CRISPR/Cas9 Approaches to Investigate the Noncoding Genome

Davide Seruggia and Lluis Montoliu

Abstract Studies of gene function in mice have been supported during the past decade by a nearly exhaustive collection of mutants, systematically obtained by homologous recombination in murine ES cells. Unfortunately, the study of the noncoding fraction of the genome did not benefit from the same valuable resources. Nevertheless, increasing evidence of the relevance of this fraction of the vertebrate genome has been accumulated in the past years. Comprehensive maps of histone modifications, methylation patterns, and DNA-binding protein occupancies have been made available to predict key regulatory elements through the work of various international collaborative consortia, such as ENCODE. Comparing these maps with data from genome-wide association studies (GWAS) suggested that variants in noncoding sequence elements might be involved in several traits and disease conditions. Therefore, there is an urgent need for accurate functional tests and genetic modelling of noncoding elements. In this chapter, we propose a number of strategies to test hypothesis regarding noncoding DNA elements, by taking advantage of the most recent genome editing techniques, namely, CRISPR/Cas9 approaches.

Keywords Enhancer • Insulator • ChIP-seq • Transcription factor • Chromatin • Epigenetics

D. Seruggia
Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Campus Cantoblanco, Darwin 3, 28049 Madrid, Spain

CIBERER-ISCIII, Madrid, Spain

Department of Hematology/Oncology, Boston Children's Hospital, Karp Family Research Laboratories 7, 1 Blackfan Circle, Boston, MA 02115, USA

L. Montoliu (✉)
Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Campus Cantoblanco, Darwin 3, 28049 Madrid, Spain

CIBERER-ISCIII, Madrid, Spain
e-mail: montoliu@cnb.csic.es

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_2

Introduction

Protein-coding genes account for just a small fraction of the vertebrate genomes, estimated to be about 2 %. The remainder, 98 % of the genome, is composed of repetitive DNA elements and, most importantly, of a large variety of noncoding sequences that share a considerable degree of conservation across species [1]. Such conserved, noncoding elements are required for accurate regulation of gene expression and are involved in determining cell type identity and function. Active and inactive regulatory sequences are associated with specific biochemical marks. DNA methylation, histone modifications, and protein occupancy are predictive for the function and the state of a given element [2]. Active enhancer sequences are typically marked by mono-methylation of lysine 4 of histone H3 (H3K4me1), by acetylation of lysine 27 (H3K27ac), and localize in open chromatin regions that show hypersensitivity to DNase I digestion (Fig. 1). In addition, the distribution of two enhancer-binding proteins, CHD7 and P300 [3], can be used with confidence to identify enhancers. Binding profiles of transcription factors in gene neighbourhoods allow inferring the underlying networks of gene regulation [4]. Interestingly, the distribution of chromatin marks and transcription factors that decorates the noncoding fraction of the genome varies between different cell types and changes dynamically during development and differentiation [5]. A direct correlation between transcription profile and chromatin signatures has been described in many cell types or tissues, including embryonic stem (ES) cells [6], further triggering the interest for noncoding, regulatory elements in the fields of development and stem cell biology. Several sequence variants in noncoding elements have already been found associated with human traits as well as with disease conditions [7–11]. For example, genome-wide association studies (GWAS) highlighted that a common trait such as eye colour is strongly associated with a DNA polymorphism lying 21 kb upstream of the pigmentation-related OCA2 gene. Molecular analyses indicated that a particular single-nucleotide polymorphism (SNP) is located within a OCA2 enhancer and, interestingly, the rs12913832 C-allele is associated with decreased OCA2 expression, reduced transcription factor recruitment, and chromatin looping [11]. Hair colour also results, in part, from variants at noncoding sequences [12]. Thus, increasing evidence denotes the functional role of noncoding variants in both human traits and disease. Therefore, there is a pressing need of adequate modelling of noncoding variants and mutations. This chapter aims to provide an overview of the experimental approaches that can be used to study the role of noncoding DNA elements and to obtain their inactivation in model systems using the CRISPR/Cas9 system.

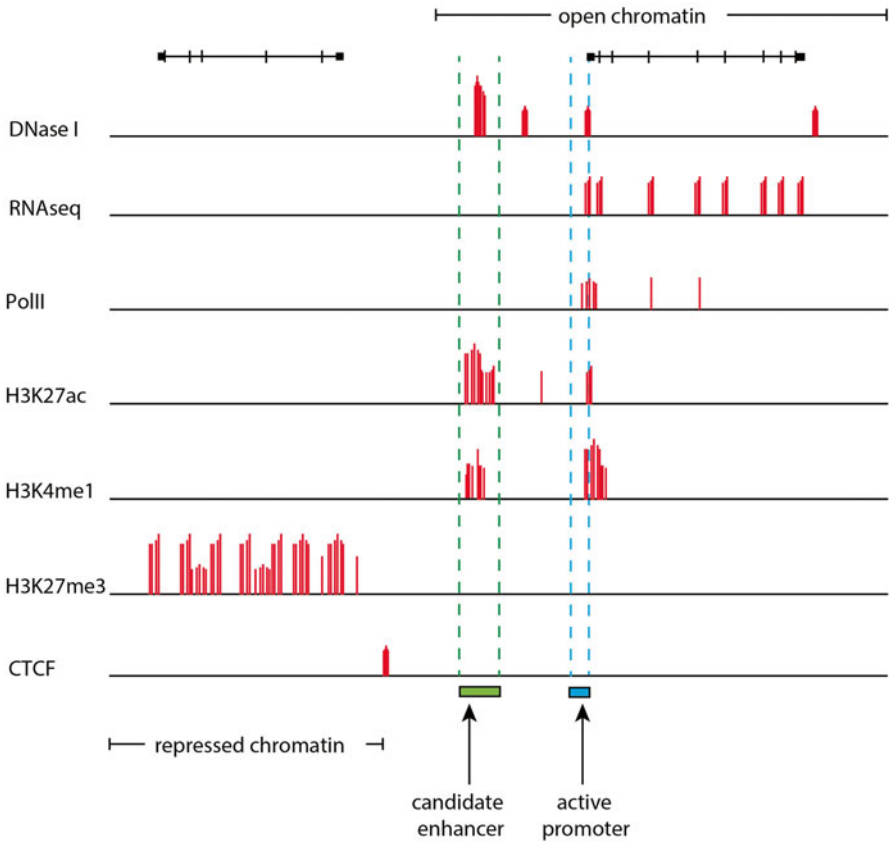

Fig. 1 Chromatin marks guide the identification of active regulatory elements. ENCODE and Roadmap Epigenomics datasets can be browsed to identify putative regulatory elements. DNaseI-seq tracks highlight open chromatin regions. RNA-seq tracks can be used to identify actively transcribing genes. Similarly, RNA polymerase II (PolII) ChIP-seq tracks marks transcription. H3K27ac is a typical mark of active enhancers, whereas H3K4me1 can be used to identify active promoters. In contrast, H3K27me3 marks repressed chromatin. CTCF protein is associated with chromatin boundaries and marks the transition between open and closed chromatin

Genetic Manipulations of Noncoding Sequences in the Prenuclease Era

Loss-of-function alleles of nearly all protein-coding genes have been obtained by homologous recombination in mouse ES cells [13]. On paper, the same experimental approach could be applied to inactivate noncoding elements. However, only a handful of noncoding elements have been inactivated using this strategy in ES cells [14–16]. In fact, the noncoding fraction of the genome is particularly enriched in repetitive elements, a characteristic that poses difficulties and challenges the design of optimal targeting vectors, because the presence of non-unique DNA sequences

within the homology arms is obviously detrimental for targeting efficiency. Hence, most targeting events would be expected to occur elsewhere, outside the desired noncoding DNA element, but at genomic locations rich in similar repetitive sequences.

Furthermore, a recent study highlighted a bias associated with gene targeting in ES cells. In fact, by comparing the genome of 129/Sv and C57BL/6J, the mouse strains that are most commonly used as source of ES cells and as recipient embryos, respectively, more than 1000 passenger mutations in the vicinity of coding genes have been identified, which can confound the interpretation of the associated mutant phenotypes [17].

An alternative strategy that has been largely explored in previous years involves the use of large, genomic-type, transgenes [18]. Instead of targeting regulatory elements at their endogenous locus, several laboratories reproduced inactivating mutations within large genomic constructs, such as those included with artificial chromosome type of transgenes (i.e., bacterial artificial chromosomes, BACs; or yeast artificial chromosomes, YACs). There, the desired mutation is built in the context of a large genomic-type transgene. Using bacteria or yeast, the efficiency of homologous recombination is much higher and the handling of large numbers of clones is easier and also less expensive [19]. In addition, the recombined constructs are delivered to cells or model organisms, and the effect of the mutation in noncoding DNA elements is often read through the activity of a reporter gene included in the transgene [20]. Alternatively, the modified BAC or YAC can be introduced in mice where the endogenous gene was previously inactivated [21]. Nevertheless, this approach suffers from a number of drawbacks. First, large constructs are not easy to manipulate and not all laboratories succeeded in establishing the required protocols. Second, modelling a mutation in a large transgene would in turn introduce several other non-isogenic variants into the model. Finally, copy number, site of integration, and transgene integrity may severely affect the phenotype, leading to complex phenotypes resulting from a mixture of variables, where variegated expression of the transgenes can be also a confounding factor [22, 23].

The CRISPR/Cas9 system, with its high efficiency and flexibility, seems to be the ideal candidate to fill the existing gap in the genetic modelling and functional assessment of noncoding sequences in vitro and in nearly all model organisms [24, 25]. By using this system, the limitations that we described associated with gene targeting in ES cells and large transgenes can be easily overcome. In particular, with the use of CRISPR/Cas9 approaches the requirement for sequence homology can be reduced to just 20 base pairs, relatively easy to find even within stretches of repetitive DNA sequences. And, most importantly, the modification occurs precisely at the endogenous locus, hence avoiding any chromosomal position effect [23].

In this chapter we illustrate three distinct strategies for the genetic perturbation of noncoding elements using CRISPR/Cas9 approaches. First, we introduce the inactivation by chromosomal deletion (Fig. 2a). Next, we discuss a novel technique known as *epigenome editing* (Fig. 2b). Finally, we discuss strategies to target particular non-protein-coding genes, such as miRNA and lincRNAs.

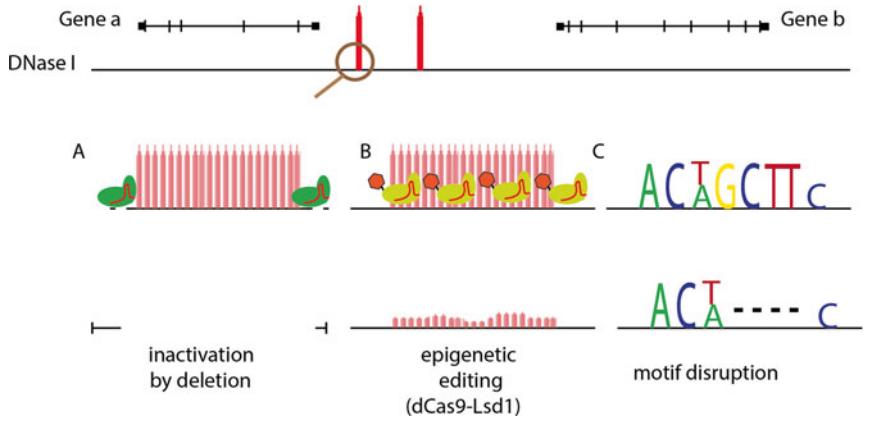

Fig. 2 Different CRISPR/Cas9 approaches. (**a**) Inactivation by deletion: two single guide (sg)RNAs are designed to flank a chromatin region of interest. When simultaneous double-stranded breaks are produced, the intervening DNA is deleted. (**b**) Epigenetic editing: sgRNAs are designed across the target sequence and delivered with a dCas9 protein fused with the desired chromatin remodelling catalytic domain. For example, the Lsd1 catalytic domain is able to inactivate active enhancers. (**c**) Motif disruption. When the DNA-binding specificities of a transcription factor are well characterized and defined, a single sgRNA can be devised to target specifically the nucleotides that constitute a specific DNA-binding site. Upon NHEJ DNA repair, *indel* mutations will be introduced at the DNA-binding motif, potentially abolishing the binding of the cognate transcription factor

Editing by Chromosomal Deletion

Noncoding regulatory elements are contained within arrays of transcription factor-binding sites. Their mechanism of action often relies on the cooperative occupancy of multiple DNA-binding proteins. Binding of a specific transcription factor to a DNA sequence can be anticipated by the presence of a particular motif or consensus sequence. However, DNA-binding motifs are not always completely predictive because many consensus-binding sequences are poorly defined. Position weight matrices (PWM)-based prediction of DNA-binding consensus highlighted that the majority of transcription factors tolerate multiple point mutations at their target sequence [26]. Hence, different from protein-coding genes that can be disrupted introducing frameshift mutations at their sequence, point mutations might be insufficient to fully inactivate complex regulatory elements. One thoughtful approach would be to remove the noncoding DNA element from the genome, within a larger DNA sequence, including multiple predicted DNA-binding sites, that is hypothetically linked with a function (typically, a 1- to 2-kb large sequence). Next, in a second round of experiments, smaller overlapping deletions can be produced to fine-map core elements or to dissect the differential relevance associated with each of the individual DNA sequences. The nonhomologous end-joining DNA repair route (NHEJ), triggered by the CRISPR/Cas9 approach, can be exploited to obtain such alleles, promoting religation of two distal DNA ends generated by adjacent and

simultaneous double-strand breaks (DSB); this would in turn lead to the loss of the intervening DNA sequences existing between the two DSBs (Fig. 2a). This method represents a simple and effective approach to obtain the targeted inactivation of noncoding elements, in their endogenous genomic context, by designing two targeted nucleases (i.e., in a CRISPR/Cas9 approach: two single guide RNAs, sgRNA, to drive the DSB caused by Cas9). Through surgical removal of the putative regulatory element, alteration in gene expression level of nearby and distal genes can be measured [27], as well as alterations in chromatin marks at the surrounding sequences.

In our laboratory, we have used a CRISPR/Cas9 deletion approach to test the in vivo relevance of a regulatory element found upstream from the mouse *Tyr* gene (encoding tyrosinase, the first and fundamental enzyme in the biosynthetic pathway of melanin) that we had previously investigated using YAC-based and standard transgenes [21, 28, 29]. The *Tyr* 5′-upstream region contains multiple DNA-binding motifs and it is marked by EP300, a DNA-binding protein that often decorates active enhancers, and by H3K4me1 [23]. Through delivery of two sgRNAs flanking this element and the Cas9 mRNA to mouse fertilized eggs, we have generated several deletion alleles with high frequency [23]. When homozygous deletions are produced, a similar loss of coat-colour pigmentation is observed in several independent lines, indicating that such an element is indeed required to achieve wild-type *Tyr* expression. As stated before, double-CRISPR/Cas9 deletions are generated through the error-prone NHEJ DNA repair route. Thus, each allele we obtained carries a typical and unique scar at the DNA sealing point. Another consequence of error-prone DNA repair is the production of partial and larger deletions, and even inversions, probably the result of rearrangements during DNA repair, favoured by the number of repetitive DNA elements that flank the target sequences. Such additional alleles can be used for genetic mapping. In fact, by comparing the phenotype of distinct alleles, we have assembled a genetic map of this mouse *Tyr* 5′-enhancer/boundary element [23].

A previous study in cultured cells highlighted that the efficiency of induced chromosomal deletions decreases with the size of the desired deletion [30]. Nevertheless, deletions in the megabase order are indeed possible [31], opening the possibility of precise modelling large chromosomal deletions, structural variants, or copy number variation (CNV) that are often found in families associated with genetic conditions or diseases.

One striking example has been reported by Lupiáñez and colleagues upon modelling mutations at genomic boundaries. Genomic boundaries, or insulators, are regulatory elements involved in the regulation of multiple genes domains. Typically, boundaries are located in between two different but adjacent topological expression domains [32–34]. Genomic boundaries physically separate chromatin territories containing genes with distinct expression patterns and restrict the activity of enhancer clusters to the cognate gene set [34]. Studying such sequences in an ectopic manner, using DNA constructs linked to reporter genes that mimic the natural conformation, might reveal the insulating activity of these boundary elements but does not reveal the true function of such elements in the endogenous context [35]. In contrast, this can be achieved by inactivating them at the endogenous site by

CRISPR/Cas9 deletion [23]. Interestingly, imposing such genomic alterations in mice reproduced the phenotype observed in humans carrying similar chromosomal aberrations [36].

The CRISPR/Cas9-deletion approach can also be applied to the genetic dissection of super- or stretch-enhancers [37]. In fact, sgRNAs can be easily designed to flank the full super-enhancer as well as individual regions associated with discrete chromatin immunoprecipitation resolved by DNA sequencing (ChIP-seq) peaks. With this approach, the contribution of each module of the super-enhancer can be elucidated [38]. For some specific cases, inactivation of noncoding elements can also be achieved using one sgRNA, by disrupting a previously characterized DNA-binding motif (Fig. 2c); this is the case of well-characterized transcription factor-binding motifs, such as GATA-2. Recently, Bresnick and colleagues described the relevance of a number of GATA-2-binding sites in hematopoiesis [39]. By targeting a CRISPR/Cas9 approach to the consensus DNA sequence, *indels* associated with the DNA repair scar will mutate some of the nucleotides constituting such a motif. As a caveat, one must always take into account that several transcription factor-binding sites can still be recognized by the corresponding nuclear factors, even though their sequences might differ significantly from the observed consensus. Because the frequency of CRISPR/Cas9 target sites in the mammalian genome approaches one every eight nucleotides, tiled sgRNAs can also be generated to interrogate a regulatory element under saturating conditions. Using this approach, Canver and colleagues identified vulnerabilities within an erythroid-specific BCL11A enhancer that can be used for therapeutical fetal globin re-induction in the context of sickle cell disease and β-thalassemias [40].

Recently, the mutational signatures of several types of malignancies have been described. Interestingly, mutations at noncoding sequences have been identified in a significant number of cancer patients [41–43]. In contrast to mutations at protein-coding genes, whose impact can be predicted by taking advantage of already existing animal and cellular models, the effect of a mutation outside the coding sequence can be difficult to anticipate. Hence, CRISPR/Cas9 mutagenesis represents a valid experimental approach to rapidly identify the effect of those noncoding mutations. For example, mutations at a noncoding sequence were found accumulated in a number of chronic lymphocytic leukemia (CLL) patients. In fact, a hypermutated region was found to display enhancer-like features such as H3K4 and H3K27ac enrichment. By CRISPR/Cas9-mediated chromosomal deletion of this putative enhancer, Puente and colleagues proved that loss of that enhancer resulted in a 40 % decrease in *Pax5*, a gene located 330 kb upstream [44].

Epigenome Editing

Epigenetic modifications are defined as chemical modifications of either DNA or histone proteins that affect chromatin structure and accessibility to DNA-binding proteins. These modifications are dynamically remodelled during development, differentiation, and aging. In fact, specific classes of enzymes exist that edit back and

forth such biochemical marks, including DNA and histone methylases, demethylases, histone acetyl transferases (HAT), and deacetylases (HDAC) [45]. Each of these enzymes can be targeted by specific drugs to inhibit their activities. In the past years, these inhibitors were used to interfere with chromatin-remodelling enzymes. Furthermore, the use of these molecules was also proposed and has already been explored as potential chemotherapy agents [46–48]. Unfortunately, these inhibitors act globally, nonspecifically. Advances in genome editing and synthetic biology led to the ability of building designer DNA-binding proteins over distinct targeting nuclease platforms. Artificial DNA-binding scaffolds can be coupled with a variety of catalytic domains, including those from the chromatin remodelling proteins listed here. This method allows imposing locally, and in a targeted manner, specific chromatin marks in the absence of any genetic, irreversible manipulation (Fig. 2b). For example, a TALE (transcription activators-like effectors) DNA-binding array was coupled with Tet1, an enzyme that promotes DNA demethylation. Targeting such TALE-Tet1 fusion to methylated DNA targets resulted in the loss of local DNA methylation. For example, by targeting the *RHOXF2* promoter in Hela and HEK 293 cells, this resulted in a 50- to 1000-fold gene activation [49]. Recently, the CRISPR/Cas9 system has been also adapted for applications in epigenome editing, by using a nuclease-dead Cas9 variant (dCas9). By fusing dCas9 with a catalytic domain of LSD1, a lysine-specific demethylase produced a tool to target active enhancers. By programming dCas9-LSD1 with enhancer-specific gRNAs, Kearns and colleagues could induce downregulation of target genes [50]. By targeting enhancers relevant for the expression of pluripotency factors, these authors induced morphological changes in murine ES cells. Interestingly, these changes depend on the depletion of H3K27ac from key enhancers.

It is also possible to turn the switch in the opposite direction. By coupling dCas9 with EP300, a histone acetyltransferase, a strong activation of target genes can be achieved by using enhancer- and promoter-specific guide RNAs [51].

This approach is not associated with irreversible DNA sequence alteration, a feature that is highly desirable for the study of chromatin dynamics. In fact, only the epigenetic features, and not the DNA sequences themselves, are altered. This factor allows performing a sequence-based assay, such as chromosome conformation capture, in the epigenetic-edited cells, a possibility that is obviously lost if a deletion occurs at the targeted sequence in the genome. In addition, reversibility and noninheritability traits are highly desirable features for therapeutical applications. Therefore, epigenome editing is one of the most promising applications derived from the use of the CRISPR/Cas9 tools [52].

Targeting Noncoding RNAs

Recent evidence supported the role of noncoding RNAs, including micro-RNAs (miRNAs) and long intergenic noncoding RNAs (lincRNAs) in regulating gene expression at multiple levels [53]. Catalogues of these transcripts account for more

than 56,000 human lncRNAs [54] and a similar number of small miRNAs. Thus, appropriate genetic models are required to define the role of this emerging class of regulatory RNA molecules. Dissimilar to protein-coding genes, these transcripts cannot be inactivated by triggering frameshift mutations at their corresponding exons, as point mutations are not likely to fully inactivate their regulatory potential. Rather, they need to be removed from the genome by a deletion strategy, similarly to what was previously described in this chapter for noncoding DNA elements such as enhancers and insulators. For example, Han and colleagues generated a large deletion in mice, encompassing the 23 kb of the imprinted lncRNA *Rian* [55]. Interestingly, the authors detected an increase in the expression of genes adjacent to the lncRNA *Rian*, providing evidence of the fact that lncRNA can regulate transcription of nearby genes.

In a study aimed to characterize the function of the lncRNA *Haunt*, Yin and colleagues deployed a number of distinct CRISPR/Cas9-based strategies. To dissect the effect of the *Haunt* transcript from that of the *Haunt* locus, the authors produced a series of deletions of different sizes and directed to functional elements within the element. By deleting small sequences constituting the *Haunt* promoter, resulting in the loss of *Haunt* lncRNA expression, the authors detected an increase in the expression of *Haunt* target genes, suggesting that this transcript has a suppressive role. In contrast, by producing larger deletions, the authors observed a decrease in the expression of the same set of target genes, indicating that the *Haunt* locus acts as a HOXA enhancer. To confirm these data, the authors induced *Haunt* overexpression by knocking in the sequence of a strong constitutive promoter, CAG, just upstream the *Haunt* TSS, using CRISPR/Cas9 to trigger homologous recombination [56]. With this strategy, overexpression of lncRNA could be obtained at the endogenous locus. As miRNAs are often organized into large gene clusters, we could consider that removal of the full cluster could be interesting. In this regard, by injecting two adjacent sgRNA in zebrafish embryos, Xiao and colleagues achieved a very large deletion encompassing a miRNA cluster on zebrafish chromosome 9 [57].

Conclusion

The efficiency, flexibility, and reproducibility of CRISPR/Cas9 approaches triggered many researchers to engage in challenging experiments that were technically very difficult to achieve, or nearly impossible to undertake, just a few years ago. CRISPR/Cas9-mediated genome mutagenesis has provided an outstandingly simple solution to functionally assess both coding and noncoding DNA sequences at their endogenous locations. Currently, with all these target nuclease experimental approaches, the entire mammalian genome can be investigated, in vivo, to decipher the role of coding and noncoding DNA elements in physiology and pathology.

Acknowledgments This review was supported by grants from MINECO projects BIO2012-39980, BIO2015-70978-R, CIBERER (ACCI-2015) and BMBS COST action BM1308 SALAAM to L.M. D.S. was supported by the La Caixa International PhD and EMBO Short Term Fellowship (AST140-2013) programs.

References

1. Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, Sundaram V, Xing X, Dogan N, Li J, Euskirchen G, Lin S, Lin Y, Visel A, Kawli T, Yang X, Patacsil D, Keller CA, Giardine B; Mouse ENCODE Consortium, Kundaje A, Wang T, Pennacchio LA, Weng Z, Hardison RC, Snyder MP. Principles of regulatory information conservation between mouse and human. Nature. 2014;515(7527):371–5.
2. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009;457(7231):854–8.
3. Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Pereira CF, Fisher AG, Adams DJ, Flicek P, Crawford GE, Laframboise T, Tesar P, Wei CL, Scacheri PC. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. PLoS Genet. 2010;6(7), e1001023.
4. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJ, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539): 317–30.
5. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov VV, Ecker JR, Thomson JA, Ren B. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518(7539):331–6.
6. Ouyang Z, Zhou Q, Wong WH. ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. Proc Natl Acad Sci U S A. 2009; 106(51):21521–6.
7. Bauer DE, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, Sabo PJ, Vierstra J, Voit RA, Yuan GC, Porteus MH, Stamatoyannopoulos JA, Lettre G, Orkin SH. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science. 2013;342(6155):253–7.
8. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43.
9. Lidral AC, Liu H, Bullard SA, Bonde G, Machida J, Visel A, Uribe LM, Li X, Amendt B, Cornell RA. A single nucleotide polymorphism associated with isolated cleft lip and palate, thyroid cancer and hypothyroidism alters the activity of an oral epithelium and thyroid enhancer near FOXE1. Hum Mol Genet. 2015;24(14):3895–907.
10. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviindran V, Tam D, Shen M, Son JE, Vakili NA, Sung HK, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui CC, Gomez-Skarmeta JL, Nóbrega MA. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014;507(7492):371–5.

11. Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Res. 2012;22(3):446–55.
12. Guenther CA, Tasic B, Luo L, Bedell MA, Kingsley DM. A molecular basis for classic blond hair color in Europeans. Nat Genet. 2014;46(7):748–52.
13. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005;6(6):507–12.
14. Bender MA, Bulger M, Close J, Groudine M. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol Cell. 2000;5(2):387–93.
15. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature. 2010;464(7287):409–12.
16. Volpi SA, Verma-Gaur J, Hassan R, Ju Z, Roa S, Chatterjee S, Werling U, Hou Jr H, Will B, Steidl U, Scharff M, Edelman W, Feeney AJ, Birshtein BK. Germline deletion of Igh 3′ regulatory region elements hs 5, 6, 7 (hs5-7) affects B cell-specific regulation, rearrangement, and insulation of the Igh locus. J Immunol. 2012;188(6):2556–66.
17. Vanden Berghe T, Hulpiau P, Martens L, Vandenbroucke RE, Van Wonterghem E, Perry SW, Bruggeman I, Divert T, Choi SM, Vuylsteke M, Shestopalov VI, Libert C, Vandenabeele P. Passenger mutations confound interpretation of all genetically modified congenic mice. Immunity. 2015;43(1):200–9.
18. Giraldo P, Montoliu L. Size matters: use of YACs, BACs and PACs in transgenic animals. Transgenic Res. 2001;10(2):83–103.
19. Giraldo P, Giménez E, Montoliu L. The use of yeast artificial chromosomes in transgenic animals: expression studies of the tyrosinase gene in transgenic mice. Genet Anal. 1999;15(3-5):175–8.
20. Murisier F, Guichard S, Beermann F. Distinct distal regulatory elements control tyrosinase expression in melanocytes and the retinal pigment epithelium. Dev Biol. 2007;303(2):838–47.
21. Montoliu L, Umland T, Schütz G. A locus control region at −12 kb of the tyrosinase gene. EMBO J. 1996;15(22):6026–34.
22. Giménez E, Giraldo P, Jeffery G, Montoliu L. Variegated expression and delayed retinal pigmentation during development in transgenic mice with a deletion in the locus control region of the tyrosinase gene. Genesis. 2001;30(1):21–5.
23. Seruggia D, Fernández A, Cantero M, Pelczar P, Montoliu L. Functional validation of mouse tyrosinase non-coding regulatory DNA elements by CRISPR-Cas9-mediated mutagenesis. Nucleic Acids Res. 2015;43(10):4855–67.
24. Harms DW, Quadros RM, Seruggia D, Ohtsuka M, Takahashi G, Montoliu L, Gurumurthy CB. Mouse genome editing using the CRISPR/Cas system. Curr Protoc Hum Genet. 2014;83:15.7.1–27.
25. Seruggia D, Montoliu L. The new CRISPR-Cas system: RNA-guided genome engineering to efficiently produce any desired genetic alteration in animals. Transgenic Res. 2014;23(5):707–16.
26. Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Münch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 2003;31:374–8.
27. Li Y, Rivera CM, Ishii H, Jin F, Selvaraj S, Lee AY, Dixon JR, Ren B. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. PLoS One. 2014;9(12), e114485.
28. Giraldo P, Martínez A, Regales L, Lavado A, García-Díaz A, Alonso A, Busturia A, Montoliu L. Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. Nucleic Acids Res. 2003;31(21):6290–305.
29. Giraldo P, Montoliu L. Artificial chromosome transgenesis in pigmentary research. Pigment Cell Res. 2002;15(4):258–64.

30. Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH, Orkin SH. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. J Biol Chem. 2014;289(31):21312–24.
31. Essletzbichler P, Konopka T, Santoro F, Chen D, Gapp BV, Kralovics R, Brummelkamp TR, Nijman SM, Bürckstümmer T. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res. 2014;24(12):2059–65.
32. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485(7398):376–80.
33. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289–93.
34. Moltó E, Fernández A, Montoliu L. Boundaries in vertebrate genomes: different solutions to adequately insulate gene expression domains. Brief Funct Genomic Proteomic. 2009;8(4): 283–96.
35. Wang J, Vicente-García C, Seruggia D, Moltó E, Fernandez-Miñán A, Neto A, Lee E, Gómez-Skarmeta JL, Montoliu L, Lunyak VV, Jordan IK. MIR retrotransposon sequences provide insulators to the human genome. Proc Natl Acad Sci U S A. 2015;112(32):E4428–37.
36. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, Santos-Simarro F, Gilbert-Dussardier B, Wittler L, Borschiwer M, Haas SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielmann M, Visel A, Mundlos S. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell. 2015;161(5):1012–25.
37. Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N; NISC Comparative Sequencing Program, Black BL, Visel A, Pennacchio LA, Collins FS; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A. 2013;110(44):17921–17926.
38. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell. 2015;58(2):362–70.
39. Hewitt KJ, Kim DH, Devadas P, Prathibha R, Zuo C, Sanalkumar R, Johnson KD, Kang YA, Kim JS, Dewey CN, Keles S, Bresnick EH. Hematopoietic signaling mechanism revealed from a stem/progenitor cell cistrome. Mol Cell. 2015;59(1):62–74.
40. Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O, Chen DD, Schupp PG, Vinjamur DS, Garcia SP, Luc S, Kurita R, Nakamura Y, Fujiwara Y, Maeda T, Yuan GC, Zhang F, Orkin SH, Bauer DE. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature. 2015;527(7577):192–7.
41. Chiba K, Johnson JZ, Vogan JM, Wagner T, Boyle JM, Hockemeyer D. Cancer-associated TERT promoter mutations abrogate telomerase silencing. Elife. 2015;4. doi:10.7554/eLife.07918.
42. Katainen R, Dave K, Pitkänen E, Palin K, Kivioja T, Välimäki N, Gylfe AE, Ristolainen H, Hänninen UA, Cajuso T, Kondelin J, Tanskanen T, Mecklin JP, Järvinen H, Renkonen-Sinisalo L, Lepistö A, Kaasinen E, Kilpivaara O, Tuupanen S, Enge M, Taipale J, Aaltonen LA. CTCF/cohesin-binding sites are frequently mutated in cancer. Nat Genet. 2015;47(7):818–21.
43. Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, Etchin J, Lawton L, Sallan SE, Silverman LB, Loh ML, Hunger SP, Sanda T, Young RA, Look AT. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science. 2014;346(6215):1373–7.

44. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, Munar M, Rubio-Pérez C, Jares P, Aymerich M, Baumann T, Beekman R, Belver L, Carrio A, Castellano G, Clot G, Colado E, Colomer D, Costa D, Delgado J, Enjuanes A, Estivill X, Ferrando AA, Gelpí JL, González B, González S, González M, Gut M, Hernández-Rivas JM, López-Guerra M, Martín-García D, Navarro A, Nicolás P, Orozco M, Payer ÁR, Pinyol M, Pisano DG, Puente DA, Queirós AC, Quesada V, Romeo-Casabona CM, Royo C, Royo R, Rozman M, Russiñol N, Salaverría I, Stamatopoulos K, Stunnenberg HG, Tamborero D, Terol MJ, Valencia A, López-Bigas N, Torrents D, Gut I, López-Guillermo A, López-Otín C, Campo E. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526(7574):519–524.
45. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications—writers that read. EMBO Rep. 2015;16(11):1467–81.
46. Ahuja N, Sharma AR, Baylin SB. Epigenetic therapeutics: a new weapon in the war against cancer. Annu Rev Med. 2016;67:73–89.
47. Falahi F, van Kruchten M, Martinet N, Hospers GA, Rots MG. Current and upcoming approaches to exploit the reversibility of epigenetic mutations in breast cancer. Breast Cancer Res. 2014;16(4):412.
48. Maes T, Carceller E, Salas J, Ortega A, Buesa C. Advances in the development of histone lysine demethylase inhibitors. Curr Opin Pharmacol. 2015;23:52–60.
49. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat Biotechnol. 2013;31(12):1137–42.
50. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, Maehr R. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat Methods. 2015;12(5):401–3.
51. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol. 2015;33(5):510–7.
52. Kungulovski G, Jeltsch A. Epigenome editing: state of the art, concepts, and perspectives. Trends Genet. 2015; pii: S0168-9525(15)00212-7.
53. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell. 2013; 154(1):26–46.
54. Xie C, Yuan J, Li H, Li M, Zhao G, Bu D, Zhu W, Wu W, Chen R, Zhao Y. NONCODEv4: exploring the world of long non-coding RNA genes. Nucleic Acids Res. 2014;42(Database issue):D98–103.
55. Han J, Zhang J, Chen L, Shen B, Zhou J, Hu B, Du Y, Tate PH, Huang X, Zhang W. Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol. 2014;11(7):829–35.
56. Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H, Orkin SH, Shen X. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. Cell Stem Cell. 2015;16(5):504–16.
57. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S, Zhang B. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 2013;41(14), e141.

http://www.springer.com/978-3-319-34146-0

Genome Editing
Turksen, K. (Ed.)
2016, IX, 166 p. 24 illus. in color., Hardcover
ISBN: 978-3-319-34146-0