Genomics 106 (2015) 145–150

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Review

STARR-seq — Principles and applications

Felix Muerdter *[,1], Łukasz M. Boryń, Cosmas D. Arnold *[,1]

Research Institute of Molecular Pathology IMP, Vienna Biocenter VBC, Dr Bohr-Gasse 7, 1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 12 March 2015
Revised 19 May 2015
Accepted 8 June 2015
Available online 11 June 2015

Keywords:
Enhancer
STARR-seq
Reporter
Reporter assay
High-throughput
Transcription

ABSTRACT

Differential gene expression is the basis for cell type diversity in multicellular organisms and the driving force of development and differentiation. It is achieved by cell type-specific transcriptional enhancers, which are genomic DNA sequences that activate the transcription of their target genes. Their identification and characterization is fundamental to our understanding of gene regulation. Features that are associated with enhancer activity, such as regulatory factor binding or histone modifications can predict the location of enhancers. Nonetheless, enhancer activity can only be assessed by transcriptional reporter assays. Over the past years massively parallel reporter assays have been developed for large scale testing of enhancers. In this review we focus on the principles and applications of STARR-seq, a functional assay that quantifies enhancer strengths in complex candidate libraries and thus allows activity-based enhancer identification in entire genomes. We explain how STARR-seq works, discuss current uses and give an outlook to future applications.

Contents

1. What are transcriptional enhancers?

How can a single genome give rise to the diverse cell types of an animal? Differential gene expression is the basic process creating this diversity throughout development and differentiation [1–6]. Gene transcription starts at core promoters, which are the sites where the transcription machinery assembles [7]. As core promoters typically only support low-level basal transcription, so called *cis*-regulatory modules (CRMs) or enhancers [8] carry most of the regulatory information in gene expression. These regions act as binding platforms for transcription factors (TFs) and co-factors, which together activate productive transcription. They can be located near the core promoter, often referred to as promoter-proximal, or distal to the target gene and are typically

* Corresponding authors.
E-mail addresses: felix.muerdter@imp.ac.at (F. Muerdter), cosmas.arnold@imp.ac.at (C.D. Arnold).
[1] The authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ygeno.2015.06.001

found within accessible chromatin [1,3,4,9,10]. The combined input of activating and repressing transcription factors bound to an enhancer determines its overall regulatory output in a cell type or tissue-specific manner (e.g. the *even-skipped* stripe 2 enhancer [11], *sparkling* [12], or the interferon-beta enhanceosome [13]; reviewed in [10]). Furthermore, genes with complex expression patterns are often regulated in a modular fashion via multiple enhancers that contribute individual patterns additively [3,14–16]. The modularity of gene regulation and the context-independent functions of enhancers become evident when they are tested *in vivo* using reporter assays. Enhancers typically retain their endogenous activity patterns and, in combination, recapitulate their target genes' endogenous expression patterns. This has been demonstrated for several examples, including the stripe pattern of the *even-skipped* [3] and *blimp1* [17] genes in fruit fly embryos or *sonic hedgehog (shh)* in mice [18], and holds for most tested enhancers genome-wide [16]. Correspondingly, deletion or disruption of a single enhancer can cause domain-specific loss of gene expression, e.g. as shown for *blimp1* in specific stripes [17] or *shh* in limb buds [18]. Overall, this demonstrates that single enhancers can be context independent and are sufficient to recapitulate endogenous expression patterns, even when tested in an ectopic position. For the above reasons, the regulation of gene expression can be studied using functional enhancer assays.

2. How can we identify enhancers?

Enhancers display features at the sequence and chromatin level that distinguish them from non-regulatory genomic regions [5,19]. Active enhancers are bound by TFs and co-factors (eg. CBP/p300; [20]) and are therefore typically found in accessible chromatin regions (DNaseI hypersensitive sites; [21]). Adjacent nucleosomes often contain histones that are post-translationally modified at their C-terminal tails. Active enhancers have been shown to correlate with acetylation of lysine 27 of the adjacent histone H3 (H3K27ac) and methylation of lysine 4 (H3K4me1) [22–25]. In addition, enhancers can be bound and transcribed by RNA polymerase II, which gives rise to enhancer RNAs or eRNAs [26,27] and additionally are often found to be hypomethylated in mammalian cells [28,29]. The expansion of next-generation sequencing technologies during the past decade made it possible to assay these diverse sequence and chromatin properties genome-wide [19], and thus has allowed extensively predicting putative enhancer regions in the genomes of many organisms.

Even though these features correlate with enhancer activity, they cannot assess reporter activity nor measure enhancer strength [19]. Furthermore, these features are not exclusive to active enhancers, which leads to false positive predictions [19,30]. Thus, to validate and functionally characterize enhancers, assays that report on enhancer activity directly are of paramount importance.

3. How can we assess enhancer activity in high throughput?

Traditionally, enhancers have been studied using reporter assays that assess enhancer activity through the expression of various reporter genes: When introduced into a reporter construct upstream of a minimal promoter, enhancers will activate the transcription of a reporter gene, whose expression levels can be either visualized (by LacZ staining, fluorescence or *in situ* hybridization) or quantified using bioluminescence (e.g. in luciferase assays) [19]. The abundance of reporter transcript or protein is in direct relation to the strength of the enhancer. These assays are therefore considered the gold standard in the study of enhancer activity of individual candidate sequences. Nonetheless, such reporter assays require that each candidate be tested individually, which strongly limits the number of candidates.

Recent advances in functional approaches for enhancer activity have overcome the low throughput of these assays and have led to the development of massively parallel reporter assays (MPRAs). Here, individual candidate fragments are tested in parallel instead of one-by-one in separate experiments, creating the challenge to couple the assay's readout of activity to the individual candidates' identities. This has been achieved mainly by two different means with different advantages and drawbacks: in the first scenario individual cells contain only a single candidate fragment each, such that the cellular levels of a GFP reporter can be used to separate cells that contain active or inactive fragments (coupling at the cellular level [31–34]). Alternatively, the reporter transcripts contain information that allows their unique assignment to the respective candidate fragments, e.g. in the form of molecular barcodes (coupling at the sequence/plasmid level [30,35–40]).

The former concept allows for the testing not only of pre-selected candidates, but also the screening of randomly sheared fragments from different sources of DNA such as BACs or DNase I hypersensitive regions [31–34].

The fact that only a single candidate can be tested per cell, however, couples the throughput of such an approach to the number of cells. Additionally, positive cells have to be separated from negative cells by fluorescence-activated cell sorting (FACS). To achieve integration into a large number of cells with a controlled number of integration events per cell, lentiviral strategies have been successfully applied [33]. The fact that the site of integration is random, however, can have strong positional effects on the reporter activity [41]. In order to alleviate such positional effects, site-specific integration into a more defined landing site can be employed [31,32,34]. Here, the low efficiency of site-directed integration further constrains the throughput by demanding a high number of input cells or tissue to generate a certain number of integration-positive cells. Interestingly, whether or not the integration of reporter constructs into chromosomes has a differential effect on reporter activity, in comparison to plasmid-based constructs, has not been systematically assessed. Two studies in flies and yeast even suggest that the measured activity for enhancers agree, when tested in integrated or episomal constructs [34,42]. Finally, the binary readout (active vs. inactive) of above methodologies makes it hard to quantitatively assess the strength of the tested candidate.

Reporter activity can also be measured as abundance of the reporter transcript through RNA-seq, which quantitatively reflects the strength of the enhancer. Here, candidate sequences can drive the transcription of a reporter construct, which harbors a unique barcode in its transcript that informs on the identity of the enhancer candidate. Massively parallel sequencing of the barcodes encoded as cDNA and subsequent paring with upstream candidate sequences thus reflects the activity of the candidate sequence [30,35–40]. In this scenario, several candidate sequences can be interrogated per cell, making high throughput more feasible. One challenge of these approaches, however, is to match the candidates and their respective barcodes. If the candidates are paired randomly with barcodes during the cloning step, the resulting libraries have to be sequenced first in order to create the assignments [37]. Synthesizing the candidate-barcode-pair in a single reaction, thereby creating known pairs, can circumvent this complication [36]. Nonetheless, synthetic oligonucleotides are still limited in size and are costly when produced in high numbers [43], which would be required for the production of highly complex libraries. Given above limitations, these assays are most often used to systematically probe candidate sequences rather than for the identification of enhancers.

To screen entire genomes for enhancers based on their activity, we recently developed a quantitative method that directly couples the identity of a candidate sequence to its activity for millions of sequences in parallel: Self-transcribing active regulatory region sequencing (STARR-seq) (Fig. 1A).

Testing of defined enhancer candidates are discussed elsewhere [19], which is why we will focus more on the aspect of high throughput screening for enhancer activity for the remainder of this review.

4. What is STARR-seq?

STARR-seq is a massively parallel reporter assay to identify transcriptional enhancers directly based on their activity in entire genomes and to assess their activity quantitatively [42].

5. How does STARR-seq work?

Enhancer activity is directly linked to the underlying DNA sequence and measured as presence of the resulting reporter transcripts among cellular RNA by deep sequencing.

Specifically, DNA fragments are cloned downstream of a core promoter and into the 3′ UTR of a reporter gene. Active enhancers will transcribe themselves and become part of the resulting reporter transcripts (Fig. 1A). This setup allows the simultaneous testing of millions of DNA sequences in a highly complex reporter library and also ensures that the identified sequences act as *bona fide* enhancers (rather than for example promoters) as they activate transcription from a remote position (Fig. 1A) [42].

6. What are the features of the reporter library?

Candidate DNA fragments can be obtained from arbitrary sources of DNA, including genomic DNA [42,44–46], targeted regions via bacterial artificial chromosomes (BACs) [42,46], DNA fragments enriched for regions of interest such as open chromatin, TF binding sites, or predicted enhancers [47] as well as synthetic DNA. The size of the candidate DNA fragments can be of a wide range including sizes that fully cover known enhancers.

7. What are the features of STARR-seq?

As STARR-seq is an ectopic, plasmid-based assay, the measured activity directly reflects the regulatory capacity of enhancer sequences. This measurement is not affected by the location of the candidate sequences within the transcript or their orientation [42] and accurately reflects activity changes after cellular signaling such as hormone treatment highlighting its episomal responsiveness [44]. Due to its episomal nature, STARR-seq is unlikely to suffer from position effects resulting from random genomic integration as has been observed for integrated reporter assays [41]. Furthermore, integration of the reporter constructs in the genome in order to propagate them to daughter cell populations is not necessary due to the short timeframe of STARR-seq, which is in the order of a single cell cycle or less for most cell types. The activity of candidate sequences tested independently in luciferase assays and STARR-seq is linearly correlated in flies and humans. Thus, STARR-seq reports quantitatively on enhancer activity and constitutes a genome-wide equivalent of luciferase assays. In addition, enhancers identified by STARR-seq are active after random integration into the genome of cell lines, as well as *in vivo* in transgenic flies after site-specific integration [42]. Taken together, STARR-seq draws genome-wide cell type-specific quantitative enhancer activity maps of any cell type that allows efficient delivery of the reporter library [42].

Although STARR-seq does not measure the enhancer's activity in its endogenous chromatin environment, the majority of enhancers identified by STARR-seq overlap accessible chromatin and carry enhancer-typical histone modifications [42]. This indicates that they are functional in their endogenous context. In addition, STARR-seq can also measure enhancer activities of sequences that lie within inaccessible chromatin. These “closed enhancers” are marked by H3K4me1, suggesting that they are recognized as *bona fide* enhancers by the cell's *trans*-regulatory environment. As they are marked by H3K27me3, they are likely actively silenced by Polycomb, presumably at the chromatin level. These regions would be invisible to methods that predict enhancers based on chromatin features (DHS-seq, ChIP-seq etc.) alone, yet provide interesting starting points to investigate chromatin-mediated silencing and the extent to which this form of silencing is involved in gene regulation [42].

8. What has STARR-seq been used for?

8.1. Enhancer identification, validation and characterization in flies and human

STARR-seq can be used to ask fundamental questions of transcriptional regulation and enhancer biology. Cell type-specific enhancers drive differential gene expression, hence, identifying such enhancers can be of utmost importance to our understanding of development and differentiation. We generated quantitative enhancer activity maps that describe the enhancer activity landscape of three *Drosophila melanogaster* cell types of developmentally different origin, one derived from embryos (S2), one from larval brain (BG3), and one from adult ovaries (OSC) by screening the entire fly genome [42]. This revealed thousands of enhancers, exhibiting an activity spectrum ranging from strictly cell type-specific to equally active across cell types and their genomic locations. Importantly, cell type-specific gene expression levels were correlated to the combined activities of the flanking enhancers within each cell type and across cell types [42], which links differential gene expression to differences in enhancer activity and demonstrates that ectopic assays can accurately assess cell type-specific enhancer activities (Fig. 1B).

In human, the 20-times larger genome poses considerable challenges to the cloning of high-coverage libraries, the screening of large numbers of cells and processing of high levels of RNA. Thus, STARR-seq has been applied to libraries of reduced complexity to test selected enhancer candidates [47] or to screen defined genomic regions in an unbiased fashion [42]. The agreement of luciferase assays and STARR-seq signal further showed that STARR-seq is also quantitative in mammalian cells [42,47]. Spicuglia and colleagues captured fragments from sonicated genomic DNA that corresponded to candidate regulatory regions predicted based on chromatin accessibility and TF binding. This allowed them to test 7152 candidate regions in parallel in their murine T-cell model (Fig. 1B), revealing 2279 weak and 433 strong enhancers, but also demonstrating that many predicted candidates were negative [47], emphasizing the need for functional validation of such predictions.

8.2. DNA sequence features of active enhancers

The genome-wide enhancer activity maps from STARR-seq provided large collections of functional enhancer sequences and negative controls. This unique set of enhancers enabled us to derive *cis*-regulatory sequence rules as well as functionally important TF binding motifs by computational analysis, thus explaining cell type-specific enhancer function from the DNA sequence. In addition, STARR-seq enabled identification of a novel class of enhancer sequence elements (dinucleotide repeat motifs — DRM), which are important for the activity of broadly active enhancers [48].

8.3. Dynamic changes of enhancer activity induced by signaling

Signaling pathways often cause changes in gene expression levels. Signaling-responsive enhancers direct these changes. STARR-seq allows assessing the underlying differences in enhancer activities by comparing genome-wide enhancer activity maps before and after signal induction. For example, steroid hormones regulate gene expression through their nuclear receptors, which bind to specific DNA motifs and act as transcription factors, thus influencing the activity of the bound enhancer. Using ecdysone signaling in *D. melanogaster* S2 and OSCs hundreds of hormone dependent, cell type-specific enhancers could be identified. In addition this allowed extracting the underlying *cis*-regulatory logic responsible for a cell type-specific hormone response from these enhancer sequences (Fig. 1B) [44]. This study demonstrates how STARR-seq

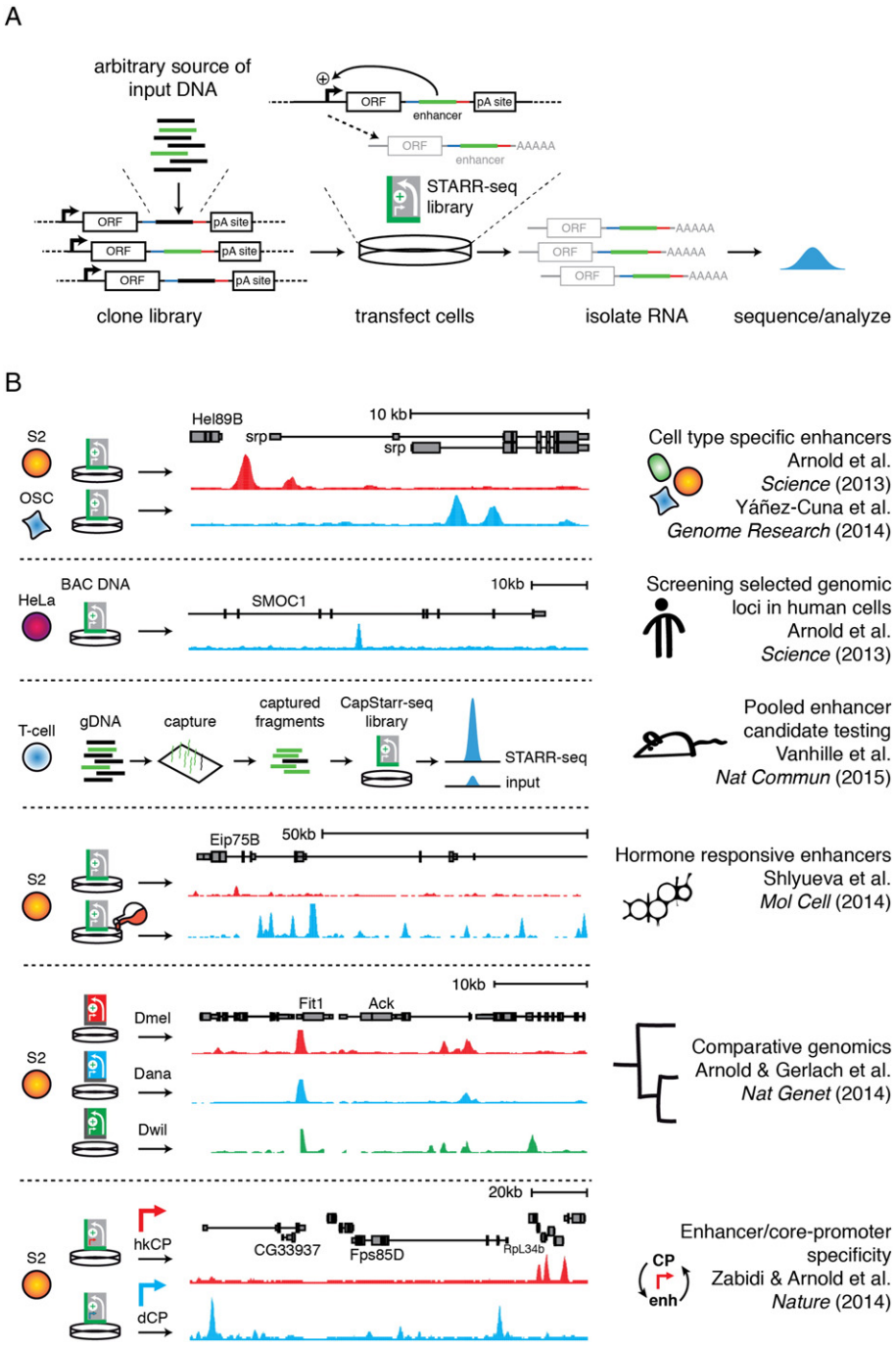

Fig. 1. STARR-seq — Principles and Applications. (A) Overview of the STARR-seq pipeline. First a reporter library is cloned from an arbitrary DNA source, which can be sonicated genomic DNA for comprehensive genome-wide screens. The reporter library is transfected into cultured cells and the reporter transcripts are isolated from the pool of total RNA after 24 h. After cDNA synthesis and PCR amplification of the self-transcribing sequences, deep sequencing is conducted and the resulting reads are mapped to the reference genome. The enrichment of reporter cDNA over input directly and quantitatively reflects enhancer activity [42]. (B) Shown are graphical abstracts of STARR-seq applications displaying original data (for details refer to the original publication and the main text) and a schematic of STARR-seq in a murine T-cell model using a capture approach (CapStarr-seq) [47].

can be used to study the principles of signal-dependent gene expression, which are important in many aspects of cellular transitions and differentiation.

8.4. Impact of cis-regulatory sequence variation on enhancer activity and evolution

Furthermore, STARR-seq allows screening multiple *cis*-regulatory genomes in a single cell type, i.e. in the same *trans*-regulatory environment, enabling powerful comparative analyses of differential enhancer activities that arise from sequence variation.

Screening the genomes of 5 *Drosophila* species (spanning an evolutionary distance of 30–40 Ma) in a single *D. melanogaster* cell type (S2) revealed that a large portion of *D. melanogaster* orthologous enhancers is functionally conserved, presumably due to stabilizing turnover of TF motifs. Interestingly, functional enhancers can also be gained within relatively short evolutionary timespans without apparent adaptive selection, yet can be involved in changes of gene expression *in vivo* (Fig. 1B) [45]. The above approach can be extended to studying

sequence variation across different species, within selected populations, or between DNA from healthy versus disease-tissue (e.g. in cancer) with respect to phenotypic variation or disease.

8.5. Mechanistic aspects of transcriptional regulation

Finally, the functional readout and the defined setup of STARR-seq makes it possible to ask basic mechanistic questions of enhancer biology. For example, screening the entire fly genome by STARR-seq using different core promoters derived from either ubiquitously expressed housekeeping genes or from developmentally regulated and cell type-specific genes revealed thousands of enhancers that are specific to either of these two classes of core promoters. This suggests the existence of two major transcriptional programs, one for ubiquitous expression of housekeeping genes through ubiquitously active, promoter proximal enhancers and one for the expression of developmental and cell type-specific genes. The latter involves more distally located and often cell type-specific enhancers. Thus, STARR-seq revealed the separation of housekeeping and developmental gene regulation through global enhancer–core-promoter specificity and motif analysis combined with ChIP-seq data identified the responsible TFs (Fig. 1B) [46]. This work demonstrates that STARR-seq is a powerful tool to address longstanding and fundamental questions of gene regulation [49].

9. What comes next?

In recent years, advances in sequencing technology have allowed large-scale predictions of enhancers in many cell types and tissues [19]. There is a need to validate and test these predictions using the functional assays described above. Additionally, the comprehensiveness of functional enhancer activity assays at a genome-wide scale will allow the field to re-evaluate and refine the models currently used in enhancer prediction [30]. These assays can further assess the modularity and sufficiency of enhancers and it will be exciting to see recent advances in genome-editing tools such as the CRISPR/Cas9 system applied to test their requirement for gene regulation in genomic loss-of-function models [50–52].

The quantitative assessment of enhancer activity is also especially interesting in biological systems that exhibit strong changes in gene expression. For example, it will be exciting to see the dynamics of enhancer strength during cellular differentiation and identify the causal regulatory regions. Another intriguing cellular transition that is extensively studied at the enhancer level is malignant transformation in mammalian cells. STARR-seq in combination with co-factor disruption by CRISPRi [53] or inhibition with small molecule inhibitors could shed light on the involvement of certain co-factors in cancer. Some of these factors are especially interesting, as they might become important targets in cancer therapy [54]. Finally, it will be exciting to see STARR-seq and similar approaches to be applied to diverse tissues and cell types *in vivo*.

In this review, we highlight a powerful functional assay that enables characterization of enhancers directly based on their activity. We foresee that this and related assays will continue to advance our understanding of transcriptional regulation and are excited to see future developments in this area of research.

Acknowledgments

The authors thank A. Stark (Research Institute of Molecular Pathology IMP, Vienna, Austria) for discussions. We apologize to all colleagues whose work could not be discussed or referenced owing to space limitations. F.M. is an EMBO long-term fellow (EMBO ALTF 491–2014). L.M.B. and C.D.A. are supported by a European Research Council (ERC) Starting grant (no. 242922) awarded to A. Stark. Basic research at the IMP is supported by Boehringer Ingelheim GmbH.

References

[1] M. Levine, R. Tjian, Transcription regulation and animal diversity, Nature 424 (2003) 147–151, http://dx.doi.org/10.1038/nature01763.
[2] M. Levine, Transcriptional enhancers in animal development and evolution, Curr. Biol. 20 (2010) R754–R763, http://dx.doi.org/10.1016/j.cub.2010.06.070.
[3] M. Levine, C. Cattoglio, R. Tjian, Looping back to leap forward: transcription enters a new era, Cell 157 (2014) 13–25, http://dx.doi.org/10.1016/j.cell.2014.02.009.
[4] C. Buecker, J. Wysocka, Enhancers as information integration hubs in development: lessons from genomics, Trends Genet. 28 (2012) 276–284, http://dx.doi.org/10.1016/j.tig.2012.02.008.
[5] E. Calo, J. Wysocka, Modification of enhancer chromatin: what, how, and why? Mol. Cell 49 (2013) 825–837, http://dx.doi.org/10.1016/j.molcel.2013.01.038.
[6] M. Bulger, M. Groudine, Functional and mechanistic diversity of distal transcription enhancers, Cell 144 (2011) 327–339, http://dx.doi.org/10.1016/j.cell.2011.01.024.
[7] S.T. Smale, J.T. Kadonaga, The RNA polymerase II core promoter, Annu. Rev. Biochem. 72 (2003) 449–479, http://dx.doi.org/10.1146/annurev.biochem.72.121801.161520.
[8] J. Banerji, S. Rusconi, W. Schaffner, Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences, Cell 27 (1981) 299–308.
[9] C.-T. Ong, V.G. Corces, Enhancer function: new insights into the regulation of tissue-specific gene expression, Nat. Rev. Genet. 12 (2011) 283–293, http://dx.doi.org/10.1038/nrg2957.
[10] F. Spitz, E.E.M. Furlong, Transcription factors: from enhancer binding to developmental control, Nat. Rev. Genet. 13 (2012) 613–626, http://dx.doi.org/10.1038/nrg3207.
[11] S. Small, A. Blair, M. Levine, Regulation of even-skipped stripe 2 in the *Drosophila* embryo, EMBO J. 11 (1992) 4047–4057.
[12] C.I. Swanson, N.C. Evans, S. Barolo, Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer, Dev. Cell 18 (2010) 359–370, http://dx.doi.org/10.1016/j.devcel.2009.12.026.
[13] D. Thanos, T. Maniatis, Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome, Cell 83 (1995) 1091–1100.
[14] M.I. Arnone, E.H. Davidson, The hardwiring of development: organization and function of genomic regulatory systems, Development 124 (1997) 1851–1864.
[15] A. Visel, J.A. Akiyama, M. Shoukry, V. Afzal, E.M. Rubin, L.A. Pennacchio, Functional autonomy of distant-acting human enhancers, Genomics (2009) 1–5, http://dx.doi.org/10.1016/j.ygeno.2009.02.002.
[16] E.Z. Kvon, T. Kazmar, G. Stampfel, J.O. Yáñez-Cuna, M. Pagani, K. Schernhuber, et al., Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo, Nature 512 (2014) 91–95, http://dx.doi.org/10.1038/nature13395.
[17] E.Z. Kvon, G. Stampfel, J.O. Yáñez-Cuna, B.J. Dickson, A. Stark, HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature, Genes Dev. 26 (2012) 908–913, http://dx.doi.org/10.1101/gad.188052.112.
[18] L.A. Lettice, S.J.H. Heaney, L.A. Purdie, L. Li, P. de Beer, B.A. Oostra, et al., 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.
[19] D. Shlyueva, G. Stampfel, A. Stark, Transcriptional enhancers: from properties to genome-wide predictions, Nat. Rev. Genet. 15 (2014) 272–286, http://dx.doi.org/10.1038/nrg3682.
[20] A. Visel, M.J. Blow, Z. Li, T. Zhang, J.A. Akiyama, A. Holt, et al., ChIP-seq accurately predicts tissue-specific activity of enhancers, Nature 457 (2009) 854–858, http://dx.doi.org/10.1038/nature07730.
[21] A.P. Boyle, S. Davis, H.P. Shulha, P. Meltzer, E.H. Margulies, Z. Weng, et al., High-resolution mapping and characterization of open chromatin across the genome, Cell 132 (2008) 311–322, http://dx.doi.org/10.1016/j.cell.2007.12.014.
[22] N.D. Heintzman, R.K. Stuart, G. Hon, Y. Fu, C.W. Ching, R.D. Hawkins, et al., Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome, Nat. Genet. 39 (2007) 311–318, http://dx.doi.org/10.1038/ng1966.
[23] N.D. Heintzman, G.C. Hon, R.D. Hawkins, P. Kheradpour, A. Stark, L.F. Harp, et al., Histone modifications at human enhancers reflect global cell-type-specific gene expression, Nature 459 (2009) 108–112, http://dx.doi.org/10.1038/nature07829.
[24] A. Rada-Iglesias, R. Bajpai, T. Swigut, S.A. Brugmann, R.A. Flynn, J. Wysocka, A unique chromatin signature uncovers early developmental enhancers in humans, Nature 470 (2011) 279–283, http://dx.doi.org/10.1038/nature09692.
[25] S. Bonn, R.P. Zinzen, C. Girardot, E.H. Gustafson, A. Perez-Gonzalez, N. Delhomme, et al., Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development, Nat. Genet. 44 (2012) 148–156, http://dx.doi.org/10.1038/ng.1064.
[26] T.-K. Kim, M. Hemberg, J.M. Gray, A.M. Costa, D.M. Bear, J. Wu, et al., Widespread transcription at neuronal activity-regulated enhancers, Nature 465 (2010) 182–187, http://dx.doi.org/10.1038/nature09033.
[27] R. Andersson, C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, et al., An atlas of active enhancers across human cell types and tissues, Nature 507 (2014) 455–461, http://dx.doi.org/10.1038/nature12787.
[28] F. Schlesinger, A.D. Smith, T.R. Gingeras, G.J. Hannon, E. Hodges, De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements, Genome Res. 23 (2013) 1601–1614, http://dx.doi.org/10.1101/gr.157271.113.
[29] M.B. Stadler, R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Schöler, et al., DNA-binding factors shape the mouse methylome at distal regulatory regions, Nature 480 (2011) 490–495, http://dx.doi.org/10.1038/nature10716.
[30] J.C. Kwasnieski, C. Fiore, H.G. Chaudhari, B.A. Cohen, High-throughput functional testing of ENCODE segmentation predictions, Genome Res. 24 (2014) 1595–1602, http://dx.doi.org/10.1101/gr.173518.114.

[31] S.S. Gisselbrecht, L.A. Barrera, M. Porsch, A. Aboukhalil, P.W. Estep, A. Vedenko, et al., Highly parallel assays of tissue-specific enhancers in whole *Drosophila* embryos, Nat. Methods 10 (2013) 774–780, http://dx.doi.org/10.1038/nmeth.2558.
[32] D.E. Dickel, Y. Zhu, A.S. Nord, J.N. Wylie, J.A. Akiyama, V. Afzal, et al., Function-based identification of mammalian enhancers using site-specific integration, Nat. Methods 11 (2014) 566–571, http://dx.doi.org/10.1038/nmeth.2886.
[33] M. Murtha, Z. Tokcaer-Keskin, Z. Tang, F. Strino, X. Chen, Y. Wang, et al., FIREWACh: high-throughput functional detection of transcriptional regulatory modules in mammalian cells, Nat. Methods 11 (2014) 559–565, http://dx.doi.org/10.1038/nmeth.2885.
[34] E. Sharon, Y. Kalma, A. Sharp, T. Raveh-Sadka, M. Levo, D. Zeevi, et al., Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters, Nat. Biotechnol. 30 (2012) 521–530, http://dx.doi.org/10.1038/nbt.2205.
[35] J. Nam, E.H. Davidson, Barcoded DNA-tag reporters for multiplex cis-regulatory analysis, PLoS ONE 7 (2012) e35934, http://dx.doi.org/10.1371/journal.pone.0035934.
[36] A. Melnikov, A. Murugan, X. Zhang, T. Tesileanu, L. Wang, P. Rogov, et al., Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay, Nat. Biotechnol. 30 (2012) 271–277, http://dx.doi.org/10.1038/nbt.2137.
[37] R.P. Patwardhan, J.B. Hiatt, D.M. Witten, M.J. Kim, R.P. Smith, D. May, et al., Massively parallel functional dissection of mammalian enhancers in vivo, Nat. Biotechnol. 30 (2012) 265–270, http://dx.doi.org/10.1038/nbt.2136.
[38] R.P. Smith, L. Taher, R.P. Patwardhan, M.J. Kim, F. Inoue, J. Shendure, et al., Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model, Nat. Genet. 45 (2013) 1021–1028, http://dx.doi.org/10.1038/ng.2713.
[39] J.C. Kwasnieski, I. Mogno, C.A. Myers, J.C. Corbo, B.A. Cohen, Complex effects of nucleotide variants in a mammalian cis-regulatory element, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 19498–19503, http://dx.doi.org/10.1073/pnas.1210678109.
[40] I. Mogno, J.C. Kwasnieski, B.A. Cohen, Massively parallel synthetic promoter assays reveal the in vivo effects of binding site variants, Genome Res. 23 (2013) 1908–1915, http://dx.doi.org/10.1101/gr.157891.113.
[41] W. Akhtar, J. de Jong, A.V. Pindyurin, L. Pagie, W. Meuleman, J. de Ridder, et al., Chromatin position effects assayed by thousands of reporters integrated in parallel, Cell 154 (2013) 914–927, http://dx.doi.org/10.1016/j.cell.2013.07.018.
[42] C.D. Arnold, D. Gerlach, C. Stelzer, Ł.M. Boryń, M. Rath, A. Stark, Genome-wide quantitative enhancer activity maps identified by STARR-seq, Science 339 (2013) 1074–1077, http://dx.doi.org/10.1126/science.1232542.
[43] S. Kosuri, G.M. Church, Large-scale de novo DNA synthesis: technologies and applications, Nat. Methods 11 (2014) 499–507, http://dx.doi.org/10.1038/nmeth.2918.
[44] D. Shlyueva, C. Stelzer, D. Gerlach, J.O. Yáñez-Cuna, M. Rath, L.M. Boryń, et al., Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin, Mol. Cell 54 (2014) 180–192, http://dx.doi.org/10.1016/j.molcel.2014.02.026.
[45] C.D. Arnold, D. Gerlach, D. Spies, J.A. Matts, Y.A. Sytnikova, M. Pagani, et al., Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution, Nat. Genet. (2014)http://dx.doi.org/10.1038/ng.3009.
[46] M.A. Zabidi, C.D. Arnold, K. Schernhuber, M. Pagani, M. Rath, O. Frank, et al., Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation, Nature (2014)http://dx.doi.org/10.1038/nature13994.
[47] L. Vanhille, A. Griffon, M.A. Maqbool, J. Zacarias-Cabeza, L.T.M. Dao, N. Fernandez, et al., High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq, Nat. Commun. 6 (2015) 6905, http://dx.doi.org/10.1038/ncomms7905.
[48] J.O. Yáñez-Cuna, C.D. Arnold, G. Stampfel, L.M. Boryń, D. Gerlach, M. Rath, et al., Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features, Genome Res. 24 (2014) 1147–1156, http://dx.doi.org/10.1101/gr.169243.113.
[49] J. van Arensbergen, B. van Steensel, H.J. Bussemaker, In search of the determinants of enhancer–promoter interaction specificity, Trends Cell Biol. 24 (2014) 695–702, http://dx.doi.org/10.1016/j.tcb.2014.07.004.
[50] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, et al., Multiplex genome engineering using CRISPR/Cas systems, Science 339 (2013) 819–823, http://dx.doi.org/10.1126/science.1231143.
[51] P. Mali, L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. DiCarlo, et al., RNA-guided human genome engineering via Cas9, Science 339 (2013) 823–826, http://dx.doi.org/10.1126/science.1232033.
[52] S. Gröschel, M.A. Sanders, R. Hoogenboezem, E. de Wit, B.A.M. Bouwman, C. Erpelinck, et al., A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia, Cell 157 (2014) 369–381, http://dx.doi.org/10.1016/j.cell.2014.02.019.
[53] L.A. Gilbert, M.A. Horlbeck, B. Adamson, J.E. Villalta, Y. Chen, E.H. Whitehead, et al., Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation, Cell 159 (2014) 647–661, http://dx.doi.org/10.1016/j.cell.2014.09.029.
[54] P. Valent, J. Zuber, BRD4: a BET(ter) target for the treatment of AML? Cell Cycle 13 (2014) 689–690, http://dx.doi.org/10.4161/cc.27859.