



Contents lists available at ScienceDirect

## Genomics

journal homepage: [www.elsevier.com/locate/ygeno](http://www.elsevier.com/locate/ygeno)

## Functional autonomy of distant-acting human enhancers

Axel Visel<sup>a</sup>, Jennifer A. Akiyama<sup>a</sup>, Malak Shoukry<sup>a</sup>, Veena Afzal<sup>a</sup>,  
Edward M. Rubin<sup>a,b</sup>, Len A. Pennacchio<sup>a,b,\*</sup><sup>a</sup> Genomics Division, MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA<sup>b</sup> U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598 USA

## ARTICLE INFO

## Article history:

Received 6 November 2008

Accepted 19 February 2009

Available online 5 March 2009

## Keywords:

Enhancer  
Cis-regulatory  
Combinatorial  
Evolution

## ABSTRACT

Many human genes are associated with dispersed arrays of transcriptional enhancers that regulate their expression in time and space. Studies in invertebrate model systems have suggested that these elements could function as discrete and independent regulatory units, but the *in vivo* combinatorial properties of vertebrate enhancers remain poorly understood. To explore the modularity and regulatory autonomy of human developmental enhancers, we experimentally concatenated up to four enhancers from different genes and used a transgenic mouse assay to compare the *in vivo* activity of these compound elements with that of the single modules. In all of the six different combinations of elements tested, the reporter gene activity patterns were additive without signs of interference between the individual modules, indicating that regulatory specificity was maintained despite the presence of closely-positioned heterologous enhancers. Even in cases where two elements drove expression in close anatomical proximity, such as within neighboring subregions of the developing limb bud, the compound patterns did not show signs of cross-inhibition between individual elements or novel expression sites. These data indicate that human developmental enhancers are highly modular and functionally autonomous and suggest that genomic enhancer shuffling may have contributed to the evolution of complex gene expression patterns in vertebrates.

Published by Elsevier Inc.

## Introduction

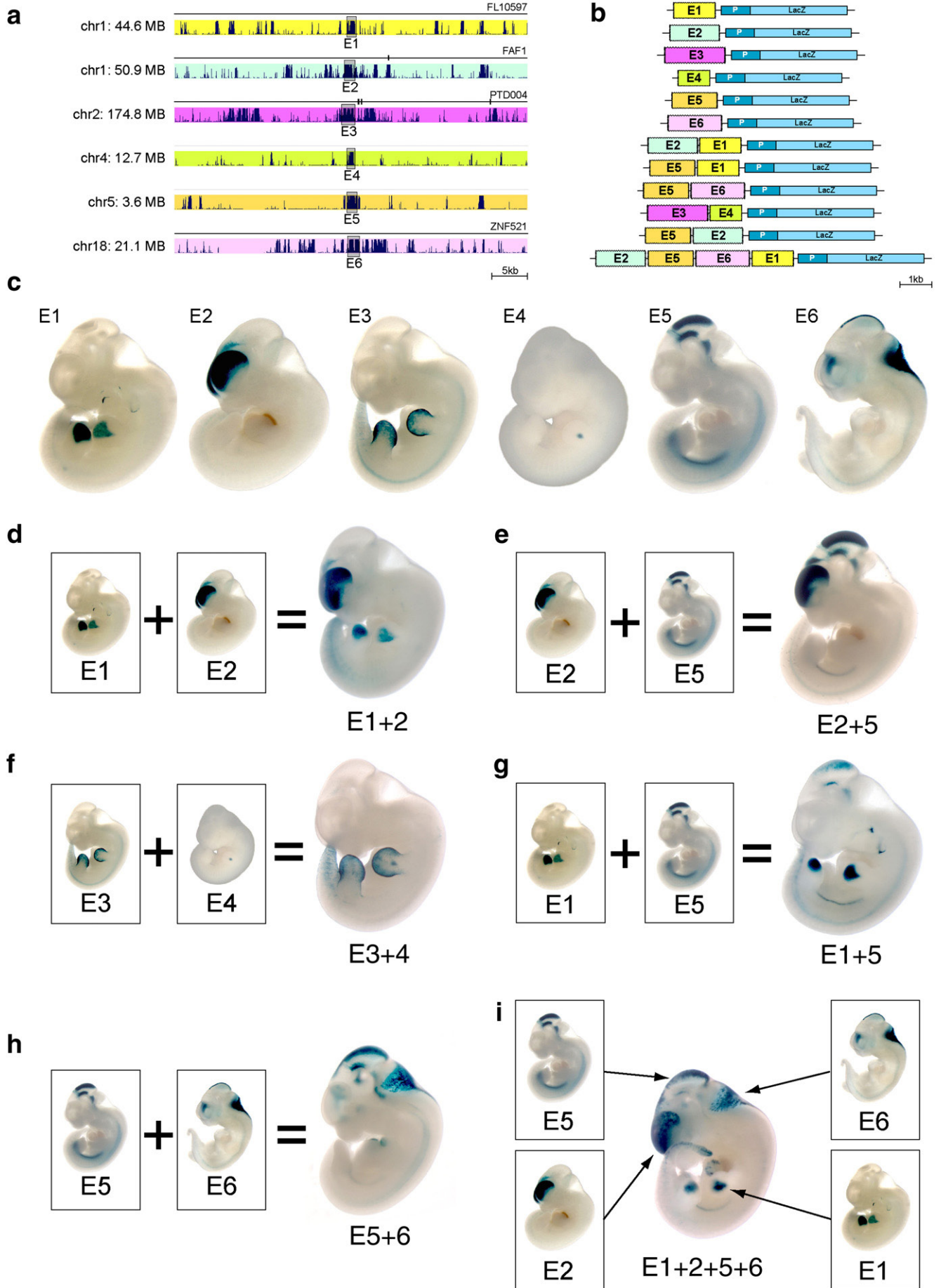
The regulation of many human genes is controlled by multiple discrete enhancer sequences with different tissue specificities (e.g. references [1–6]). Such enhancers activate gene expression independent of their orientation [7] and are commonly scattered across large noncoding intervals [8,9], in some extreme cases functioning >1 Mb from their gene promoter target [10,11]. While progress towards genome-wide annotation of developmental enhancers has been made by coupling comparative genomic and chromatin immunoprecipitation-sequencing approaches to experimental studies in mice and fish [12–15], the functional and evolutionary significance of their dispersed arrangement remains unclear. Structural modularity of enhancer architecture may facilitate evolutionary fine tuning of distinct aspects of expression patterns [16,17], but observations in invertebrate models [18,19] have also raised the possibility that intergenic translocation of preformed enhancer modules may have contributed to the evolution of complex gene expression patterns in vertebrates. However, an open question remains whether this proposed mechanism of regulatory evolution is feasible since it requires that enhancers accurately retain their individual activities when placed into a new genomic context.

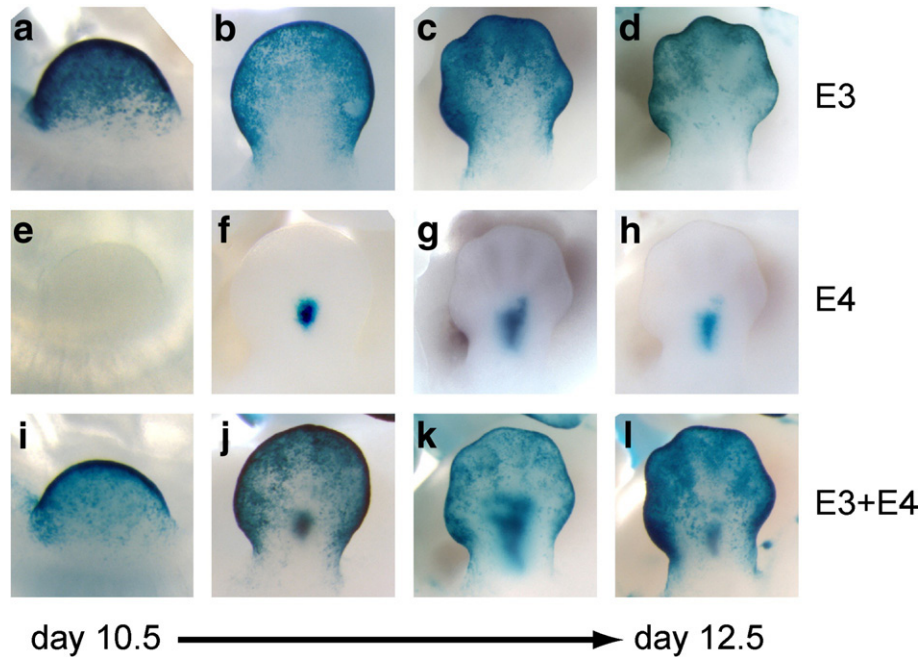
## Results

*Heterologous pairs of enhancers have additive regulatory properties*

To explore the prevalence of possible positive or negative interactions among human developmental enhancers, we recombined enhancer modules from different, functionally unrelated genes and studied their regulatory *in vivo* properties during embryonic development in transgenic mice. We selected for this purpose six *in vivo* validated enhancers [12,13] (E1–E6, Fig. 1a; Suppl. Table 1). When individually coupled to a minimal heat shock protein 68 promoter linked to a LacZ reporter gene (Fig. 1b; [13,20]), each of these elements drove reproducible tissue-specific expression in transgenic mouse embryos. Representative patterns observed with these single enhancers are shown in Fig. 1c. Of note, the enhancers were selected for analysis based on their expression patterns which are easily distinguished at the resolution of whole-mount staining, yet also include features that are located in close spatial proximity. In particular, activity patterns comprise different subregions of the developing forebrain (E2 and E6), the midbrain (E5 and E6), the hindbrain (E5 and E6), and three different subregions of the developing limb (E1, E3, and E4). A detailed overview of the strong reproducibility of these features in independent transgenic embryos, as well as additional reproducible features that were observed for the individual patterns, is provided in Suppl. Table 2. All elements are located on different human chromosomes and thereby are expected to

\* Corresponding author. Genomics Division, One Cyclotron Road, MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Fax: +510 486 4229.  
E-mail address: [LAPennacchio@lbl.gov](mailto:LAPennacchio@lbl.gov) (L.A. Pennacchio).





**Fig. 2.** Temporal and spatial additivity of individual enhancer activities within the developing limb. (a–l) Dorsal surface view of forelimb buds of individual embryos transgenic for E3 (a–d), E4 (e–h) and the compound enhancer E3 + E4 (i–l). For each construct, embryos were collected at e10.5 (a, e, i), e11.5 (b, f, j), and e12.5 (c, d, g, h, k, l) and representative limbs were stage-matched based on morphology. The transgenic status of the embryo shown in (e) was confirmed by genotyping.

regulate independent gene(s), with the exception of E1 and E2 which are within introns of different genes on chromosome 1 but are located more than 5 Mb apart from each other. All elements are located at least 30 kb away from the closest known promoter, with 4 elements found within genes (E1, E2, E3 and E6) and the remaining 2 elements found between genes (E4 and E5).

These enhancers are experimentally defined by their ability to drive reporter gene expression in transgenic mouse embryos, but (as for most human developmental enhancers) it is unknown whether additional *cis*-regulatory cues such as general or tissue-specific repressor/silencer activities are also embedded in these elements that might affect the activity of other regulatory elements in their vicinity. To determine the combinatorial properties of these human enhancers, we generated five constructs containing pair-wise tandem fusions of the heterologous enhancers described above (Fig. 1b). For each construct, we obtained multiple transgenic embryos at e11.5 (representing 8 to 15 independent transgenic integration events, see Suppl. Table 2). In each of the cases studied we observed reproducible patterns that were a direct superimposition of the two individual patterns (Fig. 1d–h). For instance, as one representative example, a construct containing E2 (forebrain) coupled to E5 (medial–dorsal and lateral cell populations of the midbrain and ventral hindbrain) targets reporter gene expression to the same respective subregions of the fore-, mid- and hindbrain as observed for E2 and E5 alone (Fig. 1c, e). Using the presence of reproducible staining in individual anatomical structures as a qualitative measure, we observed no disruption of enhancer activities in the tandem fusion constructs in comparison to each enhancer element alone (Suppl. Table 2). Conversely, the concatenation of heterologous enhancers did not result in any reproducible staining in additional anatomical structures or domains outside of those observed for the individual enhancer constructs. These data indicate that in all instances tested, the two enhancers

retained their specificity independent of each other despite the artificial coupling of enhancer modules that regulate different genes.

#### *Complex but predictable regulatory output from a heterologous multi-enhancer array*

To test whether more complex combinations of enhancer modules would result in positive or negative interactions, we concatenated four different enhancers from unrelated genes and tested the activity of this compound construct at embryonic day 11.5. Similar to the additive results observed in the combination of two discrete enhancers, independent transgenic embryos for the compound 4-mer construct had highly reproducible patterns that included all the major features of the individual patterns, while not introducing additional reproducibly stained structures (Fig. 1i, Suppl. Table 2). Thus, even when multiple distant-acting enhancers were combined in close proximity in a single construct, the individual enhancer units retained their distinct spatial specificities at this time-point.

#### *Both spatial and temporal regulatory specificity are preserved*

In addition to spatial properties, we also examined temporal aspects of expression driven by concatenated enhancers. We selected for this purpose construct E3 + E4, containing a pair of enhancers that drove expression in the developing limb (Fig. 1f), where morphological changes enable precise developmental stage matching of independently generated transgenic embryos. When tested individually at stages ranging from e10.5 to e12.5, enhancer E3 alone targets expression to the apical ectodermal ridge (AER) and the surface ectoderm of the limb bud (Fig. 2a–d). In contrast, E4 alone does not drive limb staining at e10.5, but targets expression to a sharply restricted central cell population in the limb at e11.5 that continues

**Fig. 1.** Spatial additivity of tissue-specific enhancers fused from different genes. (a) Genomic environment of six conserved enhancers used in this study. A 50 kb genomic interval bracketing each enhancer is shown, including intron/exon structure of overlapping genes (black) and conservation in 17 vertebrates (color shaded boxes; [26]). All enhancers included an ultraconserved core region (Suppl. Table 1; [27]). (b) Single and compound enhancer constructs for *in vivo* testing. (c) Enhancer activity of single elements at mouse embryonic day 11.5. (d–i) *In vivo* activity of heterologous compound enhancers. (d) E1 + E2, (e) E2 + E5, (f) E3 + E4, (g) E1 + E5, (h) E5 + E6, (i) E1 + E2 + E5 + E6. Only one representative embryo is shown for each single and compound pattern, see Suppl. Table 2 for reproducibility across independent transgenic animals.

throughout e12.5 (Fig. 2e–h). In order to compare the developmental progression of expression driven by constructs E3, E4, and E3 + E4, we collected multiple transgenic embryos at e10.5, e11.5 and e12.5 for each construct. We found that the compound construct E3 + E4 drove AER expression at all time-points examined, whereas the medial expression domain was first observed at e11.5 as with E4 alone, indicating that the developmental onset of expression driven by E4 is not affected by the presence of E3 in its immediate proximity. Taken together, these results indicate that the functional independence of human developmental enhancers and the absence of obvious regulatory interference among them could allow the generation of complex spatiotemporal expression patterns through modular intergenic recombination of enhancers.

## Discussion

The high degree of regulatory autonomy observed in this study suggests that functional independence and spatiotemporal additivity are common features of human distant-acting enhancer modules. Only a limited number of permutations of a small subset of human enhancers were tested in the present study, therefore it remains to be determined whether these properties are a universal feature applying to all developmental enhancers. Moreover, the LacZ-based approach used in this study provides direct evidence for qualitative additivity of patterns, yet we cannot exclude the possibility that subtle quantitative differences in enhancer activities between single and compound constructs were not detectable. These potential limitations notwithstanding, our results indicate that emerging collections of human and other vertebrate enhancers [12–15] provide a toolbox enabling the design of regulatory composites driving customized, complex *in vivo* expression patterns in a predictable manner due to the additive nature of the individual components.

These observations also have potential evolutionary implications, as it has been proposed that duplication of regulatory elements into new genomic locations may have contributed to the emergence of complex gene expression patterns [17,21,22]. *Bona fide* examples of intergenic enhancer shuffling in the human genome remain to be identified, but recent comparative genomic evidence suggests that exaptation of transposable genome elements occurred on a pervasive scale [23]. While some of these mobile elements gave rise to functional enhancers [24,25], it remains uncertain whether such transposon-derived elements typically arose *de novo* or if partially preformed *cis*-regulatory functions were already embedded at the time of their translocation.

The remarkable modular additivity of spatial and temporal enhancer activities indicates that functional distant-acting enhancers, if translocated into new genomic environments, have the potential to transfer aspects of expression patterns between genes without disrupting the function of pre-existing enhancers. We expect that the combinatorial properties of human enhancers demonstrated through our experiments will provide confidence and a conceptual basis for constructing increasingly complex arrays of enhancers by homologous recombination in mice, which is expected to provide further insights into enhancer shuffling as a potential mechanism of vertebrate genome evolution.

## Materials and methods

### Enhancer reporter constructs

All enhancer sequences were PCR amplified from human genomic DNA (Clontech) using the primers listed in Suppl. Table 1. PCR fragments were cloned into the pENTR plasmid (Invitrogen), transferred into an Hsp68–LacZ reporter vector containing a Gateway cassette using LR recombination (Invitrogen; [13,20]) and sequence validated.

### Compound enhancers

To generate compound enhancers, inserts from individual constructs were subcloned by standard molecular cloning techniques. Sequence and orientation of enhancers in the final constructs are indicated in Suppl. Table 1. Residual multiple cloning site fragments of up to 48 bp residing between enhancers are also listed in Suppl. Table 1.

### Transgenic mice

Transgenic mouse embryos were generated by pronuclear injection in accordance with protocols reviewed and approved by the Lawrence Berkeley National Laboratory Animal Welfare and Research Committee. Zygotes at 0.5 dpc for pronuclear injection were collected from FVB strain donor females (Charles River) and, after injection, transferred into pseudopregnant CD-1 strain recipient females (Charles River). Embryos were collected and stained for LacZ activity as previously described [8].

### Assessment of reporter gene expression

Only anatomical structures in which reporter gene expression was present in at least three embryos resulting from independent transgene integration events were considered reproducible. Reproducibilities for all patterns observed with individual and compound enhancers are listed in Suppl. Table 2.

## Acknowledgments

The authors wish to thank Shyam Prabhakar, Rotem Sorek, Marcelo Nobrega, James Noonan and Nadav Ahituv for critical comments on the manuscript; Keith Lewis, Amy Holt, Ingrid Plajzer-Frick, Sengthavy Phouanavong, and Sumita Bhardwaj for technical support. L.A.P./E.M.R. were supported by the Berkeley-PGA, under the Programs for Genomic Applications, funded by National Heart, Lung, and Blood Institute, and L.A.P. by the National Human Genome Research Institute. Research was performed under Department of Energy Contract DE-AC02-05CH11231, University of California, E.O. Lawrence Berkeley National Laboratory. A.V. was supported by an American Heart Association postdoctoral fellowship.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2009.02.002.

## References

- [1] W.S. Simonet, N. Bucay, R.E. Pitas, S.J. Lauer, J.M. Taylor, Multiple tissue-specific elements control the apolipoprotein E/C-I gene locus in transgenic mice, *J. Biol. Chem.* 266 (1991) 8651–8654.
- [2] R.J. Schwartz, E.N. Olson, Building the heart piece by piece: modularity of *cis*-elements regulating *Nkx2-5* transcription, *Development* 126 (1999) 4187–4192.
- [3] J.B. Burch, Regulation of GATA gene expression during vertebrate development, *Semin. Cell. Dev. Biol.* 16 (2005) 71–81.
- [4] A.A. Abbasi, Z. Paparidis, S. Malik, D.K. Goode, H. Callaway, G. Elgar, K.H. Grzeschik, Human GLI3 intragenic conserved non-coding sequences are tissue-specific enhancers, *PLoS ONE* 2 (2007) e366.
- [5] M.A. Bender, J.N. Roach, J. Halow, J. Close, R. Alami, E.E. Bouhassira, M. Groudine, S.N. Fiering, Targeted deletion of 5'HS1 and 5'HS4 of the beta-globin locus control region reveals additive activity of the DNaseI hypersensitive sites, *Blood* 98 (2001) 2022–2027.
- [6] J.R. Landry, N. Bonadies, S. Kinston, K. Knezevic, N.K. Wilson, S.H. Oram, M. Janes, S. Piltz, M. Hammett, J. Carter, T. Hamilton, I.J. Donaldson, G. Lacaud, J. Frampton, G. Follows, V. Kouskoff, B. Gottgens, Expression of the leukaemia oncogene *Lmo2* is controlled by an array of tissue specific elements dispersed over 100 kb and bound by *Tal1/Lmo2*, *Ets* and *Gata* factors, *Blood* (2009), doi:10.1182/blood-2008-11-187757 Electronic publication ahead of print.
- [7] J. Banerji, S. Rusconi, W. Schaffner, Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences, *Cell* 27 (1981) 299–308.

- [8] M.A. Nobrega, I. Ovcharenko, V. Afzal, E.M. Rubin, Scanning human gene deserts for long-range enhancers, *Science* 302 (2003) 413.
- [9] E. de la Calle-Mustienes, C.G. Feijoo, M. Manzanares, J.J. Tena, E. Rodriguez-Seguel, A. Letizia, M.L. Allende, J.L. Gomez-Skarmeta, A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts, *Genome Res.* 15 (2005) 1061–1072.
- [10] L.A. Lettice, S.J. Heaney, L.A. Purdie, L. Li, P. de Beer, B.A. Oostra, D. Goode, G. Elgar, R.E. Hill, E. de Graaff, A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly, *Hum. Mol. Genet.* 12 (2003) 1725–1735.
- [11] T. Sagai, M. Hosoya, Y. Mizushima, M. Tamura, T. Shiroishi, Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb, *Development* 132 (2005) 797–803.
- [12] A. Visel, S. Prabhakar, J.A. Akiyama, M. Shoukry, K.D. Lewis, A. Holt, I. Plajzer-Frick, V. Afzal, E.M. Rubin, L.A. Pennacchio, Ultraconservation identifies a small subset of extremely constrained developmental enhancers, *Nat. Genet.* 40 (2008) 158–160.
- [13] L.A. Pennacchio, N. Ahituv, A.M. Moses, S. Prabhakar, M.A. Nobrega, M. Shoukry, S. Minovitsky, I. Dubchak, A. Holt, K.D. Lewis, I. Plajzer-Frick, J. Akiyama, S. De Val, V. Afzal, B.L. Black, O. Couronne, M.B. Eisen, A. Visel, E.M. Rubin, In vivo enhancer analysis of human conserved non-coding sequences, *Nature* 444 (2006) 499–502.
- [14] A. Woolfe, M. Goodson, D.K. Goode, P. Snell, G.K. McEwen, T. Vavouri, S.F. Smith, P. North, H. Callaway, K. Kelly, K. Walter, I. Abnizova, W. Gilks, Y.J. Edwards, J.E. Cooke, G. Elgar, Highly conserved non-coding sequences are associated with vertebrate development, *PLoS Biol.* 3 (2005) e7.
- [15] A. Visel, M.J. Blow, Z. Li, T. Zhang, J.A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, B. Ren, E.M. Rubin, L.A. Pennacchio, ChIP-seq accurately predicts tissue-specific activity of enhancers, *Nature* 457 (2009) 854–858.
- [16] E.H. Davidson, *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*, Academic Press, Burlington, MA, 2006.
- [17] S.B. Carroll, J. Grenier, S. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*, Wiley-Blackwell, Oxford, 2005.
- [18] C.V. Kirchhamer, L.D. Bogarad, E.H. Davidson, Developmental expression of synthetic cis-regulatory systems composed of spatial control elements from two different genes, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 13849–13854.
- [19] S. Gray, P. Szymanski, M. Levine, Short-range repression permits multiple enhancers to function autonomously within a complex promoter, *Genes Dev.* 8 (1994) 1829–1838.
- [20] R. Kothary, S. Clapoff, S. Darling, M.D. Perry, L.A. Moran, J. Rossant, Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice, *Development* 105 (1994) 707–714.
- [21] R.J. Britten, E.H. Davidson, Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty, *Q. Rev. Biol.* 46 (1971) 111–138.
- [22] M. Kermekchiev, M. Pettersson, P. Matthias, W. Schaffner, Every enhancer works with every promoter for all the combinations tested: could new regulatory pathways evolve by enhancer shuffling? *Gene. Expr.* 1 (1991) 71–81.
- [23] C.B. Lowe, G. Bejerano, D. Haussler, Thousands of human mobile element fragments undergo strong purifying selection near developmental genes, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 8005–8010.
- [24] A.M. Santangelo, F.S. de Souza, L.F. Franchini, V.F. Bumashny, M.J. Low, M. Rubinstein, Ancient exaptation of a CORE-SINE retroposon into a highly conserved mammalian neuronal enhancer of the proopiomelanocortin gene, *PLoS. Genet.* 3 (2007) 1813–1826.
- [25] G. Bejerano, C.B. Lowe, N. Ahituv, B. King, A. Siepel, S.R. Salama, E.M. Rubin, W.J. Kent, D. Haussler, A distal enhancer and an ultraconserved exon are derived from a novel retroposon, *Nature* 441 (2006) 87–90.
- [26] A. Siepel, G. Bejerano, J.S. Pedersen, A.S. Hinrichs, M. Hou, K. Rosenbloom, H. Clawson, J. Spieth, L.W. Hillier, S. Richards, G.M. Weinstock, R.K. Wilson, R.A. Gibbs, W.J. Kent, W. Miller, D. Haussler, Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes, *Genome. Res.* 15 (2005) 1034–1050.
- [27] G. Bejerano, M. Pheasant, I. Makunin, S. Stephen, W.J. Kent, J.S. Mattick, D. Haussler, Ultraconserved elements in the human genome, *Science* 304 (2004) 1321–1325.