

# CTCF-dependent co-localization of canonical Smad signaling factors at architectural protein binding sites in *D. melanogaster*

Kevin Van Bortle<sup>1,†</sup>, Aidan J Peterson<sup>2</sup>, Naomi Takenaka<sup>1</sup>, Michael B O'Connor<sup>2</sup>, and Victor G Corces<sup>1,\*</sup>

<sup>1</sup>Department of Biology; Emory University, Atlanta, GA USA; <sup>2</sup>Department of Genetics; Cell Biology; and Development; University of Minnesota, Minneapolis, MN USA

<sup>†</sup>Current affiliation: Department of Genetics; Stanford University, Stanford, CA USA

**Keywords:** BMP, chromatin, epigenetics, growth factor, Transcription, TGF- $\beta$  signaling

The transforming growth factor  $\beta$  (TGF- $\beta$ ) and bone morphogenetic protein (BMP) pathways transduce extracellular signals into tissue-specific transcriptional responses. During this process, signaling effector Smad proteins translocate into the nucleus to direct changes in transcription, but how and where they localize to DNA remain important questions. We have mapped *Drosophila* TGF- $\beta$  signaling factors Mad, dSmad2, Medea, and Schnurri genome-wide in Kc cells and find that numerous sites for these factors overlap with the architectural protein CTCF. Depletion of CTCF by RNAi results in the disappearance of a subset of Smad sites, suggesting Smad proteins localize to CTCF binding sites in a CTCF-dependent manner. Sensitive Smad binding sites are enriched at low occupancy CTCF peaks within topological domains, rather than at the physical domain boundaries where CTCF may function as an insulator. In response to Decapentaplegic, CTCF binding is not significantly altered, whereas Mad, Medea, and Schnurri are redirected from CTCF to non-CTCF binding sites. These results suggest that CTCF participates in the recruitment of Smad proteins to a subset of genomic sites and in the redistribution of these proteins in response to BMP signaling.

## Introduction

Transforming growth factor  $\beta$  (TGF- $\beta$ ) and bone morphogenetic protein (BMP) signaling effector proteins, called Smads, direct the transcriptional response to signaling pathways involved in controlling cellular homeostasis, proliferation, differentiation, and apoptosis.<sup>1,2</sup> In the canonical TGF- $\beta$  signaling pathway, serine-threonine kinase transmembrane receptors bind to extracellular TGF- $\beta$  ligands, and in turn phosphorylate receptor-regulated Smad proteins, which are then able to form complexes that translocate into the nucleus to promote transcriptional activation and repression.<sup>3</sup> Several recent studies have focused attention on uncovering the chromatin determinants of Smad localization. Master regulatory transcription factors, which control the transcription of key cellular identity genes and are themselves expressed in a cell-type specific manner, were shown to direct the localization of BMP and Wnt signaling factors in haematopoietic progenitor cells, and of TGF- $\beta$  effector Smad proteins in embryonic stem cells (ESCs).<sup>4,5</sup> The co-localization of Smads at cell-type specific master transcription factor binding sites in multi-potent cells provides an attractive model for how TGF- $\beta$  signaling events might produce diverse, tissue-specific responses.

In addition to master transcription factors, TRIM33, a multi-functional Smad-interacting protein that recognizes the dual

histone mark motif of H3K9me3 and H3K18ac, also creates a platform for Smad localization in response to signaling events in ESCs.<sup>6</sup> Nevertheless, the mechanisms driving recruitment of Smad proteins to DNA in non-stem cells remain largely unexplored. Studies probing the mammalian Alzheimer amyloid precursor protein (APP) gene promoter<sup>7</sup> and the H19 imprinting control region<sup>8</sup> have identified sites in which Smad recruitment to DNA is mediated by the architectural protein CTCF, previously characterized for its ability to mediate long-range interactions and to function as an insulator.<sup>9</sup> To what degree CTCF recruits Smad proteins to DNA on a global scale, whether these interactions are related to the TGF- $\beta$  response, and whether CTCF-directed Smad localization is conserved in other organisms remains unknown.

CTCF and other architectural proteins establish high occupancy architectural protein binding sites (APBSs) at the borders of topologically associating domains (TADs), which represent highly self-interacting regions of eukaryotic chromosomes.<sup>10–15</sup> Both TADs and high occupancy APBSs appear to be largely tissue-invariant,<sup>14</sup> suggesting most modular chromatin domains are conserved throughout development. However, intra-TAD interactions, which may be facilitated by low occupancy APBSs within domains, often consist of enhancer-promoter and promoter-promoter interactions that are likely to be dynamically regulated in

\*Correspondence to: Victor G Corces; Email vcorces@emory.edu

Submitted: 04/06/2015; Revised: 05/08/2015; Accepted: 05/16/2015

http://dx.doi.org/10.1080/15384101.2015.1053670

response to cell signaling events and between cell types to produce cell-type specific transcription patterns.<sup>16-19</sup> We have previously demonstrated that *Drosophila* architectural proteins exhibit moderate changes in genome-wide localization that correlate with dynamic chromosome organization in response to the insect steroid hormone 20-hydroxyecdysone,<sup>20</sup> which is bound by the nuclear ecdysone receptor, a ligand-activated transcription factor that in turn activates specific genes.<sup>21</sup> However, to what degree architectural protein binding contributes to additional signaling events remains poorly addressed.

Here we report the genome-wide landscape of the receptor-regulated Smad proteins Mothers against DPP (Mad) and dSmad2, the co-Smad Medea, and the co-repressor Schnurri in *D. melanogaster*. Indeed, we identify numerous sites in which Smad binding co-localizes with the *Drosophila* homolog of CTCF (dCTCF), and further demonstrate that dCTCF binding is required for co-localization at a subset of sites. These data support a role for CTCF as a conserved determinant of Smad localization from *Drosophila* to humans. Finally, we demonstrate that dynamic binding of Smad proteins in response to the *Drosophila* bone morphogenetic protein Decapentaplegic (DPP), a member of the TGF- $\beta$  superfamily, occurs in the context of dCTCF-independent binding sites and that dCTCF binding itself remains unchanged. Together, these results suggest that whereas architectural protein occupancy is not dynamically regulated in response to TGF- $\beta$ , signaling events otherwise redirect Smad localization to promote changes in transcription independent of CTCF.

## Results

### Genome-wide mapping of *Drosophila* Smad proteins

TGF- $\beta$  superfamily signaling in *D. melanogaster* is traditionally broken into 2 branches based on the ligand-receptor interaction and the class of receptor-regulated Smad proteins (R-Smads) that are subsequently phosphorylated.<sup>43</sup> BMP ligands bind the type I serine-threonine receptors Thickveins (Tkv) and Saxophone (Sax) to induce C-terminal phosphorylation of the R-Smad Mothers against DPP (Mad).<sup>44,45</sup> TGF- $\beta$ /Activin ligands signal through splice isoforms of the Baboon receptor to activate the R-Smad dSmad2 (Smad on X).<sup>46-48</sup> The phosphorylated forms of dSmad2 and Mad are able to form heterotrimeric complexes with the co-Smad Medea, and upon which translocate into the nucleus to direct changes in transcription.<sup>49</sup> The Mad/Medea complex has been shown to target activation regulatory elements in response to DPP,<sup>50</sup> as well as repressive regulatory elements to which the transcriptional repressor Schnurri is recruited.<sup>51,52</sup> The expression of several genes depends on Activin signaling components,<sup>53-55</sup> but dSmad2 regulatory elements have not been identified. Therefore, in order to effectively capture the chromatin landscape of TGF- $\beta$  signaling pathways, we carried out ChIP-seq experiments against dSmad2, Mad, Medea, and Schnurri in the late-embryonic *Drosophila* hemocyte cell line Kc167 (Fig. 1, Supplemental Table 1).

Initial genome-wide analysis of Smad binding reveals extensive overlap between R-Smads and the co-Smad Medea (Mad overlap

with Medea – 85%; dSmad2 overlap with Medea – 81%;  $p < 0.0001$ , permutation test) as well as with the co-repressor Schnurri (Mad overlap with Schnurri – 67%; dSmad2 overlap with Schnurri – 70%;  $p < 0.0001$  permutation test) (Fig. 1A). Interestingly, nearly half of all binding sites for either Mad or dSmad2 also overlap with the distinct R-Smad (Mad overlap with dSmad2 – 52%; dSmad2 overlap with Mad – 47%;  $p < 0.0001$ , permutation test), suggesting a potential interaction between TGF- $\beta$ /Activin and BMP signaling pathways at the genomic level. In support of this observation, recent studies have shown complex formation between TGF- $\beta$  and BMP Smad proteins at BMP response elements in human epithelial cells,<sup>56</sup> suggesting cross-talk between TGF- $\beta$  and BMP signaling factors may be a conserved phenomenon. Ligand stimulation of the TGF- $\beta$  receptor Baboon can also induce phosphorylation of the BMP R-Smad Mad,<sup>27</sup> and mutation of the Activin R-Smad dSmad2 leads to altered BMP pathway signaling,<sup>57</sup> suggesting a potentially strong degree of interaction between TGF- $\beta$  and BMP signaling pathways both upstream and downstream of nuclear translocation. Surprisingly, numerous Medea and Schnurri binding sites are present independently of either Mad or dSmad2, though a majority of Schnurri binding sites overlap with Medea (Fig. 1A), suggesting that additional regulatory factors may be involved in recruiting Medea and Schnurri to DNA, or that Medea and Schnurri independently localize to specific regulatory elements.

Identification of genomic loci co-bound by Medea, Schnurri, and either R-Smad reveals even greater overlap between the BMP and TGF- $\beta$  signaling factors Mad and dSmad2 (Fig. 1B; 69% for Mad; 66% for dSmad2;  $p < 0.0001$ , permutation test), suggesting both R-Smads co-localize with Medea and Schnurri at similar target sequences. The 1,220 overlapping Schnurri, R-Smad, Medea target sites bound by both Mad and dSmad2 provide a high confidence list of signaling target regions commonly identified in 4 independent ChIP-seq experiments that we consider for further analysis. We hereafter refer to these sites co-bound by dSmad2, Mad, Medea, and Schnurri as dSMMS modules analogous to previously defined target elements.<sup>52</sup> Visual inspection of high confidence dSMMS modules confirms co-localization of these signaling proteins at well characterized target genes. For example, *brinker*, a gene that encodes a nuclear repressor that antagonizes DPP signaling by binding similar Mad/Medea target sequences, is transcriptionally repressed by Mad, Medea, and Schnurri in response to DPP signaling events. Accordingly, the *brinker* gene was recently shown to include several promoter modules targeted by Mad, Medea, and Schnurri.<sup>58</sup> Although *brinker* is modestly expressed in *Drosophila* Kc167 cells, our ChIP-seq data provide evidence that these dSMMS modules are highly occupied in cell culture even prior to induction with DPP (Fig. 1C), suggesting either a minimal level of signaling occurs in Kc167 cells or that Smad-mediated repression is regulated downstream of sequence binding.

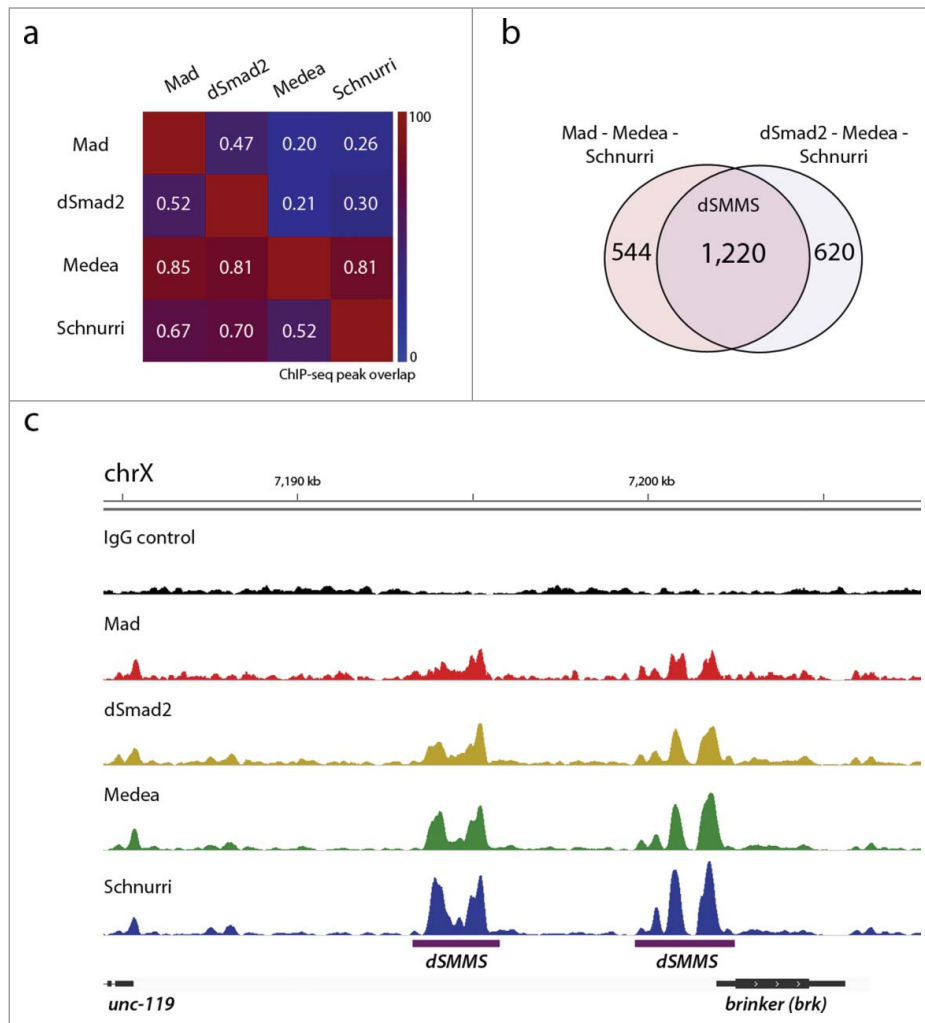
### dSMMS modules overlap *Drosophila* CTCF and other architectural proteins

*De novo* motif analysis on dSMMS modules using MEME-ChIP<sup>59</sup> identifies enrichment for 2 sequences that show

similarity to putative Mad and Medea consensus sequences (Fig. 2). Previous comparisons of specific Mad/Medea-binding elements suggest that Smads target GCCGNC as a consensus binding sequence,<sup>60,61</sup> and, similarly, MEME-ChIP identifies enrichment for GCYGSC at dSMMS modules (Fig. 2A). dSMMS modules are nearly equally enriched for a GC rich consensus sequence with strong similarity to putative Medea binding sites, together suggesting that our ChIP-seq profiles for dSmad2, Mad, Medea, and Schnurri provide an accurate means for identifying Smad-signaling response elements.

Interestingly, motif analysis also identifies a consensus sequence previously identified as being enriched at *Drosophila* CTCF binding sites that overlap with additional architectural proteins, such as Boundary Element Associated Factor of 32 kDa (BEAF-32), Centrosomal Protein 190 (CP190), Modifier of mdg4 (Mod (mdg4)), Chromator, and the cohesin and condensin complex kleisen subunits Rad21 and CAP-H2 respectively.<sup>10,11</sup> In fact, we find significant motif enrichment for DNA binding architectural protein consensus sequences targeted by dCTCF, BEAF-32, and CP190, whereas sequences targeted by Suppressor of Hairy-wing (Su(Hw)) are depleted in dSMMS modules (Fig. 2B). This result indicates that genomic elements bound by R-Smads Medea and Schnurri often occur in close spatial proximity with APBSs, and suggests that architectural proteins may play an important role in either creating an accessible chromatin landscape for dSMMS occupancy, or may themselves directly recruit Smads to DNA.

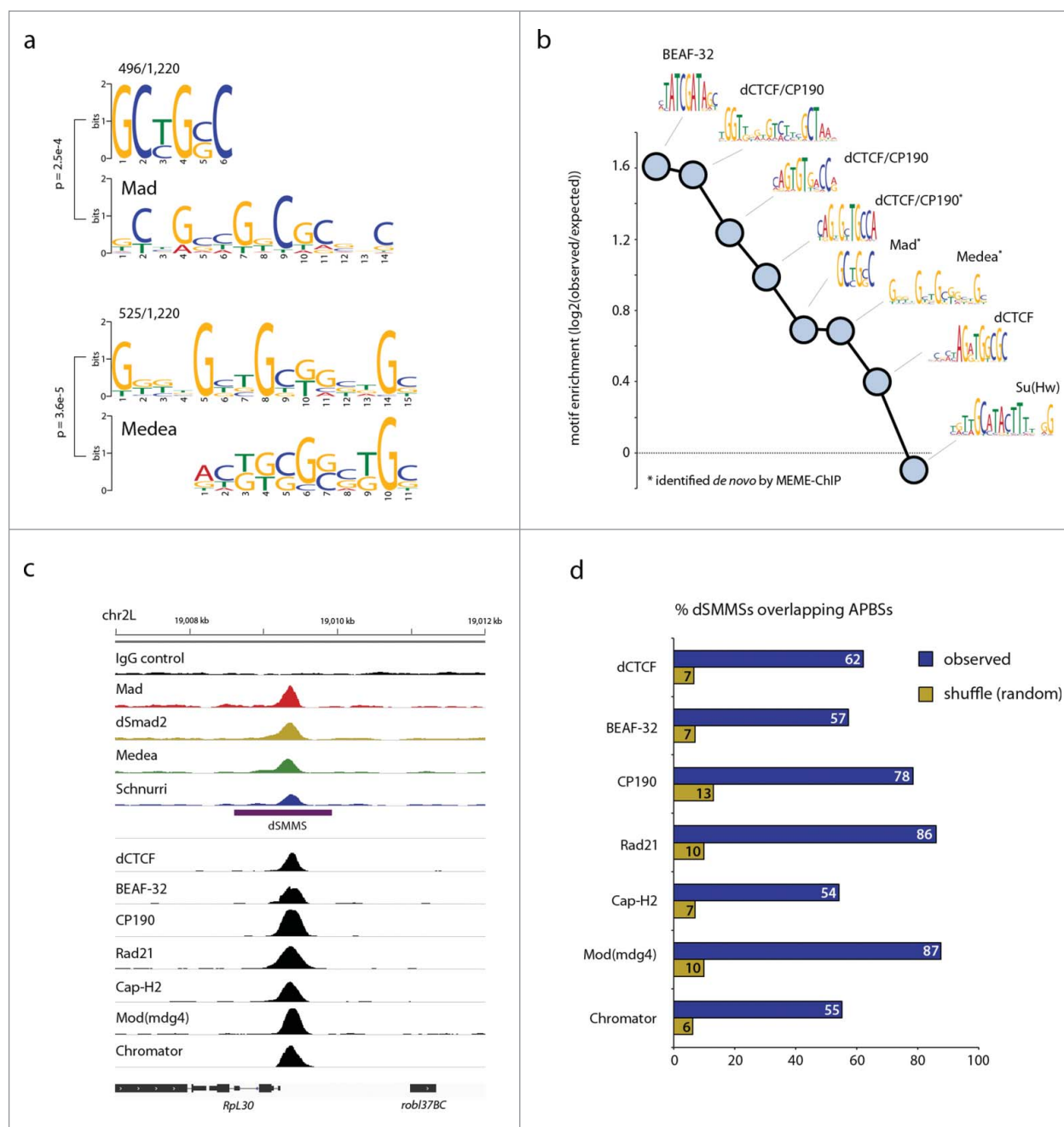
Visualization of dSMMS modules overlapping APBSs illustrates that ChIP-seq read densities for BMP and TGF- $\beta$  Smad proteins tightly intersect those of dCTCF and other architectural proteins (Fig. 2C). Genome-wide, 62% of dSMMS modules co-localize with dCTCF ( $p < 0.0001$ , permutation test), and a majority of modules also co-localize with additional architectural proteins (Fig. 2D). These data, together with motif enrichment profiles for architectural proteins in Smad binding sites, suggest that architectural proteins may play an important role in Smad localization.



**Figure 1.** Genome-wide mapping of BMP and TGF- $\beta$  signaling proteins Mad, dSmad2, Medea, and Schnurri in *Drosophila* Kc167 cells. **(A)** Heatmap representation of ChIP-seq peak overlap between individual Smad proteins. Corresponding values represent percentage of total binding sites for factors along the horizontal axis that overlap peaks identified for factors along the vertical axis, ranging from 0 (blue) to 100 (red). **(B)** Overlap of Schnurri-Mad-Medea modules individually identified for receptor-regulated Smad proteins Mad or dSmad2. Overlap is statistically significant ( $p < 0.0001$ , permutation test). **(C)** Illustration of ChIP-seq experiments and Smad protein overlap at dSMMS regulatory elements present in previously characterized promoter-proximal modules of the *brinker* gene locus.

### dCTCF-dependent co-localization of Smad proteins at APBSs

The strong overlap of BMP and TGF- $\beta$  effectors with *Drosophila* CTCF suggests that CTCF-dependent recruitment of Smad proteins may be an important, highly conserved genome-wide phenomenon. In order to test whether Smad proteins target *Drosophila* APBSs in a dCTCF-dependent manner, we repeated ChIP-seq experiments for Mad, dSmad2, and Medea in Kc167 cells depleted of dCTCF by RNAi.<sup>10,11</sup> In all cases, disruption of dCTCF levels significantly perturbs the levels of Smad ChIP-seq read density at a subset of sites (Figs. 3A–C). For example, Mad occupancy significantly decreases at more than 200 binding sites



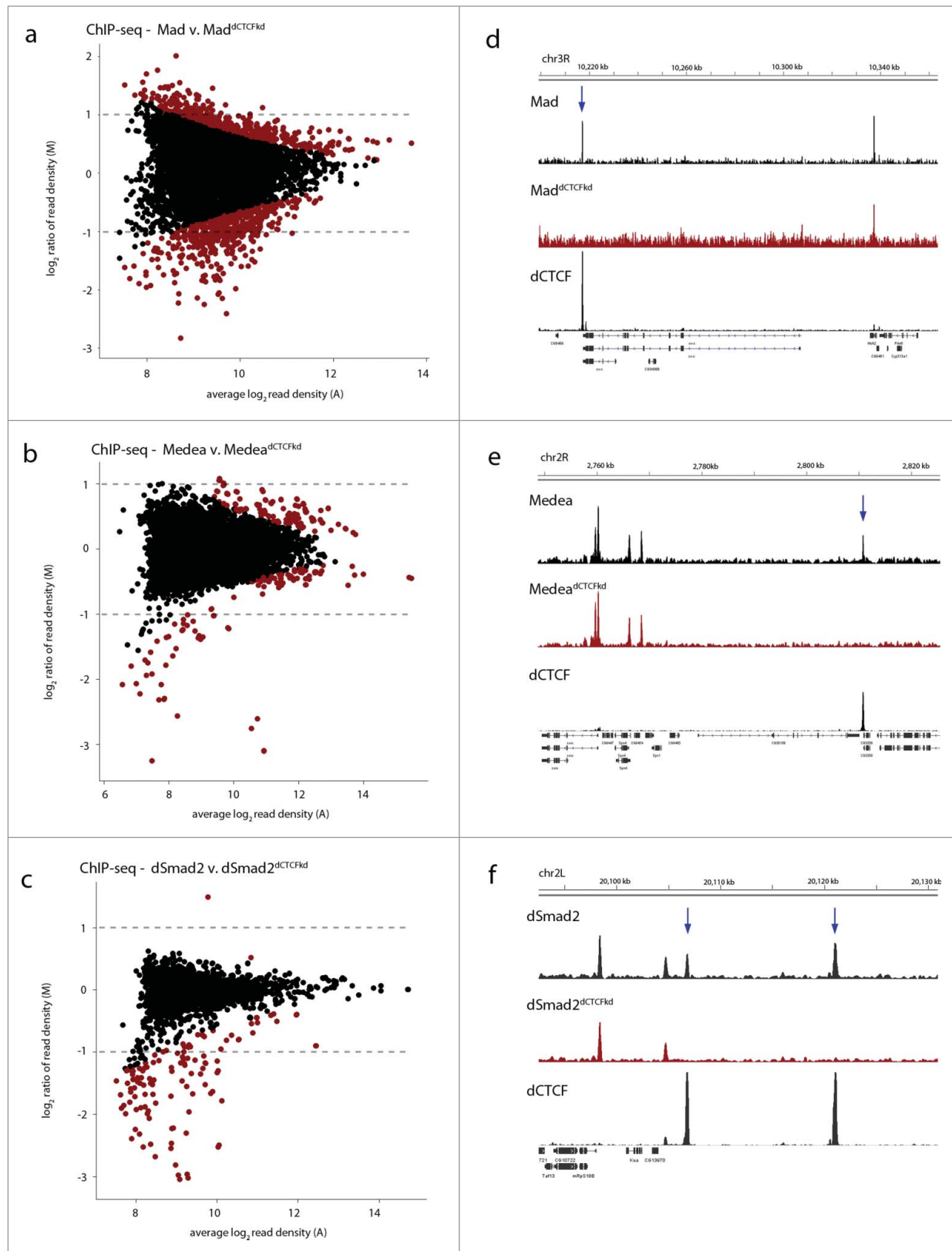
**Figure 2.** dSMMS module motif enrichment and overlap at architectural protein binding sites. **(A)** Putative Mad and Medea motifs identified *de novo* by MEME-ChIP and the fraction of dSMMS modules containing the identified consensus sequence. Sequences are aligned with database motifs for Mad (top) and Medea (bottom) predicted based on previous DNA-binding experiments. Alignment and p-values defined by TOMTOM motif comparison tool. **(B)** Consensus sequence enrichment ( $\log_2(\text{observed/expected frequency})$ ) for Mad, Medea, and dCTCF/CP190 motifs identified *de novo* by MEME-ChIP (\*) along with previously characterized DNA-binding architectural protein consensus sequences. **(C)** Example genomics viewer illustration of dSMMS module overlap at architectural protein binding sites at the Ribosomal Protein L30 (*Rpl30*) gene promoter. **(D)** Percentage of dSMMS modules overlapping architectural protein binding sites for dCTCF, BEAF-32, CP190, Rad21, Cap-H2, Mod(mdg4), and Chromator. Comparison of observed overlap frequencies (blue) with randomized shuffle control peaks (yellow) illustrates the strong enrichment of dSMMS modules at APBSs.

(Fig. 3A), and these sites are significantly enriched for the dCTCF motif. Whereas fewer dSmad2 and Medea peaks are significantly downregulated using the same threshold (absolute M value cutoff of 1,  $p < 0.02$ ), all Medea and dSmad2 peaks identified as being lost after dCTCF RNAi indeed

overlap dCTCF ( $p < 0.0001$ , permutation test), suggesting loss of Smad binding is a direct consequence of dCTCF depletion.

Strikingly, Mad, Medea, and dSmad2 binding sites that are affected by dCTCF depletion are most often entirely lost after





**Figure 3.** RNAi depletion of *Drosophila* CTCF abrogates Smad localization to a subset of dCTCF binding sites. **(A–C)** MA plots depicting changes in ChIP-seq read density as a function of average peak read densities for Mad, Medea, and dSmad2 respectively. dCTCF RNAi and control ChIP-seq experiments were normalized using the MANorm package for quantitative comparison of ChIP-seq data sets.<sup>39</sup> An M value (log<sub>2</sub> ratio of ChIP-seq read density) threshold cutoff of 1 (dotted line) and a p value cutoff of 0.02 (red) were chosen for consideration of significantly changing peaks. **(D–F)** Genomic viewer comparison of dCTCF RNAi (red) and control ChIP-seq experiments (black) for Mad, Medea, and dSmad2 respectively. In all cases, several significant ChIP-seq peaks overlapping dCTCF are lost in response to knockdown of dCTCF (blue arrow), whereas non-overlapping Smad binding sites remain comparatively unchanged.

dCTCF RNAi (Figs. 3D-F), suggesting recruitment of Smad proteins to these loci is entirely dependent on dCTCF. These results mirror the dependence of Smad localization to both the H19 imprinting control region and the Alzheimer  $\beta$ -amyloid precursor protein proximal promoter region on human CTCF. However, despite the significant drop in Smad occupancy at a subset of sites in *Drosophila* Kc167 cells, ChIP-seq read density for Mad, Medea, and dSmad2 is not otherwise depleted at all dCTCF sites (Supplemental Fig. 1), suggesting dSMMS elements are differentially affected by, and perhaps differentially dependent on, dCTCF.

#### **dCTCF-dependent Smad binding occurs at low occupancy APBSs within topological domains**

*Drosophila* CTCF and other architectural proteins target thousands of regulatory elements that play unique roles in shaping chromosome organization and transcription. For example, long-range interactions mediated by architectural proteins can facilitate either active enhancer-promoter interactions or repressive Polycomb response element interactions, whereas sites bound by numerous architectural proteins are involved in establishing discrete topological domains and are capable of robust insulator function for which these proteins were originally characterized. We therefore asked whether Smad binding sites that are sensitive or insensitive to dCTCF RNAi occur within unique contexts. Here we compare Mad, Medea, and dSmad2 binding sites overlapping dCTCF that are lost in response to dCTCF depletion with overlapping dSMMS modules that do not change in response to RNAi.

Motif analysis reveals significant enrichment for the dCTCF core consensus sequence at sites in which Mad, Medea, and/or dSmad2 are lost after dCTCF knockdown (Fig. 4A). In contrast, Smad binding sites in which dCTCF co-localization occurs independently of dCTCF are enriched for the BEAF-32 motif and consensus sequences known to be enriched at high occupancy APBSs. *Drosophila* CTCF-independent, overlapping Smad peaks are also enriched for the putative Mad and Medea motifs identified by MEME-ChIP, whereas sites sensitive to dCTCF knockdown are depleted for Mad and Medea consensus sequences. These differences in motif enrichment suggest potentially unique binding modes for Smad proteins at dCTCF-independent versus dCTCF-dependent peaks. Intuitively, Smad peaks containing direct target sequences are not sensitive to dCTCF depletion, whereas sites lacking direct binding motifs appear to be entirely dependent on dCTCF for recruitment to these loci.

Enrichment of the BEAF-32 consensus sequence and dCTCF/CP190 motifs present at high occupancy APBSs also suggests that, in addition to dCTCF, other architectural proteins may provide some redundancy in recruiting Smad proteins to APBSs. Comparison of DNase I hypersensitivity at dCTCF-sensitive and dCTCF-insensitive Smad peaks by DNase-seq in *Drosophila* Kc167 cells provides additional evidence that dCTCF-independent binding occurs at high occupancy APBSs. For example, whereas Smad peaks that are lost after dCTCF knockdown exhibit sharp but comparatively weaker DNase I

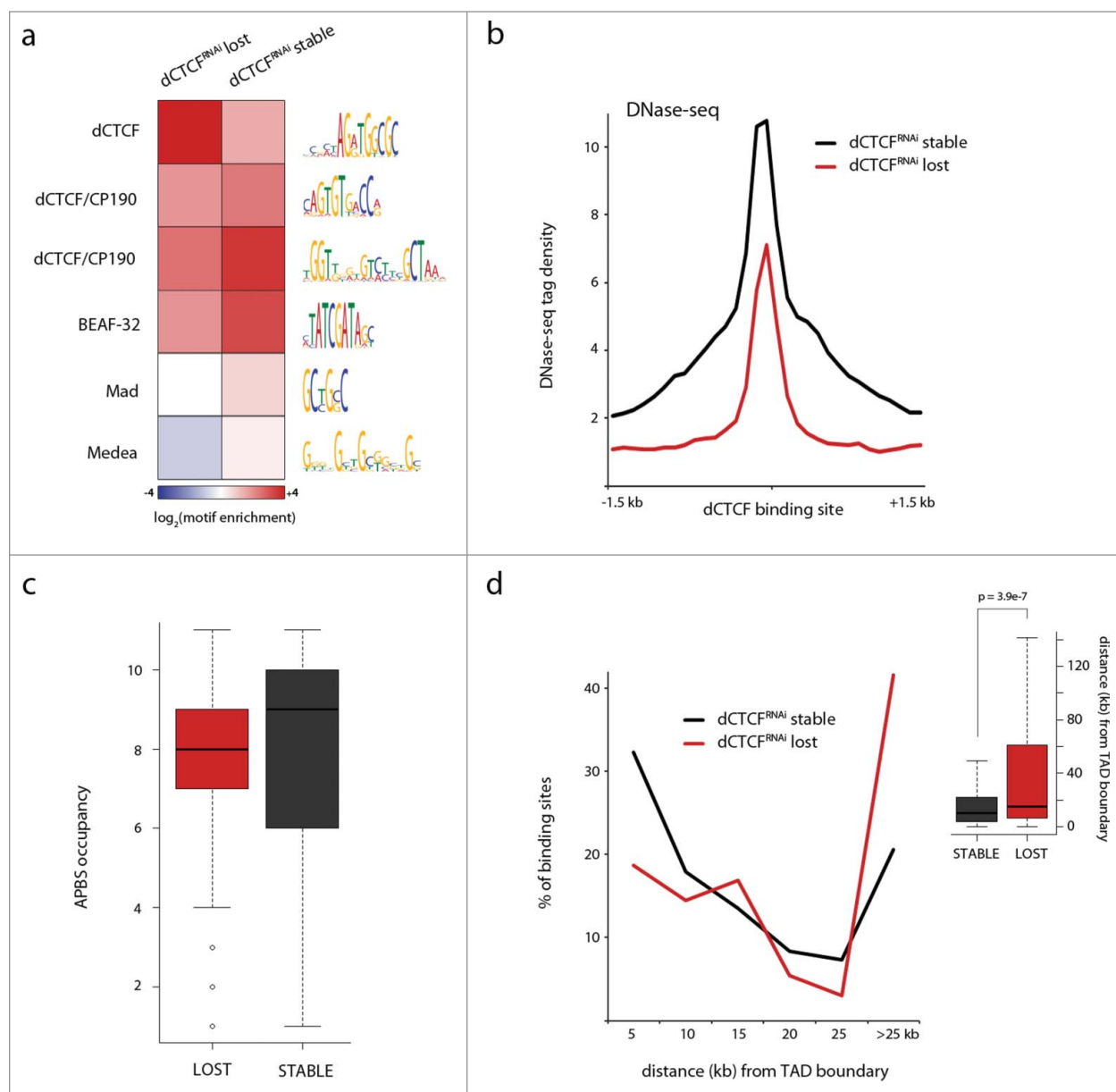
hypersensitivity, dCTCF-independent Smad binding sites are characterized by broad, robust DNase sensitivity commonly observed at high occupancy APBSs (Fig. 4B). Accordingly, dCTCF-independent Smad peaks overlap on average more architectural proteins than dCTCF-dependent loci (Fig. 4C). Though the difference in overlap does not reach statistical significance ( $p = 0.061$ , wilcoxon rank-sum test), this comparison draws on binary peak identification by MACS and does not take into account the relative ChIP-seq tag enrichment.

The observed differences in consensus sequence enrichment, DNase I hypersensitivity, and total architectural protein occupancy at differentially affected Smad binding sites suggest that dCTCF-dependent and dCTCF-independent Smad peaks are likely present in unique chromosomal contexts. Integration of ChIP-seq with genome-wide chromosomal interaction mapping has recently shown that interphase chromosomes are organized into discrete, self-interacting domains that are separated by high occupancy APBSs.<sup>10</sup> Accordingly, we find that Mad, Medea, and dSmad2 peaks that overlap dCTCF but are not significantly affected by dCTCF depletion are enriched near the boundaries of topological domains (Fig. 4D). In contrast, dCTCF-dependent Smad binding sites do not predominantly localize near TAD borders, consistent with the nature of low occupancy dCTCF binding sites residing within topological domains. Recent enhancer-trap assays suggest that individual topological domains functionally limit the activity of enhancers to genes that reside within the same TAD.<sup>62</sup> We speculate that, whereas high occupancy APBSs mediate long-range higher order chromosomal domain organization, low occupancy dCTCF binding sites may facilitate short-range enhancer-promoter interactions relevant to transcription. However, to what degree dCTCF binding plays a role in the TGF- $\beta$  signaling response has not been previously addressed.

#### **dCTCF binding remains constant in response to DPP-induced signaling events**

In order to assay whether dCTCF binding is influenced by TGF- $\beta$  signaling events, we repeated ChIP-seq experiments for dCTCF as well as Mad, Medea, and Schnurri in response to the *Drosophila* bone morphogenetic protein DPP, a standard approach for stimulating Mad phosphorylation in cell culture. Indeed, treatment with 30 nM DPP for 6 hr robustly activates phosphorylation of Mad in *Drosophila* Kc167 cells, whereas phospho-Mad is undetectable in untreated control cells (Fig. 5A). Surprisingly, ligand-mediated activation of the BMP cascade does not dynamically alter the DNA-binding profile for dCTCF (Fig. 5B, less than 1% of CTCF binding sites increase or decrease significantly), suggesting the architectural protein occupancy landscape remains relatively static. However, DPP does induce upregulation and downregulation of Mad (488 sites down, 161 sites up), Medea (9 sites down, 63 sites up), and Schnurri (109 sites down, 90 sites up) across the genome (Figs. 5C-E, Supplemental Fig. 2), suggesting Smad localization is dynamically altered in response to signal transduction.

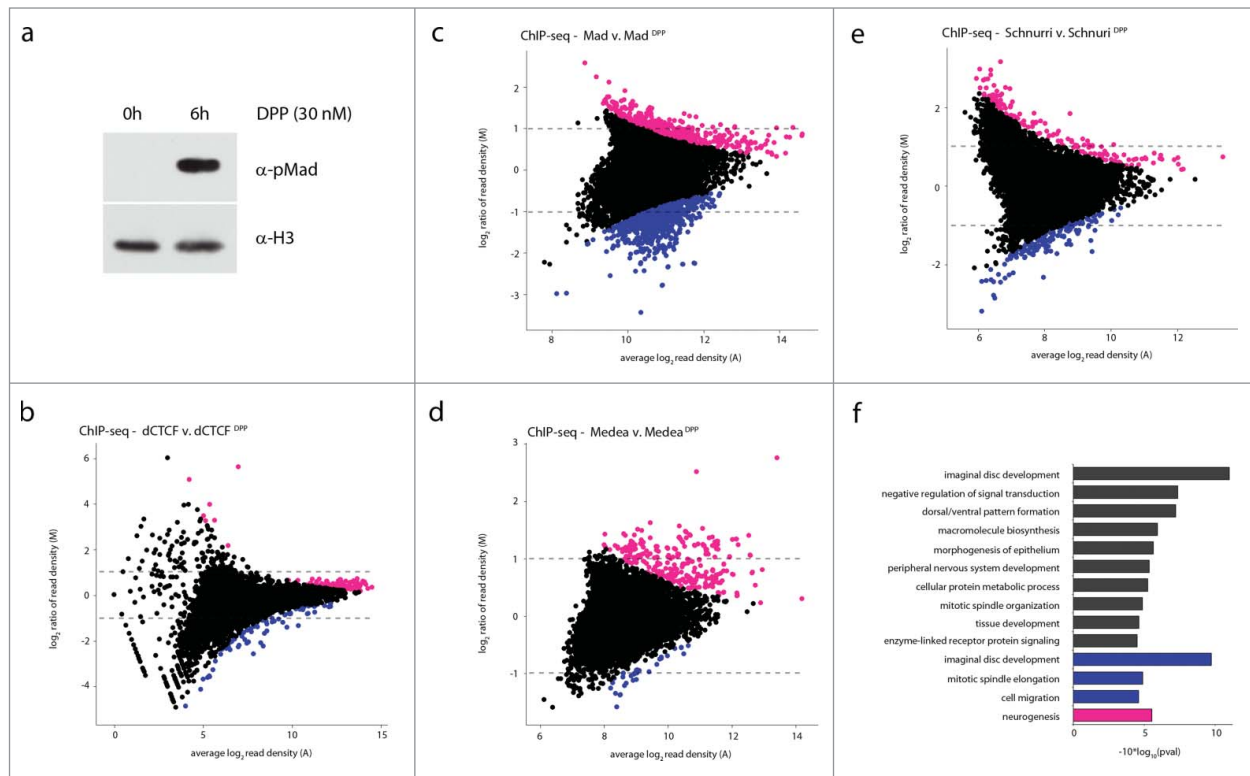
We next performed gene ontology (GO) analysis on genes that are in close spatial proximity to dSMMS modules and near sites that increase or decrease in response to DPP (<4 kb from



**Figure 4.** Differential motif enrichment, architectural protein occupancy, and chromosomal location of dCTCF-dependent versus independent Smad binding sites. **(A)** Motif enrichment ( $\log_2(\text{observed/expected frequency})$ ) for architectural protein consensus sequences and putative Mad/Medea motifs at dCTCF<sup>RNAi</sup> lost (left) and dCTCF<sup>RNAi</sup> stable Smad binding sites. Comparison includes only Smad peaks which overlap dCTCF. **(B)** DNase-seq read densities centered on dCTCF binding sites overlapping Mad, Medea, and dSmad2 that are stable (black) or lost in response to dCTCF RNAi (red). **(C)** APBS occupancy, determined as the number of overlapping MACS-called peaks for architectural proteins, at Smad binding sites that are stable (black) or lost after dCTCF knockdown (red) ( $p = 0.061$ , wilcoxon rank-sum test). **(D)** Distance (kb) from topologically associating domain (TAD) boundaries, at which high occupancy APBSs are highly enriched.<sup>10</sup> dCTCF<sup>RNAi</sup> lost Smad peaks are most abundant within domains, whereas dCTCF<sup>RNAi</sup> stable peaks are enriched near TAD boundaries (median distance from TAD borders: 14.9 kb and 9.9 kb respectively,  $p = 3.9 \times 10^{-7}$ , wilcoxon rank-sum test).

TSS). Whereas Mad, Medea, and Schnurri peaks that decrease are enriched near genes associated with imaginal disc development, consistent with dSMMS module occupancy prior to induction with DPP, dynamically upregulated binding sites occur in close spatial proximity to genes associated with neurogenesis (Fig. 5F). Though DPP is known to influence neuronal development in a context-specific manner, whether these dynamics are at all meaningful for Kc167 cellular development

*in vivo* remains unknown. Nevertheless, comparison of upregulated Smad peaks with respect to dCTCF binding suggests that this signaling transcriptional response occurs in a non-dCTCF context. For example, upregulated Mad binding sites are depleted for dCTCF, and co-localization of Medea and Schnurri upregulated peaks with dCTCF is no greater than expected by random chance (Supplemental Fig. 3). DPP-induced downregulation of Smad binding, on the other hand, does occur at dCTCF-binding



**Figure 5.** Dynamic Smad localization in response to DPP activated phosphorylation of Mad. **(A)** Western blot analysis of phosphorylated Mad levels before and after 6 h treatment with DPP (30 nM), using phospho-specific p-Mad antibody. Loading control staining of histone H3. **(B)** MA plot depicting changes in dCTCF ChIP-seq read density as a function of average peak read density. DPP treatment and control ChIP-seq experiments were normalized using the MANorm package for quantitative comparison of ChIP-seq data sets.<sup>39</sup> An M value ( $\log_2$  ratio of ChIP-seq read density) threshold cutoff of 1 (dotted line) and a p value cutoff of 0.02 (upregulated peaks – pink; downregulated peaks – blue) were chosen for consideration of significantly changing peaks. Only 40 peaks decrease and 7 peaks increase significantly in response to DPP. **(C–E)** Analogous MA plots for Mad, Medea, and Schnurri ChIP-seq experiments before and after treatment with DPP. Number of significantly changing peaks: Mad 161 increasing, 488 decreasing; Medea: 63 increasing, 9 decreasing; Schnurri: 90 increasing, 109 decreasing. **(F)** Gene ontology (GO) analysis for dSMMS modules and significantly changing peaks in response to DPP treatment. GO term enrichment was performed for closest genes (TSS) within 4 kb of Mad, Medea, and Schnurri binding sites before treatment (black), that decrease (blue), or increase (pink) significantly after treatment with DPP. Whereas decreasing Smad binding sites are enriched near genes associated with imaginal disc development and signaling, increasing binding sites occur near genes involved in neurogenesis.

sites, many of which coincide with low occupancy dCTCF APBSs, suggesting loss of Smad proteins occurs at architectural protein-bound regulatory elements within topological domains. The dynamic localization of Mad, Medea, and Schnurri from dCTCF sites to non-dCTCF sites suggests that DPP signaling may actively redirect Smad localization to unique chromatin contexts relevant to cell differentiation.

## Discussion

TGF- $\beta$  effector proteins have been shown to co-localize with mammalian CTCF in a CTCF-dependent manner at just 2 individual loci. We now extend this observation to *Drosophila* using a genome-wide approach, providing evidence that architectural protein CTCF and canonical Smad signaling proteins, both highly conserved from fly to humans, co-localize on a global scale. We further uncover context-specific features in which Smad localization is dependent or independent of CTCF

binding. Interestingly, our genome-wide analysis identifies Mad, dSmad2, Medea, and Schnurri binding to previously characterized response elements even in the absence of DPP ligand, in which levels of phosphorylated Mad are undetectable. This signal-independent clustering of signaling proteins suggests that the genomic TGF- $\beta$  signaling response is not as simple as regulating binary “off vs. on” states, dependent on phosphorylated Mad. However, our attempts to map the genomic landscape of phosphorylated-Mad before and after DPP stimulation were unsuccessful, likely due to issues with currently available p-Mad antibodies. Though we could not determine the role of phosphorylation as a determinant in Mad localization, it is conceivable that phosphorylation of Mad might play a role in regulating the resident time of DNA-binding, the recruitment of additional regulatory partners, or the ability to establish functional long-range interactions.

We find that Smad co-binding at dCTCF sites is sensitive to dCTCF depletion at low occupancy dCTCF target sequences for which Smad consensus sequences are depleted, whereas high



occupancy dCTCF binding sites co-bound by additional architectural proteins remain unaffected. The dCTCF-independent recruitment of Smads to high occupancy APBSs suggests that additional architectural proteins may redundantly recruit Smads, or simply provide an accessible chromatin landscape to which Mad, Medea, and dSmad2 can associate. Nevertheless, dCTCF-dependent localization of Smad proteins to specific low occupancy elements is consistent with the CTCF-dependent nature of Smad binding at both the APP and H19 promoters in humans.<sup>7,8</sup> We speculate that dCTCF-dependent Smad localization to low occupancy APBSs within topological domains may represent regulatory elements involved in enhancer-promoter interactions, whereas dCTCF-independent high occupancy APBSs are involved in establishing higher-order chromosome organization. What role Smads might play in establishing or maintaining such long-range interactions relevant to chromosome architecture, or whether Smads and other transcription factors simply localize to high occupancy APBSs due to chromatin accessibility, remains difficult to address. However, we have recently shown that high occupancy APBSs are distinct from analogous transcription factor hotspots, suggesting some level of specificity, most likely governed by protein-protein interactions, decides which factors can associate and where. Alternatively, the enrichment of ChIP-seq signal at high occupancy APBSs may, to some degree, reflect indirect association via long-range interactions with regulatory elements directly bound by Smad proteins. This possibility raises a potential explanation for why Smad ChIP signal is independent of dCTCF binding at high occupancy APBSs.

Surprisingly, DPP-activated phosphorylation of Mad does not lead to significant changes in dCTCF binding, whereas Mad, Medea, and Schnurri levels increase at regulatory elements away from dCTCF. These results suggest that TGF- $\beta$  signaling in Kc167 cells redirects Smad binding to genomic loci independent of architectural proteins, and that architectural proteins may facilitate binding of nuclear Smad proteins in the absence of signaling. The complete loss of Smad ChIP signal at numerous dCTCF binding sites enriched for the core dCTCF consensus sequence nevertheless provides compelling evidence that recruitment of Smad proteins is directly governed by *Drosophila* CTCF at a subset of binding sites. These results establish CTCF as an important determinant of Smad localization and, depending on the cell-type specific binding patterns of CTCF, suggest that CTCF might also influence the tissue-specific localization of Smad proteins analogous to master regulatory transcription factors in multi-potent stem cells.

## Materials and Methods

### Cell Culture and DPP Treatment

*Drosophila* Kc167 cells were grown in HyClone SFX cell culture medium. For treatment with DPP, cells were split overnight to a density of  $0.5 \times 10^6$  cells/mL and treated with 30 nM DPP the following morning. Recombinant *Drosophila* Decapentaplegic was obtained from R&D systems (Cat#159-DP-20). DPP

treatment and control lysates were extracted after 6 hr incubations. RNAi knockdown of *Drosophila* CTCF in cell culture was conducted as per the *Drosophila* RNAi Screening Center (DRSC) protocol<sup>22</sup> with the exception that dsRNA was added every day for 3 days and the cells were then collected on the 4<sup>th</sup> day. Chromatin isolation from dCTCF depleted cell culture were performed as part of previously published knockdown experiments and described protocols.<sup>11</sup> For a list of primer sequences used for dCTCF RNAi as well as demonstration of knockdown efficiency see **Table S4** and **Figure S7**, respectively, at [http://genome.cshlp.org/content/suppl/2012/08/01/gr.136788.111.DC1/Supplemental\\_materials.pdf](http://genome.cshlp.org/content/suppl/2012/08/01/gr.136788.111.DC1/Supplemental_materials.pdf).

### Immunoprecipitation and Western Analysis

Chromatin immunoprecipitation was performed as previously described.<sup>23</sup> Protein A Sepharose beads (GE Healthcare) were pre-washed in PBSMT and pre-incubated with 6  $\mu$ L of rabbit polyclonal anti-Mad,<sup>24</sup> rabbit polyclonal anti-Medea,<sup>25</sup> affinity-purified rabbit polyclonal anti-Schnurri,<sup>26</sup> or rabbit polyclonal IgG (Santa Cruz sc-2027) for 1 hr prior to pull-down of sheared DNA. Affinity-purified sheep polyclonal anti-dSmad2 (R&D systems cat#AF7948) antibodies were pre-incubated with Protein Sepharose G beads (GE Healthcare). To generate sequencing libraries, ChIP DNA was prepared for adaptor ligation by end repair (End-It DNA End Repair Kit, Epicenter Cat# ER0720) and addition of "A" base to 3' ends (Klenow 3'-5' exo-, NEB Cat# M0212S). Illumina adaptors (Illumina Cat# PE-102-1001) were titrated according to prepared DNA ChIP sample concentration and ligated with T4 ligase (NEB Cat# M0202S). Ligated ChIP samples were PCR-amplified using Illumina primers and Phusion DNA polymerase (NEB Cat# F-530L) and size selected for 200–300 bp by gel extraction. ChIP libraries were sequenced at the HudsonAlpha Institute for Biotechnology, using an Illumina HiSeq 2000.

For Western blotting, membranes were blocked in TBST (20 mM Tris, pH7.4, 150 mM NaCl, 0.05% Tween 20) with 5% nonfat milk powder and incubated overnight with phospho-specific Mad antibodies.<sup>27</sup> Membranes were washed 3 times with TBST and probed with secondary antibodies-conjugated to HRP (1:5000, Jackson ImmunoResearch Laboratories) for 1 hr. After three more washes, the presence of different proteins was detected using SuperSignal West Pico/Dura Chemiluminescent substrate (Thermo Scientific).

### ChIP-seq and reference data

In addition to novel Mad, Medea, and Schnurri ChIP experiments, architectural protein ChIP-seq data was obtained from previously published sources.<sup>10,11</sup> Sequences were mapped to the dm3 genome with Bowtie 0.12.3,<sup>28</sup> using default settings. Peaks were then called with MACS 1.4.0alpha2<sup>29</sup> using equal numbers of unique reads for IgG control and ChIP samples, using a p value cutoff of  $1 \times 10^{-10}$ . For classification of overlapping binding sites, MACS-identified peaks were intersected using publicly available tools on Galaxy.<sup>30-32</sup> For visualization, mapped sequence reads were loaded on to the Integrated Genomics Viewer.<sup>33,34</sup> Previously published ChIP-seq data for *Drosophila*

architectural proteins were obtained from GEO accessions GSE30740<sup>20</sup> and GSE36944.<sup>11</sup> Raw DNase-seq in Kc167 cells was obtained from published resources.<sup>35</sup>

### Bioinformatics analyses

DNA sequence motifs were identified by MEME-ChIP set to identify 12 motifs and otherwise using default settings [28]. Sequences were aligned with database motifs and p-values determined using the TOMTOM motif comparison tool.<sup>36</sup> Comparison of APBSs with respect to Pol II-transcribed genes employed gene structure (TSSs) obtained using the UCSC genome browser.<sup>37,38</sup> Significantly changing ChIP-seq peaks in response to dCTCF RNAi and DPP signaling were determined using the MAnorm package for quantitative comparison of ChIP-seq data sets.<sup>39</sup> An M value ( $\log_2$  ratio of ChIP-seq read density) threshold cutoff of 1 and a p value cutoff of 0.02 were chosen for consideration of significantly changing peaks. Enrichment profiles for Mad, Medea, Schnurri, and dSmad2 protein co-occurrence and overlap with architectural proteins were defined as the observed overlapping frequencies over expected frequencies determined by shuffling datasets, while controlling for the number of peaks and start/stop location of peaks on each chromosome. P-values were determined as the chance of observing an equal or greater co-occurrence across 100,000 Monte Carlo Permutation tests. dSMMS modules were defined as MACS-called R-smad peaks that overlap all 4 signaling proteins (dSmad2, Mad, Medea, and Schnurri), and does not require the presence of specific DNA sequences. Unless otherwise noted, read intensity plots were generated by binning ChIP-seq reads into 100 bp bins and extracting read counts in bins surrounding described anchor points (eg. dSMMS modules or dCTCF peaks), and visualized using Java Treeview.<sup>40</sup> Gene Ontology analysis was performed by

running functional annotation tools on DAVID<sup>41</sup> and biological process (level 5) GO terms were subsequently summarized using REVIGO.<sup>42</sup>

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Acknowledgments

We are especially thankful to Drs. Lan Xu, Laurel Raftery, and Mark Biggin for graciously sharing antibodies against Mad, Medea, and Schnurri. We thank the Genomic Services Lab at the HudsonAlpha Institute for Biotechnology for help in performing Illumina sequencing of ChIP-Seq samples.

### Funding

This work was supported by US. Public Health Service Awards 5R01GM035463 to VGC and 5R01 GM095746 to MBO from the National Institutes of Health. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

### Supplemental Material

Supplemental data for this article can be accessed on the publisher's website.

### Accession Numbers

ChIP-seq data are deposited in NCBI's Gene Expression Omnibus (GEO) ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)) under accession number **GSE68654**.

### References

- Liu F. Receptor-regulated Smads in TGF-beta signaling. *Front Biosci* 2003; 8:s1280-303; PMID:12957874; <http://dx.doi.org/10.2741/1149>
- Ross S, Hill CS. How the Smads regulate transcription. *Int J Biochem Cell Biol* 2008; 40:383-408; PMID:18061509; <http://dx.doi.org/10.1016/j.biocel.2007.09.006>
- Massague J. How cells read TGF-beta signals. *Nat Rev Mol Cell Biol* 2000; 1:169-78; PMID:11252892; <http://dx.doi.org/10.1038/35043051>
- Trompouki E, Bowman TV, Lawton LN, Fan ZP, Wu DC, DiBiase A, Martin CS, Cech JN, Sessa AK, Leblanc JL, et al. Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell* 2011; 147:577-89; PMID:22036566; <http://dx.doi.org/10.1016/j.cell.2011.09.044>
- Mullen AC, Orlando DA, Newman JJ, Loven J, Kumar RM, Bilodeau S, Reddy J, Guenther MG, DeKoter RP, Young RA. Master transcription factors determine cell-type-specific responses to TGF-beta signaling. *Cell* 2011; 147:565-76; PMID:22036565; <http://dx.doi.org/10.1016/j.cell.2011.08.050>
- Xi Q, Wang Z, Zaromytidou AI, Zhang XH, Chow-Tsang LF, Liu JX, Kim H, Barlas A, Manova-Todorova K, Kaartinen V, et al. A poised chromatin platform for TGF-beta access to master regulators. *Cell* 2011; 147:1511-24; PMID:22196728; <http://dx.doi.org/10.1016/j.cell.2011.11.032>
- Burton T, Liang B, Dibrov A, Amara F. Transforming growth factor-beta-induced transcription of the Alzheimer beta-amyloid precursor protein gene involves interaction between the CTCF-complex and Smads. *Biochem Biophys Res Commun* 2002; 295:713-23; PMID:12099698; [http://dx.doi.org/10.1016/S0006-291X\(02\)00725-8](http://dx.doi.org/10.1016/S0006-291X(02)00725-8)
- Bergstrom R, Savary K, Moren A, Guibert S, Heldin CH, Ohlsson R, Moustakas A. Transforming growth factor beta promotes complexes between Smad proteins and the CCCTC-binding factor on the H19 imprinting control region chromatin. *J Biol Chem* 2010; 285:19727-37; PMID:20427289
- Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell* 2009; 137:1194-211; PMID:19563753; <http://dx.doi.org/10.1016/j.cell.2009.06.001>
- Van Bortle K, Nichols MH, Li L, Ong CT, Takenaka N, Qin ZS, Corces VG. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* 2014; 15:R82; PMID:24981874; <http://dx.doi.org/10.1186/gb-2014-15-5-r82>
- Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi J, Corces V. Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome research* 2012; 22(11):2176-87; PMID:22722341
- Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. *Mol Cell* 2012; 48:471-84; PMID:23041285; <http://dx.doi.org/10.1016/j.molcel.2012.08.031>
- Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* 2012; 148(3):458-72; PMID:22265598; <http://dx.doi.org/10.1016/j.cell.2012.01.010>
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012; 485:376-80; PMID:22495300; <http://dx.doi.org/10.1038/nature11082>
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 2012; 485:381-5; PMID:22495304; <http://dx.doi.org/10.1038/nature11049>
- Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* 2011; 43:630-8; PMID:21685913; <http://dx.doi.org/10.1038/ng.857>
- Phillips-Cremens JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, Ong CT, Hookway TA, Guo C, Sun Y, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 2012; 48:471-84; PMID:23041285; <http://dx.doi.org/10.1016/j.molcel.2012.08.031>

- Cell 2013; 153:1281-95; PMID:23706625; <http://dx.doi.org/10.1016/j.cell.2013.04.053>
18. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, Ing-Simmons E, Lenhard B, Giorgetti L, Heard E, Fisher AG, et al. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res* 2013; 23(12):2066-77; PMID:24002784
19. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. *Nature* 2012; 489:109-13; PMID:22955621; <http://dx.doi.org/10.1038/nature11279>
20. Wood AM, Van Bortle K, Ramos E, Takenaka N, Rohrbach M, Jones BC, Jones KC, Corces VG. Regulation of chromatin organization and inducible gene expression by a *Drosophila* insulator. *Mol Cell* 2011; 44:29-38; PMID:21981916; <http://dx.doi.org/10.1016/j.molcel.2011.07.035>
21. Li TR, White KP. Tissue-specific gene expression and ecdysone-regulated genomic networks in *Drosophila*. *Dev Cell* 2003; 5:59-72; PMID:12852852; [http://dx.doi.org/10.1016/S1534-5807\(03\)00192-8](http://dx.doi.org/10.1016/S1534-5807(03)00192-8)
22. Armknecht S, Boutros M, Kiger A, Nybakken K, Mathey-Prevot B, Perrimon N. High-throughput RNA interference screens in *Drosophila* tissue culture cells. *Methods Enzymol* 2005; 392:55-73; PMID:15644175
23. Bushey AM, Ramos E, Corces VG. Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Gen Dev* 2009; 23:1338-50; PMID:19443682; <http://dx.doi.org/10.1101/gad.1798209>
24. Xu L, Yao X, Chen X, Lu P, Zhang B, Ip YT. Msk is required for nuclear import of TGF- $\beta$ /BMP-activated Smads. *J Cell Biol* 2007; 178:981-94; PMID:17785517; <http://dx.doi.org/10.1083/jcb.200703106>
25. Sutherland DJ, Li M, Liu XQ, Stefancsik R, Raftery LA. Stepwise formation of a SMAD activity gradient during dorsal-ventral patterning of the *Drosophila* embryo. *Development* 2003; 130:5705-16; PMID:14534137; <http://dx.doi.org/10.1242/dev.00801>
26. MacArthur S, Li XY, Li J, Brown JB, Chu HC, Zeng L, Grondona BP, Hechmer A, Simirenko L, Keranen SV, et al. 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* 2009; 10:R80; PMID:19627575; <http://dx.doi.org/10.1186/gb-2009-10-7-r80>
27. Peterson AJ, Jensen PA, Shimell M, Stefancsik R, Wijayatunge R, Herder R, Raftery LA, O'Connor MB. R-Smad competition controls activin receptor output in *Drosophila*. *PLoS one* 2012; 7:e36548; PMID:22563507; <http://dx.doi.org/10.1371/journal.pone.0036548>
28. Langmead B. Aligning short sequencing reads with Bowtie. *Current protocols in bioinformatics* / editorial board, Andreas D Baxevanis; et al 2010; Chapter 11:Unit 11.7.
29. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008; 9:R137; PMID:18798982; <http://dx.doi.org/10.1186/gb-2008-9-9-r137>
30. Goecks J, Nekrutenko A, Taylor J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 2010; 11:R86; PMID:20738864; <http://dx.doi.org/10.1186/gb-2010-11-8-r86>
31. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. Galaxy: a web-based genome analysis tool for experimentalists. *Current protocols in molecular biology* / edited by Frederick M Ausubel; et al 2010; Chapter 19:Unit 19.0 1-21
32. Giardine B, Riemer C, Hardison RC, Burhans R, Elitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, et al. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* 2005; 15:1451-5; PMID:16169926; <http://dx.doi.org/10.1101/gr.4086505>
33. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol* 2011; 29:24-6; PMID:21221095; <http://dx.doi.org/10.1038/nbt.1754>
34. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief bioinform* 2013; 14:178-92; PMID:22517427; <http://dx.doi.org/10.1093/bib/bbs017>
35. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 2011; 471:480-5; PMID:21179089; <http://dx.doi.org/10.1038/nature09725>
36. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol* 2007; 8:R24; PMID:17324271; <http://dx.doi.org/10.1186/gb-2007-8-2-r24>
37. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002; 12:996-1006; PMID:12045153; <http://dx.doi.org/10.1101/gr.229102>
38. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, et al. The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res* 2013; 41:D64-9; PMID:23155063; <http://dx.doi.org/10.1093/nar/gks1048>
39. Shao Z, Zhang Y, Yuan GC, Orkin SH, Waxman DJ. MAAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol* 2012; 13:R16; PMID:22424423; <http://dx.doi.org/10.1186/gb-2012-13-3-r16>
40. Saldanha AJ. Java Treeview—extensible visualization of microarray data. *Bioinformatics* 2004; 20:3246-8; PMID:15180930; <http://dx.doi.org/10.1093/bioinformatics/bth349>
41. Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, Lempicki RA. Extracting biological meaning from large gene lists with DAVID. *Current protocols in bioinformatics* / editorial board, Andreas D Baxevanis; et al 2009; Chapter 13:Unit 13.1
42. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS one* 2011; 6:e21800; PMID:21789182; <http://dx.doi.org/10.1371/journal.pone.0021800>
43. Peterson AJ, O'Connor MB. Strategies for exploring TGF- $\beta$  signaling in *Drosophila*. *Methods* 2014; 68:183-93; PMID:24680699; <http://dx.doi.org/10.1016/j.ymeth.2014.03.016>
44. Brummel TJ, Twombly V, Marques G, Wrana JL, Newfield SJ, Attisano L, Massague J, O'Connor MB, Gelbart WM. Characterization and relationship of Dpp receptors encoded by the saxophone and thick veins genes in *Drosophila*. *Cell* 1994; 78:251-61; PMID:8044839; [http://dx.doi.org/10.1016/0092-8674\(94\)90295-X](http://dx.doi.org/10.1016/0092-8674(94)90295-X)
45. Letsou A, Arora K, Wrana JL, Simin K, Twombly V, Jamal J, Staehling-Hampton K, Hoffmann FM, Gelbart WM, Massague J, et al. *Drosophila* Dpp signaling is mediated by the punt gene product: a dual ligand-binding type II receptor of the TGF  $\beta$  receptor family. *Cell* 1995; 80:899-908; PMID:7697720; [http://dx.doi.org/10.1016/0092-8674\(95\)90293-7](http://dx.doi.org/10.1016/0092-8674(95)90293-7)
46. Brummel TJ, Abdollah S, Haerry TE, Shimell MJ, Merriam J, Raftery L, Wrana JL, O'Connor MB. The *Drosophila* activin receptor baboon signals through dSmad2 and controls cell proliferation but not patterning during larval development. *Gen Dev* 1999; 13:98-111; PMID:9887103; <http://dx.doi.org/10.1101/gad.13.1.98>
47. Jensen PA, Zheng X, Lee T, O'Connor MB. The *Drosophila* Activin-like ligand Dawdle signals preferentially through one isoform of the Type-I receptor Baboon. *Mech Dev* 2009; 126:950-7; PMID:19766717; <http://dx.doi.org/10.1016/j.mod.2009.09.003>
48. Awasaki T, Huang Y, O'Connor MB, Lee T. Glia instruct developmental neuronal remodeling through TGF- $\beta$  signaling. *Nat Neurosci* 2011; 14:821-3; PMID:21685919; <http://dx.doi.org/10.1038/nn.2833>
49. Gao S, Steffen J, Laughon A. Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. *J Biol Chem* 2005; 280:36158-64; PMID:16109720; <http://dx.doi.org/10.1074/jbc.M506882200>
50. Weiss A, Charbonnier E, Ellertsdottir E, Tsigros A, Wolf C, Schuh R, Pyrowolakis G, Affolter M. A conserved activation element in BMP signaling during *Drosophila* development. *Nat Struct Mol Biol* 2010; 17:69-76; PMID:20010841; <http://dx.doi.org/10.1038/nsmb.1715>
51. Muller B, Hartmann B, Pyrowolakis G, Affolter M, Basler K. Conversion of an extracellular Dpp/BMP morphogen gradient into an inverse transcriptional gradient. *Cell* 2003; 113:221-33; PMID:12705870; [http://dx.doi.org/10.1016/S0092-8674\(03\)00241-1](http://dx.doi.org/10.1016/S0092-8674(03)00241-1)
52. Pyrowolakis G, Hartmann B, Muller B, Basler K, Affolter M. A simple molecular complex mediates widespread BMP-induced repression during *Drosophila* development. *Dev Cell* 2004; 7:229-40; PMID:15296719; <http://dx.doi.org/10.1016/j.devcel.2004.07.008>
53. Zheng X, Wang J, Haerry TE, Wu AY, Martin J, O'Connor MB, Lee CH, Lee T. TGF- $\beta$  signaling activates steroid hormone receptor expression during neuronal remodeling in the *Drosophila* brain. *Cell* 2003; 112:303-15; PMID:12581521; [http://dx.doi.org/10.1016/S0092-8674\(03\)00072-2](http://dx.doi.org/10.1016/S0092-8674(03)00072-2)
54. Gibbens YY, Warren JT, Gilbert LI, O'Connor MB. Neuroendocrine regulation of *Drosophila* metamorphosis requires TGF $\beta$ /Activin signaling. *Development* 2011; 138:2693-703; PMID:21613324; <http://dx.doi.org/10.1242/dev.063412>
55. Chng WB, Sleiman MS, Schupfer F, Lemaître B. Transforming Growth Factor  $\beta$ /Activin Signaling Functions as a Sugar-Sensing Feedback Loop to Regulate Digestive Enzyme Expression. *Cell Rep* 2014; 9:336-48; PMID:25284780; <http://dx.doi.org/10.1016/j.celrep.2014.08.064>
56. Gronroos E, Kingston JJ, Ramchandran A, Randall RA, Vizan P, Hill CS. Transforming growth factor beta inhibits bone morphogenetic protein-induced transcription through novel phosphorylated Smad1/5-Smad3 complexes. *Mol Cell Biol* 2012; 32:2904-16; PMID:22615489; <http://dx.doi.org/10.1128/MCB.00231-12>
57. Peterson AJ, O'Connor MB. Activin receptor inhibition by Smad2 regulates *Drosophila* wing disc patterning through BMP-response elements. *Development* 2013; 140:649-59; PMID:23293296; <http://dx.doi.org/10.1242/dev.085605>
58. Yao LC, Phin S, Cho J, Rushlow C, Arora K, Warrior R. Multiple modular promoter elements drive graded brinker expression in response to the Dpp morphogen gradient. *Development* 2008; 135:2183-92; PMID:18506030; <http://dx.doi.org/10.1242/dev.015826>
59. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* 2011; 27:1696-7; PMID:21486936; <http://dx.doi.org/10.1093/bioinformatics/btr189>
60. Xu X, Yin Z, Hudson JB, Ferguson EL, Frasch M. Smad proteins act in combination with synergistic and antagonistic regulators to target Dpp responses to the *Drosophila* mesoderm. *Gen Dev* 1998; 12:2354-70; PMID:9694800; <http://dx.doi.org/10.1101/gad.12.15.2354>
61. Kusanagi K, Inoue H, Ishidou Y, Mishima HK, Kawabata M, Miyazono K. Characterization of a bone morphogenetic protein-responsive Smad-binding element. *Mol Biol Cell* 2000; 11:555-65; PMID:10679014; <http://dx.doi.org/10.1091/mbc.11.2.555>
62. Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, Schwarzer W, Ettwiller L, Spitz F. Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 2014; 24:390-400; PMID:24398455; <http://dx.doi.org/10.1101/gr.163519.113>