LETTER

doi:10.1038/nature25461

Enhancer redundancy provides phenotypic robustness in mammalian development

Marco Osterwalder[1], Iros Barozzi[1], Virginie Tissières[2,3], Yoko Fukuda-Yuzawa[1], Brandon J. Mannion[1], Sarah Y. Afzal[1], Elizabeth A. Lee[1], Yiwen Zhu[1], Ingrid Plajzer-Frick[1], Catherine S. Pickle[1], Momoe Kato[1], Tyler H. Garvin[1], Quan T. Pham[1], Anne N. Harrington[1], Jennifer A. Akiyama[1], Veena Afzal[1], Javier Lopez-Rios[2,3], Diane E. Dickel[1], Axel Visel[1,4,5] & Len A. Pennacchio[1,4,6]

Distant-acting tissue-specific enhancers, which regulate gene expression, vastly outnumber protein-coding genes in mammalian genomes, but the functional importance of this regulatory complexity remains unclear[1,2]. Here we show that the pervasive presence of multiple enhancers with similar activities near the same gene confers phenotypic robustness to loss-of-function mutations in individual enhancers. We used genome editing to create 23 mouse deletion lines and inter-crosses, including both single and combinatorial enhancer deletions at seven distinct loci required for limb development. Unexpectedly, none of the ten deletions of individual enhancers caused noticeable changes in limb morphology. By contrast, the removal of pairs of limb enhancers near the same gene resulted in discernible phenotypes, indicating that enhancers function redundantly in establishing normal morphology. In a genetic background sensitized by reduced baseline expression of the target gene, even single enhancer deletions caused limb abnormalities, suggesting that functional redundancy is conferred by additive effects of enhancers on gene expression levels. A genome-wide analysis integrating epigenomic and transcriptomic data from 29 developmental mouse tissues revealed that mammalian genes are very commonly associated with multiple enhancers that have similar spatiotemporal activity. Systematic exploration of three representative developmental structures (limb, brain and heart) uncovered more than one thousand cases in which five or more enhancers with redundant activity patterns were found near the same gene. Together, our data indicate that enhancer redundancy is a remarkably widespread feature of mammalian genomes that provides an effective regulatory buffer to prevent deleterious phenotypic consequences upon the loss of individual enhancers.

Enhancers are a principal class of *cis*-regulatory elements that orchestrate precise gene expression patterns, which are essential for numerous processes including embryonic development[2]. They are now routinely predicted by genome-wide chromatin profiling methods, which identify positions of open chromatin or enhancer-associated histone marks[3]. Enhancers predicted by these high-throughput approaches outnumber genes by approximately an order of magnitude[1], raising the question of their functional significance. In particular, it remains unclear whether mammalian enhancers typically regulate complementary spatiotemporal aspects of gene expression in an additive fashion[4–7], or if this regulatory complexity more commonly results in functional redundancy among enhancers associated with the same gene[8–10].

Using the developing limb as a model for gene regulation during morphogenetic processes[11,12], we investigated the functional importance of enhancers *in vivo*. We used CRISPR–Cas9 genome editing to

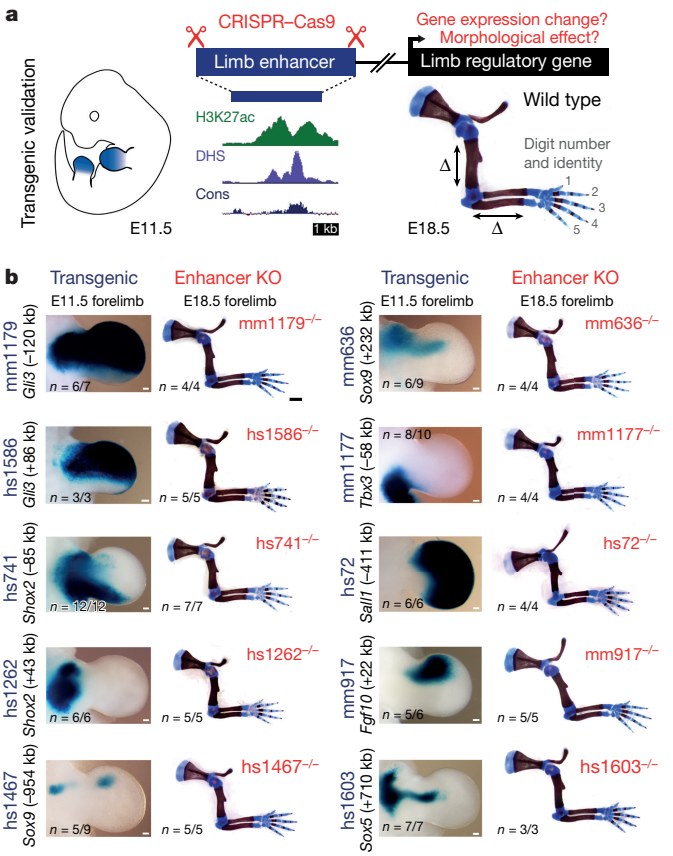


Figure 1 | Lack of limb morphological abnormalities in ten enhancer deletion lines. a, All selected enhancers are active in the limb mesenchyme (blue shading) at E11.5, are marked by epigenomic H3K27 acetylation and DNase I hypersensitivity (DHS) at E11.5, and contain a conserved core sequence (Cons). Target gene expression and limb morphology were assessed following deletion of individual enhancers (Extended Data Fig. 1a–j). **b**, None of the individual enhancer deletions caused obvious defects in the structure of skeletal elements. Enhancers are identified by VISTA ID numbers. Enhancer activities (left, E11.5) and forelimb skeletons of enhancer knockout (KO) embryos (right, E18.5) are shown (see Extended Data Fig. 3 for wild-type controls). Predicted target gene and enhancer distance (+, downstream; −, upstream) from the transcriptional start site (TSS) are indicated. *n* represents independent biological replicates with similar results. Scale bars, 100 μm (E11.5), 1 mm (E18.5).

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA. [2]Department of Biomedicine, University of Basel, 4058 Basel, Switzerland. [3]Centro Andaluz de Biología del Desarrollo, CSIC/JA/Universidad Pablo de Olavide, 41013 Seville, Spain. [4]US Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA. [5]School of Natural Sciences, University of California, Merced, California 95343, USA. [6]Comparative Biochemistry Program, University of California, Berkeley, California 94720, USA.

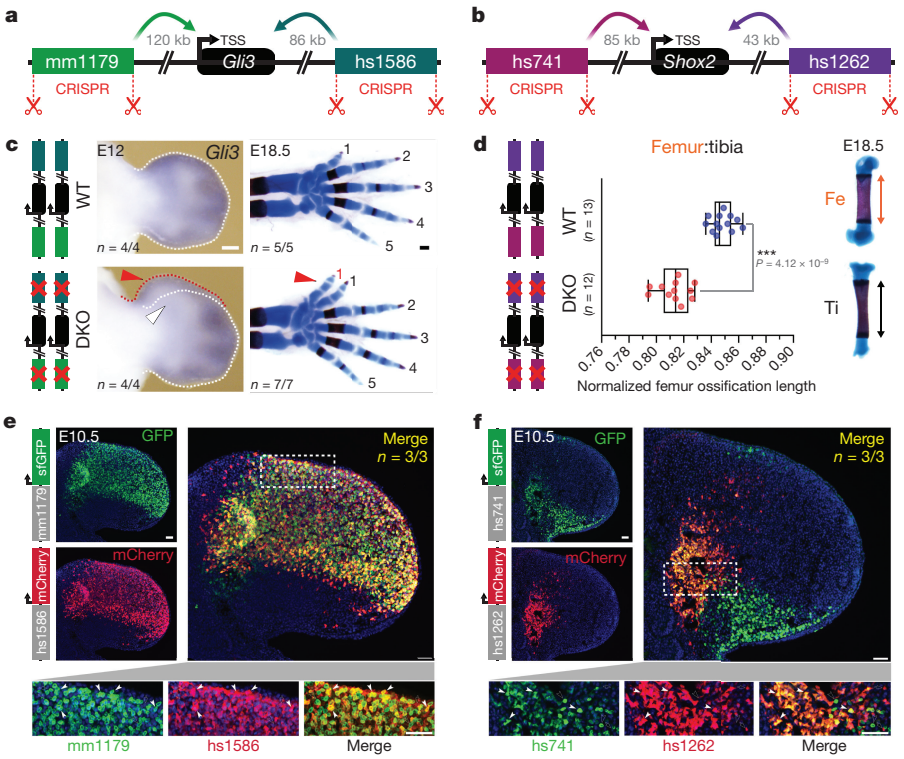


Figure 2 | Pairwise loss of limb enhancers with overlapping activities results in morphological abnormalities. **a**, **b**, CRISPR-deleted enhancers and their distance to the TSSs of predicted target genes (*Gli3*, *Shox2*). **c**, Left, RNA *in situ* hybridization reveals reduced *Gli3* expression in anterior hand plates of mm1179/hs1586 DKO embryos (white arrowhead). Red arrowhead, local expansion of anterior mesenchyme, a hallmark of *Gli3* deficiency. Right, forelimb skeletons with digits labelled 1 to 5, from anterior to posterior. DKO embryos exhibit duplication of digit 1 (red arrowhead). Scale bars, 200 μm. WT, wild type. **d**, Reduced femur ossification length in hs741/hs1262 DKO embryos (normalized to tibia ossification length). Box plot indicates median, interquartile values, range and individual biological replicates. ***$P < 0.001$ (two-tailed, unpaired *t*-test). **e**, **f**, Co-localization of *Gli3* (**e**; mm1179, green; hs1586, red) and *Shox2* (**f**; hs741, green; hs1262, red) enhancer activities via enhancer-reporter transgenes and immunofluorescence in forelimb buds of double transgenic embryos. White arrowheads indicate examples of double-positive cells. Empty arrowheads or arrows indicate cells marked by single enhancers. Nuclei are stained blue. Scale bars, 50 μm. *n* represents independent biological replicates with similar results.

individually delete ten embryonic enhancers, each with strong evolutionary conservation and robust limb activity in transgenic mouse reporter assays[13–17] (VISTA Enhancer Browser: https://enhancer.lbl.gov/) (Fig. 1a, Extended Data Fig. 1a–j and Supplementary Table 1). Each enhancer (identified by VISTA ID number) is located in the vicinity of a gene associated with human congenital limb malformations, and deletion of these genes in mice results in limb phenotypes ranging from polydactyly (*Gli3*) to complete loss of limbs (*Fgf10*) (Extended Data Fig. 1 and Supplementary Table 2). In all cases, the limb activity pattern of the enhancer at embryonic day 11.5 (E11.5) overlaps spatial RNA expression of the associated target gene, suggesting that these enhancers are part of the regulatory architecture that controls the expression of these genes[16–21] (Extended Data Fig. 2). Capture-C chromatin conformation data from embryonic limbs[22] confirmed that at least six of these enhancers physically interacted with their predicted target genes (Extended Data Fig. 1k). This framework enabled us to investigate the functional contribution of each enhancer by comparing the potential limb skeletal abnormalities caused by enhancer loss to the phenotypes observed in gene knockout mice.

Unexpectedly, we did not detect any abnormalities in bone number, shape, length, position or mineralization in mice in which any of the ten single enhancers was deleted (Fig. 1b and Extended Data Fig. 3). Similarly, we observed neither significant differences in predicted target gene expression in embryonic limbs for nine out of ten individual enhancer deletions, nor obvious changes in local H3K27ac (acetylation of lysine 27 on histone H3) signatures outside the deleted enhancers (Extended Data Figs 2, 4). Together, these results suggest that a substantial proportion of limb enhancers, even if highly conserved in evolution, are not individually essential for normal limb morphogenesis.

One possible explanation for the lack of an obvious phenotype in individual limb enhancer knockout lines is that different enhancers associated with the same gene may have spatiotemporally redundant, rather than unique, activity. Our selected panel of enhancers (Fig. 1b and Extended Data Fig. 1a–j) included three enhancer pairs with overlapping limb activity domains and the same predicted target gene (mm1179–hs1586, hs741–hs1262, and hs1467–mm636; Extended Data Fig. 5a–c). Using iterative CRISPR–Cas9 genome editing, we generated double enhancer knockout (DKO) mice for each enhancer pair (Extended Data Fig. 5a–d, g, j), such that both deletions occurred in *cis*. In two out of three cases, involving enhancer pairs near *Gli3* and *Shox2*, homozygous DKO embryos showed phenotypic abnormalities affecting skeletal limb morphology (Fig. 2a–d and Extended Data Fig. 5f, i, j). Mice lacking both enhancers near *Gli3* (mm1179 and hs1586) had substantially reduced *Gli3* expression in the embryonic hand plate and exhibited forelimb-specific polydactyly (Fig. 2c and Extended Data Fig. 5e, f), a phenotypic hallmark of diminished *Gli3* expression[23,24]. In addition, combined deletion of the two enhancers near *Shox2* (hs741 and hs1262) reduced *Shox2* expression, predominantly in embryonic hindlimbs, and resulted in a marked reduction in femur ossification (Fig. 2d and Extended Data Fig. 5h, i), consistent with the stylopod reductions observed when the *Shox2* gene is inactivated[18,25]. Together, these results show that although each of the four enhancers near *Gli3* and *Shox2* is individually dispensable for limb morphology, the respective pairs of enhancers are collectively required for normal limb development.

To examine the degree of overlap between the activity patterns of phenotypically redundant enhancers at the cellular level, we generated transgenic mouse lines expressing fluorescent reporters under the

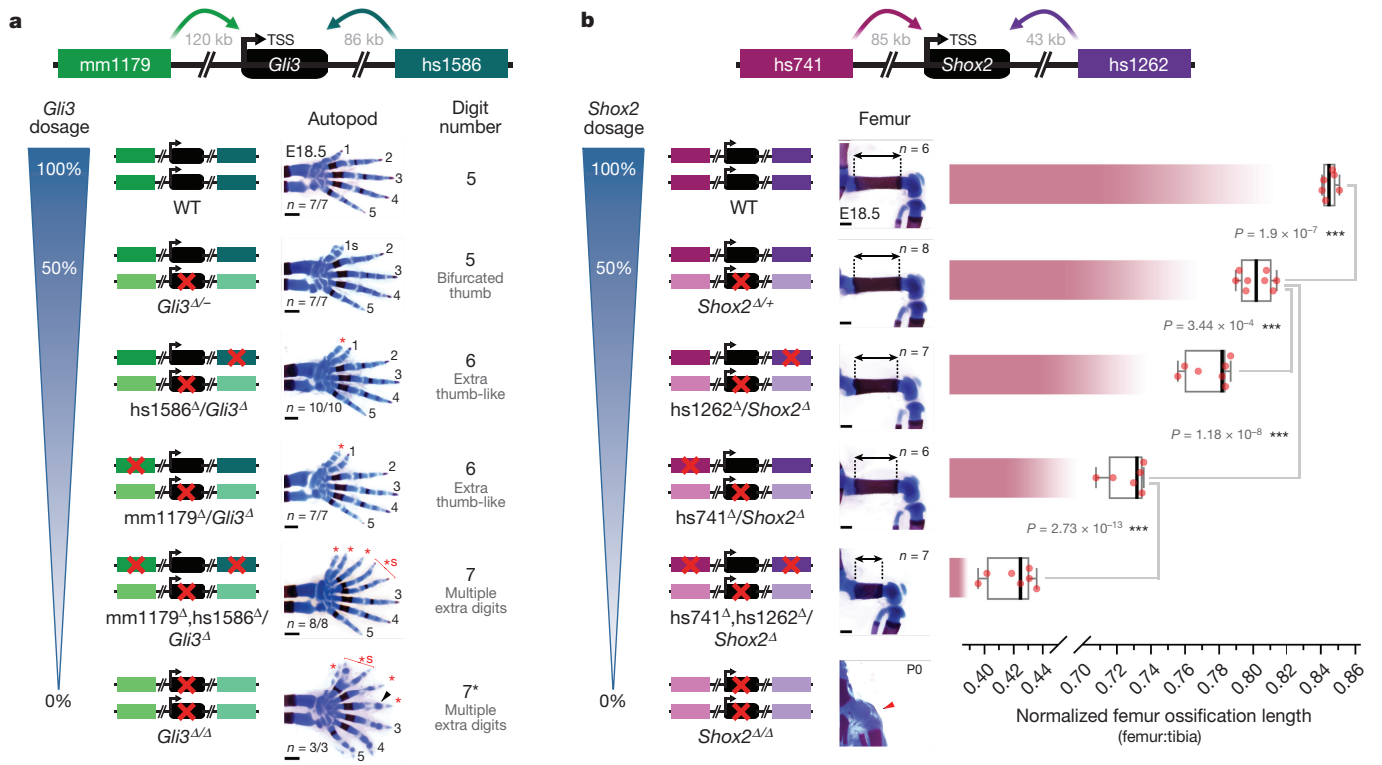


Figure 3 | Normally dispensable individual enhancers are required for limb morphology in a sensitized background. Individual and combined enhancer deletions in the presence of only one copy of the *Gli3* (**a**) or *Shox2* (**b**) target genes and the resulting limb morphology at E18.5. Wedges indicate inferred gene dosage. **a**, Skeletal forelimb autopod phenotypes at E18.5 resulting from mm1179 and hs1586 enhancer deletions in the presence of reduced *Gli3* dosage. 1–5, normal digits. Red asterisk, extra digits with unclear identity. *s, 'split' digit. Black asterisk and arrowhead indicate the presence of hypoplastic distal phalanges. **b**, Progressive reduction in femur ossification length (double arrows) due to hs741 and hs1262 enhancer loss in a *Shox2*-sensitized background. The relative length of femur ossification, normalized to the tibia ossification length, is shown. For comparison, the bottom panel shows absence of femur ossification in *Shox2*-deficient limbs at P0 (red arrowhead, reproduced with permission from ref. 25). *n* represents number of independent biological replicates with similar results. Box plots indicate median, interquartile values, range and individual biological replicates. ***$P < 0.001$ (two-tailed, unpaired *t*-test). Scale bars, 500 μm.

control of each of the *Gli3* or *Shox2* enhancers (mm1179–GFP, hs1586–mCherry, hs741–GFP and hs1262–mCherry). Using immunofluorescence on limb sections from double transgenic embryos, we tracked the activity of each of the four enhancers during limb development (Fig. 2e, f and Extended Data Fig. 6). Consistent with the preaxial polydactyly observed in *Gli3* DKO embryos, limb progenitor cells marked by both *Gli3* enhancers were observed at high density in the anterior limb mesenchyme (Fig. 2e and Extended Data Fig. 6c, d). In *Shox2* double enhancer reporter embryos, a major accumulation of cells with dual *Shox2* enhancer activities is present in a proximal limb mesenchymal cell population known to harbour stylopod progenitors[12] (Fig. 2f). In conjunction with our deletion studies, these results illustrate the degree of functional overlap between pairs of enhancers near the same gene at the cellular level.

Considering the apparent contrast between the morphological redundancy of pairs of enhancers and the strong evolutionary conservation of each individual enhancer, we studied the phenotypic effect of single and combinatorial enhancer deletions in sensitized genetic backgrounds carrying heterozygous deletions of the presumptive target genes (Fig. 3). We used CRISPR–Cas9 to engineer *Gli3* and *Shox2* gene loss-of-function alleles, which recapitulated expected gene dosage reductions and previously published phenotypes (Extended Data Figs 7, 8). We then used these alleles to generate compound heterozygous mice harbouring one or more disrupted enhancers with a wild-type gene on one allele and a disrupted gene but wild-type enhancers on the other allele (Fig. 3). For *Gli3*, the absence of either enhancer (mm1179 or hs1586) in the presence of only one functional *Gli3* allele resulted in a supernumerary anterior digit (Fig. 3a and Extended Data Fig. 8a), which is more severe than the terminally bifurcated thumb observed in *Gli3* heterozygotes (Fig. 3a). Similarly, for *Shox2* the removal of either neighbouring enhancer (hs1262 or hs741) in combination with compound heterozygous deletion of the *Shox2* gene resulted in a more pronounced reduction in femur length than observed in *Shox2* heterozygotes (Fig. 3b). For both pairs of enhancers, compound heterozygous mice carrying deletions of both enhancers on one allele and a deletion of the gene on the other allele showed even more severe phenotypes. In the case of *Gli3,* loss of both enhancers over a *Gli3* null allele resulted in greatly reduced expression of *Gli3* (Extended Data Fig. 7b, c) and severe pre-axial polydactyly in forelimbs, similar in severity to homozygous loss of the *Gli3* gene[24] (Fig. 3a and Extended Data Fig. 8a). Likewise, compound heterozygous deletion of enhancers hs741 and hs1262 over a *Shox2* gene deletion strongly reduced *Shox2* expression (Extended Data Fig. 7e, f) and resulted in a severe reduction in femur length and substantial shortening of the humerus (Fig. 3b and Extended Data Fig. 8b, c), consistent with the phenotypes that result from homozygous *Shox2* gene loss[18,25]. Together, our data demonstrate that these developmental enhancers, although seemingly dispensable under non-sensitized conditions, show individual functional contributions to limb development under conditions of reduced genetic robustness.

The lack of phenotypic change upon deletion of individual enhancers, and the functional redundancy observed among enhancer pairs, raises the question of how commonly such redundancy occurs in mammalian gene regulatory landscapes. To explore this question systematically, we devised a genome-wide, correlation-based computational approach to estimate the number of enhancers that regulate each gene

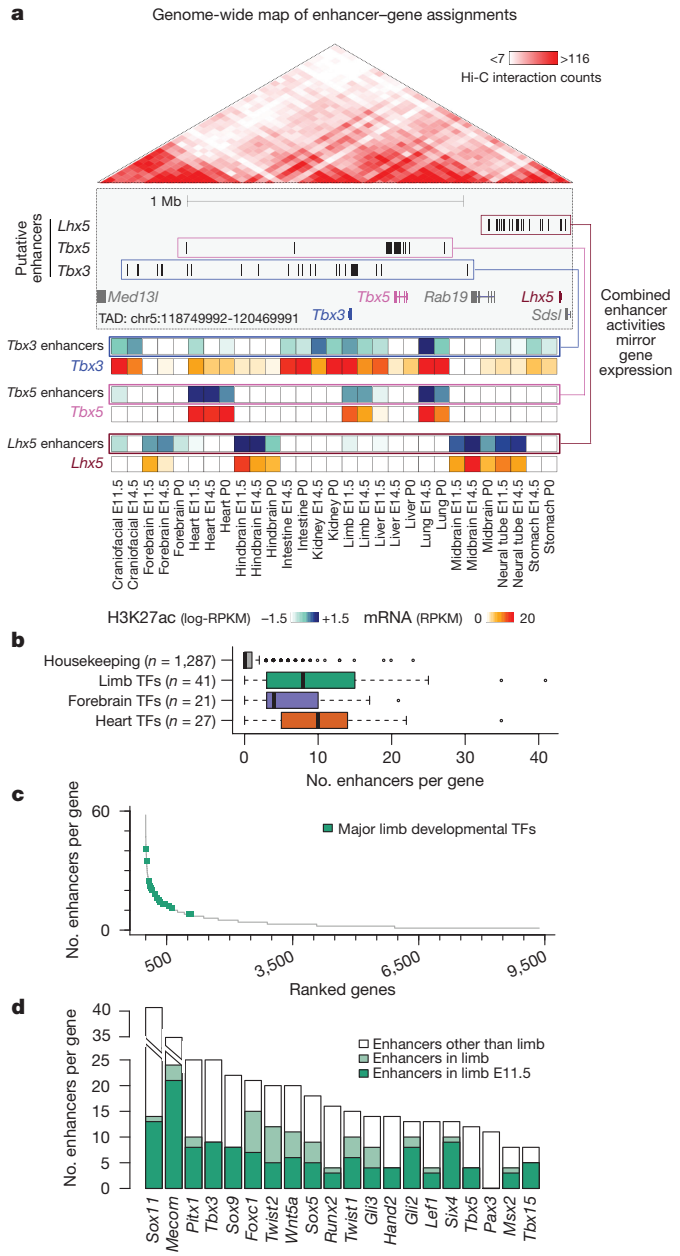


Figure 4 | Enhancers with redundant signatures are prevalent near developmental genes. a, Enhancer–gene assignments based on correlation of H3K27ac and mRNA profiles across a wide array of tissues (Extended Data Fig. 9a). Top, at an example locus encompassing *Tbx3*, *Tbx5*, and *Lhx5*, up to 25 enhancers are assigned to each of these three genes (blue, pink and brown boxes, Extended Data Fig. 9c). Genes showing fewer than five assigned enhancers are shown in grey. Bottom, heat maps showing meta-profiles of each gene's expression profile across tissues (red shades), along with the cumulative activity profile of its assigned enhancers (blue shades). **b**, Distribution of the number of enhancers assigned to developmental transcription factors (TFs) with biased expression in limb ($P = 5 \times 10^{-19}$ versus housekeeping), forebrain ($P = 8 \times 10^{-15}$), and heart ($P = 3 \times 10^{-25}$) (two-sided Mann–Whitney tests). Box plots show median, interquartile values, range, and outliers (individual points). **c**, Complete spectrum of genes with at least one assigned enhancer, sorted by decreasing enhancer number. Limb-biased transcription factors are highlighted in green. **d**, Total number of enhancers (in all tissues analysed) assigned to each transcription factor in **c**, with the number of assigned enhancers predicted specifically in limb at E11.5 (dark green) or any other stage analysed (light green).

during development, taking advantage of chromatin signatures of distal enhancers and gene transcription measured across multiple tissues and time points of mouse development (Fig. 4 and Extended Data Figs 9, 10). We analysed correlations between H3K27ac chromatin immunoprecipitation followed by sequencing (ChIP–seq) and RNA sequencing (RNA-seq) datasets from twelve different mouse tissues at two or three embryonic or perinatal time points per tissue (https://www.encodeproject.org/) to assign each enhancer to its most likely target gene within the same topologically associated domain (TAD)[26] (Fig. 4, Extended Data Fig. 9a–c and Methods). We then used this framework to examine the average number of enhancers associated with genes expressed in three developmental tissues (limb, heart, and forebrain). Genes with limb-biased expression showed a median of three associated distal enhancers, versus a median of zero for housekeeping genes (Extended Data Fig. 9d, e). For the specific class of limb-biased genes encoding transcription factors, we observed an even more complex enhancer landscape, with a median of eight distinct enhancers per gene (Fig. 4b). Notably, some of these transcription factor genes were associated with more than ten tissue-specific limb enhancers with highly overlapping activity patterns in the same tissue (Fig. 4c, d and Methods). We observed similarly large numbers of potentially redundant enhancers near brain- and heart-specific transcription factor genes (Extended Data Fig. 10a, b). Even under stringent correlation thresholds, our analysis uncovered 1,058 genes associated with five or more enhancers showing putatively redundant activity patterns—that is, enhancers that are active in the same tissue (Extended Data Fig. 10c–f). These results indicate that developmentally expressed genes are commonly associated with multiple enhancers that show overlapping activity patterns, supporting the widespread existence of functionally redundant enhancers in mammalian genomes.

Studies of individual loci have identified examples of mammalian enhancers near the same gene with remarkably similar spatiotemporal activity patterns or functions[15,27–32], reminiscent of invertebrate 'shadow enhancers'[8,9,33–35]. The lack of marked morphological phenotypes in our enhancer deletion mouse models suggests that panels of mammalian enhancers with large degrees of redundancy act as a regulatory buffer for key developmental processes, thereby reducing the likelihood of severe consequences resulting from genetic or environmental challenges[8]. Although individual examples of enhancers whose loss leads to severe phenotypes have been described[4,36], our findings suggest that redundancy is far more common. As indicated by the phenotypes observed in sensitized genetic backgrounds, our results suggest that pairs of enhancers act redundantly in organismal patterning, but additively in establishing gene expression levels. This observation is consistent with high-throughput loss-of-function screens in cultured cells, in which the disruption of individual enhancers leads to measurable gene expression changes but rarely results in the complete loss of target gene expression[37]. It appears plausible to assume that limited but specific contributions to overall gene expression levels are relevant for organismal fitness under specific pressures, thus subjecting enhancers to purifying selection over evolutionary time. Alternatively, additional tissue-specific functions may also explain the evolutionary constraints on these loci.

Our observations have implications for the interpretation of noncoding regulatory variants in relation to human phenotypes. Our findings suggest that many loss-of-function enhancer mutations will cause, at most, subtle phenotypes in humans. Thus, for many genetic loci, enhancer-associated disease phenotypes may be more likely to result from gain-of-function mutations that either expand enhancer activity[38] or alter the positions of enhancers relative to genes[39].

Online Content Methods, along with any additional 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 27 March; accepted 18 December 2017.
Published online 31 January 2018.

1. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
2. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167,** 1170–1187 (2016).
3. Andrey, G. & Mundlos, S. The three-dimensional genome: regulating gene expression during pluripotency and development. *Development* **144,** 3646–3658 (2017).
4. Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. & Shiroishi, T. Elimination of a long-range cis-regulatory module causes complete loss of limb-specific *Shh* expression and truncation of the mouse limb. *Development* **132,** 797–803 (2005).
5. Menke, D. B., Guenther, C. & Kingsley, D. M. Dual hindlimb control elements in the *Tbx4* gene and region-specific control of bone size in vertebrate limbs. *Development* **135,** 2543–2553 (2008).
6. Shim, S., Kwan, K. Y., Li, M., Lefebvre, V. & Sestan, N. *Cis*-regulatory control of corticospinal system development and evolution. *Nature* **486,** 74–79 (2012).
7. Hay, D. *et al.* Genetic dissection of the α-globin super-enhancer *in vivo*. *Nat. Genet.* **48,** 895–903 (2016).
8. Frankel, N. *et al.* Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466,** 490–493 (2010).
9. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow enhancers foster robustness of *Drosophila* gastrulation. *Curr. Biol.* **20,** 1562–1567 (2010).
10. Montavon, T. *et al.* A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147,** 1132–1145 (2011).
11. Petit, F., Sears, K. E. & Ahituv, N. Limb development: a paradigm of gene regulation. *Nat. Rev. Genet.* **18,** 245–258 (2017).
12. Zeller, R., López-Ríos, J. & Zuniga, A. Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nat. Rev. Genet.* **10,** 845–858 (2009).
13. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35,** D88–D92 (2007).
14. Pennacchio, L. A. *et al. In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444,** 499–502 (2006).
15. Attanasio, C. *et al.* Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342,** 1241006 (2013).
16. Osterwalder, M. *et al.* HAND2 targets define a network of transcriptional regulators that compartmentalize the early limb bud mesenchyme. *Dev. Cell* **31,** 345–357 (2014).
17. Rosin, J. M., Abassah-Oppong, S. & Cobb, J. Comparative transgenic analysis of enhancers from the human *SHOX* and mouse *Shox2* genomic regions. *Hum. Mol. Genet.* **22,** 3063–3076 (2013).
18. Cobb, J., Dierich, A., Huss-Garcia, Y. & Duboule, D. A mouse model for human short-stature syndromes identifies *Shox2* as an upstream regulator of *Runx2* during long-bone development. *Proc. Natl Acad. Sci. USA* **103,** 4511–4515 (2006).
19. Akiyama, H., Chaboissier, M. C., Martin, J. F., Schedl, A. & de Crombrugghe, B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of *Sox5* and *Sox6*. *Genes Dev.* **16,** 2813–2828 (2002).
20. Kawakami, Y. *et al.* Sall genes regulate region-specific morphogenesis in the mouse limb by modulating Hox activities. *Development* **136,** 585–594 (2009).
21. Min, H. *et al.* Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to *Drosophila* branchless. *Genes Dev.* **12,** 3156–3161 (1998).
22. Andrey, G. *et al.* Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. *Genome Res.* **27,** 223–233 (2017).
23. Hui, C. C. & Joyner, A. L. A mouse model of Greig cephalopolysyndactyly syndrome: the extra-toesJ mutation contains an intragenic deletion of the *Gli3* gene. *Nat. Genet.* **3,** 241–246 (1993).
24. Lopez-Rios, J. *et al.* GLI3 constrains digit number by controlling both progenitor proliferation and BMP-dependent exit to chondrogenesis. *Dev. Cell* **22,** 837–848 (2012).
25. Ye, W. *et al.* A unique stylopod patterning mechanism by *Shox2*-controlled osteogenesis. *Development* **143,** 2548–2560 (2016).
26. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376–380 (2012).
27. 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).
28. Visel, A. *et al.* A high-resolution enhancer atlas of the developing telencephalon. *Cell* **152,** 895–908 (2013).
29. Lam, D. D. *et al.* Partially redundant enhancers cooperatively maintain mammalian pomc expression above a critical functional threshold. *PLoS Genet.* **11,** e1004935 (2015).
30. Yao, Y. *et al. Cis*-regulatory architecture of a brain signaling center predates the origin of chordates. *Nat. Genet.* **48,** 575–580 (2016).
31. Antosova, B. *et al.* The gene regulatory network of lens induction is wired through Meis-dependent shadow enhancers of *Pax6*. *PLoS Genet.* **12,** e1006441 (2016).
32. Will, A. J. *et al.* Composition and dosage of a multipartite enhancer cluster control developmental expression of *Ihh* (Indian hedgehog). *Nat. Genet.* **49,** 1539–1545 (2017).
33. Hong, J. W., Hendrix, D. A. & Levine, M. S. Shadow enhancers as a source of evolutionary novelty. *Science* **321,** 1314 (2008).
34. Barolo, S. Shadow enhancers: frequently asked questions about distributed *cis*-regulatory information and enhancer redundancy. *BioEssays* **34,** 135–141 (2012).
35. Cannavò, E. *et al.* Shadow enhancers are pervasive features of developmental regulatory networks. *Curr. Biol.* **26,** 38–51 (2016).
36. Yanagisawa, H., Clouthier, D. E., Richardson, J. A., Charité, J. & Olson, E. N. Targeted deletion of a branchial arch-specific enhancer reveals a role of dHAND in craniofacial development. *Development* **130,** 1069–1078 (2003).
37. Fulco, C. P. *et al.* Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354,** 769–773 (2016).
38. Lettice, L. A., Hill, A. E., Devenney, P. S. & Hill, R. E. Point mutations in a distant sonic hedgehog *cis*-regulator generate a variable regulatory output responsible for preaxial polydactyly. *Hum. Mol. Genet.* **17,** 978–985 (2008).
39. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161,** 1012–1025 (2015).

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

Acknowledgements This work was supported by National Institutes of Health grants R01HG003988, U54HG006997, R24HL123879 and UM1HL098166 (to A.V. and L.A.P.) and the University of Basel and the Novartis Foundation for Biomedical Research (to J.L.-R.). M.O. was supported by a Swiss National Science Foundation (SNSF) fellowship. We thank B. Ren for providing access to the ChIP–seq and RNA-seq data from ENCODE; J. Doudna for providing a plasmid containing a human-optimized *Cas9* gene; W. Ye and Y. Chen for sharing the image of a *Shox2*-deficient limb skeleton (Fig. 3b); and the members of the L.A.P., A.V. and D.E.D. groups for technical advice and comments on the manuscript, in particular C. Spurrell and E. Kvon. Research was conducted at the E. O. Lawrence Berkeley National Laboratory and performed under Department of Energy Contract DE-AC02-05CH11231, University of California.

Author Contributions M.O., D.E.D., A.V., and L.A.P. conceived the study. M.O., D.E.D., B.J.M., S.Y.A., E.A.L., Y.Z., I.P.-F., C.S.P., M.K., T.H.G., Q.T.P., A.N.H., J.A.A., and V.A. performed the genome editing and mouse phenotyping studies. I.B. and M.O. devised the computational framework, and I.B. performed the correlative analysis. V.T. performed *in situ* hybridization under the supervision of J.L.-R. Y.F.-Y. conducted the ChIP–seq and RNA-seq data analysis. M.O., D.E.D., A.V., and L.A.P. wrote the manuscript with input from the remaining authors.

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. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to D.E.D. (dedickel@lbl.gov), A.V. (avisel@lbl.gov), or L.A.P. (lapennacchio@lbl.gov).

METHODS

Experimental design. All animal work was reviewed and approved by the Lawrence Berkeley National Laboratory (LBNL) Animal Welfare Committee. All mice used in this study were housed at the Animal Care Facility (ACF) at LBNL. Mice were monitored daily for food and water intake, and animals were inspected weekly by the Chair of the Animal Welfare and Research Committee and the head of the animal facility in consultation with the veterinary staff. The LBNL ACF is accredited by the American Association for the Accreditation of Laboratory Animal Care (AAALAC). Transgenic mouse assays and enhancer knockouts were performed in *Mus musculus* FVB strain mice. The following developmental stages were used in this study: embryonic day E10.5, E11.5, E12.5 and E18.5 mice. Animals of both sexes were used in the analysis. Sample size selection and randomization strategies were conducted as described below.

Transgenic mouse assay selection and randomization. Sample sizes were selected empirically on the basis of our previous experience of performing transgenic mouse assays for more than 2,000 total putative enhancers (VISTA Enhancer Browser: https://enhancer.lbl.gov/). Mouse embryos were excluded from further analysis if they did not contain the reporter transgene or if the developmental stage was not correct. All transgenic mice were treated with identical experimental conditions. Randomization and experimenter blinding were unnecessary and not performed.

Enhancer knockout selection and randomization. Sample sizes were selected empirically on the basis of our previous studies[15]. All phenotypic characterization of knockout mice used a matched littermate selection strategy. All phenotyped mice described in the paper resulted from crossing heterozygous enhancer deletion mice together to allow the comparison of matched littermates of different genotypes. Embryonic samples used for *in situ* hybridizations, RNA-seq, and skeletal preparations were dissected blinded to genotype.

***In vivo* transgenic reporter assays.** Enhancer names in this study are the unique identifiers used in the VISTA Enhancer Browser (https://enhancer.lbl.gov/; mm: originally identified in mouse; hs: originally identified in human). Transgenic results for most enhancers have been reported previously[13–16]. Newly tested enhancers (hs1586 at E10.5 and hs1262) were amplified from human genomic DNA and cloned into an hsp68-*lacZ* expression vector as previously described[14]. Genomic coordinates of all enhancers are listed in Supplementary Table 1. LacZ transgenic mouse assays were conducted as previously described[14,40]. To directly compare the activity domains between apparently redundant enhancers, enhancers were cloned, using Gateway (Thermo Fisher Scientific) or Gibson[41] methods, into an hsp68-based reporter vector similar to that described above, with the exception of a fluorescent reporter replacing *LacZ*. The enhancer–reporter combinations were generated as follows: mm1179–sfGFP, hs1586–mCherry, hs741–sfGFP and hs1262–mCherry. sfGFP is a fusion of Sun1 and 2×sfGFP as described[42] and localizes to the nuclear membrane. Mice carrying the individual fluorescent reporter transgenes were then generated via pronuclear injection (using FVB strain zygotes), and stable lines were established from founders showing reproducible reporter activity in the embryonic limb.

Generation of enhancer knockout mice using CRISPR–Cas9. Mouse strains lacking limb enhancer(s) or harbouring gene loss-of-function alleles were generated using *in vivo* CRISPR–Cas9 editing, as previously described, with only minor modifications[43,44]. Pairs of single guide RNAs (sgRNAs) targeting genomic sequence 5′ and 3′ to the sequence to be deleted were designed using CHOPCHOP[45] (see Supplementary Table 1 for sgRNA sequences and coordinates of deleted regions). Knockout mice were engineered as described previously[46] using a mix containing Cas9 mRNA (final concentration of 100 ng/μl) and two sgRNAs (25 ng/μl each) in injection buffer (10 mM Tris, pH 7.5; 0.1 mM EDTA). This mix was injected into the cytoplasm of single-cell FVB strain mouse embryos. Founder (F0) mice were genotyped using PCR with High Fidelity Platinum Taq Polymerase (Thermo Fisher Scientific) to identify those with the desired non-homologous end joining (NHEJ)-generated deletion breakpoints (see Extended Data Figs 1a–j, 5a–c, 7a, d and Supplementary Table 3 for genotyping strategy, primer sequences and PCR amplicons). Sanger sequencing was used to identify and confirm deletion breakpoints in F0 and F1 mice (Extended Data Figs 1a–j, 5a–c, 7a, d). Unless noted otherwise, mice homozygous-null for the targeted limb enhancers showed normal pre- and postnatal viability and appeared outwardly normal. For iterative CRISPR–Cas9 genome editing, fertilized mouse eggs harbouring the primary deletion were collected and injected with sgRNAs targeting the secondary enhancer for deletion. Only those founder lines harbouring both deletions on the same haplotype were analysed further.

***In situ* hybridization and skeletal preparations.** To assess spatial changes in gene expression in mouse embryonic limbs, whole mount *in situ* hybridization using digoxigenin-labelled antisense riboprobes was carried out as previously described[46]. Forelimbs and hindlimbs from at least three independent embryos were analysed for each genotype (including wild-type littermate controls). Mouse embryonic skeletons at E18.5 were stained with Alcian blue and Alizarin red to differentiate cartilage (blue) and bone (red) using standard methods[47]. For comparison of limb skeletons from enhancer knockout embryos and wild-type littermates, general parameters such as bone number, shape, length, position or mineralization were assessed. Embryonic limbs and limb skeletons were imaged, and skeletal elements were measured, using a Leica MZ16 stereo-microscope coupled to a Leica DFC300Fx or DFC420 digital camera. Brightness and contrast were adjusted uniformly using Photoshop CS5. Measurements of the ossified portions of humerus and femur (stylopodial elements) were normalized to those of the ulna and tibia (related zeugopodial elements), respectively (as shown in Figs 2d, 3b and Extended Data Figs 5i, 8c).

Quantitative real-time PCR (qPCR) and RNA-seq. RNA was isolated from microdissected forelimbs or hindlimbs of mouse embryos at E11.5 using the Ambion RNAqueous Total RNA Isolation Kit (Life Technologies) according to the manufacturer's instructions. For qPCR, RNA was treated with RNase-free DNase (Promega) and reverse transcribed using SuperScript III (Life Technologies) with random hexamer or poly-dT priming according to the manufacturer's instructions. qPCR was performed on a LightCycler 480 (Roche) using KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) according to the manufacturer's instructions. qPCR primers (listed in Supplementary Table 4) were designed *in silico* using Primer3 (http://primer3.wi.mit.edu/), and amplicons span exon–exon junctions in order to prevent amplification of genomic DNA. Relative gene expression levels were calculated using the $2^{-\Delta\Delta C_t}$ method[48], normalized to the *Actb* housekeeping gene, and the mean of wild-type control samples was set to 1.

For RNA-seq, RNA samples were treated with DNase (TURBO DNA-free Kit, Life Technologies), and RNA quality was verified using a 2100 Bioanalyzer (Agilent) with an RNA 6000 Nano Kit (Agilent). RNA-seq libraries were generated using the TruSeq Stranded mRNA Sample Prep Kit (Illumina), following the manufacturer's instructions, and purified, eluted, and quantified as described previously[49]. RNA-seq libraries were pooled (four per lane) and sequenced using single end 50-bp reads on a HiSeq 4000 (Illumina).

Immunofluorescence. Mouse embryonic limbs at E10.5, E11.5 or E12.5 were dissected in cold PBS and fixed in 4% PFA for 2–3 h. Following incubation in a sucrose gradient and embedding in a 1:1 mixture of 30% sucrose and optimum cutting temperature solution, sagittal 10-μm frozen sections were cut using a cryostat. Cryosections were incubated overnight with the following primary antibodies: chicken anti-GFP (1:500, Thermo Fisher Scientific, A10262), rabbit anti-mCherry (1:1,000, Thermo Fisher Scientific, PA5-34974) and goat anti-Sox9 (1:500, R&D Systems, AF3075). Goat-anti chicken, goat anti-rabbit and donkey anti-goat secondary antibodies conjugated to Alexa Fluor 488, 568, 594 or 647 (1:1,000, Thermo Fisher Scientific) were used for detection. Hoechst 33258 (Sigma-Aldrich) was used to counterstain nuclei. Fluorescent images were acquired using a Zeiss AxioImager fluorescence microscope in combination with a Hamamatsu Orca-03 camera. Brightness and contrast were adjusted uniformly using Photoshop CS5.

ChIP–seq. For each of six single enhancer knockout lines, ChIP–seq to H3K27ac was performed using a protocol optimized for mouse embryonic tissues[50]. In brief, forelimb buds from ten wild-type embryos (four biological replicates) and ten enhancer knockout embryos (at least two biological replicates) were dissected at E11.5, formaldehyde crosslinked, and sheared using a Diagenode Bioruptor Sonicator. After pre-clearing, chromatin was incubated with anti-H3K27ac antibody (Active Motif cat no. 39133) for 2 h at 4 °C. Freshly rinsed Dynabeads (1:1 protein A:protein G mix) were then added to the antibody-treated chromatin, and immunoprecipitation was performed on a rotator for 30 min at 4 °C. Libraries were prepared using the Illumina Truseq DNA sample prep kit following the manufacturer's instructions with minor modifications. Library quality was assessed using a 2100 Bioanalyzer with the High Sensitivity DNA Kit (Agilent), and quantification was performed using a Qubit Fluorometer with the dsDNA HS Assay Kit (Life Technologies). ChIP–seq and input libraries were pooled and sequenced via single-end 50-bp reads on a HiSeq 2000 or 4000 (Illumina).

RNA-seq and ChIP–seq analysis. Analysis of ChIP–seq and RNA-seq data from limb enhancer knockout and related wild-type control samples was performed as follows: CASAVA v1.8.0 (Illumina) was used to demultiplex data, and reads with CASAVA 'Y' flag (purity filtering) were discarded. For each sample, between 12 million and 55 million (ChIP–seq) or 23 million and 71 million (RNA-seq) reads were obtained following quality filtering and adaptor trimming using cutadapt_v1.1 (https://cutadapt.readthedocs.io/) with parameter '-m 25 -q 25'. Mouse genome sequence (mm9) and gene annotations were retrieved from the iGenomes repository (https://support.illumina.com/sequencing/sequencing_software/igenome.html).

To align the RNA-seq reads to the mouse reference genome and transcriptome, we used Tophat v2.0.6[51], and the reads mapping to UCSC known genes were counted by HTSeq[52]. Genes with counts per million (CPM) >1 in at least two samples were processed for further differential gene expression analysis comparing enhancer knockout and wild-type control samples using edgeR[53]. In each case, the

top 100 differentially expressed genes, sorted by false discovery rate (FDR), are listed in Supplementary Tables 5–7.

For read mapping and peak calling of ChIP–seq datasets, bowtie[54] (version 0.12.8) with parameter '-m 1 -v 2' and MACS[55] (version 1.4.2) with parameter '-mfold = 10,30 -nomodel -p 0.0001' were used, respectively. Biological replicates were combined using MSPC[56], with the following parameters: -r biological -s 1E-10 -W 1E-6 -m Highest -c 2. The predicted enhancer intervals were assigned the best *P* value (as defined by MACS[55]) among the overlapping peaks.

ENCODE ChIP–seq data analysis. Raw data were downloaded from the Data Coordination Center of the ENCODE project (https://www.encodeproject.org/, see Supplementary Table 8 for the complete list of sample identifiers). Short reads were aligned to the mm10 assembly of the mouse genome using bowtie[54], with the following parameters: -a -m 1 -n 2 -l 32 -e 3001. Peak calling was performed using MACS v1.4[55], with the following arguments: –gsize = mm–bw = 300–nomodel–shiftsize = 100. Experiment-matched input DNA was used as a control.

ENCODE RNA-seq data analysis. Raw data were downloaded from the ENCODE Data Coordination Center (https://www.encodeproject.org/, see Supplementary Table 8 for the complete list of sample identifiers). Short reads were aligned to the mm10 assembly of the mouse genome using Tophat v2.0.8[57] and Gencode vM3[58] as the reference transcriptome. Cuffnorm v2.2.1[51] was run to quantify transcripts across conditions using the Gencode vM3[58] transcriptome as the reference and setting -library-norm-method to geometric. Only genes with a level of expression of at least one RPKM (reads per kilobase of exons per million mapped reads) in at least one of the considered conditions were included in further analyses. Small and non-coding RNAs were excluded by retaining only those genes with a Gencode biotype[58] supporting protein-coding functionality.

Classifying genes by tissue-biased patterns of expression. For each protein-coding gene in the mouse genome, the expression variability across the twenty-nine ENCODE RNA-seq experiments from multiple tissues and developmental time points was evaluated using two metrics: a measure of tissue-specificity (τ)[59] ranging from 0 (consistent expression across all conditions) to 1 (expression in one single condition); and a measure of relative expression in a condition of interest (for example, limb at E11.5). Given a gene, the latter was defined as the difference between the percentile of expression of the gene in the given condition and the median percentile of expression across all the samples. A large positive number indicates a gene that is much more expressed in the condition of interest than the average.

Tissue-biased genes were defined as showing $\tau \geq 0.7$ and relative expression higher than the 95th percentile. Housekeeping genes were defined as having $\tau \leq 0.4$ and relative expression between the 5th and 95th percentiles. The complete lists of genes assigned to each category are available in Supplementary Table 9.

Gene classification based on pre-specified functional categories. Tissue-biased developmental transcription factors (sometimes referred to as tissue-specific transcription factors) were defined as genes with biased expression in a given tissue (see previous section), associated with abnormal developmental phenotypes in the same tissue (terms extracted from the Mouse Genome Informatics (MGI) database[60], listed in Supplementary Table 10) and annotated as a transcription factor under the terms GO:0003700 or GO:0003705 in the Gene Ontology (GO)[61]. Annotations were downloaded from GO and MGI on July 7, 2016.

Topologically associated domains. TAD coordinates[26] estimated from mouse embryonic stem cell Hi-C data were downloaded from http://chromosome.sdsc.edu/mouse/hi-c/download.html. Coordinates were converted from mm9 to mm10 using liftOver[62].

A statistical framework defining enhancer–promoter associations genome-wide. A list of putative enhancer regions was first defined as follows: after excluding any region annotated to the mitochondrial or any random chromosome, the BED coordinates of the H3K27ac peaks across the twenty-nine conditions (different combinations of tissue and developmental stage as defined by the ENCODE consortium, see 'ENCODE ChIP–seq data analysis' above) were merged using the mergeBed utility from BEDTools v.2.17.0[63]. For a more robust signal estimation (see below), regions shorter than 500 bp were enlarged to 1 kb from their central coordinate. Promoters, defined as regions within 2.5 kb of the transcriptional start sites of genes annotated in Gencode vM3[58], were then excluded using subtractBed from BEDTools v.2.17.0[63]. After that, any remaining region shorter than 1 kb was excluded. Uniquely aligned, de-duplicated reads were then used to quantify the H3K27ac signals at each region, for each one of the 29 conditions. These signals were measured using the coverageBed utility from BEDTools v.2.17.0[63], normalized to RPKM (according to the sequencing depth of each specific sample), and $\log_2$-transformed. The resulting list of 74,366 predicted enhancers and their corresponding H3K27ac signal quantifications, along with the mRNA expression measurements for the protein-coding genes (as defined in 'Classifying genes by tissue-biased patterns of expression'), were used as input for the statistical framework described below. The main steps of the approach are also outlined in Extended Data Fig. 9b.

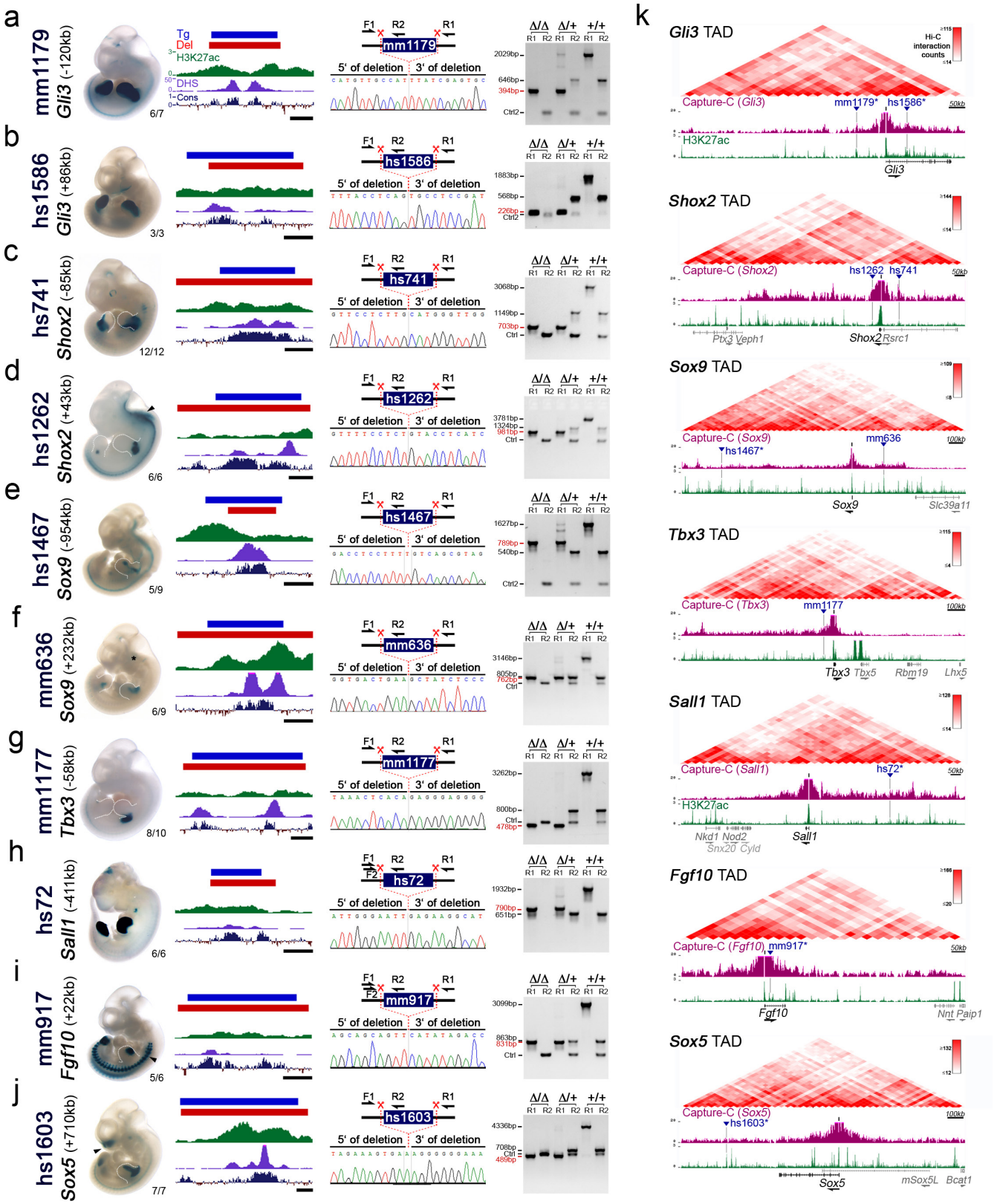
For each previously defined TAD in the mouse genome[26], we retrieved all of the enhancers predicted and the genes expressed in at least one of the twenty-nine conditions considered that fell within that TAD. Pairwise correlations between all possible enhancer–gene combinations within the TAD were then evaluated by calculating Spearman's rank correlation coefficient (SCC) between the H3K27ac pattern of enrichment at the enhancer and the mRNA expression of the gene across the conditions. Each putative enhancer was initially assigned to the gene showing the highest SCC value (in the very rare case of ties, all of the genes showing the same SCC value were assigned to the enhancer). After that, a null distribution of SCC values was estimated empirically, by pairing the enhancer with 1,000 randomly picked genes from the same chromosome. The *z*-score for the correlation coefficient was then calculated by subtracting the mean and dividing by the standard deviation estimated from the empirical null. The corresponding *P* value was calculated using the pnorm function in R. Finally, only those putative enhancers showing a *P* value ≤ 0.05 and a SCC $>= 0.25$ were retained, resulting in a set of 34,882 enhancers with an assigned target (Supplementary Table 11). Considering the entire, genome-wide set of pairwise associations, a $P = 0.05$ corresponds to a Benjamini–Hochberg corrected FDR of 0.087. This analysis resulted in the assignment of one or more putative enhancers to 9,365 protein-coding genes (Supplementary Table 12). To define a set of genes with many redundant enhancers, we considered enhancers as redundant only if they were associated with the same gene by correlation and showed a strong peak of H3K27ac in the same exact tissue under examination (for example, both enhancers are active in limb and linked to the *Gli3* gene). Although this correlative approach may result in a subset of false-positive assignments for individual genes, it enables an approximation of both regulatory complexity and potential enhancer redundancy across the entire genome. We found 1,276 genes that showed multiple assigned enhancers such that at least five of the enhancers were all active in the same tissue (limb, heart or brain). We then used a permutation scheme to directly evaluate the statistical robustness of this conclusion (that is, 1,276 genes with 5 or more redundant enhancers in either developing limbs, heart or forebrain), which considered increasingly higher correlation values between the activity of putative enhancers and expression of genes (Extended Data Fig. 10c–f). By re-shuffling the expression values of each gene across conditions (100 genome-wide permutations), we estimated the FDR of observing a gene with five or more enhancers attached to it, for increasingly larger correlation coefficients. Each permutation consisted of the same enhancers and genes, in which the H3K27ac values were left as in the actual data whereas the RNA expression values of the genes across the different samples were randomly reshuffled. For each genome-wide permuted matrix, the entire statistical approach described above was re-run and a map of enhancer–promoter associations was generated. For each value of Spearman's correlation coefficient (0.25 to 0.75, with a 0.01 step) the number of genes showing five or more enhancers in the permuted data was calculated. The average across the 100 iterations was then computed and used for FDR estimation. This was calculated as the average number of genes showing five or more enhancers across the permuted data, over the number of genes derived from the actual data.

Statistical analysis. Statistical analyses are described in detail in the Methods sections above. Whenever a *P* value is reported in the text, the statistical test is also indicated. Unless specified otherwise, all the statistics were estimated and plots drawn using the statistical computing environment R (https://www.r-project.org) or GraphPad Prism 7 software.

Data availability. ChIP–seq and RNA-seq datasets are available in the NCBI GEO database with the accession code GSE93730. Additional data supporting the findings of this study are available from the corresponding authors upon reasonable request.

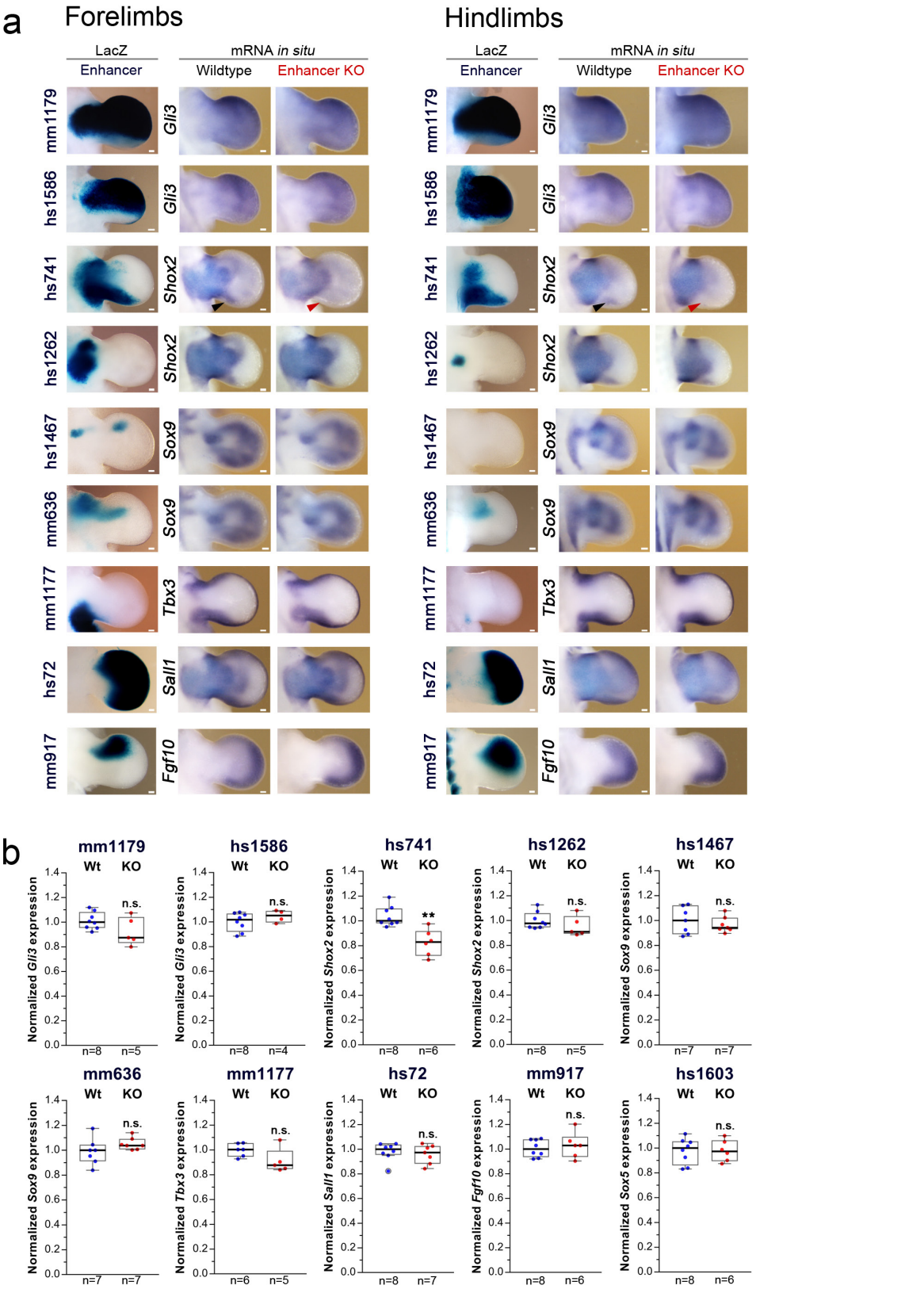
40. Kothary, R. *et al.* Inducible expression of an *hsp68-lacZ* hybrid gene in transgenic mice. *Development* **105,** 707–714 (1989).
41. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).
42. Mo, A. *et al.* Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86,** 1369–1384 (2015).
43. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154,** 1370–1379 (2013).
44. Yang, H., Wang, H. & Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protocols* **9,** 1956–1968 (2014).
45. Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **42,** W401–W407 (2014).
46. Kvon, E. Z. *et al.* Progressive loss of function in a limb enhancer during snake evolution. *Cell* **167,** 633–642 (2016).

47. Ovchinnikov, D. Alcian blue/alizarin red staining of cartilage and bone in mouse. *Cold Spring Harb. Protoc.* **2009,** prot5170 (2009).
48. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C_T method. *Nat. Protocols* **3,** 1101–1108 (2008).
49. Dickel, D. E. *et al.* Genome-wide compendium and functional assessment of in vivo heart enhancers. *Nat. Commun.* **7,** 12923 (2016).
50. Nord, A. S. *et al.* Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155,** 1521–1531 (2013).
51. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515 (2010).
52. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
53. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).
54. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
55. Zhang, Y. *et al.* Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9,** R137 (2008).
56. Jalili, V., Matteucci, M., Masseroli, M. & Morelli, M. J. Using combined evidence from replicates to evaluate ChIP–seq peaks. *Bioinformatics* **31,** 2761–2769 (2015).
57. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).
58. Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22,** 1760–1774 (2012).
59. Yanai, I. *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21,** 650–659 (2005).
60. Bult, C. J., Eppig, J. T., Blake, J. A., Kadin, J. A. & Richardson, J. E. Mouse genome database 2016. *Nucleic Acids Res.* **44,** D840–D847 (2016).
61. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* **43,** D1049–D1056 (2015).
62. Speir, M. L. *et al.* The UCSC Genome Browser database: 2016 update. *Nucleic Acids Res.* **44,** D717–D725 (2016).
63. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
64. Gordon, C. T. *et al.* Identification of novel craniofacial regulatory domains located far upstream of SOX9 and disrupted in Pierre Robin sequence. *Hum. Mutat.* **35,** 1011–1020 (2014).
65. Li, Q. *et al.* A Gli silencer is required for robust repression of gremlin in the vertebrate limb bud. *Development* **141,** 1906–1914 (2014).

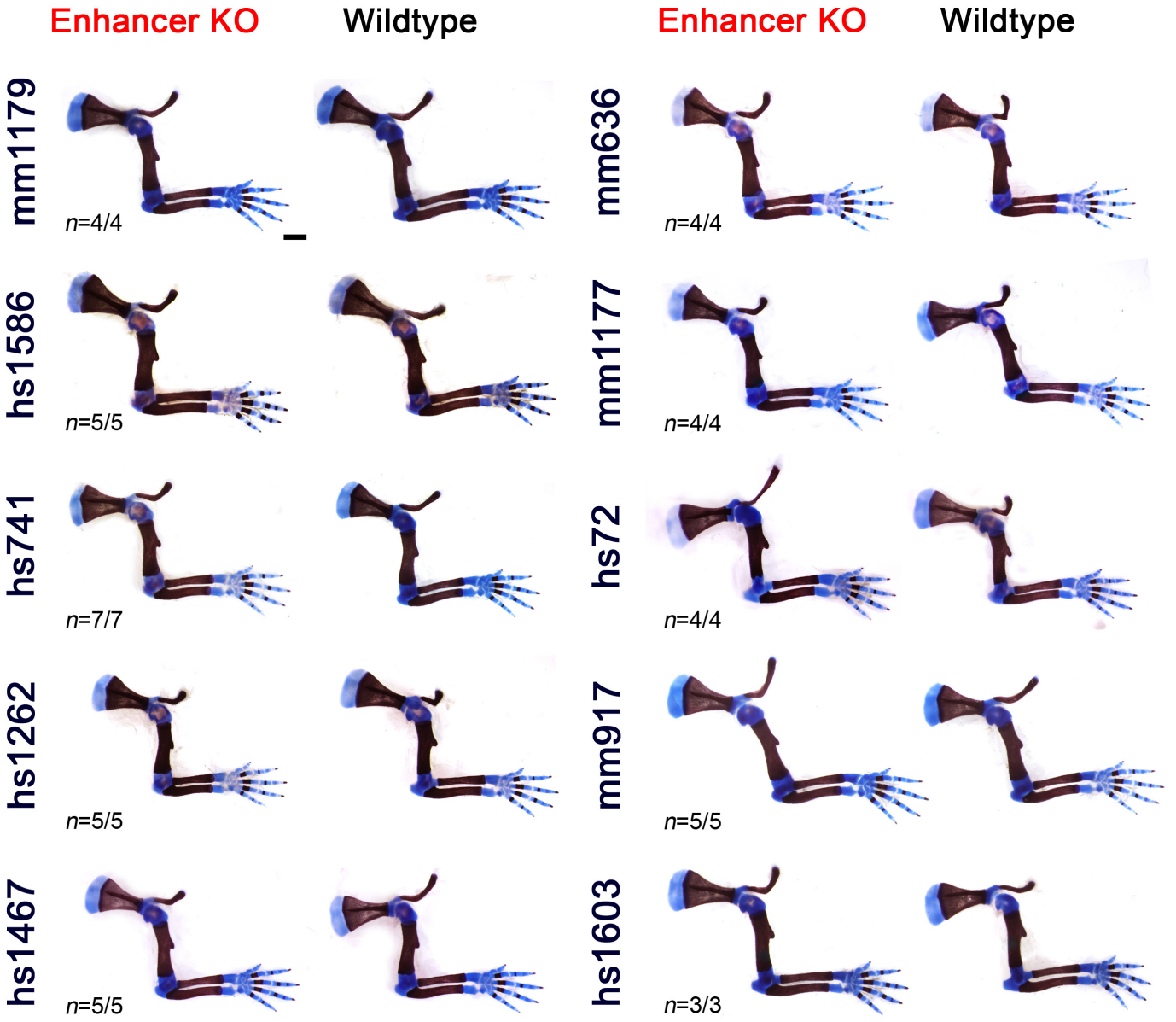


Extended Data Figure 1 | See next page for caption.

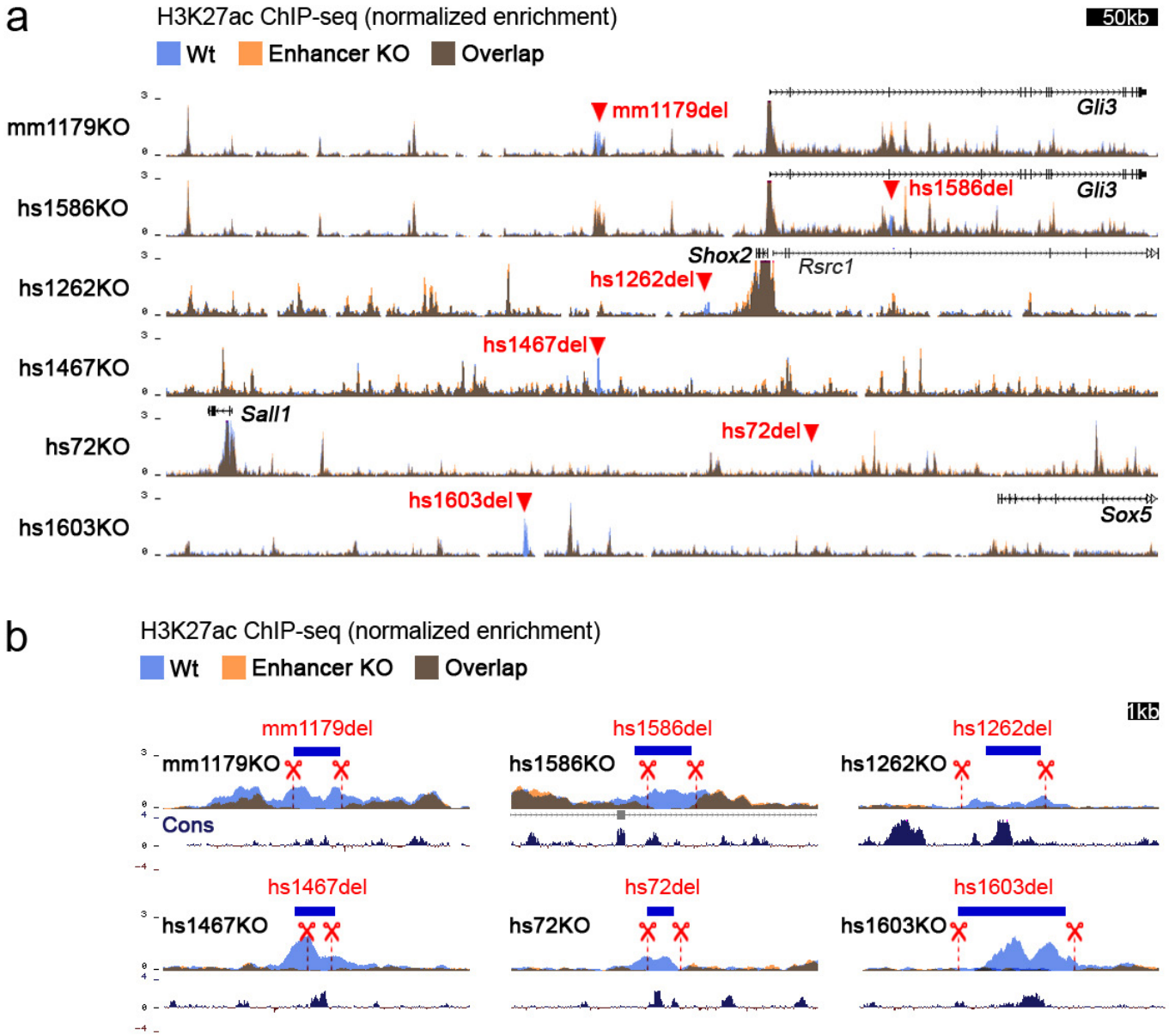
Extended Data Figure 1 | CRISPR deletion of ten limb enhancers and regulatory interaction landscape of associated target genes. **a–j**, Left, representative activity patterns of the selected enhancers in mouse embryos at E11.5 (VISTA enhancer browser)[13] and the respective genomic enhancer regions tested in transgenic assays (Tg, blue bar), along with the regions deleted in enhancer knockout mice (Del, red bar). Corresponding H3K27 acetylation patterns (green) in wild-type mouse embryonic forelimbs at E11.5 (this study) are depicted with open chromatin (ENCODE DHS in forelimbs at E11.5, purple) and the Placental Mammal basewise conservation track by PhyloP (Cons, blue/red). Scale bars, 500 bp. VISTA enhancer IDs (mm and hs numbers) are indicated on the left, with the distance of the enhancer from the transcriptional start site of the predicted target gene in the mouse genome. Numbers at the bottom right of each embryo indicate the reproducibility of the enhancer reporter assay. Arrowheads mark additional activity domains (other than limb): hs1262 (hindbrain, reproducibility: 5/6, also shown previously[17]), mm917 (dorsal root ganglion, 7/7) and hs1603 (nose, 7/7; and branchial arch, 5/7). Asterisk indicates potential craniofacial enhancer activity for mm636, which was observed in 3 of 9 embryos[64]. Right, PCR validation strategy and results for enhancer knockout lines. Red scissors indicate CRISPR-mediated deletion breakpoints. PCR was used to detect the wild-type (+) and enhancer deletion (Δ) alleles. Below, Sanger sequencing traces show the deletion breakpoints (indicated by the dashed line) for the enhancer knockout alleles. PCR genotyping results are shown with amplicon sizes indicated on the left (enhancer deletion allele in red). Primers (Ctrl or Ctrl2) amplifying an unrelated genomic region were included as a PCR positive control. See Supplementary Table 3 for all primer sequences and related PCR product sizes. **k**, Top, Hi-C interaction heat maps of topologically associated chromatin domains (mouse embryonic stem cell TADs)[26]. Bottom, selected enhancers (blue triangles) and their predicted target genes (TSS indicated as black bar). The Capture-C UCSC browser track (purple) illustrates three-dimensional chromatin interaction profiles from E11.5 embryonic limbs (3-kb window) using promoters of the predicted enhancer target genes as anchor points[22]. H3K27ac enrichment (green) in wild-type forelimbs at E11.5 (this study) is shown below. Six of the ten enhancers selected for deletion analysis display local Capture-C enrichment (*), indicating physical interaction with the predicted target gene promoter at E10.5 or E11.5, based on the stringent statistical approach (95th percentile threshold) applied in the original study[22]. Other genes present in the TAD are shown in grey.



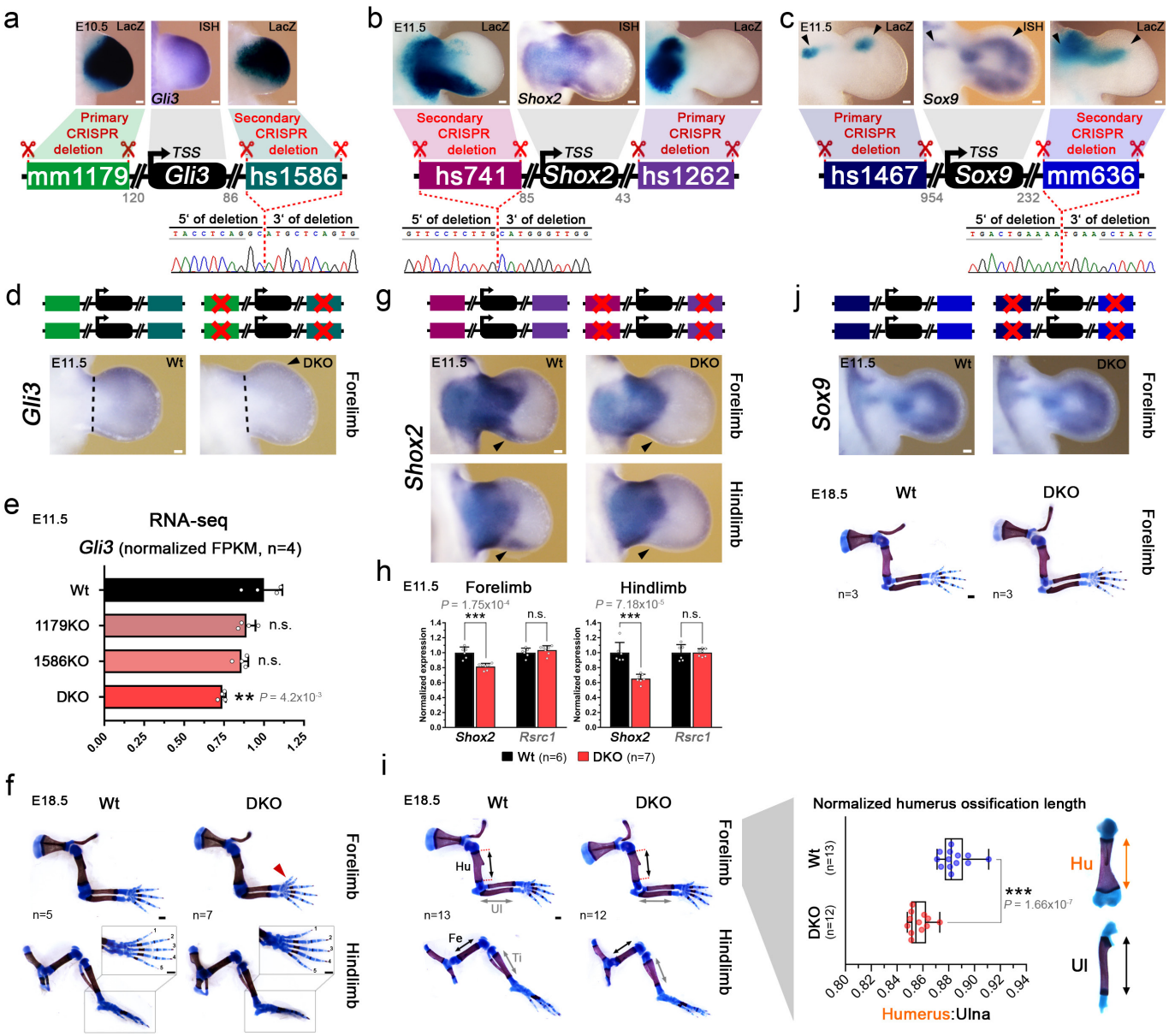
Extended Data Figure 2 | No major differences in expression of predicted target genes in individual enhancer knockouts. a, Spatial enhancer activity domains (LacZ, see also Fig. 1b) are compared to mRNA expression domains (by *in situ* hybridization) of the predicted target genes in embryonic forelimbs and hindlimbs at E11.5. No significant changes in expression patterns were observed in enhancer knockouts compared to wild-type limbs, except in limbs lacking hs741, where a small subdomain of target gene expression was lost (red arrowhead marks loss of the posterior *Shox2* domain in the distal limb, compared with black arrowhead in wild type). Transcript distribution was reproduced in at least $n = 3$ independent biological replicates. **b**, Quantitative real-time PCR using limbs of homozygous null (KO, red dots) and wild-type (Wt, blue dots) embryos at E11.5 reveals lack of significantly downregulated transcript levels of predicted enhancer target genes in nine out of ten cases. Box plots indicate median, interquartile values, range and individual biological replicates. Outliers are shown as circled data points. $^{**}P = 0.0012$, unpaired, two-tailed *t*-test. n.s., not significant. Scale bars, 100 μm.



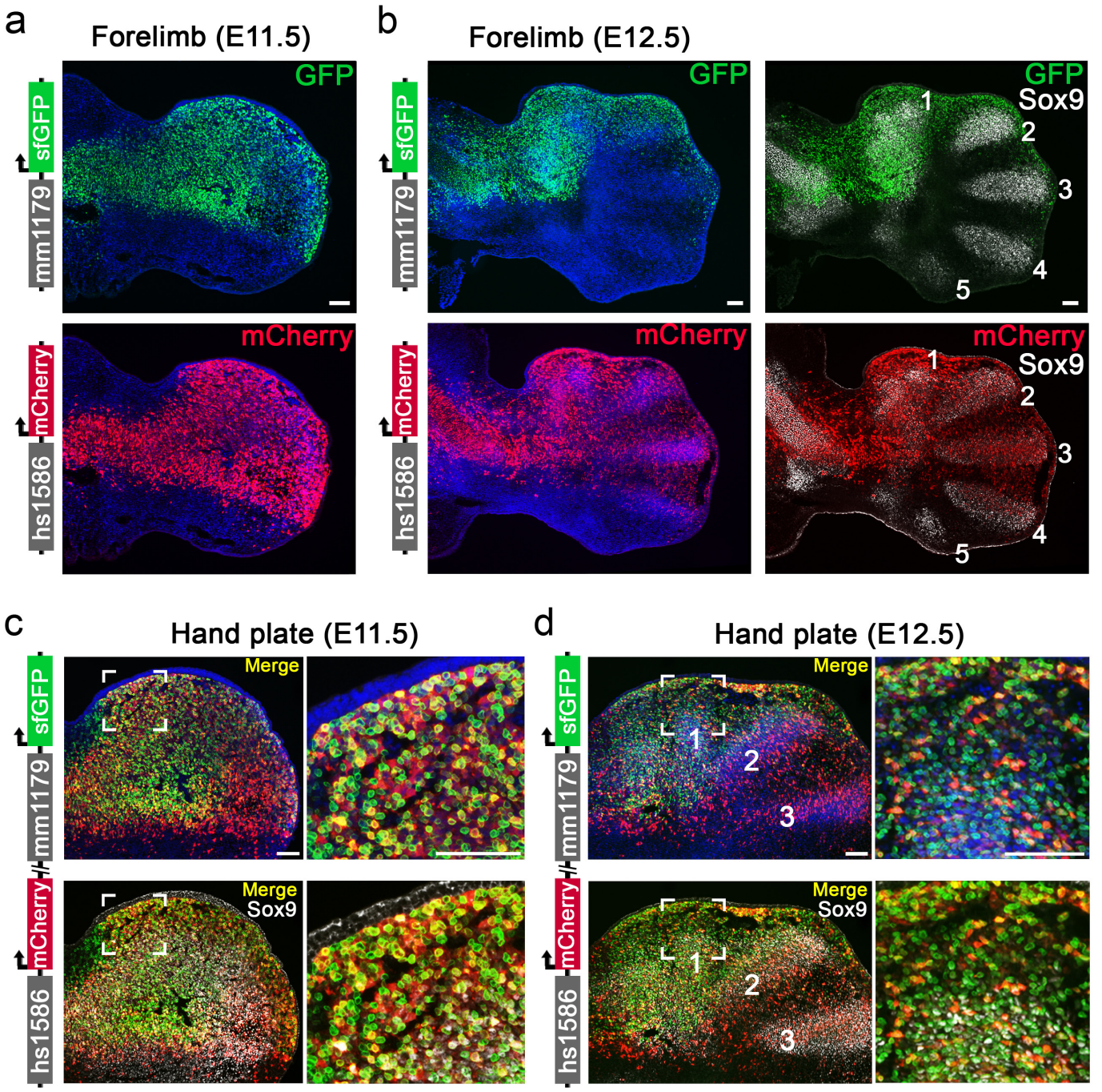
Extended Data Figure 3 | Absence of obvious morphological abnormalities in limb enhancer knockouts. Side-by-side comparison of enhancer knockout limb skeletons and wild-type littermate controls at E18.5. Neither forelimbs (this figure) nor hindlimbs (data not shown) of the enhancer knockout lines revealed any obvious morphological differences in comparison to wild-type littermates. Cartilage is stained blue and bone dark red. The number of embryos with normal limb phenotypes over the total number of homozygous-null embryos examined is shown in the bottom left. *n* represents number of independent biological replicates with similar results. Scale bar, 1 mm.



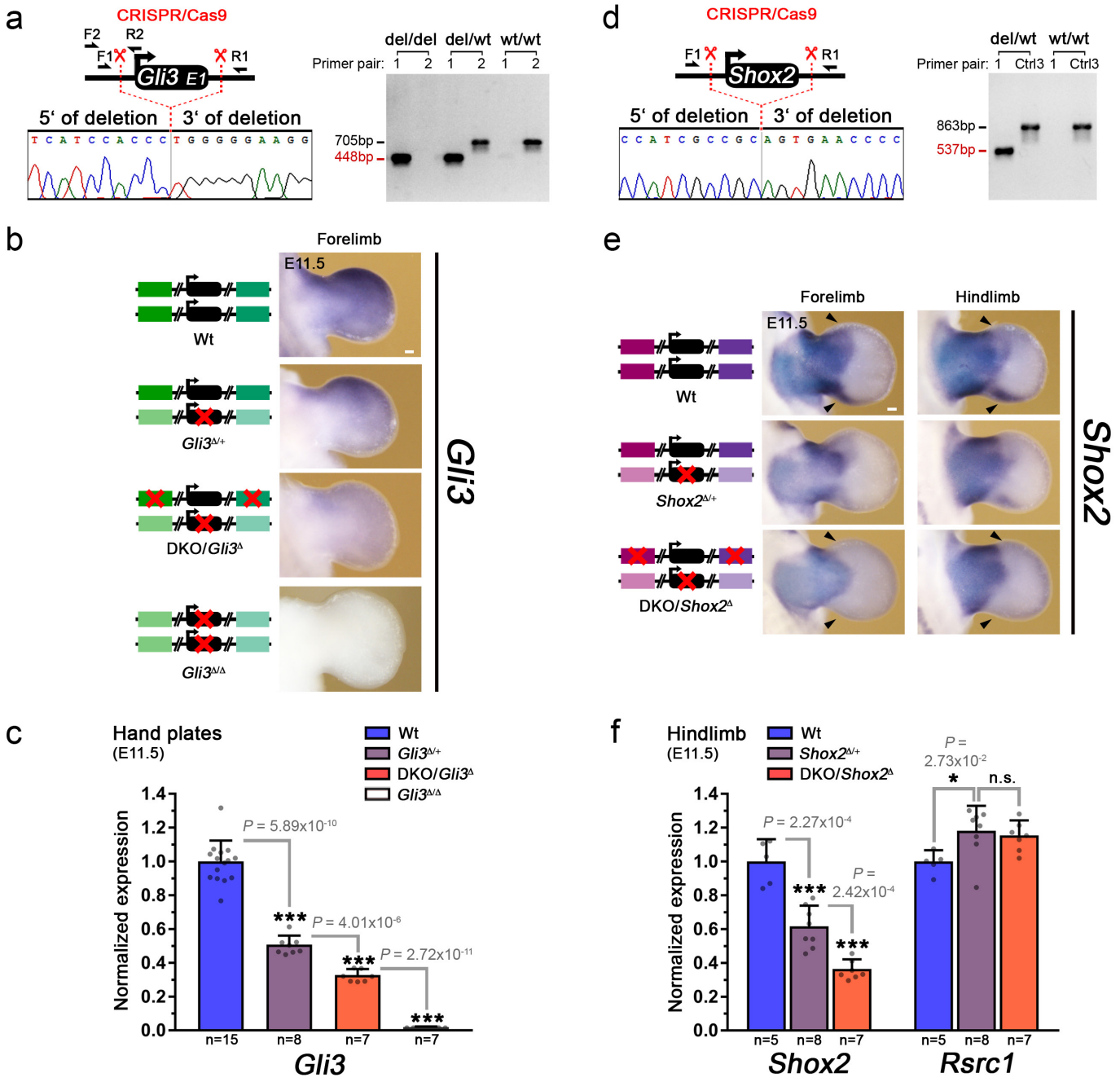
Extended Data Figure 4 | Absence of compensatory enhancer signatures in limbs of enhancer knockout embryos. a, Layered ChIP–seq H3K27 acetylation (ac) profiles surrounding the deleted enhancers and from wild-type (blue, $n = 4$ independent biological replicates) and enhancer knockout embryos (orange, at least $n = 2$ biological replicates). For all samples, E11.5 forelimb was profiled. For display, replicates were merged using bigWigMerge (UCSC tools) and normalized. Red triangles indicate the positions of individual enhancer deletions. **b**, H3K27ac enrichments in targeted regions marked by red triangles in **a**, showing the absence of H3K27ac at the deletion site in individual enhancer knockout (orange) compared to wild-type (blue) samples. Blue bars indicate locations of enhancer sequences. Dashed red lines demarcate the regions deleted by CRISPR. Vertebrate basewise conservation track by PhyloP (Cons) is shown.



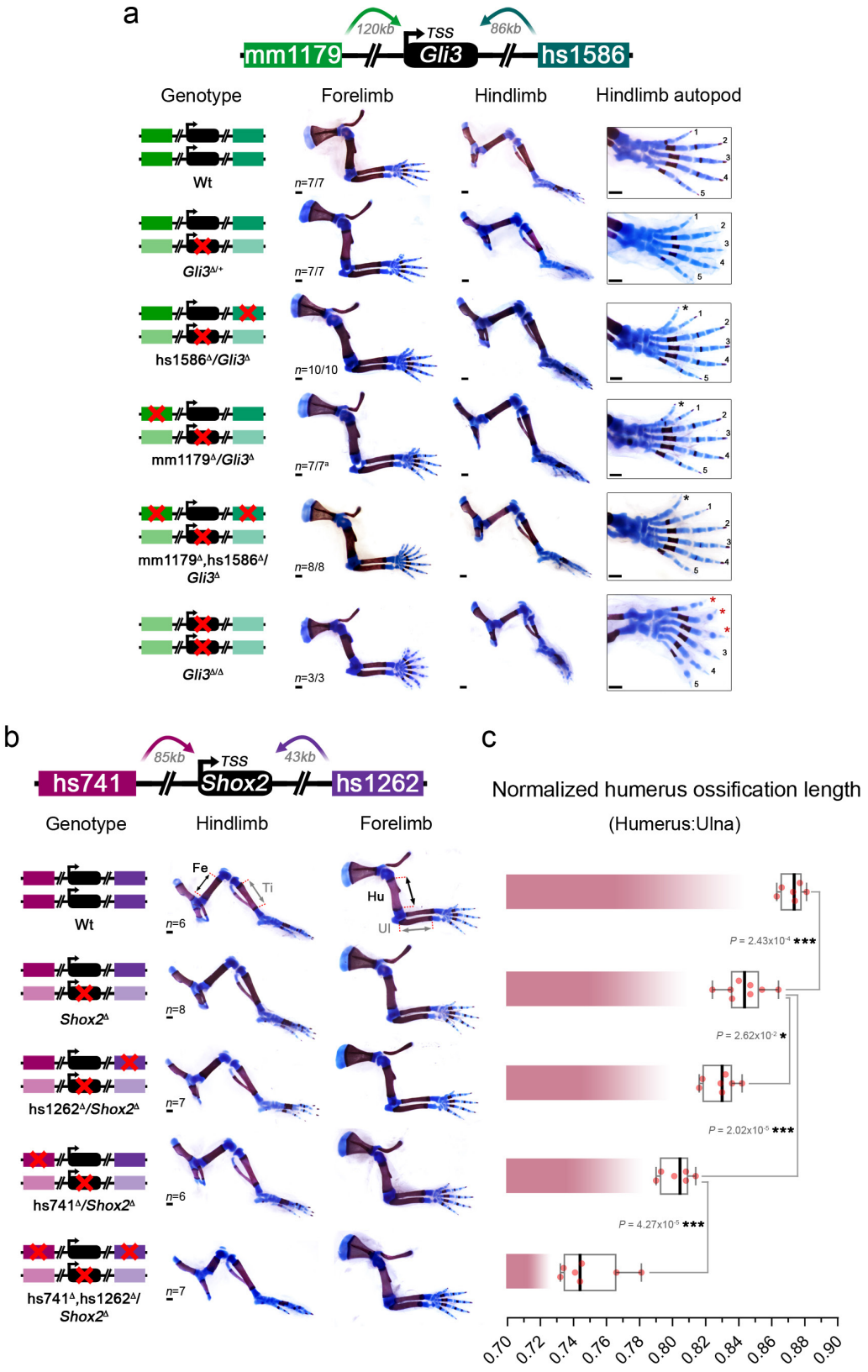
Extended Data Figure 5 | Transcriptional and phenotypic impact of dual enhancer deletions engineered by iterative CRISPR–Cas9 genome editing. a–c, Top, enhancer pairs with overlapping limb activities (LacZ), coinciding with domains of predicted target gene expression visualized by *in situ* hybridization (ISH). For *Sox9* enhancers, black arrowheads indicate overlapping domains. Schematics, double enhancer deletion strategy to delete the three enhancer pairs with overlapping activity (see Methods). Grey numbers indicate enhancer distance (kb) from the TSS. Bottom, Sanger sequencing verification of the secondary enhancer deletion. Deletion breakpoint is marked by the dashed line. Grey horizontal bars indicate bases present in the primary deletions (single enhancer knockout lines, see Extended Data Fig. 1a–j). *Shox2*- and *Sox9*-associated LacZ panels are also used in Extended Data Fig. 2. **d**, *Gli3* transcript distribution *in situ* hybridization in wild-type (Wt) and mm1179/hs1586 DKO embryos. Arrowhead points to reduced *Gli3* transcript in the anterior limb mesenchyme. Dashed line indicates dissected hand plate for RNA-seq. **e**, RNA-seq confirmed significantly reduced *Gli3* expression in hand plates of DKO embryos but not individual enhancer knockout embryos (compared to wild-type hand plates). **f**, Unaffected hindlimb morphology in mm1179/hs1586 DKO embryos. Red arrowhead points to digit 1 duplication in forelimbs (see also Fig. 2). **g**, *Shox2* expression (*in situ* hybridization) in forelimbs and hindlimbs of hs741/hs1262 DKO embryos. The distal-posterior domain (arrowhead) is dependent on hs741 (Extended Data Fig. 2a). **h**, Reduced *Shox2* expression in forelimbs and hindlimbs of hs741/hs1262 DKO embryos (qPCR). Expression of the nearby *Rsrc1* gene was unchanged. **i**, Left, representative limb skeletons of wild-type and hs741/hs1262 DKO embryos. Hu, humerus; Ul, ulna; Fe, femur; Ti, tibia. Right, mild but significant reduction in humerus ossification length (double arrows) in hs741/hs1262 DKO limb skeletons. $***P = 1.66 \times 10^{-7}$ (two-tailed, unpaired *t*-test). **j**, Absence of evident differences in *Sox9* expression or skeletal abnormalities in embryos lacking both the hs1467 and mm636 enhancers near *Sox9*. For *in situ* hybridization, transcript distribution was reproduced in at least $n = 3$ independent biological replicates. *n* represents number of independent biological replicates with similar results. For bar graphs and boxplots, individual biological replicates are shown as data points. Bar graphs illustrate mean and s.d. Box plot indicates median, interquartile values and range. $***P < 0.001$; $**P < 0.01$ (two-tailed, unpaired *t*-test). n.s., not significant. Scale bars, 100 μm (white) and 500 μm (black).



Extended Data Figure 6 | Cellular resolution of redundant *Gli3* enhancer activities at the onset of digit formation. **a**, **b**, Individual *Gli3* enhancer activities as detected by immunofluorescence (mm1179, green; hs1586, red) in forelimbs of transgenic reporter embryos. Sox9 (grey) marks chondrogenic progenitors of the mesenchymal condensations forming digit primordia (digits 1–5, from anterior to posterior). **c**, **d**, Co-localization of mm1179 and hs1586 enhancer activities in hand plates of double enhancer transgenic embryos. Close-ups (right) show that the anterior mesenchyme (Fig. 2c) harbours many cells with dual enhancer activities (yellow). A fraction of double enhancer-positive cells carries the signature of Sox9 digit progenitors (white, bottom). $n = 3$ independent embryos per genotype were analysed, with similar results. Nuclei, detected via Hoechst staining, are blue. Scale bars, 100 μm (**a**, **b**); 50 μm (**c**, **d**).

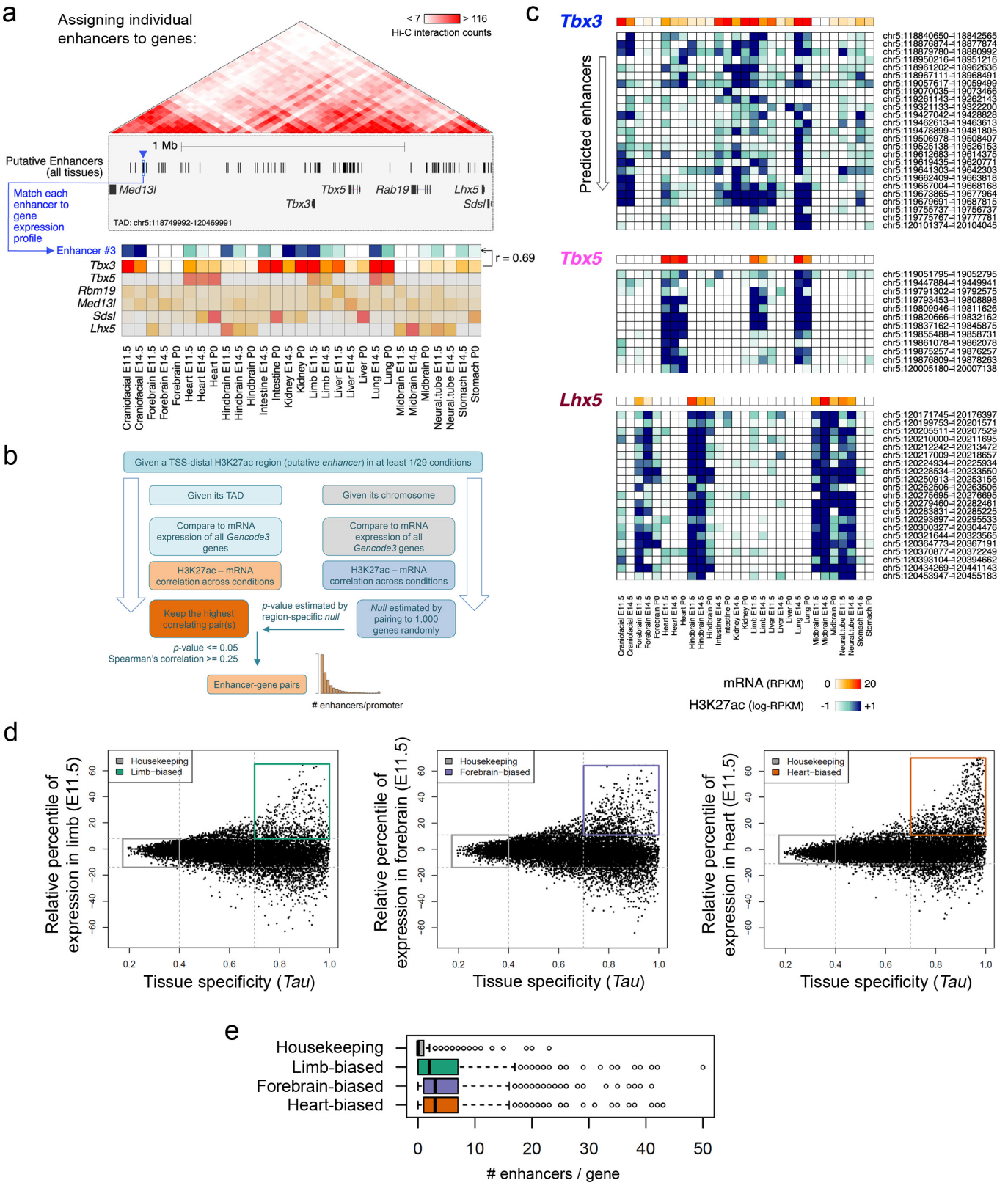


Extended Data Figure 7 | Generation of *Gli3* and *Shox2* knockout alleles and characterization of enhancer deletions in a sensitized background. **a**, **d**, Top, schematic showing CRISPR–Cas9-mediated deletions used to generate *Gli3* and *Shox2* loss-of-function alleles. Genotyping primers used to validate targeted deletion events are indicated. Bottom, Sanger sequencing confirmation of deletion event, with grey and red dashed lines indicating breakpoints. Right, PCR genotyping examples with the size of the product specific for the deletion allele depicted in red (primers listed in Supplementary Table 3). **b**, *In situ* hybridization showing the gradual decrease in anterior *Gli3* transcript in forelimbs of wild-type, $Gli3^{\Delta/+}$ and sensitized mm1179/hs1586 DKO (DKO/$Gli3^{\Delta}$) embryos. **c**, qPCR validation of *Gli3* mRNA levels in forelimb hand plates from the genotypes shown in **b**. **e**, *Shox2* expression (*in situ* hybridization) in forelimbs and hindlimbs of wild-type, $Shox2^{\Delta/+}$ and sensitized hs741/hs1262 DKO (DKO/$Shox2^{\Delta}$) embryos. Arrowheads point to the domains where *Shox2* expression is nearly abolished in enhancer DKO/$Shox2^{\Delta}$ embryos. **f**, qPCR revealing significantly downregulated *Shox2* mRNA levels in hindlimbs of DKO/$Shox2^{\Delta}$ compared to $Shox2^{\Delta/+}$ embryos. *n* indicates the number of independent biological replicates with similar results. Bar plots illustrate mean and s.d., with individual biological replicates shown. $^{***}P < 0.001$; $^{*}P < 0.05$ (two-tailed, unpaired *t*-test). n.s., not significant. For *in situ* hybridization, transcript distribution was reproduced in at least $n = 3$ independent biological replicates. Scale bars, 100 μm.



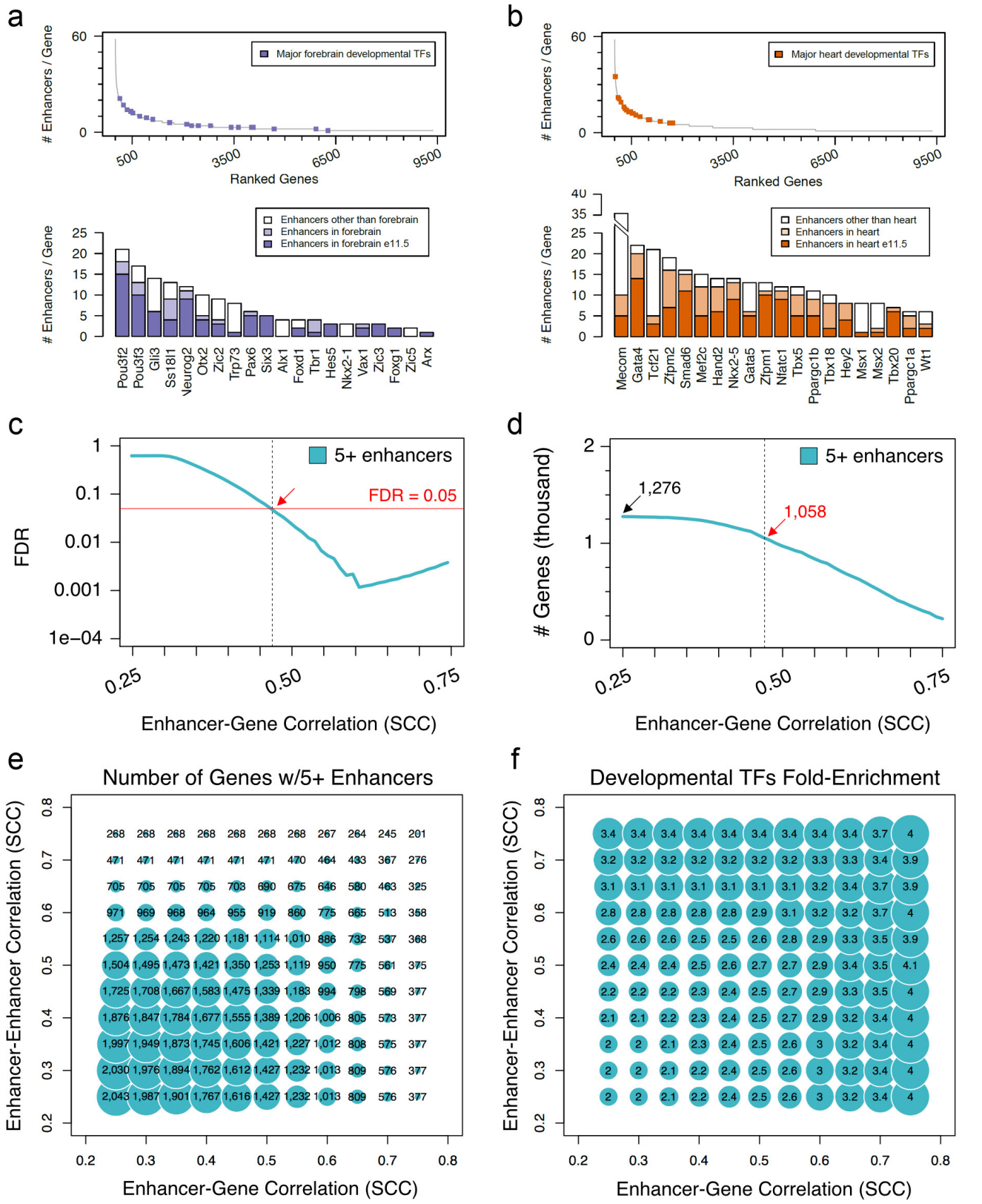
Extended Data Figure 8 | See next page for caption.

Extended Data Figure 8 | Limb phenotypes of individual and combinatorial *Gli3* and *Shox2* enhancer knockouts in the presence of reduced target gene dosage. a, Skeletal phenotypes resulting from mm1179 and hs1586 enhancer deletions in combination with reduction to one copy of the *Gli3* gene at E18.5. Genotypes are shown on the left with red crosses indicating elements deleted by CRISPR–Cas9. While forelimbs of $Gli3^{\Delta/+}$ embryos displayed bifurcated digit 1 terminal phalanges[65], hindlimbs showed an extra toe structure but without detectable cartilage template. Four out of seven mm1179$^{\Delta}$/$Gli3^{\Delta}$ embryos displayed additional bifurcation of digit 2 of the right forelimb (**a**), which suggests that removal of mm1179 reduces *Gli3* levels in the anterior forelimb more than deletion of hs1586. An almost complete anterior extra toe formed in hindlimbs of embryos with single or dual enhancer deletions in the sensitized background (black asterisks). Loss of both *Gli3* copies resulted in anterior hindlimb polydactyly with altered digit identities (red asterisks)[24]. **b**, Allelic series depicting shortening of the stylopod (humerus and femur) in limb skeletons with individual or combined hs741 and hs1262 enhancer deletions in a *Shox2* sensitized condition (see also Fig. 3b). Stylopod ossification length (double arrows) appears less reduced in forelimbs (humerus, Hu) than in hindlimbs (femur, Fe) of embryos lacking the activity of both enhancers (hs741$^{\Delta}$, hs1262$^{\Delta}$/$Shox2^{\Delta}$). Tibia (Ti) and ulna (Ul) were normal in all genotypes examined. **c**, Humerus ossification length (normalized to ulna ossification length) is significantly reduced in embryos lacking either hs741 or hs1262 in the presence of only one copy of *Shox*2. In embryos lacking both enhancers in the sensitized background, significant shortening of humerus ossification is observed (compared to all other genotypes). *n* indicates the number of independent biological replicates with similar results. Box plots indicate median, interquartile values, range and individual biological replicates. $***P < 0.001$; $*P < 0.05$ (two-tailed, unpaired *t*-test). Scale bars, 500 μm.



Extended Data Figure 9 | See next page for caption.

Extended Data Figure 9 | A correlative framework to define enhancer–promoter associations across the mouse genome. a, The TAD including the transcriptional regulators *Tbx3*, *Tbx5* and *Lhx5* illustrates the statistical framework to define enhancer–promoter associations genome-wide. For each predicted enhancer, correlation between its H3K27ac signal (blue arrowhead, blue-shaded heat map) with the mRNA expression profiles of every gene in the TAD (red-shaded heat map) across all available tissues and developmental stages was assessed. The enhancer was then assigned to the most highly correlated gene, *Tbx3* in the case of enhancer 3. **b**, Schematic depicting the underlying statistical framework used to determine genome-wide enhancer–promoter interactions (see Methods). **c**, Activity pattern for the enhancers assigned to *Tbx3*, *Tbx5* and *Lhx5*. Genomic coordinates are listed on the right. For each predicted enhancer–gene pair, Spearman's correlation coefficient (SCC, $n = 29$) and the corresponding empirically estimated *P* value (from 1,000 random enhancer–gene pairings) are shown in Supplementary Table 11. **d**, Identifying genes with biased expression in embryonic limb, forebrain, or heart. Expression variability across 29 RNA-seq datasets from multiple tissues and developmental time points, measures of tissue specificity (Tau (τ), *x*-axis) and specific tissue-biased expression at E11.5 (*y*-axis) for each protein-coding gene were calculated (see Methods). Housekeeping genes were defined as displaying $\tau \leq 0.4$ and relative expression in the limb between the 5th and 95th percentiles. Tissue-biased genes were defined as showing $\tau \geq 0.7$ and relative expression higher than the 95th percentile. **d**, Distribution of enhancer numbers assigned to each gene, for the different gene categories. Genes with tissue-biased expression profiles were associated with a significantly higher number of enhancers than housekeeping genes. $P = 4 \times 10^{-121}$ ($n = 553$), $P = 7 \times 10^{-97}$ ($n = 626$) and $P = 6 \times 10^{-83}$ ($n = 826$) for limb, forebrain and heart biased genes, respectively (two-sided Mann–Whitney tests). $n = 1{,}287$ for housekeeping genes. Box plots indicate median, interquartile values and range. Outliers are shown as individual points.



Extended Data Figure 10 | See next page for caption.

Extended Data Figure 10 | Enhancer redundancy as a widespread feature of developmental genes and robustness to the choice of thresholds used in the correlative approach. **a**, **b**, Top, number of enhancers assigned to each gene through the correlative framework, with developmental transcription factors (TFs) showing biased expression in forebrain (**a**, blue dots) or heart (**b**, orange dots) indicated. Classification of tissue-biased developmental transcription factors is described in Methods. Genes with at least one assigned enhancer are displayed and sorted according to the number of assigned enhancers (left to right). Bottom, bar plot showing the total number of enhancers assigned to each of the transcription factors highlighted in the top panels. For each gene, a colour code shows the number of predicted enhancers assigned to that gene in the relevant tissue (**a**, heart; **b**, forebrain) at E11.5 (dark colour), in the relevant tissue at any other developmental stage included in the analysis (light colour), or in any other tissue (white). **c**, Estimated FDR (based on genome-wide permutations, see Methods) of observing a gene with five or more enhancers assigned to it, for increasingly larger correlation coefficients (0.25 to 0.75). The red solid line indicates an FDR of 0.05. The red arrow and the black dashed line highlight the lowest correlation coefficient (0.47, considering a step of 0.01) with an FDR $\leq$ 0.05 (FDR = 0.0495). **d**, Number of genes showing five or more enhancers assigned to them, for increasingly larger correlation coefficients (0.25 to 0.75). The total number of genes (SCC $\geq$ 0.25) along with the number of genes identified using the threshold set in **c** (SCC >= 0.47) is indicated (1,276 and 1,058, respectively; see Supplementary Tables 11, 12). **e**, Bubble plot showing the number of genes with five or more enhancers assigned to them, at increasingly higher correlation between enhancer and target gene expression (*x*-axis) and between enhancers assigned to the same gene (*y*-axis). **f**, Bubble plot displaying the fold-enrichment (linear) for developmental transcription factor genes among each set in **c**.

nature research

Corresponding author(s): Diane Dickel, Axel Visel, Len Pennacchio

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Methods, section "Experimental Design" (pages 23 and 24).
Specific information related to the different experimental approaches can be obtained in the individual methods sections.

2. Data exclusions

Describe any data exclusions.

Methods, section "Experimental Design" (pages 23 and 24).
Specific information related to the different experimental approaches can be obtained in the individual methods sections.

3. Replication

Describe whether the experimental findings were reliably reproduced.

Attempts at replication were successful and statistical parameters reproducibility numbers are indicated in figure panels/text as required.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Methods, section "Experimental Design" (pages 23 and 24).

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Methods, section "Experimental Design" (pages 23 and 24).

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Statistics were estimated and the plots were drawn using the statistical computing environment R (www.r-project.org). Bar plots and some of the box plots were drawn using the GraphPad Prism software (www.graphpad.com).

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	Materials used are readily available from the authors or from standard commercial sources (see Methods for specific reagents).

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	For ChIP-seq the anti-H3K27ac antibody from Active Motif (cat no. 39133) was used: http://www.activemotif.com/catalog/details/39133/histone-h3-acetyl-lys27-antibody-pab It has been validated and used e.g. in the following study: e.g. Kuwahara, M., Ise, W., et al. (2016), 'Bach2-Batf interactions control Th2-type immune response by regulating the IL-4 amplification loop.', Nat Commun, 7, pp. 12596 For Immunofluorescence the following primary antibodies have been used (see Methods, page 33): chicken 15 anti-GFP (1:500, Thermo Fisher Scientific, A10262), rabbit anti-mCherry (1:1,000, Thermo Fisher Scientific, PA5-34974) and goat anti-Sox9 (1:500, R&D Systems, AF3075). These antibodies were validated in several publications and are all widely used for immuno-detection of the corresponding proteins in mouse tissues (see websites of vendors). For in situ hybridization a standard protocol involving the Anti-Digoxigenin-AP antibody (Roche, 11093274910) was used (see Methods page 32 for reference, described in detail in Kvon et al., Cell 2016).

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used in this study.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used in this study.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used in this study.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eukaryotic cell lines were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	See Methods, section "Experimental Design" (page 23). Experiments were performed in Mus musculus FVB strain mice. The following developmental stages were used in this study: embryonic day E10.5, E11.5, E12.5 and E18.5 mice. Animals of both sexes were used in the analysis and adult mice involved in breedings were up to 1 year of age.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	This study did not involve human research participants.

nature research

Corresponding author(s): Diane Dickel, Axel Visel, Len Pennacchio

☐ Initial submission ☐ Revised version ☒ Final submission

ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data deposition

1. For all ChIP-seq data:

☒ a. Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ b. Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

2. Provide all necessary reviewer access links.
The entry may remain private before publication.

All ChIP-seq files (raw data and BED files) have been deposited in the GEO database (accession GSE93730) as stated in the Methods section of the manuscript. The secure token to access these files is the following: wlkbwuicztklnoz

3. Provide a list of all files available in the database submission.

GSM2461145: Wt_replicate_01_ChIP_H3K27ac
GSM2461146: Wt_replicate_02_ChIP_H3K27ac
GSM2461153: Wt_replicate_03_ChIP_H3K27ac
GSM2461154: Wt_replicate_04_ChIP_H3K27ac
GSM2461149: Wt_replicate_01_input
GSM2461150: Wt_replicate_02_input
GSM2461159: Wt_replicate_03_input
GSM2461160: Wt_replicate_04_input
GSM2461147: mm1179KO_replicate_01_ChIP_H3K27ac
GSM2461148: mm1179KO_replicate_02_ChIP_H3K27ac
GSM2461157: mm1179KO_replicate_03_ChIP_H3K27ac
GSM2461158: mm1179KO_replicate_04_ChIP_H3K27ac
GSM2461151: mm1179KO_replicate_01_input
GSM2461152: mm1179KO_replicate_02_input
GSM2461163: mm1179KO_replicate_03_input
GSM2461164: mm1179KO_replicate_04_input
GSM2461155: hs72KO_replicate_01_ChIP_H3K27ac
GSM2461156: hs72KO_replicate_02_ChIP_H3K27ac
GSM2461161: hs72KO_replicate_01_input
GSM2461162: hs72KO_replicate_02_input
GSM2461129: hs1586KO_replicate_01_ChIP_H3K27ac
GSM2461137: hs1586KO_replicate_02_ChIP_H3K27ac
GSM2461133: hs1586KO_replicate_01_input
GSM2461141: hs1586KO_replicate_02_input
GSM2461130: hs1262KO_replicate_01_ChIP_H3K27ac
GSM2461138: hs1262KO_replicate_02_ChIP_H3K27ac
GSM2461134: hs1262KO_replicate_01_input
GSM2461142: hs1262KO_replicate_02_input
GSM2461131: hs1467KO_replicate_01_ChIP_H3K27ac
GSM2461139: hs1467KO_replicate_02_ChIP_H3K27ac
GSM2461135: hs1467KO_replicate_01_input
GSM2461143: hs1467KO_replicate_02_input
GSM2461132: hs1603KO_replicate_01_ChIP_H3K27ac
GSM2461140: hs1603KO_replicate_02_ChIP_H3K27ac
GSM2461136: hs1603KO_replicate_01_input
GSM2461144: hs1603KO_replicate_02_input

4. If available, provide a link to an anonymized genome browser session (e.g. UCSC).	n.a.

Methodological details

5. Describe the experimental replicates.	Tissue used per sample: pooled forelimbs from wild-type or enhancer knockout mouse embryos at e11.5 (Mus musculus, FVB strain) Replicates: 4 replicates for samples from wild-type and mm1179 enhancer KO embryos, 2 replicates for samples from hs72, hs1586, hs1262, hs1467 and hs1603 enhancer KO embryos.
6. Describe the sequencing depth for each experiment.	Between 12 and 55 million reads were retained per library following quality filtering (details provided in Methods, page 35). Libraries were prepared using the Illumina Truseq DNA sample prep kit with the following PCR conditions for enrichment of DNA fragments: Denaturing, 98C for 10 seconds; Annealing, 60C for 30 seconds; Extension, 72C for 30 seconds (13 cycles). ChIP-seq and input libraries were sequenced via single end 50 bp reads (see details in Methods, page 34).
7. Describe the antibodies used for the ChIP-seq experiments.	H3K27ac antibody: Active Motif cat no. 39133 http://www.activemotif.com/catalog/details/39133/histone-h3-acetyl-lys27-antibody-pab Citation: e.g. Kuwahara, M., Ise, W., et al. (2016), 'Bach2-Batf interactions control Th2-type immune response by regulating the IL-4 amplification loop.', Nat Commun, 7, pp. 12596
8. Describe the peak calling parameters.	see Methods, section "RNA-seq and ChIP-seq analysis" (page 29)
9. Describe the methods used to ensure data quality.	see Methods, section "RNA-seq and ChIP-seq analysis" (page 29)
10. Describe the software used to collect and analyze the ChIP-seq data.	see Methods, section "RNA-seq and ChIP-seq analysis" (page 29)