

Cooperative recruitment of Yan via a high-affinity ETS supersite organizes repression to confer specificity and robustness to cardiac cell fate specification

Jean-François Boisclair Lachance,¹ Jemma L. Webber,¹ Lu Hong,^{2,3} Aaron R. Dinner,² and Ilaria Rebay¹

¹Ben May Department for Cancer Research, ²Department of Chemistry, ³Graduate Program in Biophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

Cis-regulatory modules (CRMs) are defined by unique combinations of transcription factor-binding sites. Emerging evidence suggests that the number, affinity, and organization of sites play important roles in regulating enhancer output and, ultimately, gene expression. Here, we investigate how the *cis*-regulatory logic of a tissue-specific CRM responsible for *even-skipped* (*eve*) induction during cardiogenesis organizes the competing inputs of two E-twenty-six (ETS) members: the activator Pointed (Pnt) and the repressor Yan. Using a combination of reporter gene assays and CRISPR–Cas9 gene editing, we suggest that Yan and Pnt have distinct syntax preferences. Not only does Yan prefer high-affinity sites, but an overlapping pair of such sites is necessary and sufficient for Yan to tune *Eve* expression levels in newly specified cardioblasts and block ectopic *Eve* induction and cell fate specification in surrounding progenitors. Mechanistically, the efficient Yan recruitment promoted by this high-affinity ETS supersite not only biases Yan–Pnt competition at the specific CRM but also organizes Yan-repressive complexes in three dimensions across the *eve* locus. Taken together, our results uncover a novel mechanism by which differential interpretation of CRM syntax by a competing repressor–activator pair can confer both specificity and robustness to developmental transitions.

[*Keywords:* *Drosophila* embryogenesis; ETS transcription factor; *cis*-regulatory syntax; heart development; even-skipped; receptor tyrosine kinase]

Supplemental material is available for this article.

Received September 12, 2017; revised version accepted February 14, 2018.

Development of a multicellular organism relies on tissue-specific gene expression programs to establish distinct cell fates and morphologies. The requisite patterns of gene expression must be both spatiotemporally precise and robust in the face of genetic and environmental variation; this is achieved through the action of transcription factors (TFs), whose activating and repressive inputs are integrated at the *cis*-regulatory modules (CRMs) or enhancers of their target genes. Consequently, the sequence of each CRM provides a physical blueprint for a combinatorial regulatory code that translates upstream signaling information into downstream gene expression. While significant advances have been made in our ability to distinguish regulatory elements from background noncoding genomic DNA and identify consensus TF-binding motifs within them, our understanding of how the intrinsic logic of the *cis*-regulatory syntax (namely, the number, affinity, position, spacing, and orientation of binding sites) organ-

izes the necessary set of protein–protein and protein–DNA interactions remains poor (Siggers and Gordán 2014; Inukai et al. 2017). Because single-nucleotide polymorphisms in TF-binding sites are being increasingly correlated with altered gene expression and disease susceptibility (Oldridge et al. 2015; Soldner et al. 2016), the ability to deduce the regulatory logic of an enhancer based on its sequence is important.

The tendencies for TFs to cluster into superfamilies and for cells to coexpress multiple nonredundant members of the same superfamily imply that enhancer syntax must enable TFs with very similar DNA-binding preferences to compete, cooperate, and discriminate between binding sites to achieve appropriate gene expression output. Recent insight into these behaviors has come from studies of Hox family TFs (Crocker et al. 2015). The emerging

Corresponding author: irebay@uchicago.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.307132.117>.

© 2018 Boisclair Lachance et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

model suggests a specificity–affinity trade-off such that low-affinity sites are best discriminated, while high-affinity sites can be bound by many different Hox factors. Clustering multiple low-affinity Hox sites permits the cooperative and additive interactions needed for robust gene activation responses without compromising specificity. Whether analogous syntax rules apply to other TF superfamilies is not known, and how transcriptional repressors solve the specificity–affinity problem remains to be tested.

The E-twenty-six (ETS) superfamily includes both activators and repressors, all of which recognize the same core DNA sequence, 5'-GGAA/T-3' (Hollenhorst et al. 2011). ETS TFs are found across metazoan phyla and play key roles in regulating the gene expression programs that direct many aspects of normal development and patterning (Hollenhorst et al. 2011). Exemplifying this, the *Drosophila* transcriptional activator Pointed (Pnt) and the repressor Yan operate downstream from receptor tyrosine kinase (RTK) signaling pathways to orchestrate numerous cell fate transitions (Klambt 1993; Scholz et al. 1993; O'Neill et al. 1994; Rebay and Rubin 1995; Sopko and Perimon 2013). Much of the current understanding of Yan and Pnt stems from studying their regulation of *even-skipped* (*eve*) expression during cardiac muscle precursor specification at stage 11 of embryogenesis (Carmena et al. 1998, 2002; Halfon et al. 2000; Knirr and Frasch 2001) and *prospero* (*pros*) expression during R7 photoreceptor specification in the developing eye (Xu et al. 2000; Hayashi et al. 2008). Abrogating Pnt-mediated activation or Yan-mediated repression of *eve* or *pros* leads to respective loss or ectopic induction of the associated cell fate (Halfon et al. 2000; Xu et al. 2000; Hayashi et al. 2008; Webber et al. 2013a). Gel shift assays using probes from *eve* or *pros* CRMs revealed that most Yan-bound ETS sites are also bound by Pnt (Flores et al. 2000; Halfon et al. 2000; Xu et al. 2000), and subsequent high-throughput assays confirm their preferences for very similar sequences (Zhu et al. 2011; Nitta et al. 2015). Because none of the in vitro biochemistry has been done with full-length proteins, how accurately the results will predict the outcome of Yan–Pnt competition for ETS sites in CRMs in vivo is uncertain.

Hints that binding site syntax might influence Yan recruitment come from in vitro binding studies with TEL1, the human counterpart of *Drosophila* Yan, and from mathematical modeling of Yan's ETS site occupancy. TEL1 and Yan, unlike Pnt or its mammalian counterpart, ETS1 (Slupsky et al. 1998; Mackereth et al. 2004), self-associate via their sterile α motifs (SAMs), and this homotypic interaction is essential for transcriptional repression in both flies and humans (Qiao et al. 2004; Green et al. 2010; Zhang et al. 2010). Using gel shift assays, SAM–SAM interactions were shown to mediate cooperative binding of TEL1 at paired ETS sites (Green et al. 2010). A recent theoretical analysis of TEL1/Yan occupancy at equilibrium explains how such cooperative SAM–SAM interactions might promote preferential recruitment to tandem ETS-binding sites (Hope et al. 2017). Because neither study examined repressive output, the

question of whether preferential or cooperative binding of Yan to closely apposed ETS sites might bias Yan–Pnt competition to permit more complex discrimination of CRM syntax than current models assume remains pressing.

To evaluate how CRM syntax organizes the opposing repressive and activating inputs from Yan and Pnt to dictate precise transcriptional output, we assessed the impact of mutating the eight putative ETS-binding sites identified in the *eve* muscle heart enhancer (MHE) that drives *eve* expression in 10 segmentally arrayed clusters of pericardial and muscle cells (Halfon et al. 2000). We found that sites with strong affinity best discriminate between Yan and Pnt, with paired sites showing the strongest bias. Thus, mutating a pair of overlapping, conserved, high-affinity ETS sites significantly elevated or expanded MHE reporter expression, consistent with compromised repression. Using CRISPR/Cas9 gene editing of the endogenous MHE, we showed that mutation of this high-affinity ETS supersite reduced Yan recruitment to not only the MHE but also two other CRMs across the *eve* locus. Mesodermal *Eve* expression was elevated, consistent with compromised Yan recruitment, resulting in inadequate repression. In this compromised background, environmental and genetic stresses that would normally be buffered against were now sufficient to induce specification of ectopic *Eve*-positive (*Eve*⁺) cells and reduce survival. We conclude that the conserved high-affinity ETS pair within the MHE plays a unique and pivotal role in not just recruiting Yan-repressive complexes to the isolated enhancer but also longer-range coordination of transcriptional complex organization and function across the locus.

Results

A pair of high-affinity Yan-binding sites organizes transcriptional repression at the MHE

The 312-base-pair (bp) MHE from the *eve* locus provides a tractable system for exploring how *cis*-regulatory architecture coordinates the activating and repressive inputs that determine expression. Prior work showed that the MHE integrates spatial and temporal cues from multiple signaling pathways and mesodermal determinants to drive expression in a bilaterally symmetric pattern that matches closely the segmental pattern of 10 three-cell clusters of *Eve*⁺ cells within the cardiogenic mesoderm (Fig. 1A,B; Halfon et al. 2000). This isolated enhancer also responds appropriately to loss or gain of the TFs that regulate *eve* mesodermal expression, including Yan, Pnt, the mesodermal determinant Twist (*Twi*), and the Dpp effector Mothers against Dpp (*Mad*) (Fig. 1B'–E; Supplemental Fig. S1A–A'; Halfon et al. 2000; Webber et al. 2013a). Four of the eight GGAA/T core-containing putative ETS motifs have been shown to bind Yan and Pnt in vitro; in vivo, their individual mutation reduced reporter expression in the *Eve*⁺ clusters, while their simultaneous mutation abolished expression (Halfon et al. 2000). These results suggest that their role in Pnt-mediated

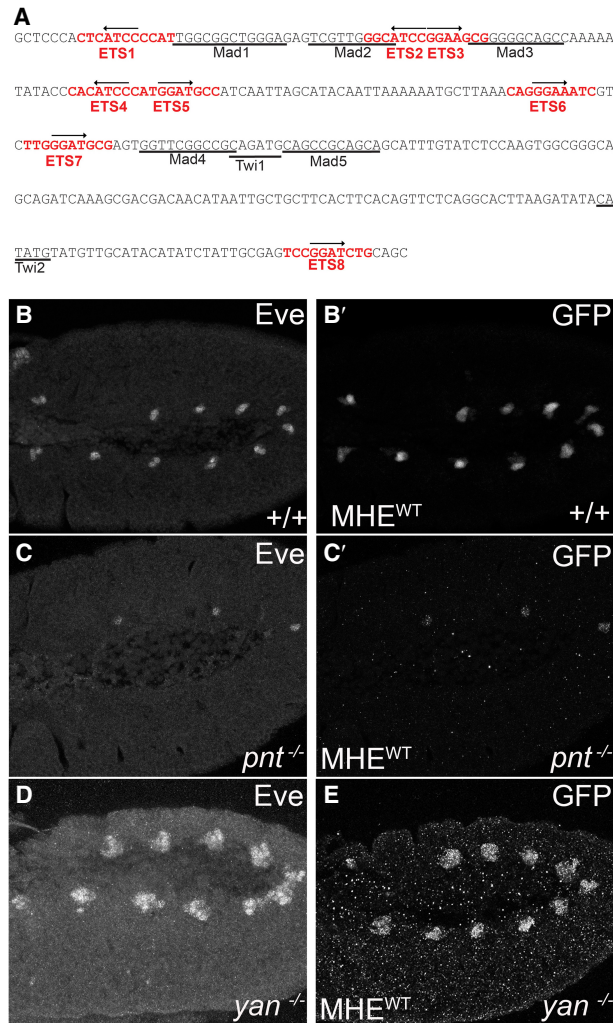


Figure 1. The MHE reliably reports the pattern of mesodermal Eve expression. (A) MHE sequence with putative ETS sites in red. Black arrows indicate site orientation. Previously characterized Mad and Twi sites are underlined. (B–E) Lateral views, oriented with anterior to the left and ventral down, of the thoracic and abdominal segments of representative stage 11 embryos expressing two copies of the MHE^{WT}-GFP reporter in *w¹¹¹⁸* (B), *pnt^{Δ88}* (C), or *yan^{ER443}* mutants (D,E). Costaining with anti-Eve (B,C) and anti-GFP (B',C') shows that the reporter-driven pattern matches closely that of endogenous Eve. (D,E) Loss of *yan* expands these Eve expressing and GFP-expressing cell clusters.

activation of MHE expression outweighs their contributions to Yan-mediated repression.

To investigate the contribution of the remaining four candidate ETS motifs to MHE expression, we first assessed their affinity qualitatively using a competitive gel shift assay (see the Materials and Methods for details; Green et al. 2010). We used Yan rather than Pnt for these experiments both because of the greater sensitivity afforded by Yan's intrinsically higher in vitro DNA-binding ability (Xu et al. 2000) and because recombinant Pnt was even more prone to degradation than Yan. To simplify the nomenclature, we relabeled the sites 1–8 (Fig. 1A); thus, sites

annotated as Ets1–4 in Halfon et al. (2000) correspond to sites 2, 3, 5, and 8 in our study. Sites 2, 3, and 8 competed effectively, while sites 1 and 4–7 did not (Fig. 2A; Supplemental Fig. S2A). In accordance with the binding preferences identified for fly, human, and mouse ETS TFs (Wei et al. 2010; Webber et al. 2013a; Nitta et al. 2015), the three higher-affinity sites (2, 3, and 8) contain C nucleotides at the –1 and –2 positions, whereas the others do not (Fig. 1A). The mix of higher- and lower-affinity ETS sites made the MHE an ideal test case for exploring how ETS TFs solve the affinity–specificity paradox.

To begin, we measured the effects of individual ETS site mutations on enhancer output using MHE-GFP reporter transgenes. Indirect immunofluorescent measurements in fixed animals were used (see the Materials and Methods) because the positions of Eve⁺ cells deep within the embryo made quantitative analysis in live embryos prohibitively difficult. In broad agreement with Halfon et al. (2000), individual mutation of the majority of ETS sites (namely, 1, 4, 5, 6, or 7) decreased reporter expression relative to MHE^{WT}-GFP control, with the effect of mutating site 6 the most pronounced (Fig. 2B). In contrast, mutation of sites 2 and 8 elevated reporter gene expression, while site 3 was neutral. When considered together with the in vitro binding data (Fig. 2A; Supplemental Fig. S2A), the single-site mutagenesis analysis suggested that high-affinity sites (2, 3, and 8) contribute more significantly to MHE repression than lower-affinity sites, while lower-affinity sites contribute more significantly to MHE activation. Site 3 appeared to be an exception: Although an effective competitor in vitro, its mutation did not alter reporter expression.

To test whether site 3 might participate in both repression and activation, we mutated it together with site 2 or 8 and compared MHE reporter expression relative to that of the individual mutations. Unexpectedly, simultaneous mutation of sites 2 and 3 induced a greater than fourfold synergistic increase in reporter expression in the Eve⁺ clusters and ectopic expression in surrounding cells (Fig. 2C,D-D'). In addition to defining a role for site 3 in Yan repression and thereby confirming its functional relevance, the above-additive expression increase within the clusters and the ectopic expression outside the clusters identified sites 2 and 3 as critical collaborative organizers of Yan-repressive complexes at the MHE.

In marked contrast to the increased reporter expression associated with mutation of sites 2 and 3, simultaneous mutation of sites 3 and 8 reduced expression below wild-type MHE levels (Fig. 2C). This result further confirmed the functional relevance of site 3 and uncovered a role in activating expression. In other words, when site 3 is mutated alone, the lack of expression change reflects a loss of balanced repressive and activating inputs; these opposing inputs are revealed only when site 3 is mutated together with site 2 or 8, respectively. The lower than wild-type expression of MHE^{mut3,8} further suggests that site 8, despite elevating reporter expression when mutated alone, also integrates Yan–Pnt inputs. Consistent with this conclusion, expression of an MHE^{mut2,8} reporter was subadditive relative to that of the single mutants, although still

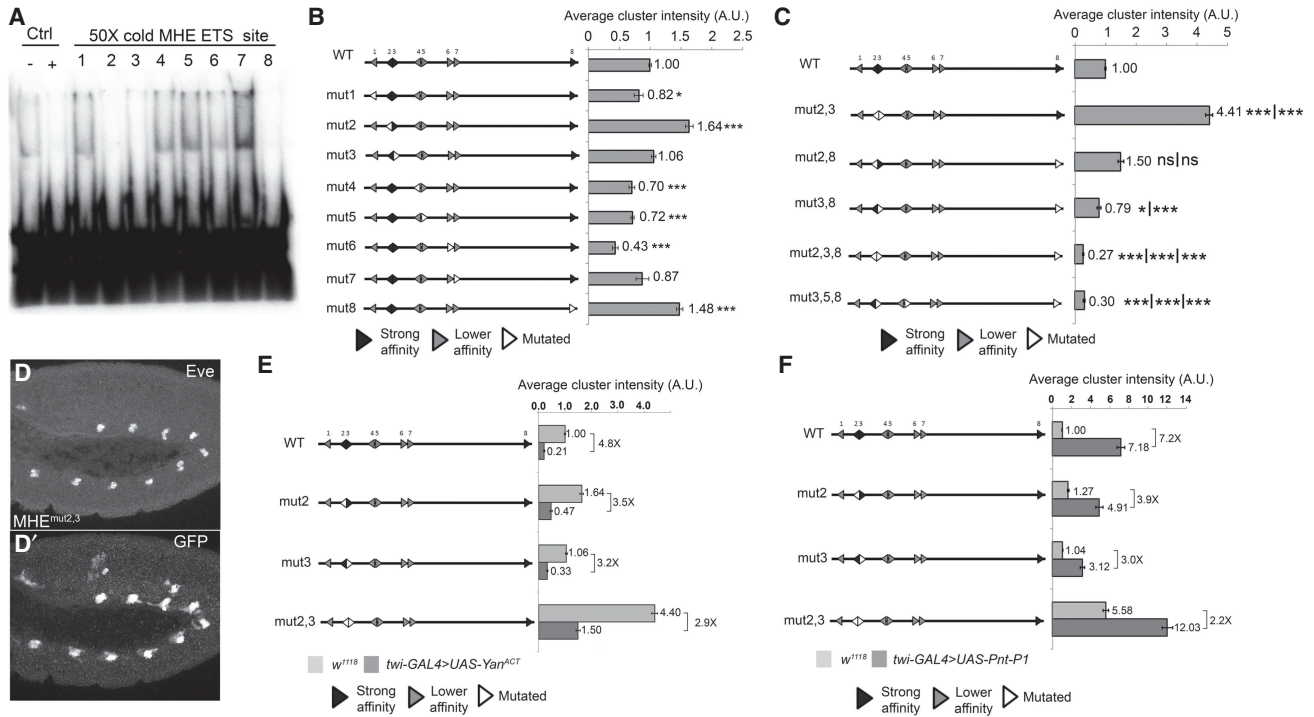


Figure 2. An overlapping pair of high-affinity ETS sites organizes MHE repression, while lower-affinity sites contribute predominantly to reporter activation. (A) Competitive gel shift assay testing the ability of MHE ETS sites 1–8 to outcompete a ³²P-labeled ETS site probe bound to a Yan^{A86D} monomer. ETS sites 2, 3, and 8 competed effectively, whereas the remaining sites did not. (B) The impact of individual ETS site mutations (mut1–mut8) on the average cluster intensity of the MHE-GFP reporter suggests a role for strong-affinity sites in repression and lower-affinity sites in activation. In this and all subsequent diagrams, ETS sites and their sense/antisense orientations are depicted with arrowheads, with lower-affinity sites in gray, strong sites in black, and mutated sites in white. Error bars show SEM. The statistical significance of single mutants relative to MHE^{WT} after Bonferroni correction is indicated. (***) *P* < 0.001. (C) The impact of double and triple mutants on MHE-GFP reporter expression. The statistical significance relative to each of the relevant single-site mutations after Bonferroni correction is provided in numerical order. (*) *P* < 0.05; (***) *P* < 0.001. The mut2,3 combination was unique in its nonadditive increase in reporter expression; although measurements were less linear at this higher range of GFP expression, all three quantification strategies tested revealed a similar trend [Supplemental Fig. S2C–E]. Mutation of all three strong sites reduced reporter expression, showing that these sites also mediate activating inputs and that effective activation requires multiple sites. (D,D') Representative stage 11 embryos showing the increased and ectopic expression driven by one copy of MHE^{mut2,3} in and outside Eve⁺ clusters. (E,F) Single and double mutations in MHE ETS sites 2 and 3 reduce the fold repression by *twi-GAL4*-driven Yan^{ACT} (E) and fold activation by *twi-GAL4*-driven PntP1 (F) relative to MHE^{WT}. All measurements were normalized to MHE^{WT}, but, in F, the intensity of the confocal laser was set on MHE^{mut2,3}/*twi-GAL4*,UAS-PntP1 embryos instead of MHE^{WT} as in E. All differences are statistically different, with *P* < 0.001. Error bars show SEM.

above wild-type levels (Fig. 2C). Apart from mut2,3 and mut2,8 and with one exception (mut2,5), all other combinations of double- or triple-mutant combinations tested, including mut2,3,8, reduced reporter expression relative to MHE^{WT} (Fig. 2C; Supplemental Fig. S2F). Taken together, these results suggest that activating inputs integrated through the full complement of ETS sites limit the strength of Yan-repressive complexes that are collaboratively organized by the high-affinity ETS2,3 pair.

We next tested the conclusion that MHE ETS sites can integrate both repressive and activating inputs by comparing the response of MHE^{WT} versus MHE^{mut2,3} to overexpression of Yan or Pnt. Whereas overexpression of a constitutively active form of Yan (Rebay and Rubin 1995) reduced MHE^{WT} expression fivefold, only a threefold reduction was achieved when one or both sites in the high-affinity pair were mutated (Fig. 2E). Thus, al-

though full repression requires an intact ETS2,3 pair, if Yan levels are sufficiently high, it can interact productively with the remaining sites to repress expression. An analogous result was obtained with Pnt overexpression such that MHE^{WT} reporter expression increased sevenfold but only twofold to fourfold when the ETS2,3 pair was mutated (Fig. 2F). The finding that increased Pnt dose increases both wild-type and mutant reporter expression suggests that Pnt levels are limiting with respect to MHE activation even in situations when repression has been compromised by mutation of the ETS2,3 pair. We conclude that Yan and Pnt can interact productively with all ETS sites within the MHE and that, under physiological conditions, the relative Pnt and Yan concentrations are likely to play important roles in shaping their interactions with the *cis*-regulatory architecture to dictate the precise expression outputs needed for cell fate specification.

Several of our results—namely, the expression increases associated with MHE^{mut2}, MHE^{mut3}, MHE^{mut8}, MHE^{mut2,3}, and MHE^{mut2,5} (Fig. 2C; Supplemental Fig. S2F)—contradicted previous findings (Halfon et al. 2000). To rule out the possibility that the 86Fb landing site used in our assays was producing artifacts, we inserted MHE^{WT} and MHE^{mut2,3} into an alternative landing site (attP2) and measured reporter expression. Although not as strong as the increase measured for MHE^{mut2,3} in the 86Fb insertion site, the trend was consistent, with an almost threefold increase in expression in the Eve⁺ clusters and significant ectopic expression outside (Supplemental Fig. S2G). We conclude that P-element position-effect differences in the earlier study (Halfon et al. 2000) may have masked the expression increases that our system more reliably detects.

Sufficiency of a high-affinity ETS pair for MHE repression

The MHE mutagenesis described above uncovered a pivotal role for the 2,3 high-affinity ETS site pair in coordinating repressive inputs. To evaluate the biological relevance of this regulatory capability, we compared ETS syntax in the MHEs of other *Drosophila* species, taking advantage of prior work that aligned *eve* enhancer sequences (Supplemental Fig. S2H,L; Hare et al. 2008). Of the three high-affinity sites, sites 2 and 3 were conserved over 40 million years of evolution, from *Drosophila melanogaster* to *Drosophila virilis* (Fig. 3A). The third high-affinity site, site 8, was conserved across the *melanogaster* subgroup but was lost in both the *obscura* and *virilis* groups (Fig. 3A). We validated the repressive function of the ETS2,3 pair in the *D. virilis* MHE using cross-species reporter transgenes (Supplemental Fig. S2I–K). Consistent with previous work (Hare et al. 2008), the control reporter DvirMHE^{WT} drove GFP expression in the 10 clusters of Eve⁺ cells; expression was weaker than that driven by the *melanogaster* MHE^{WT} reporter (cf. Fig. 1B and Supplemental Fig. S2I), presumably reflecting the low degree of overall sequence conservation between these two highly diverged species. Simultaneous mutation of the ETS2,3 pair increased reporter expression (Supplemental Fig. S2J,K), confirming its importance to MHE repression. Of the lower-affinity sites, only site 6, which our reporter analysis identified as most critical to activation (Fig. 2B), was conserved across all three *Drosophila* groups (Fig. 3A). Additional ETS sites were noted in the MHEs of all species outside the *melanogaster* subgroup, but, based on sequence, none of these are predicted to be high affinity (Supplemental Fig. S2L). Together, these comparisons suggest strong evolutionary pressure to both maintain the unique high-affinity 2,3 pair and limit the binding affinity of other ETS sequences within the enhancer.

If high-affinity Yan binding underlies the unique contribution of the 2,3 pair to MHE expression, then introducing a comparable high-affinity tandem elsewhere in the enhancer should restore repression of the MHE^{mut2,3} reporter. To test this idea, we added a ninth ETS site adjacent to and overlapping site 8 to form a new anti-parallel pair that

closely matched the endogenous 2,3 configuration (Fig. 3B). We anticipated that this manipulation would be both minimally disruptive to the enhancer and most readily interpretable because although site 8 integrates inputs from both Yan and Pnt, repression dominates (Fig. 2B,C). Quantification of GFP reporter expression showed that the newly engineered 9,8 ETS pair suppressed the elevated expression associated with MHE^{mut2,3} back to MHE^{WT} levels (Fig. 3B). The introduction of a short spacer reduced repressive efficiency, but inverting sites 9 and 8 to put them in parallel produced expression levels equivalent to those achieved with the initial anti-parallel configuration (Fig. 3B). Together, these results show that a high-affinity ETS pair can be sufficient to organize effective Yan repression at the isolated MHE.

The MHE ETS2,3 pair does not support classic cooperative Yan binding

The nonadditive expression increase that resulted from simultaneous mutation of MHE ETS sites 2 and 3 (Fig. 2B,C) suggested a partially redundant yet collaborative role in recruiting and organizing Yan-repressive complexes. In considering possible mechanisms, we were motivated by prior *in vitro* studies of the human ortholog of Yan, TEL1, that demonstrated the importance of SAM–SAM interactions to cooperative recruitment of dimers to paired ETS sites (Green et al. 2010). However, this study did not test an MHE ETS2,3-like pair in which two anti-parallel GGAA/T cores are immediately juxtaposed (Fig. 1A). We therefore used a combination of *in vitro* gel shifts and structural modeling to explore whether the synergistic contribution of ETS sites 2 and 3 to MHE repression *in vivo* could reflect classic cooperative Yan binding.

As a control, we first confirmed that Yan can bind cooperatively to the same pair of ETS consensus sites used in the TEL1 study (Green et al. 2010). As predicted, Yan monomers bound independently to the two sites, producing, at low concentrations, a faster-migrating complex corresponding to a single bound monomer and, at increasing concentrations, a slower-migrating complex corresponding to a ternary complex of two bound monomers (Fig. 4A). Even at low concentrations, Yan dimers produced primarily the slower-migrating complex, consistent with SAM–SAM interactions driving cooperative binding (Fig. 4A). Confirming that both Yan molecules in the dimer are DNA-bound in this ternary complex, mutation of one of the two ETS sites enriched for the faster-migrating form (Fig. 4B).

Quite different results were obtained using a probe containing the 2,3 pair of MHE-derived ETS sites (Fig. 4C). First, incubation with Yan monomers resulted in only a single shifted species, suggesting that the head-to-tail arrangement of MHE ETS sites 2 and 3 does not permit simultaneous occupancy of two Yan molecules. Consistent with this interpretation, structural modeling predicted a steric clash that should preclude occupancy of both sites in the ETS2,3 pair by a Yan dimer (Fig. 4D–D'). Incubation with Yan dimers produced primarily the fast-migrating species observed with Yan monomers (Fig. 4C). A

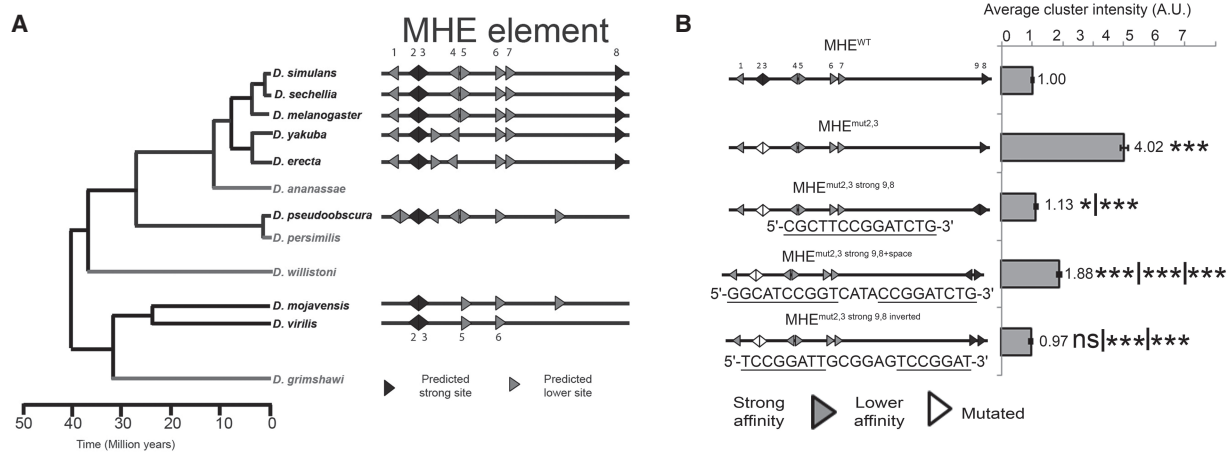


Figure 3. Evolutionary conservation and sufficiency of an ETS2,3-like pair for MHE repression. (A) Phylogenetic tree (adapted from Gramates et al. 2017) and schematic representation of MHE ETS sites in eight different *Drosophila* species; light gray indicates species whose sequences were not analyzed. Full MHE sequences, as identified by Hare et al. (2008), are in Supplemental Figure S2L. (B) The introduction of ectopic paired high-affinity ETS sites can restore effective repression to the MHE^{mut2,3} reporter. Error bars show SEM. Statistical significance after Bonferroni correction from either MHE^{WT}, MHE^{mut2,3}, or MHE^{mut2,3 strong9,8} (where relevant) is shown. (*) $P < 0.05$; (***) $P < 0.001$.

slower-migrating form was also noted, but because it was also observed with an MHE-derived probe in which one of the ETS sites was mutated, we suspect that it reflects a ternary complex in which one Yan molecule is bound to one of the two strong ETS consensus sites, and the second is either interacting nonspecifically with an adjacent sequence or is not contacting the DNA at all. Based on the combined in vitro and theoretical results, we conclude that classic cooperative recruitment of Yan dimers is unlikely to explain the strong functional synergy between ETS sites 2 and 3 revealed by our in vivo expression measurements. Instead, we suggest that the superimposition of two overlapping high-affinity ETS consensus sequences creates a “supersite” that significantly increases the probability of Yan occupancy and effective repression. Speculation about possible noncanonical cooperative regulatory mechanisms that could be mediated by the ETS2,3 supersite is presented in the Discussion.

The MHE ETS2,3 supersite confers robustness to endogenous *Eve* expression

Having established the importance of the ETS2,3 pair to MHE reporter expression and sensitivity to changes in Yan and Pnt levels, we next used CRISPR/Cas9 to explore its contribution in the context of the whole locus. To control for the impact of genetic background, we first re-engineered the wild-type MHE to create an *eve*^{MHEWT} allele (see the Materials and Methods). Specification of *Eve*⁺ cells was normal (Fig. 5B), but average *Eve* levels per cluster in *eve*^{MHEWT} homozygotes were reduced relative to *w*¹¹¹⁸ embryos (Fig. 5A; Supplemental Fig. S3). Although the sequence changes used to introduce the PAM sites are outside the minimal enhancer and do not impact known TF-binding sites (Supplemental Fig. S4), they could alter unknown activating sites or interactions. Al-

ternatively, the reduction in *Eve* levels could reflect more complicated influences of the genetic background (Jiang et al. 2015). With this in mind, we used *eve*^{MHEWT} as the reference background for all subsequent analyses.

We next mutated ETS sites 2 and 3 to generate *eve*^{MHEmut2,3}. Average *Eve* expression levels in the mesodermal clusters were 20% higher than in *eve*^{MHEWT} control embryos (Fig. 5A), but the number of *Eve*⁺ cells was unchanged, and the homozygotes eclosed with normal Mendelian expectations from balanced stocks (Fig. 5B; Supplemental Table S1). Thus, mutation of these two ETS sites in the context of the whole locus had a much more modest impact than in the context of the isolated enhancer (Fig. 2C). However, in survival assays in which stage 11 embryos were counted and placed in vials and then cultured to adulthood at elevated temperature, *eve*^{MHEmut2,3} appeared less fit than either *eve*^{MHEWT} or *w*¹¹¹⁸ controls (Fig. 5C).

To determine whether the elevated *Eve* expression and reduced fitness of *eve*^{MHEmut2,3} embryos might result from altered Yan-mediated regulation, we analyzed Yan chromatin occupancy by ChIP-qPCR (chromatin immunoprecipitation [ChIP combined with quantitative PCR [qPCR])). Yan binding was reduced at the MHE but unchanged at a positive control CRM from the *argos* (*aos*) locus (Fig. 5E). Thus, Yan recruitment to the endogenous *eve* MHE is strongly dependent on the ETS2,3 tandem.

Our previous work showed that three-dimensional (3D) interactions between the MHE and two other CRMs, termed D1 and D2, stabilize Yan occupancy and repression at *eve* such that deletion of any one CRE reduced recruitment to the other (Webber et al. 2013b). We therefore asked whether the reduced Yan recruitment to the MHE also impacted Yan binding at the D1 and D2 elements. Remarkably, significant reductions were measured at both D1 and D2, suggesting that the molecular mechanism

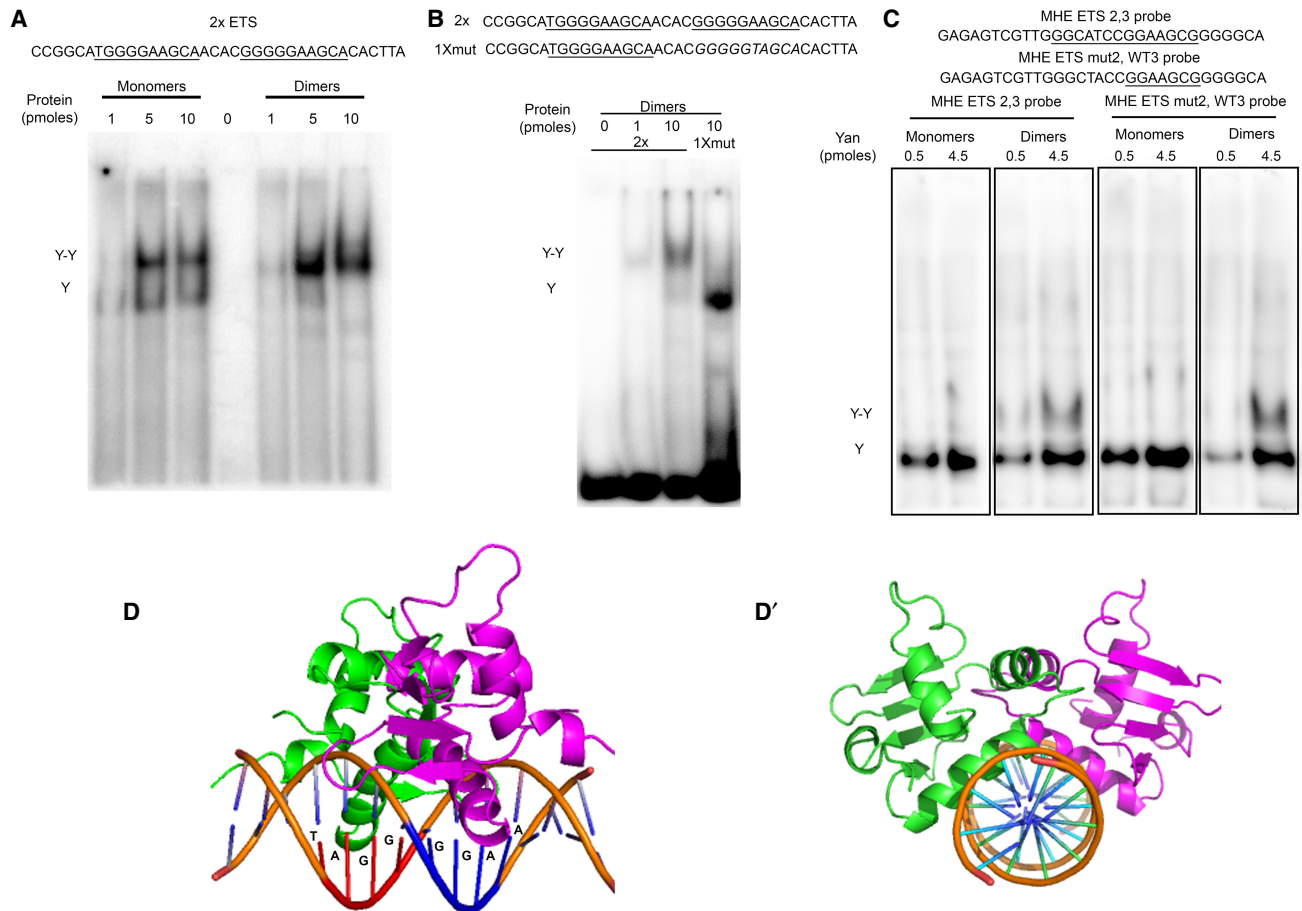


Figure 4. The overlapping anti-parallel configuration of the ETS_{2,3} pair may preclude classic cooperative recruitment of Yan dimers. (A–C) Gel shifts using ³²P-labeled probes and recombinant Yan monomers (Yan^{A86D}) or Yan dimers (1:1 Yan^{A86D}:Yan^{V105R}). Increasing concentrations of total Yan protein are indicated *above* each gel. (A) SAM-mediated dimerization induces classic cooperative binding to probe with two consensus ETS sites (2× ETS). Single-bound (Y) and double-bound (Y–Y) species are indicated. (B) Cooperative binding requires two ETS sites. (C) The MHE ETS_{2,3} sites do not support simultaneous occupancy by two Yan molecules. The slower-migrating species (Y–Y) observed with Yan dimers occurs even when one ETS site is mutated, suggesting that it reflects association of the second Yan molecule to an adjacent nonspecific sequence. (D) A structural model of a Yan ETS DNA-binding domain (DBD) dimer in complex with the MHE ETS_{2,3} sequence, viewed from the side (D) and along the axis of the DNA (D'), predicts a strong steric clash that should preclude simultaneous occupancy of both ETS sites. The two Yan DBDs are colored in green and purple.

by which Yan-repressive complexes are organized in three dimensions across the *eve* locus requires effective recruitment to MHE ETS sites 2 and 3 (Fig. 5E).

The similarity in phenotypic consequences between mutating the ETS_{2,3} pair and deleting the D1 or D2 CRM—namely, reduced Yan occupancy, increased Eve levels but no cell fate defects, and sensitivity to environmental stress—motivated us to explore further parallels. Specifically, our earlier work showed that the addition of genetic stress to a regulatory system already compromised by deletion of the D1 or D2 CRM impacted the specification of Eve⁺ cells and, ultimately, cardiac function and viability (Webber et al. 2013b). We therefore asked whether genetic manipulations that increase the Pnt:Yan ratio but do not produce phenotypes in wild-type embryos would induce extra Eve⁺ cells in the compromised repressive background of *eve*^{MHEmut2,3}.

Three different genetic interactions were tested: increased *pnt* dose, decreased *yan* dose, and decreased *aos* dose. First, using a functional *Pnt-GFP* BAC transgene (Boisclair Lachance et al. 2014) to increase Pnt levels, we found that doubling the *pnt* dose increased Eve levels, perturbed cell fate specification, and reduced survival (Fig. 5A,B,D,F,G). Although removing one copy of *yan* had no significant effect on Eve intensity or Eve⁺ cell numbers, when we simultaneously reduced *yan* and added one copy of *Pnt-GFP*, which on its own had no effect, Eve levels increased, and extra Eve⁺ cells were specified (Fig. 5A, B,H,I). Finally, we halved the dose of *aos*, which encodes an EGFR antagonist whose complete loss leads to an expansion in the number of Eve⁺ cells (Halfon et al. 2000; Carmena et al. 2002), and found increased Eve levels and extra Eve⁺ cells (Fig. 6A–D). Examining the survival of *eve*^{MHEmut2,3} embryos to adulthood in the experiments

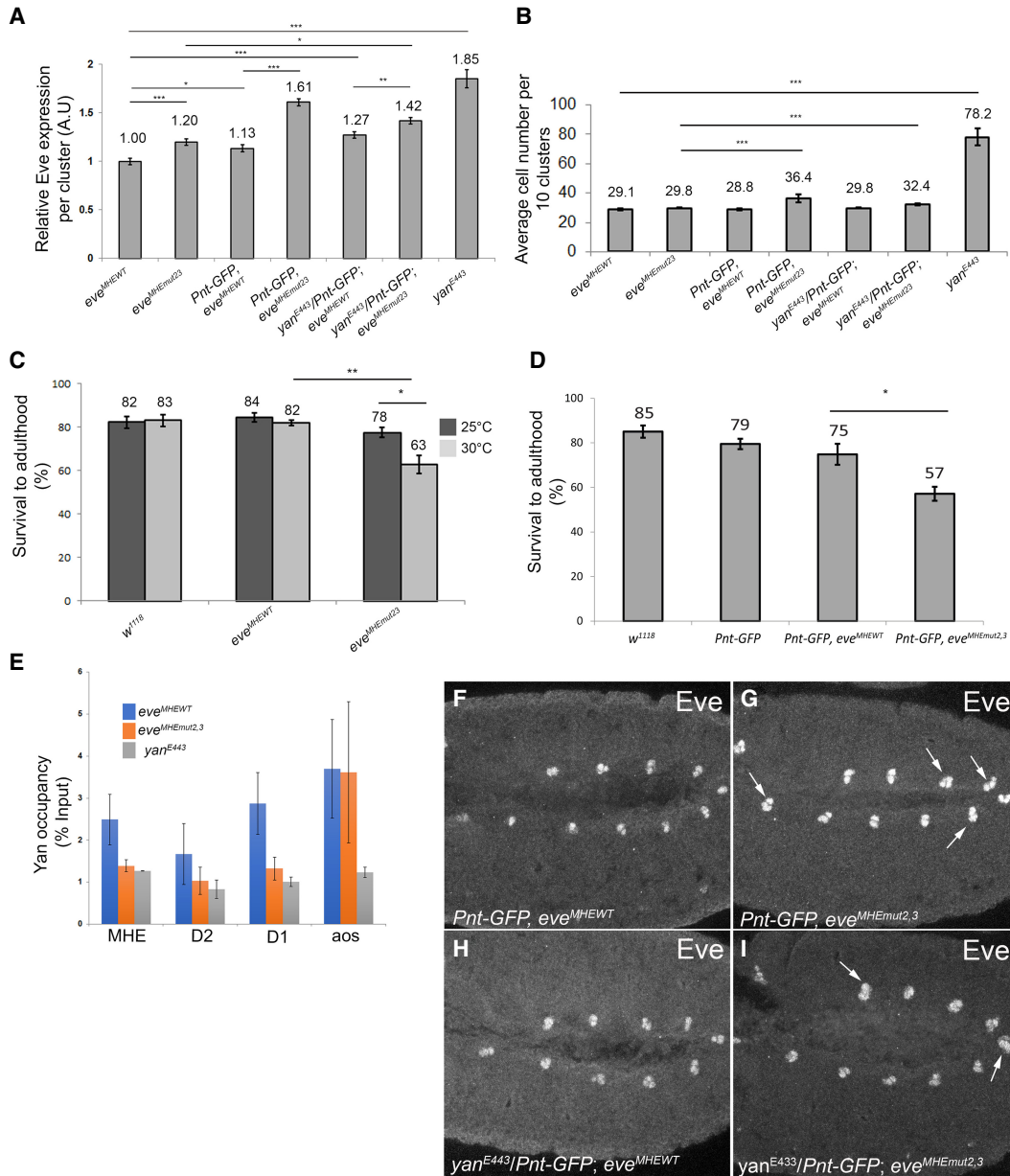


Figure 5. Loss of robustness in *eve^{mut2,3}* embryos highlights the importance of ETS2,3-mediated regulation in cardiac cell fate specification. (A) Quantification of average Eve levels per cluster shows that the *eve^{MHEmut2,3}* background is sensitized to genetic perturbations that are well buffered against in *eve^{MHEWT}* embryos. Error bars show SEM. Statistical significance after Bonferroni correction is depicted. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$. (B) The average number of Eve⁺ cells per 10 clusters in the same genetic backgrounds as in A shows that specification of ectopic Eve⁺ cells occurs in the *eve^{MHEmut2,3}* background, but not in the *eve^{MHEWT}* control, when the Pnt:Yan ratio is increased. Error bars show SEM. (C) The reduced survival of *eve^{MHEmut2,3}* embryos to adulthood was enhanced by temperature stress. (D) Increased *pnt* dose also decreased *eve^{MHEmut2,3}* survival. (E) ChIP-qPCR (chromatin immunoprecipitation [ChIP] combined with quantitative PCR [qPCR]) from stage 11 embryos. Signals were normalized to a negative control (see Supplemental Methods). Mutation of the ETS2,3 site reduced occupancy at all three *eve* CRMs but not at *argos* (*aos*). (F–I) Eve expression in representative stage 11 embryos. Examples of clusters with four or more Eve⁺ cells are indicated with white arrows. Error bars show SEM.

described above revealed a consistent inverse correlation between the extent of ectopic Eve⁺ cell fate specification and viability (Fig. 5C,D; Supplemental Table S1), implying that the specification defects observed at this stage ultimately impair cardiac function. In all of these experiments, only minor expression increases and no cell

fate changes or reduced viability were detected in the *eve^{MHEWT}* controls.

We conclude that the conserved ETS 2,3 pair organizes Yan repression at *eve* to prevent both elevated expression in the newly specified Eve⁺ signaling cells and ectopic cell fate induction in the immediately surrounding cells.

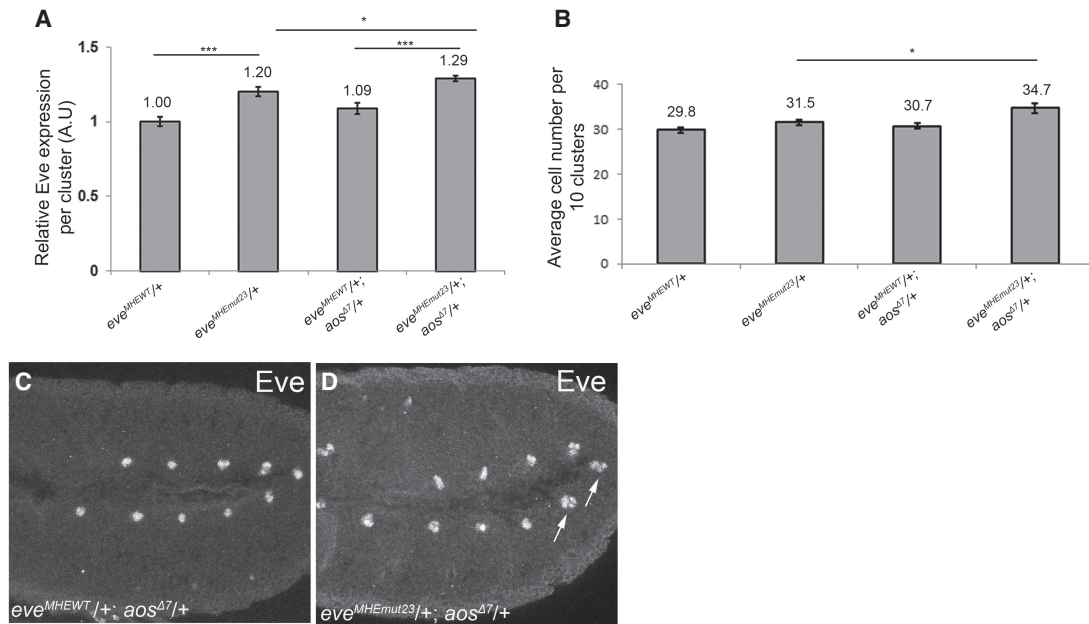


Figure 6. Mutation of the ETS2,3 site increases sensitivity to variation in upstream signaling. Quantification of average Eve levels per cluster (A) and average number of Eve⁺ cells (B) shows that *aos* heterozygosity increases the average number of Eve⁺ cells specified in *eve*^{MHEmut2,3} embryos. Error bars show SEM. (C,D) Representative stage 11 embryos, with clusters with extra Eve⁺ cells marked by a white arrow.

When these sites are mutated, the residual recruitment of Yan to the MHE and other regulatory elements confers transcriptional dynamics that support normal cell fate specification but only under ideal conditions. Finally, these results both confirm the power of the transgenic reporter analyses to pinpoint relevant features of endogenous *eve* regulation and suggest that the full regulatory complexity of an intact locus can dampen the impact of single binding site changes that dramatically alter behavior at an isolated CRM. While *eve* could be unusual in its regulatory redundancy, we suspect that regulatory complexity will be more the norm than the exception; thus, our pioneering study should perhaps provide a general warning to the field that when using genome editing to test enhancer function in vivo, more complicated assays—in our case, examination of altered TF occupancy and the addition of genetic or environmental perturbations—may be needed to reveal the phenotypic consequence.

Discussion

Focusing on ETS-binding motifs within a conserved regulatory module, the *eve* MHE, we identified a simple syntax that allows for robust qualitative and quantitative control of enhancer output. Based on extensive mutagenesis of this enhancer, we propose that Yan's and Pnt's respective preferences for high- and low-affinity ETS sites provide a mechanism for integrating their competing repressive and activating inputs at individual CRMs. In particular, we found that the use of paired strong-affinity sites appears critical to the assembly of repressive complexes

that dampen *eve* expression in newly specified cardiac precursors where Yan levels are low and prevent ectopic *eve* induction in the surrounding mesoderm where levels of activating TFs such as Twi are high. CRISPR/Cas9-mediated mutation of the endogenous MHE confirmed the importance of such optimized syntax for precise Eve expression levels and uncovered an unexpected role of the ETS2,3 supersite in longer-range organization of Yan complexes across the locus. We speculate that efficient Yan recruitment to high-affinity supersites not only influences short-range interactions at the specific enhancer but also fosters longer-range communication across multiple CRMs.

The unique repressive contribution of the ETS2,3 pair may reflect an unconventional form of cooperative recruitment that provides novel regulatory capabilities to Yan-repressive complexes. Specifically, even though the two sites in the ETS2,3 pair are probably too close to permit simultaneous occupancy, their immediate juxtaposition may significantly increase the probability of stable Yan binding. For example, although nothing is known about the kinetics and dynamics of Yan–DNA interactions, the presence of two overlapping high-affinity binding sites could promote stable occupancy by increasing the chance that a newly dissociated Yan molecule would immediately rebind. The syntax could also support a more organized dynamic in which the two molecules of a Yan dimer toggle back and forth rapidly between bound and unbound states at the two sites. Even more speculatively, because SAM-mediated dimerization is required for Yan-mediated repression (Zhang et al. 2010), if the configuration of the ETS2,3 pair ensured that one molecule of

the dimer was always bound, this could leave the second free to interact with either an adjacent nonspecific sequence, as we modeled previously (Hope et al. 2017); another high-affinity ETS site elsewhere in the MHE (for example, site 8); or, even more speculatively, an ETS motif in the D1 or D2 CRM (Webber et al. 2013b). It is also possible that the *in vivo* mechanism by which full-length Yan contacts DNA is different from that suggested by the *in vitro* assays. For example, interactions with other TFs and cofactors might somehow mitigate the steric clash to allow simultaneous occupancy by a Yan dimer. In this case, higher-order Yan complexes (for example, trimers or tetramers) could mediate the requisite longer-range contacts. Regardless of specific mechanism, the idea that high-affinity “supersites” might be used to anchor longer-range TF–TF and TF–DNA interactions will be an interesting direction for future investigations.

Our previous work exploring the *in vivo* functionality of a Yan protein in which the SAM–SAM interface has been mutated to prevent self-association further supports the importance of SAM-mediated repressor cooperativity. Specifically, we found that although Yan monomers are recruited to enhancers genome-wide in a pattern close to that of wild-type Yan, adequate repression does not occur, and phenotypes consistent with *yan* loss of function ensue (Webber et al. 2013a). This work also noted the prevalence of clustered high-affinity ETS sites across a number of Yan ChIP targets, suggesting that the mechanisms uncovered in our dissection of MHE ETS site syntax might be broadly applicable. Focusing on *eve*, we suspect that at the resolution of individual ETS sites, in the absence of SAM-mediated cooperativity, Yan occupancy of the ETS_{2,3} tandem would be insufficiently stable to either compete appropriately with Pnt at the MHE or organize the necessary 3D communication across the locus.

We also note a parallel between the consequences of mutating the high-affinity ETS_{2,3} supersite in the endogenous *eve* locus and the findings of an earlier analysis in which we deleted three different Yan-bound CRMs within a genomic *Eve*-YFP BAC transgene (Webber et al. 2013b). In this earlier study, while deleting the pattern-driving MHE almost completely ablated mesodermal *Eve*-YFP induction, deleting a “repressive” Yan-bound element (referred to as the D1) increased *Eve*-YFP expression ~1.5-fold and led to the specification of extra *Eve*⁺ cells. Additionally, deletion of either the MHE or the D1 in the BAC transgene reduced Yan occupancy at not only the deleted element but also the remaining intact CRMs. Here, we report a comparable loss of Yan occupancy across the *eve* locus upon mutation of the MHE ETS_{2,3} supersite but only a modest increase in *Eve* levels and no cell fate specification defects. The discrepancy between reduced Yan occupancy and increased *Eve* levels in the *eve*^{MHEmut2,3} mutant relative to the D1 deletion mutant suggests that deleting an entire CRM not only compromises Yan occupancy across the locus but also disrupts additional repressive inputs. Consistent with this interpretation, the *eve*^{MHEmut2,3} background appeared highly sensitized, with the increase in *Eve* levels and *Eve*⁺ cell fate specification associated with a twofold increase in

pnt dose almost exactly matching the effects of deleting an entire “repressive” CRM. Further exploration of how high-affinity ETS pairs organize Yan repression at and between CRMs and how this coordinates the competing and collaborating inputs from other TFs will be needed to test these ideas at *eve* and, more broadly, other target genes.

To conclude, we propose a working model in which Yan’s and Pnt’s differential interpretation of ETS syntax adds a “dimmer” capability to the classic on/off switch, thereby refining its sensitivity and tunability. Focusing on *eve* as an example, prior to the onset of RTK-induced cardiac cell fate specification or in cells subject to sub-maximal signaling, we suggest that Yan’s bias for high-affinity sites ensures an effectively 100% probability of occupancy at the ETS_{2,3} supersite and hence stable repression. In this regime, Yan could also outcompete Pnt at the lower-affinity sites to occupy fully the CRM, or, if Yan levels are limiting, as our data suggest, its preference for high-affinity sites and relative “distaste” for lower-affinity sites could offer Pnt an opportunity to occupy the latter and perhaps influence Yan repression. In contrast, if Yan and Pnt had identical ETS-binding preferences, we would expect a less tuned response to RTK signaling; indeed, when we removed the high-affinity 2,3 pair and hence the strong bias toward Yan occupancy and repression at the MHE, stochastic ectopic expression was induced in the surrounding mesoderm where RTK levels are submaximal. Thus, their distinct preferences ensure that only maximal RTK activation will trigger the necessary shift in Yan–Pnt occupancy and activity to activate *eve* expression. Furthermore, while previous models assumed a complete switch from total Yan occupancy to total Pnt occupancy as *Eve*⁺ cell fates are specified, our work suggests that the ETS_{2,3} supersite still recruits Yan-repressive input even in *Eve*⁺ cells with very low Yan concentration. We speculate that the ability to apply continued Yan-repressive input after cell fate induction may contribute to the robustness of certain developmental transitions by stabilizing the newly acquired cell fate. In agreement with this, in the context of the endogenous *eve* locus, disruption of the ETS_{2,3} pair sensitized *eve* to both fluctuations in upstream signaling and environmental conditions.

More broadly, we speculate that the interplay between the *cis*-regulatory logic of a CRM and the unique biophysical parameters of different TFs permits evolution to fine-tune gene expression output to a specific threshold depending on each cell’s developmental requirement. In the case of Yan–Pnt-regulated genes, the interplay between the degree of Yan SAM-mediated self-association and ETS syntax enables this repressor–activator pair to discriminate between ETS sites with unexpected precision. Furthermore, instead of RTK activation inducing a complete switch from Yan occupancy to Pnt occupancy as cell fates are induced, the cooperative recruitment of Yan to supersites may enable newly differentiating cells with lower Yan:Pnt ratios to sustain the Yan-repressive influence needed to ensure precision and robustness of the gene expression patterns. We suggest that these ideas provide an interesting new vantage point for considering how

single-nucleotide polymorphisms in TF-binding sites may heighten susceptibility to disease by compromising the robustness of gene regulatory networks.

Materials and methods

Drosophila genetics

The following chromosomes were from the Bloomington *Drosophila* Stock Center: *w¹¹¹⁸*, *pnt^{Δ88}*, *aos^{Δ7}*, *UAS-PntP1*, *nos-φC31*; *attP2*, *nos-φC31;86Fb*, *TM3,Sb-Ser,twi-GAL4,UAS-GFP(TTG)*, *Cyo-twist-GAL4,UASGFP(CTG)*, *TM6B*, *tubulin-GAL80*, and *Cyo,Tub>PBac*.

Additional strains used were *yan^{E443}* and *yan^{E833}* (Karim et al. 1996), *UAS-Yan^{ACT}* (Rebay and Rubin 1995), *Pnt-GFP* (Boisclair Lachance et al. 2014), *vasa-Cas9* [a gift from R. Fehon, GFP-negative and RFP-negative], and *eve^{MHEWT}* and *eve^{MHEmut2,3}* (this study; described below). A list of MHE-pJ20 transgenes (this study) is in the Supplemental Material.

For the CRISPR/Cas9-mediated generation of *eve^{MHEWT}* and *eve^{MHEmut2,3}* alleles, a 2.1-kb region, including the wild-type or mutated MHE sequences and mutated PAM sites, was inserted into pHD-Scarless (generated by O'Connor-Giles laboratory [Gratz et al. 2013], *Drosophila* Genomics Resource Center, 1364). Both templates were confirmed by sequencing. Guide RNAs were subcloned into the pU6-Bbs1 chiRNA plasmid (Addgene, 45946) (Gratz et al. 2013). Each template (300 ng/μL) and the two guide RNAs (75 ng/μL) were injected into a GFP/RFP-negative *vasa-Cas9* strain (a gift from Rick Fehon). G₀ adults were crossed individually to *w¹¹¹⁸*, and transformants were identified by 3X-Pax-RFP expression in the eyes of the F1 progeny. The 3X-Pax-RFP piggyBac cassette was excised, and progeny without RFP expression were crossed to CTG to establish the *eve^{MHEWT}* and *eve^{MHEmut2,3}* stocks; the desired alleles were confirmed by restriction digest and sequencing (see the Supplemental Material for additional details).

MHE reporter subcloning and transgenesis

QuikChange mutagenesis (Stratagene) of MHE^{WT}-pBluescript (Halfon et al. 2000) was used to modify ETS-binding site sequences (primers are listed in the Supplemental Material). Elements were shuttled into the pJ20 GFP reporter plasmid (Rister et al. 2015) as BamHI fragments and confirmed by sequencing prior to injection. Transgenes were inserted at the 86Fb (Bischof et al. 2007) or attP2 (Groth et al. 2004) landing sites and outcrossed to *w¹¹¹⁸* to remove the X-linked φ31 integrase.

A 484-bp *D. virilis* MHE^{WT} was amplified from genomic DNA using primers 5'-TACTCCGGCGCTCCTCGAGGTTAATGCA CCCAGCAGCC-3' and 5'-GTACCCCGCGGCCGCTAGCGT TTGCAGTGTAGCTGAAATATATGG-3' and inserted into BamHI-digested pJ20 using Gibson Assembly. A 484-bp gene block (Integrated DNA Technologies) (see the Supplemental Material for sequences) containing *D. virilis* MHE^{mut2,3} was similarly inserted into pJ20.

Embryo staining and imaging

Embryos were stained and fixed as described previously (see the Supplemental Material for details; Webber et al. 2013b). The primary antibodies used were rabbit anti-GFP (1/2000; ThermoFisher Scientific, A6455), chicken anti-GFP (1/2000; Abcam, ab13970), guinea pig anti-GFP (1:8000; gift from R. Fehon), mouse anti-Eve 3C10 (1/10; Developmental Studies Hybridoma Bank [DSHB]) (Patel et al. 1992), and mouse anti-Yan 8B12 (1/750; DSHB) (Rebay and Rubin 1995). The secondary antibodies used

were goat anti-mouse Cy3 (1/2000; Jackson ImmunoResearch), goat anti-rabbit 488 (1/2000; Jackson ImmunoResearch), and donkey anti-chicken 488 (1/2000; Jackson ImmunoResearch). All primary antibodies were validated for use in *Drosophila* embryos and are standard for the field.

Images were obtained using either a Zeiss LSM510 or Zeiss LSM880 confocal microscope using 0.45-μm optical sections. For GFP reporter expression measurements, the microscope gain was set for embryos carrying one copy of MHE^{WT}-GFP in an otherwise wild-type background, and then the same settings were used to image all other samples fixed and stained in parallel. For experiments with the CRISPR alleles, the gain was set up on the *eve^{MHEWT}* allele unless stated otherwise. GFP reporter expression and Eve expression were measured using ImageJ as described previously (Webber et al. 2013b) and normalized to MHE^{WT}/+ or *eve^{MHEWT}* expression unless stated otherwise. At least two independent experiments were analyzed per genotype. To confirm linearity, we calculated the average fluorescent intensity per cluster in embryos carrying one versus two copies of the MHE^{WT}-GFP transgene using three different methods: maximum projection, average, and sum (additional details are in the Supplemental Material). All three detected the expected twofold difference (Supplemental Fig. S2B–E). Maximal projection was selected for all further analyses to be consistent with our previously published work (Webber et al. 2013b).

Statistical analysis

All results were analyzed using two-way ANOVA analysis to test for statistical significance. When found to be significantly different, a two-way Student's *t*-test with Bonferroni correction was applied to validate statistical significance in a pairwise manner. For all experiments, at least 40 clusters (four embryos) were analyzed, but, in general, 100 clusters (10 embryos) were analyzed.

In vitro gel shift assays

Recombinant Yan^{A86D} and Yan^{V105R} (amino acids 1–499) were prepared by TEV protease cleavage of GST fusion proteins purified from the BL21 codon plus *Escherichia coli* cells (see the Supplemental Material for details). Yan protein in the eluate was assessed for quality and concentration by SDS-PAGE followed by ImageJ quantification of the intensity of Yan products. Eluates were stored for up to 5 d at 4°C.

Double-stranded probes were end-labeled with T4 polynucleotide kinase (New England Biolabs) and γ³²PATP (Perkin Elmer) for 1 h at 37°C prior to annealing. All oligonucleotide sequences are listed in the Supplemental Material.

For competition assays, ~30 pmol of Yan^{A86D} was incubated with 50× cold probe for 20 min on ice in 25-μL reactions (0.5 μg of poly-dI-dC, 5 mg of BSA, 25% glycerol in 13 mM HEPES at pH 7.9, 40 mM KCl, 0.7 mM EDTA, 0.3 mM DTT, 1 pmol of labeled probe). Cold competitor (100×) was added, and reactions were incubated an additional 20 min on ice. Cooperative binding was assessed by incubating equimolar amounts of Yan^{V105R} and Yan^{A86D} for 20 min on ice in the presence of 1 pmol of labeled probe. Samples were resolved on 10% polyacrylamide gels (competition assay) or on 4%–20% gradient gels (cooperative binding) to capture both bound and free probe run in 0.5× Tris-borate-EDTA (TBE) buffer for 45 min at 120 V.

ChIP-qPCR

ChIP from stage 11 embryos (5 h and 20 min to 7 h and 20 min at 25°C) was performed as described in Webber et al. (2013a) (see the Supplemental Material for details).

Structural modeling

The structure of the ETS domain of Yan was built using SWISS-MODEL (Arnold et al. 2006; Guex et al. 2009; Kiefer et al. 2009; Biasini et al. 2014) based on the NMR structure of the ETS domain of *Homo sapiens* transcriptional factor ETV6 (Protein Data Bank [PDB] ID 2DAO; <https://www.rcsb.org/structure/2DAO>). The B-form DNA structures containing MHE ETS sites were built using Avogadro (Hanwell et al. 2012). The protein/DNA-binding interfaces were modeled based on a crystal structure of the ETS domain of *H. sapiens* ETS2 in complex with DNA (PDB ID 4BQA) (Newman et al. 2015). Specifically, the ETS-binding sites were aligned based on DNA backbone atom positions, and the ETS domains were aligned structurally using STAMP (Russell and Barton 1992; Humphrey et al. 1996; Roberts et al. 2006).

Acknowledgments

We thank Mark Halfon for pBluescript-MHE^{WT}; Jens Rister and Claude Desplan for pJR20; Cynthia Horth, Hitoshi Matakatsu, and Rick Fehon for the generation of the GFP/RFP-negative *vasa-Cas9* strain; R. Fehon for the guinea pig anti-GFP; and Aaron Mitchell-Dick for help injecting constructs. We are also grateful for the helpful comments of the reviewers regarding our interpretation of the biochemistry. We acknowledge the Bloomington *Drosophila* Stock Center (National Institutes of Health [NIH] P40OD018537), the Developmental Studies Hybridoma Bank (created by the National Institute of Child Health and Human Development of the NIH), and the *Drosophila* Genomics Resource Center (NIH 2P40OD010949) for critical reagents. We thank Rebay and Fehon laboratory members for helpful discussions, and Trevor Davis, Matt Hope, and Richard Mann for comments on the manuscript. J.-F.B.L. and J.L.W. were supported in part by National Institutes of Health (NIH) R01 GM080372 to I.R. J.L.W. was supported by American Heart Association grants 12POST12040225 (2012–2014) and 15POST22660028 (2015). L.H. and J.-F.B.L. were supported in part by NIH R01 EY025957. Additional support came from the Genomics Core Facility through University of Chicago Cancer Center support grant P30 CA014599.

Author contributions: J.-F.B.L. and J.L.W. performed experiments. L.H. and A.R.D. performed the structural modeling. J.-F.B.L., J.L.W., and I.R. designed and interpreted experiments and wrote the manuscript.

References

- Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a Web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195–201.
- Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, et al. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**: W252–W258.
- Bischof J, Maeda RK, Hediger M, Karch F, Basler K. 2007. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci* **104**: 3312–3317.
- Boisclair Lachance JF, Peláez N, Cassidy JJ, Webber JL, Rebay I, Carthew RW. 2014. A comparative study of pointed and yan expression reveals new complexity to the transcriptional net-
- works downstream of receptor tyrosine kinase signaling. *Dev Biol* **385**: 263–278.
- Carmena A, Gisselbrecht S, Harrison J, Jiménez F, Michelson AM. 1998. Combinatorial signaling codes for the progressive determination of cell fates in the *Drosophila* embryonic mesoderm. *Genes Dev* **12**: 3910–3922.
- Carmena A, Buff E, Halfon MS, Gisselbrecht S, Jiménez F, Baylies MK, Michelson AM. 2002. Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the *Drosophila* embryonic mesoderm. *Dev Biol* **244**: 226–242.
- Crocker J, Abe N, Rinaldi L, McGregor AP, Frankel N, Wang S, Alsawadi A, Valenti P, Plaza S, Payre F, et al. 2015. Low affinity binding site clusters confer HOX specificity and regulatory robustness. *Cell* **160**: 191–203.
- Flores GV, Duan H, Yan H, Nagaraj R, Fu W, Zou Y, Noll M, Banerjee U. 2000. Combinatorial signaling in the specification of unique cell fates. *Cell* **103**: 75–85.
- Gramates LS, Marygold SJ, dos Santos G, Urbano J-M, Antonazzo G, Matthews BB, Rey AJ, Tabone CJ, Crosby MA, Emmert DB, et al. 2017. FlyBase at 25: looking to the future. *Nucleic Acids Res* **45**: D663–D671.
- Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM. 2013. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**: 1029–1035.
- Green SM, Coyne HJ, McIntosh LP, Graves BJ. 2010. DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *J Biol Chem* **285**: 18496–18504.
- Groth AC, Fish M, Nusse R, Calos MP. 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* **166**: 1775–1782.
- Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30**: S162–S173.
- Halfon MS, Carmena A, Gisselbrecht S, Sackerson CM, Jiménez F, Baylies MK, Michelson AM. 2000. Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**: 63–74.
- Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. 2012. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform* **4**: 17.
- Hare EE, Peterson BK, Iyer VN, Meier R, Eisen MB. 2008. Sepsid even-skipped enhancers are functionally conserved in *Drosophila* despite lack of sequence conservation. *PLoS Genet* **4**: e1000106.
- Hayashi T, Xu C, Carthew RW. 2008. Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the *Drosophila* eye. *Development* **135**: 2787–2796.
- Hollenhorst PC, McIntosh LP, Graves BJ. 2011. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem* **80**: 437–471.
- Hope CM, Rebay I, Reinitz J. 2017. DNA occupancy of polymerizing transcription factors: a chemical model of the ETS family factor Yan. *Biophys J* **112**: 180–192.
- Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph* **14**: 33–38, 27–28.
- Inukai S, Kock KH, Bulyk ML. 2017. Transcription factor–DNA binding: beyond binding site motifs. *Curr Opin Genet Dev* **43**: 110–119.
- Jiang P, Ludwig MZ, Kreitman M, Reinitz J. 2015. Natural variation of the expression pattern of the segmentation gene even-skipped in *melanogaster*. *Dev Biol* **405**: 173–181.

- Karim FD, Chang HC, Therrien M, Wassarman DA, Laverty T, Rubin GM. 1996. A screen for genes that function downstream of Ras1 during *Drosophila* eye development. *Genetics* **143**: 315–329.
- Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**: D387–D392.
- Klämmt C. 1993. The *Drosophila* gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**: 163–176.
- Knirr S, Frasch M. 2001. Molecular integration of inductive and mesoderm-intrinsic inputs governs even-skipped enhancer activity in a subset of pericardial and dorsal muscle progenitors. *Dev Biol* **238**: 13–26.
- Mackereth CD, Schärpf M, Gentile LN, MacIntosh SE, Slupsky CM, McIntosh LP. 2004. Diversity in structure and function of the Ets family PNT domains. *J Mol Biol* **342**: 1249–1264.
- Newman JA, Cooper CDO, Aitkenhead H, Gileadi O. 2015. Structural insights into the autoregulation and cooperativity of the human transcription factor Ets-2. *J Biol Chem* **290**: 8539–8549.
- Nitta KR, Jolma A, Yin Y, Morgunova E, Kivioja T, Akhtar J, Hens K, Toivonen J, Deplancke B, Furlong EEM, et al. 2015. Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *Elife* **4**: e04837.
- Oldridge DA, Wood AC, Weichert-Leahey N, Crimmins I, Sussman R, Winter C, McDaniel LD, Diamond M, Hart LS, Zhu S, et al. 2015. Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature* **528**: 418–421.
- O'Neill EM, Rebay I, Tjian R, Rubin GM. 1994. The activities of two Ets-related transcription factors required for *Drosophila* eye development are modulated by the Ras/MAPK pathway. *Cell* **78**: 137–147.
- Patel NH, Ball EE, Goodman CS. 1992. Changing role of even-skipped during the evolution of insect pattern formation. *Nature* **357**: 339–342.
- Qiao F, Song H, Kim CA, Sawaya MR, Hunter JB, Gingery M, Rebay I, Courey AJ, Bowie JU. 2004. Derepression by depolymerization. *Cell* **118**: 163–173.
- Rebay I, Rubin GM. 1995. Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**: 857–866.
- Rister J, Razzaq A, Boodram P, Desai N, Tsanis C, Chen H, Jukam D, Desplan C. 2015. Single-base pair differences in a shared motif determine differential Rhodopsin expression. *Science* **350**: 1258–1261.
- Roberts E, Eargle J, Wright D, Luthey-Schulten Z. 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**: 382.
- Russell RB, Barton GJ. 1992. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins Struct Funct Genet* **14**: 309–323.
- Scholz H, Deatrick J, Klaes A, Klämmt C. 1993. Genetic dissection of pointed, a *Drosophila* gene encoding two ETS-related proteins. *Genetics* **135**: 455–468.
- Siggers T, Gordán R. 2014. Protein–DNA binding: complexities and multi-protein codes. *Nucleic Acids Res* **42**: 2099–2111.
- Slupsky CM, Gentile LN, Donaldson LW, Mackereth CD, Seidel JJ, Graves BJ, McIntosh LP. 1998. Structure of the Ets-1 pointed domain and mitogen-activated protein kinase phosphorylation site. *Proc Natl Acad Sci* **95**: 12129–12134.
- Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barasa MI, Goldmann J, Myers RH, Young RA, Jaenisch R. 2016. Parkinson-associated risk variant in distal enhancer of α -synuclein modulates target gene expression. *Nature* **533**: 1–20.
- Sopko R, Perrimon N. 2013. Receptor tyrosine kinases in *Drosophila* development. *Cold Spring Harb Perspect Biol* **5**: a009050.
- Webber JL, Zhang J, Cote L, Vivekanand P, Ni X, Zhou J, Nègre N, Carthew RW, White KP, Rebay I. 2013a. The relationship between long-range chromatin occupancy and polymerization of the *Drosophila* ets family transcriptional repressor yan. *Genetics* **193**: 633–649.
- Webber JL, Zhang J, Mitchell-Dick A, Rebay I. 2013b. 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. *Genes Dev* **27**: 2293–2298.
- Wei G-H, Badis G, Berger MF, Kivioja T, Palin K, Enge M, Bonke M, Jolma A, Varjosalo M, Gehrke AR, et al. 2010. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J* **29**: 2147–2160.
- Xu C, Kauffmann RC, Zhang J, Kladny S, Carthew RW. 2000. Overlapping activators and repressors delimit transcriptional response to receptor tyrosine kinase signals in the *Drosophila* eye. *Cell* **103**: 87–97.
- Zhang J, Graham TGW, Vivekanand P, Cote L, Cetera M, Rebay I. 2010. Sterile α motif domain-mediated self-association plays an essential role in modulating the activity of the *Drosophila* ETS family transcriptional repressor Yan. *Mol Cell Biol* **30**: 1158–1170.
- Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enuameh MS, Basciotta MD, Brasefield JA, Zhu C, Asriyan Y, Lapointe DS, et al. 2011. FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res* **39**: D111–D117.