LETTER

doi:10.1038/nature13395

Genome-scale functional characterization of *Drosophila* developmental enhancers *in vivo*

Evgeny Z. Kvon[1]†, Tomas Kazmar[1], Gerald Stampfel[1]*, J. Omar Yáñez-Cuna[1]*, Michaela Pagani[1], Katharina Schernhuber[1], Barry J. Dickson[1]† & Alexander Stark[1]

Transcriptional enhancers are crucial regulators of gene expression and animal development[1] and the characterization of their genomic organization, spatiotemporal activities and sequence properties is a key goal in modern biology[2–8]. Here we characterize the *in vivo* activity of 7,705 *Drosophila melanogaster* enhancer candidates covering 13.5% of the non-coding non-repetitive genome throughout embryogenesis. 3,557 (46%) candidates are active, suggesting a high density with 50,000 to 100,000 developmental enhancers genome-wide. The vast majority of enhancers display specific spatial patterns that are highly dynamic during development. Most appear to regulate their neighbouring genes, suggesting that the *cis*-regulatory genome is organized locally into domains, which are supported by chromosomal domains, insulator binding and genome evolution. However, 12 to 21 per cent of enhancers appear to skip non-expressed neighbours and regulate a more distal gene. Finally, we computationally identify *cis*-regulatory motifs that are predictive and required for enhancer activity, as we validate experimentally. This work provides global insights into the organization of an animal regulatory genome and the make-up of enhancer sequences and confirms and generalizes principles from previous studies[1,9]. All enhancer patterns are annotated manually with a controlled vocabulary and all results are available through a web interface (http://enhancers.starklab.org), including the raw images of all microscopy slides for manual inspection at arbitrary zoom levels.

Animal development depends on differential gene expression governed by genomic regulatory elements called enhancers[1,9], which are being studied extensively[2,3,8,10,11]. Many of the basic principles of developmental gene regulation have been elucidated in the fruitfly *Drosophila melanogaster*[1,12,13], and work over the past decades has characterized gene expression, transcription factor binding, chromatin features and enhancer activity in *Drosophila* at unprecedented levels[2,5–8,14–16]. This and the ability to obtain many embryos from all developmental stages[17] make *Drosophila* an ideal model in which to characterize spatiotemporal enhancer activities at a genomic scale and throughout embryogenesis.

To systematically characterize developmental enhancers in the *D. melanogaster* genome, we made use of transgenic fly lines (Vienna Tiles (VT) library), publicly available from the Vienna *Drosophila* RNAi Center (VDRC). Each line contains a transcriptional reporter construct with a ~2 kilobase (kb) genomic DNA fragment (enhancer candidate), minimal promoter and *GAL4* reporter gene integrated into an identical position in the fly genome[16], thus allowing the direct comparison of the candidates' activities (Fig. 1a, Extended Data Fig. 1a, b and Supplementary Table 1). Together, these fragments cover about 14 million base pairs or 13.5% of the non-coding, non-repetitive genome, with little or no bias regarding the distance to transcription start sites (TSSs; Extended Data Fig. 1c) or the embryonic expression of neighbouring genes (Extended Data Fig. 1d).

We developed a high-throughput pipeline to assess transcriptional enhancer activities in fly embryos by *in situ* hybridization against the *GAL4* reporter transcript. For each transgenic line, we acquired whole-slide images, each with about 400 embryos covering all stages of embryogenesis, and manually annotated the enhancer activity patterns using a controlled vocabulary[14] at six time intervals of embryogenesis (Extended Data Fig. 1e). The pipeline reported activities independent of fragment delineation and orientation and recovered 27 out of 28 known enhancers, whereas 13 out of 13 non-*Drosophila* controls were inactive (Extended Data Fig. 2a–c and Supplementary Information section 1). Results from re-testing 34 negative and 78 positive fragments using a different genomic site (on chromosome 2L instead of 3L) and reporter gene (*lexA*) were highly similar and the majority (82%) of enhancer activity patterns matched to the expression patterns of neighbouring genes, suggesting that we predominantly measured endogenous enhancer activities (Extended Data Fig. 2d–f and Supplementary Information section 1).

3,557 of all 7,705 tested candidate fragments (46%) were active in the embryo with diverse patterns that included gap and pair-rule patterns, all primary germ layers (Extended Data Figs 3a and 4a), and all major cell types and tissues (Fig. 1b and Extended Data Figs 3b and 4b). The fraction of active fragments increased about fivefold from ~7% in early embryos (stages 4–6) to ~35% for stages 15–16, consistent with the increase in organism complexity and the number of cell types (Fig. 1c). By contrast, the number of expressed genes remains roughly constant during embryogenesis (~1.3-fold increase[18]). Enhancer activities were much sparser than gene expression patterns both temporally and spatially: while 94% of all enhancers were only transiently active and only 0.8% were ubiquitous during the entire embryogenesis, this was true for 56.7% and 20.5% of the genes, respectively (Fig. 1d, e and Extended Data Fig. 5a–c). The temporal dynamics of enhancer activity was also apparent from changes of enhancer-associated chromatin features such as DNase I hypersensitivity (DHS), binding of co-activator CBP/p300, and presence of histone H3K27 acetylation mark assessed in entire embryos or in a tissue-specific manner[2,19,20] (Fig. 1f–h, Extended Data Fig. 6 and Supplementary Information section 2). Together, this confirms and quantifies the transient and dynamic nature of enhancer function and suggests that development progresses through increasingly complex gene regulation by enhancers with temporally and spatially restricted activities.

We next identified domains in the blastoderm embryo in which enhancers appeared co-regulated (that is, were coordinately active or inactive). Automated image segmentation and reverse clustering revealed distinct regions corresponding to the presumptive anterior and posterior endoderm, head and trunk mesoderm, procephalic neuroectoderm, and others, overall strongly resembling the established fate map of the blastoderm embryo[17] (Fig. 1i and Extended Data Fig. 4c). This suggests that cells within these domains have a common developmental fate, presumably due to shared *trans*-regulatory environments. Indeed, during late stages, early mesodermal enhancers were preferentially active in mesoderm derivatives (somatic, visceral and cardiac muscles), whereas early endodermal enhancers were active in endoderm derivatives (midgut and Malpighian tubules) (Extended Data Fig. 4d). These and equivalent trends for other presumptive tissues of the early embryo (Extended Data Fig. 4e–g)

[1]Research Institute of Molecular Pathology (IMP), Vienna Biocenter VBC, Dr Bohr-Gasse 7, 1030 Vienna, Austria. †Present addresses: Howard Hughes Medical Institute, Janelia Farm Research Campus, Ashburn, Virginia 20147, USA (B.J.D.); Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA (E.Z.K.).
*These authors contributed equally to this work.

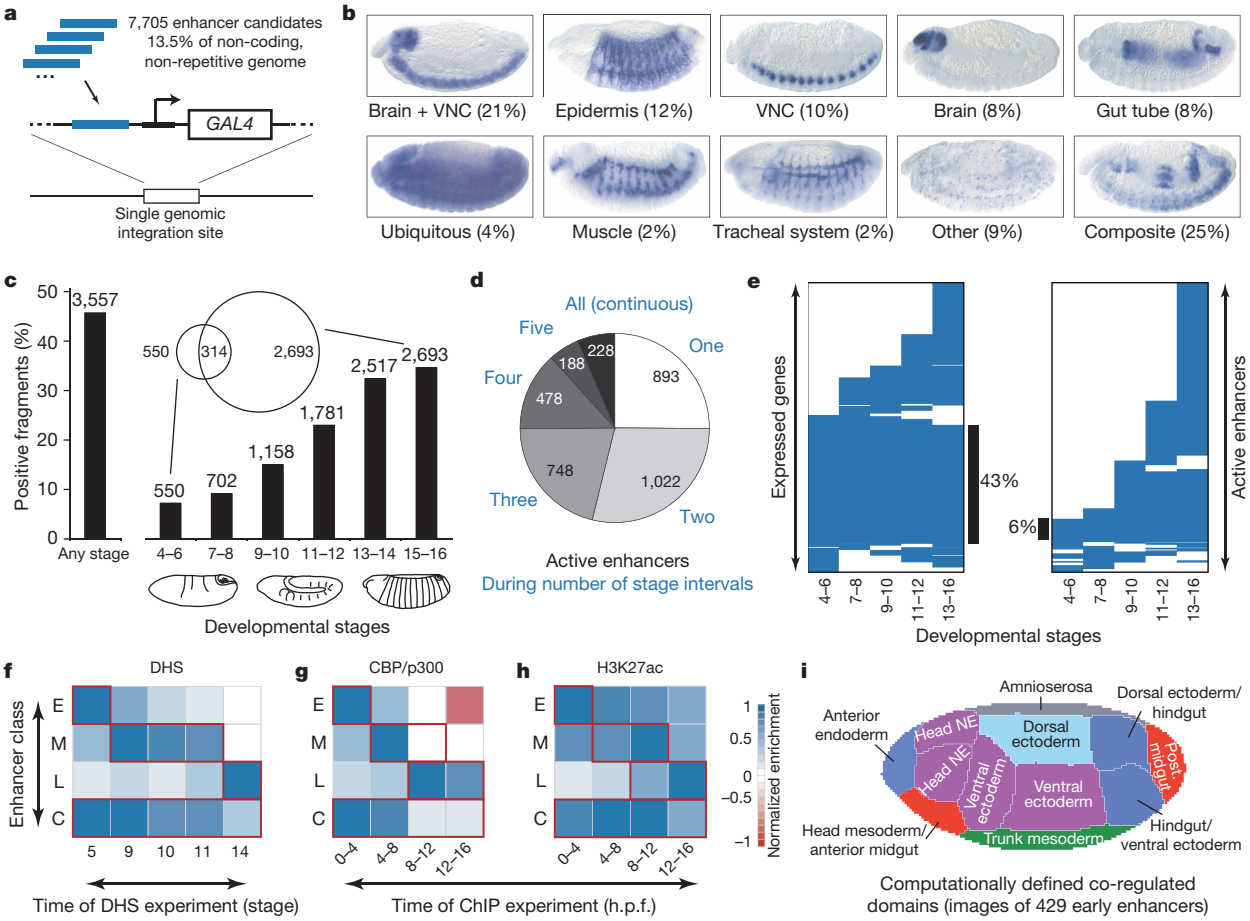

Figure 1 | Enhancers display highly diverse and dynamic activity patterns across *Drosophila* development. **a**, The VT library comprises transgenic flies with candidate fragments (blue) upstream of a transcriptional reporter (middle) in a constant genomic landing site (Extended Data Fig. 1). **b**, Proportion of enhancer activities in prominent tissues at stages 13–14 (representative embryos; Extended Data Figs 3 and 4). VNC, ventral nerve cord. **c**, The number of active enhancers increases during embryogenesis, with some overlap between early and late enhancers (Venn diagram in **c**). **d**, 3,329 (94%) embryonic enhancers are only transiently active. **e**, Temporal dynamics of gene expression (left, 5,134 genes[14]) and enhancer activity (right, 3,557 enhancers; black vertical lines indicate continuous expression or activity; Extended Data Fig. 5a). **f**–**h**, Heatmaps show the median enrichment of DNA accessibility[20] (**f**), CBP/p300 binding (**g**) and H3K27 acetylation (ac) marks[2] (**h**) on early (E), middle (M), late (L) and continuous (C) enhancers (rows) for experiments performed at different time points during *Drosophila* development (columns; red highlights coinciding time points; Extended Data Fig. 6). **i**, Co-regulated domains defined by reverse clustering of raw image data for 429 early enhancers resemble the embryo fate map[17] (Extended Data Fig. 4c). ChIP, chromatin immunoprecipitation; NE, neuroectoderm.

demonstrate that enhancer activities are consistent with the progression of development along defined cell lineages, highlighting the gene regulatory basis of development.

To analyse the locations of enhancers with respect to their putative target genes, we assigned enhancers to genes by manually matching enhancer activity and gene expression patterns (Figs 2a and 3a). For the 874 enhancers with the strongest activity patterns, we considered 3,681 genes within five genes up- and downstream of each enhancer (including host genes for intronic enhancers), that is, 9,293 enhancer–gene pairs. For 4,224 of these pairs (45%; 1,690 genes), expression patterns were available, resulting in 482 enhancer-to-gene assignments (of the enhancers for which all neighbouring genes were characterized, 82% could be assigned; Extended Data Fig. 2f and Supplementary Table 4). The assignments were supported by the location of chromosomal domain boundaries[21], binding sites of insulator proteins[22] and evolutionary chromosome breakpoints[23], all of which were depleted between enhancers and their assigned targets (Fig. 2b–d, Extended Data Fig. 7a–c and Supplementary Information section 3). Twenty-eight enhancers were assigned to and potentially regulate two genes, 23 of which were paralogues with very similar expression patterns (Supplementary Information section 4). During stages 4–6, 16 genes were assigned to enhancers with overlapping or identical activities reminiscent of shadow enhancers[24]. This is a considerable fraction (14%) among all 116 genes with multiple enhancers, in particular for developmental regulators (14 out of 16 genes are transcription factors; Supplementary Information section 5).

Along the linear genomic DNA sequence, the distances between the enhancers and the TSSs of their assigned target genes varied greatly: although many such pairs were close (21% were <4 kb), the median distance was 10 kb, and 28% of all inferred regulatory interactions were distal (>20 kb), up to more than 100 kb (Fig. 2e). However, considering the location of genes, the vast majority (88%) of all enhancers were located in the vicinity of their targets (Fig. 2f). Nevertheless, 12% of all enhancers were assigned across intervening genes and appeared to skip one (8%) or more (4%) genes to regulate a distal gene (Fig. 2f), as found for a *Sex combs reduced* (*Scr*) enhancer that lies beyond the *fushi tarazu* (*ftz*) gene[25]. Interestingly, enhancers were located almost as frequently upstream (30%) as downstream (22%) of their target genes (for example, the *SoxNeuro* (*SoxN*) locus; Extended Data Fig. 8), suggesting that no particularly preferred relative enhancer location might exist.

Thirty-six per cent of the enhancers were intragenic and appeared to predominantly (79%) regulate their host genes, as exemplified by *Thrombospondin* (*Tsp*; Fig. 3a). However, 21% were assigned to a neighbouring gene instead (Fig. 3b), including an enhancer located inside the intron of *bric a brac 1* (*bab1*) that appears to activate *bab2* over a distance of

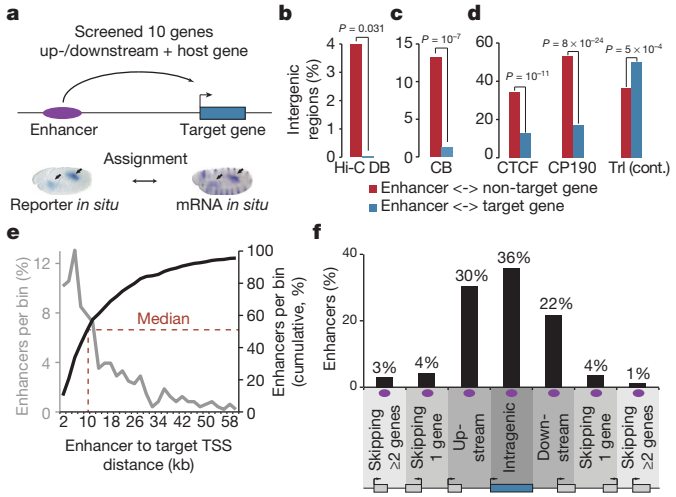

Figure 2 | The organization of the *Drosophila cis*-regulatory genome. **a**, Enhancer to target gene assignment based on enhancer activity and gene expression patterns. **b**–**d**, Chromosomal domain boundaries determined by Hi-C[21] (**b**), breakpoints during genome evolution[23] (**c**) and insulator binding sites[22] (**d**) show relative depletions between enhancers and their assigned target genes (blue) and enrichments between enhancers and non-targets (red), whereas the opposite is true for the activator Trl[22] (binomial *P* values are shown; see Extended Data Fig. 7a–c for additional insulator proteins and details). **e**, Genomic distances between enhancers and their assigned target gene TSSs in kb (grey, frequencies; black, cumulative). **f**, Frequency of enhancers (purple) at different genomic positions relative to their target genes (blue; schematic locus). Eighty-eight per cent of all enhancers are in the genes' genomic neighbourhoods within regulatory domains. CB, chromosomal breakpoints; DB, domain boundaries.

93 kb (Fig. 3c). *bab1* is not detectably expressed in the embryo during the corresponding developmental stage, which we found to be true more generally when intragenic enhancers regulated flanking genes rather than their host genes (Fig. 3d). Similarly, when intergenic enhancers were assigned to distal genes, the skipped genes were significantly less highly expressed than the target genes (Fig. 3d). Together these results support a predominantly local organization of the *Drosophila* genome into regulatory domains reminiscent of the chromosomal domains inferred from chromatin interactions[21].

The agreement of most enhancers' activities with the expression patterns of neighbouring genes (Figs 2f, 3a and Extended Data Figs 2f, 8) confirms that enhancer activity is predominantly context independent[9]. However, 18% of the enhancers could not be assigned to neighbouring genes (Extended Data Fig. 2f) and might be involved in more distal regulation (for example, ref. 26). For 19%, the activities were similar but broader and might thus be modulated in the endogenous sequence contexts in a more complex fashion (Supplementary Information section 1). Such context dependence is known for several loci in *Drosophila* (for example, the Hox locus[27]) and mouse (for example, *Fgf8* (ref. 28)), and enhancers in the bithorax complex indeed matched to gene expression patterns during early stages but appeared broader later (Extended Data Fig. 7d).

Many different enhancers showed similar or identical activity patterns in various embryonic tissues. For example, 263 were active throughout the central nervous system (CNS), 59 in midgut and 32 in macrophages (Extended Data Fig. 3), thus probably providing sufficient statistical power to discern predictive sequence signatures. Indeed, the motif content alone allowed the discrimination of enhancers from different functional classes using supervised machine learning in a cross-validated setting[29] (Extended Data Fig. 9a, b and Supplementary Table 5). The

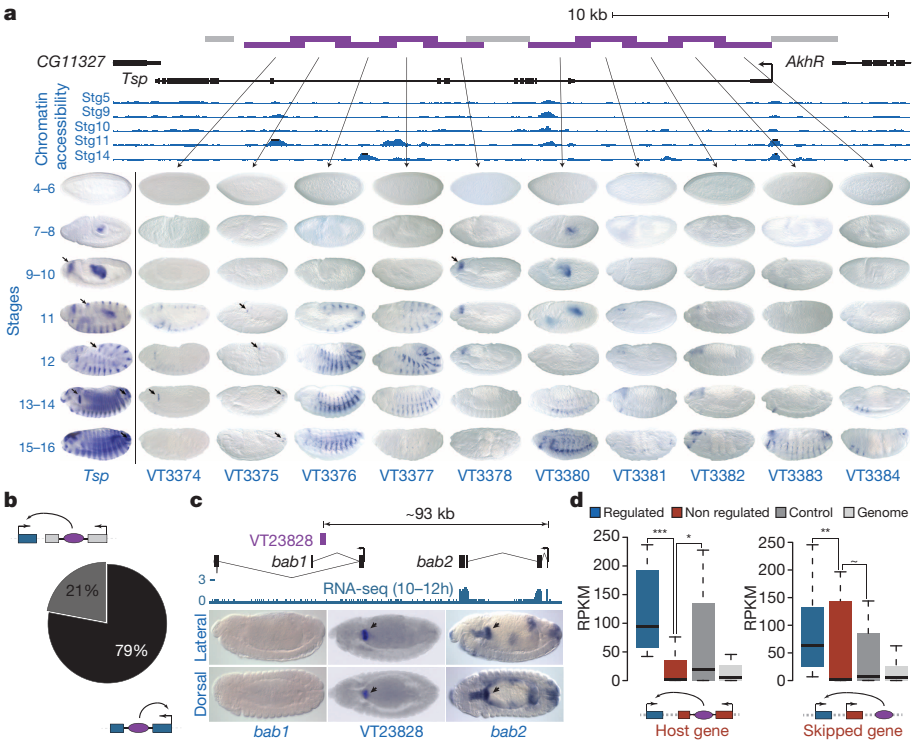

Figure 3 | Intragenic enhancers in the *Drosophila* genome. **a**, Enhancers in the *Tsp* locus. Top, UCSC Genome Browser screenshot including tested fragments (purple, positive; grey, negative) and DNA accessibility[20]. Bottom, embryos for all six time points of embryogenesis (left, *in situ* visualizing *Tsp* mRNA[14]; arrows highlight small expression/activity domains). **b**, Twenty-one per cent of intragenic enhancers are assigned to a neighbouring gene. **c**, A distal *bab2* enhancer (VT23828) in the intron of a neighbouring gene *bab1*. Top, UCSC Genome Browser screenshot including RNA-seq data for the corresponding stages[18]. Bottom, embryo images depicting the *bab1* and *bab2* expression during stages 13–14 (ref. 14) and VT23828's activity in the proventriculus (middle). **d**, Non-regulated host and skipped genes are often not expressed. Box plots show gene expression (reads per kilobase per million (RPKM)) values as measured by RNA-seq[18] for assigned target genes (blue) and non-regulated host genes (red, left) or skipped genes (red, right). Dark grey, unrelated neighbouring genes (control); light grey, all *D. melanogaster* genes. $***P = 10^{-8}$, $**P = 0.059$, $*P = 0.081$, $\sim P > 0.1$. Wilcoxon rank-sum test.

fraction of classes for which predictions were successful increased with the number of enhancers per class (Supplementary Table 5) but appeared to be independent of pattern complexity. This suggests that our understanding of regulatory sequences will benefit from the ongoing functional characterization of enhancers[4–7].

Different transcription factor motifs were strongly differentially distributed between the enhancer classes (Fig. 4a and Extended Data Fig. 9c). For example, early embryonic enhancers were enriched in motifs of the transcription factor Zelda, an important activator of embryonic gene expression[30]. Similarly, Twist (Twi) motifs were enriched in early mesodermal enhancers, Myocyte enhancing factor 2 (Mef2) motifs in late somatic muscle enhancers, and Pannier (Pnr) and Tinman (Tin) motifs in dorsal vessel enhancers, consistent with the established roles of the these transcription factors[8] (Fig. 4a and Extended Data Fig. 9c). To test whether predicted motifs are required for enhancer activity, we selected three midgut, four CNS and four anterior–posterior (A–P) enhancers (11 enhancers total), for which the successful predictions depended on GATA-like, Trithorax (Trl, also known as GAGA)-like, and Tramtrack (Ttk)-like motifs, respectively (Fig. 4a and Extended Data Fig. 9b, c). For each, we created reporter flies with an enhancer variant in which we disrupted the respective motifs by point mutations and compared the activity of the mutant and wild-type enhancers, both manually and by computational image analysis (Fig. 4b and Extended Data Fig. 10). In 10 out of 11 cases, the mutated enhancers were not active or had strongly reduced activity, validating the functional importance of the respective motifs.

Taken together this work complements efforts that study chromatin properties[2,19,20] or characterize enhancers at defined stages and in selected tissues[5–7,15,16]. Our results confirm and generalize principles and models from smaller scale studies (reviewed in refs 1, 9, 12) and suggest a high density of developmental enhancers in the *Drosophila* genome with an estimated total of ~41,000 enhancers or four enhancers per expressed protein-coding gene on average during embryogenesis alone. In addition, considering that enhancers that are exclusively active in larvae, pupae or the adult fly[5–7,15,16] (Supplementary Information section 6), we estimate between at least 50,000 to 100,000 developmental enhancers in the 170-megabase *D. melanogaster* genome. Even though the genome sequence properties (for example, repeat content and gene density) differ, this suggests that the 3-gigabase human genome could contain up to several million enhancers. In summary, the functional characterization of enhancers during the entire *Drosophila* embryogenesis adds a new level of functional annotation to the well-studied fly genome and elucidates global principles of *cis*-regulatory genome organization in animals, the importance of which for development, physiology, evolution and disease is becoming increasingly evident.

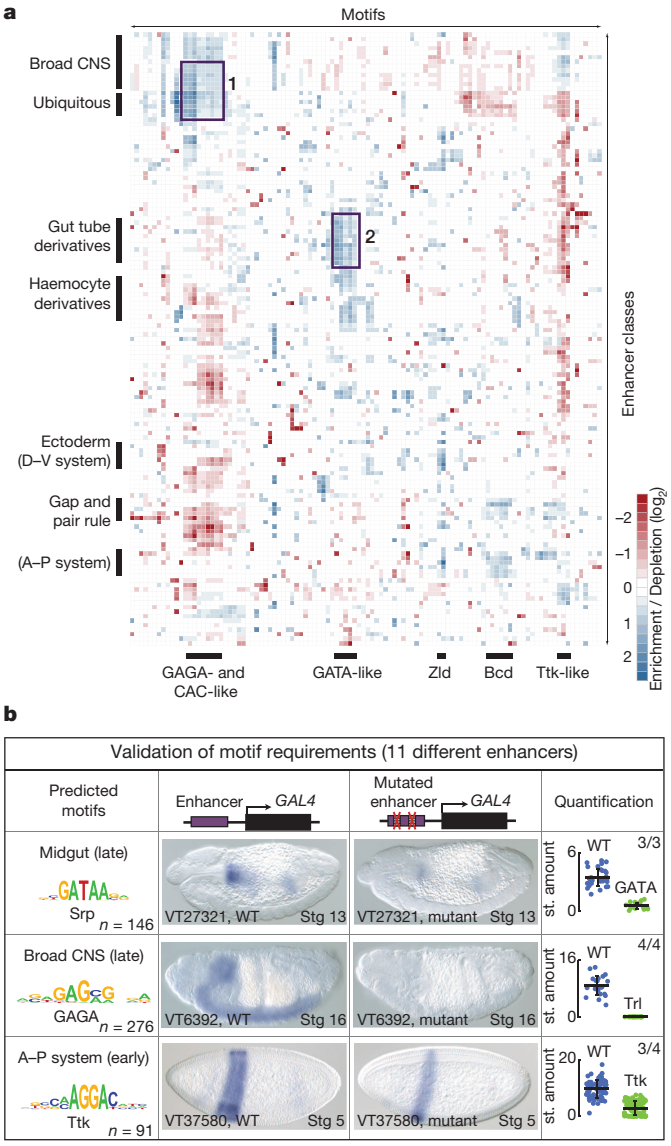

Figure 4 | Prediction and validation of *cis*-regulatory motif requirements for tissue-specific enhancer activities. **a**, Global *cis*-regulatory map of transcription factor motif enrichments in sequences of enhancers active in different tissues/cell types. Highlighted are Trl (GAGA) and CAC(N)$_N$CAC-like motifs enriched in CNS and ubiquitous enhancers (1) and GATA-like motifs enriched in midgut enhancers (2; see Extended Data Fig. 9c for the entire map). D–V, dorso–ventral. Zld, Zelda. **b**, Experimental validation of predicted *cis*-regulatory motif requirements. Shown are the most discriminative motifs (left), representative enhancers active in the midgut (stages 13–15), broad CNS (stages 15–16) and A–P system (stages 4–6) and their motif mutant variants (middle), and a quantification of the staining (st) intensities (right; all $P \leq 7 \times 10^{-10}$, Kolmogorov–Smirnov; see Extended Data Figs 9a, b and 10 for details and eight additional enhancers). WT, wild type.

METHODS SUMMARY

We assessed enhancer activities of 7,705 genomic fragments of about 2 kb in embryos of transgenic *GAL4*-reporter (VT) fly strains obtained from the VDRC (http://stockcenter.vdrc.at/) by *in situ* hybridization. Embryos of each VT strain were manually annotated with a controlled vocabulary and positive strains were imaged. Motif analyses and support vector machine (SVM) predictions were performed as described in ref. 29. All fragment coordinates and annotations are in Supplementary Table 1 and at http://enhancers.starklab.org/.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 January; accepted 17 April 2014.
Published online 1 June 2014.

1. Levine, M. Transcriptional enhancers in animal development and evolution. *Curr. Biol.* **20,** R754–R763 (2010).
2. The modENCODE Consortium. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **489,** 57–74 (2010).
3. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
4. Pennacchio, L. A. *et al. In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444,** 499–502 (2006).
5. Jenett, A. *et al.* A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2,** 991–1001 (2012).
6. Manning, L. *et al.* A resource for manipulating gene expression and analyzing *cis*-regulatory modules in the *Drosophila* CNS. *Cell Rep.* **2,** 1002–1013 (2012).
7. Jory, A. *et al.* A survey of 6,300 genomic fragments for *cis*-regulatory activity in the imaginal discs of *Drosophila melanogaster*. *Cell Rep.* **2,** 1014–1024 (2012).
8. Zinzen, R. P., Girardot, C., Gagneur, J., Braun, M. & Furlong, E. E. M. Combinatorial binding predicts spatio-temporal *cis*-regulatory activity. *Nature* **462,** 65–70 (2009).
9. Yáñez-Cuna, J. O., Kvon, E. Z. & Stark, A. Deciphering the transcriptional cis-regulatory code. *Trends Genet.* **29,** 11–22 (2013).
10. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457,** 854–858 (2009).

11. Shen, Y. *et al.* A map of the *cis*-regulatory sequences in the mouse genome. *Nature* **448,** 116–120 (2012).
12. Zeitlinger, J. & Stark, A. Developmental gene regulation in the era of genomics. *Dev. Biol.* **339,** 230–239 (2010).
13. Rubin, G. M. & Lewis, E. B. A brief history of *Drosophila*'s contributions to genome research. *Science* **287,** 2216–2218 (2000).
14. Tomancak, P. *et al.* Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol.* **3,** RESEARCH0088 (2002).
15. Gallo, S. M. *et al.* REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Res.* **39,** D118–D123 (2011).
16. Pfeiffer, B. D. *et al.* Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105,** 9715–9720 (2008).
17. Campos-Ortega, J. A. & Hartenstein, V. *The Embryonic Development of Drosophila melanogaster* (Springer, 1997).
18. Graveley, B. R. *et al.* The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471,** 473–479 (2011).
19. Bonn, S. *et al.* Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nature Genet.* **44,** 148–156 (2012).
20. Thomas, S. *et al.* Dynamic reprogramming of chromatin accessibility during *Drosophila* embryo development. *Genome Biol.* **12,** R43 (2011).
21. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148,** 458–472 (2012).
22. Nègre, N. *et al.* A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* **6,** e1000814 (2010).
23. Drosophila 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450,** 203–218 (2007).
24. Hong, J.-W., Hendrix, D. A. & Levine, M. S. Shadow enhancers as a source of evolutionary novelty. *Science* **321,** 1314 (2008).
25. Calhoun, V. C., Stathopoulos, A. & Levine, M. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila* Antennapedia complex. *Proc. Natl Acad. Sci. USA* **99,** 9243–9247 (2002).
26. Jack, J., Dorsett, D., Delotto, Y. & Liu, S. Expression of the cut locus in the *Drosophila* wing margin is required for cell type specification and is regulated by a distant enhancer. *Development* **113,** 735–747 (1991).
27. Maeda, R. K. & Karch, F. Gene expression in time and space: additive vs hierarchical organization of *cis*-regulatory regions. *Curr. Opin. Genet. Dev.* **21,** 187–193 (2011).
28. Marinić, M., Aktas, T., Ruf, S. & Spitz, F. An integrated holo-enhancer unit defines tissue and gene specificity of the *Fgf8* regulatory landscape. *Dev. Cell* **24,** 530–542 (2013).
29. Yáñez-Cuna, J. O., Dinh, H. Q., Kvon, E. Z., Shlyueva, D. & Stark, A. Uncovering *cis*-regulatory sequence requirements for context-specific transcription factor binding. *Genome Res.* **22,** 2018–2030 (2012).
30. Liang, H.-L. *et al.* The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature* **456,** 400–403 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank members of the Dickson laboratory VT project for cloning the candidate regions and generating transgenic flies and the VDRC (http://stockcenter.vdrc.at) for their maintenance and distribution. We are grateful to the IMP/IMBA Scientific Services, in particular BioOptics, Genomics, and IT for help and to C. H. Lampert (IST Austria) for advice. We thank M. Levine (UC Berkeley) and V. Hartenstein (UCLA) for their permission to reproduce figures. The Stark group is supported by a European Research Council (ERC) Starting Grant from the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 242922 awarded to A.S. and by the Austrian Science Fund (FWF, F4303-B09). Generation of transgenic lines was supported in part by an ERC Advanced Investigator Grant to B.J.D. Basic research at the IMP is supported by Boehringer Ingelheim GmbH.

Author Contributions E.Z.K. and A.S. conceived the study and wrote the paper. E.Z.K., G.S., M.P. and K.S. performed the screen, E.Z.K. annotated all patterns and performed validation experiments, E.Z.K. and T.K. performed the imaging, T.K. and G.S. analysed the image data, T.K. developed http://enhancers.starklab.org, J.O.Y.-C. performed all sequence analyses, B.J.D. provided transgenic flies prior to publication, E.Z.K., T.K., G.S., J.O.Y-C. and A.S. analysed data. A.S. supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.S. (stark@starklab.org).