai, Z.C., and Li, Y. (1999). Tramtrack69 is positively and autonomously required for Drosophila photoreceptor development. Genetics *152*, 299–305.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., et al. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301–313.

Lin, C.C., Hsiang, J.T., Wu, C.Y., Oyang, Y.J., Juan, H.F., and Huang, H.C. (2010). Dynamic functional modules in co-expressed protein interaction networks of dilated cardiomyopathy. BMC Syst. Biol. *4*, 138.

MacArthur, S., Li, X.Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondona, B.P., Hechmer, A., Simirenko, L., Keränen, S.V., et al. (2009). Developmental roles of 21 Drosophila transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. Genome Biol. *10*, R80.

Marbach, D., Roy, S., Ay, F., Meyer, P.E., Candeias, R., Kahveci, T., Bristow, C.A., and Kellis, M. (2012). Predictive regulatory models in Drosophila melanogaster by integrative inference of transcriptional networks. Genome Res. *22*, 1334–1349.

Mathur, D., Danford, T.W., Boyer, L.A., Young, R.A., Gifford, D.K., and Jaenisch, R. (2008). Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET. Genome Biol. *9*, R126.

Murali, T., Pacifico, S., Yu, J., Guest, S., Roberts, G.G., 3rd, and Finley, R.L., Jr. (2011). DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. Nucleic Acids Res. *39*, D736–D743.

Näär, A.M., Lemon, B.D., and Tjian, R. (2001). Transcriptional coactivator complexes. Annu. Rev. Biochem. *70*, 475–501.

Pagans, S., Ortiz-Lombardía, M., Espinás, M.L., Bernués, J., and Azorín, F. (2002). The Drosophila transcription factor tramtrack (TTK) interacts with Trithorax-like (GAGA) and represses GAGA-mediated activation. Nucleic Acids Res. *30*, 4406–4413.

Pai, C.Y., Lei, E.P., Ghosh, D., and Corces, V.G. (2004). The centrosomal protein CP190 is a component of the gypsy chromatin insulator. Mol. Cell *16*, 737–748.

Parks, A.L., Cook, K.R., Belvin, M., Dompe, N.A., Fawcett, R., Huppert, K., Tan, L.R., Winter, C.G., Bogart, K.P., Deal, J.E., et al. (2004). Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. Nat. Genet. *36*, 288–292.

Pfreundt, U., James, D.P., Tweedie, S., Wilson, D., Teichmann, S.A., and Adryan, B. (2010). FlyTF: improved annotation and enhanced functionality of the Drosophila transcription factor database. Nucleic Acids Res. *38*, D443–D447.

Przytycka, T.M., Singh, M., and Slonim, D.K. (2010). Toward the dynamic interactome: it's about time. Brief. Bioinform. *11*, 15–29.

Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744–752.

Reece-Hoyes, J.S., Deplancke, B., Shingles, J., Grove, C.A., Hope, I.A., and Walhout, A.J. (2005). A compendium of Caenorhabditis elegans regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol. *6*, R110.

Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., et al.; modENCODE Consortium (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science *330*, 1787–1797.

Sen, A., Dimlich, D.N., Guruharsha, K.G., Kankel, M.W., Hori, K., Yokokura, T., Brachat, S., Richardson, D., Loureiro, J., Sivasankaran, R., et al. (2013). Genetic circuitry of Survival motor neuron, the gene underlying spinal muscular atrophy. Proc. Natl. Acad. Sci. USA *110*, E2371–E2380.

Siggers, T., Duyzend, M.H., Reddy, J., Khan, S., and Bulyk, M.L. (2011). Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. Mol. Syst. Biol. *7*, 555.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., and Mann, R.S. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell *147*, 1270–1282.

Smibert, P., Miura, P., Westholm, J.O., Shenker, S., May, G., Duff, M.O., Zhang, D., Eads, B.D., Carlson, J., Brown, J.B., et al. (2012). Global patterns of tissue-specific alternative polyadenylation in Drosophila. Cell Rep. *1*, 277–289.

Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the human deubiquitinating enzyme interaction landscape. Cell *138*, 389–403.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. Nat. Rev. Genet. *3*, 176–188.

Stanyon, C.A., Liu, G., Mangiola, B.A., Patel, N., Giot, L., Kuang, B., Zhang, H., Zhong, J., and Finley, R.L., Jr. (2004). A Drosophila protein-interaction map centered on cell-cycle regulators. Genome Biol. *5*, R96.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J., et al.; FANTOM Consortium; Riken Omics Science Center (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat. Genet. *41*, 553–562.

Tang, X., Wang, J., Liu, B., Li, M., Chen, G., and Pan, Y. (2011). A comparison of the functional modules identified from time course and static PPI network data. BMC Bioinformatics *12*, 339.

Thibault, S.T., Singer, M.A., Miyazaki, W.Y., Milash, B., Dompe, N.A., Singh, C.M., Buchholz, R., Demsky, M., Fawcett, R., Francis-Lang, H.L., et al. (2004). A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. Nat. Genet. *36*, 283–287.

Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2007). Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol. *8*, R145.

van den Heuvel, S., and Dyson, N.J. (2008). Conserved functions of the pRB and E2F families. Nat. Rev. Mol. Cell Biol. *9*, 713–724.

Veraksa, A., Bauer, A., and Artavanis-Tsakonas, S. (2005). Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. Dev. Dyn. *232*, 827–834.

Wang, Z., and Mann, R.S. (2003). Requirement for two nearly identical TGIF-related homeobox genes in Drosophila spermatogenesis. Development *130*, 2853–2865.

Weaver, Z.A., and Kane, C.M. (1997). Genomic characterization of a testis-specific TFIIS (TCEA2) gene. Genomics *46*, 516–519.

Wu, L., Aster, J.C., Blacklow, S.C., Lake, R., Artavanis-Tsakonas, S., and Griffin, J.D. (2000). MAML1, a human homologue of Drosophila mastermind, is a transcriptional co-activator for NOTCH receptors. Nat. Genet. *26*, 484–489.

Xu, T., and Artavanis-Tsakonas, S. (1990). deltex, a locus interacting with the neurogenic genes, Notch, Delta and mastermind in Drosophila melanogaster. Genetics *126*, 665–677.

Yao, T.P., Forman, B.M., Jiang, Z., Cherbas, L., Chen, J.D., McKeown, M., Cherbas, P., and Evans, R.M. (1993). Functional ecdysone receptor is the product of *EcR* and *Ultraspiracle* genes. Nature *366*, 476–479.

Yu, J., Pacifico, S., Liu, G., and Finley, R.L., Jr. (2008). DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. BMC Genomics *9*, 461.

Yu, C., Wan, K.H., Hammonds, A.S., Stapleton, M., Carlson, J.W., and Celniker, S.E. (2011). Development of expression-ready constructs for generation of proteomic libraries. Methods Mol. Biol. *723*, 257–272.