GENE REGULATION

Making the cut in the dark genome

CRISPR screens will reveal important regulatory elements in the noncoding genome

By **Jaclyn M. Einstein**[1,2] *and* **Gene W. Yeo**[1,2,3,4]

Noncoding elements encompass more than 98% of the human genome and have been linked to regulatory sequences that contribute to human health and disease (*1*). Since the publication of the human genome sequence, considerable effort has been made to annotate functional elements, including noncoding regulatory regions—i.e., cis-regulatory regions and noncoding RNAs (ncRNAs) that are involved in transcriptional regulation. Transcription factors often associate with hundreds to thousands of binding sites throughout the genome, and identifying which sites regulate gene expression often requires time-consuming and complex enhancer studies, or parallel assays in which short enhancer or promoter sequences are cloned into non-native contexts (*2*, *3*). A recent study by Sanjana *et al.* (*4*) and a report by Fulco *et al.* (*5*) on page 769 of this issue address this obstacle using clustered regularly interspaced short palindromic repeats (CRISPR) screens to functionally characterize noncoding elements in their native context.

Previously, large-scale biochemical efforts have discovered potential regulatory sequences that are targeted by hundreds of DNA-binding proteins. In particular, the development of methods to identify deoxyribonuclease I hypersensitive sites (DHSs) and large-scale chromatin immunoprecipitation–sequencing (ChIP-seq) profiling have enabled a genome-wide view of the protein-bound chromatin landscape (*2*, *3*). However, despite these advances, linking these molecular associations with functional regulation has remained challenging.

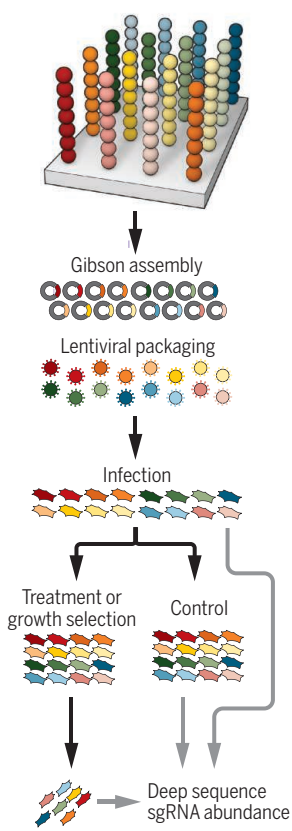
CRISPR screens offer a powerful approach for the unbiased removal of protein-coding genes using a pool of CRISPR vectors that target different genomic loci. The CRISPR system effectively generates mutations at specific genomic loci, guided by single guide RNAs (sgRNAs) that are homologous to genomic regions upstream of an NGG sequence (protospacer adjacent motif) in human cells (*6*, *7*). The first CRISPR "knockout" screens were on the genome scale, inducing full gene knockout, and overcame many limitations experienced in RNA interference screens—i.e., off-target effects and incomplete protein depletion (*8*, *9*) (see the figure).

Sanjana *et al.* used this system to investigate noncoding regulation of vemurafenib resistance in melanoma cells [vemurafenib is an inhibitor of the serine-threonine protein kinase domain of B-Raf (BRAF) proteins carrying the V600E mutation, in which valine at position 600 is substituted by glutamic acid]. Noncoding regions were targeted 5′ and 3′ of the major variants of resistance genes, neurofibromatosis type 1 (*NF1*), *NF2*, and cullin 3 (*CUL3*), using oligonucleotide arrays tiling across 713 kb of sequence. CRISPR-associated protein 9 nuclease (Cas9) from *Streptococcus pyogenes* (spCas9) generates frameshift, loss-of-function mutations by inducing double-stranded breaks in DNA and creating insertion and deletion (indel) mutations at loci guided by an sgRNA characterized by a CRISPR targeting RNA (crRNA)–trans-activating crRNA (tracrRNA) fusion (*6*, *7*). After infection and treatment with BRAF inhibitor, sgRNAs targeting CUL3 noncoding regions 5′ of the transcription start site were the most highly enriched compared with the control. Enriched sgRNAs occurred with *CUL3* depletion and were associated with regulatory regions that experienced chromatin looping, changes in posttranslational histone modifications, and disruptions in canonical transcription factor binding sites.

Fulco *et al.* used a proliferation-based CRISPR screen with sgRNAs tiling across 1.29 Mb of sequence in noncoding regions 5′ and 3′ of globin transcription factor 1 (*GATA1*)

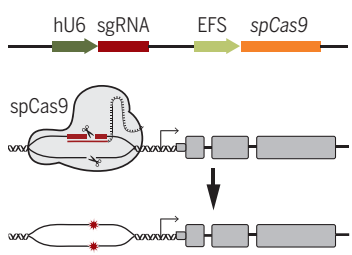
CRISPR-Cas9 screening methods: dead or alive

Pooled screening approaches identify noncoding regulatory elements using CRISPR-mediated gene knockout methods that employ Cas9 nucleases exhibiting various levels of control.

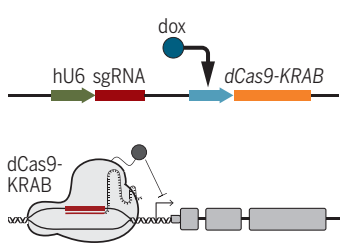


CRISPR screening method
Lentiviral vectors are created by cloning an oligonucleotide array containing sgRNA sequences into a lentiviral vector backbone. The cells are transduced and treatment selected, and sgRNA abundance is quantified by deep sequencing.

CRISPR-spCas9
A CRISPR-spCas9 vector backbone exhibiting constitutive spCas9 expression. Mutations are generated by double-strand breaks.



CRISPRi
A CRISPRi vector backbone exhibiting doxycycline-induced expression of dCas9-KRAB. The KRAB effector represses transcription of the target gene.



GRAPHIC: V. ALTOUNIAN/*SCIENCE*

[1]*Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA.* [2]*Stem Cell Program and Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA 92093, USA.* [3]*Department of Physiology, National University of Singapore, Singapore 117597.* [4]*Molecular Engineering Laboratory, A*STAR, Singapore 138673. Email: geneyeo@ucsd.edu*

and Myc proto-oncogene protein (*MYC*), which regulate proliferation of K562 erythroleukemia cells. The authors used the CRISPR interference (CRISPRi) method, which uses the catalytically inactive version of Cas9 (dCas9), providing an inducible system for RNA-guided DNA targeting without inducing mutations (*10*). When coupled to a Krüppel-associated box (KRAB) repressor domain, stable and effective transcriptional repression is achieved at specific loci guided by an sgRNA, characterized by crucial dCas9 and *S. pyogenes* terminator hairpins (*10*, *11*). After infection and doxycycline-induced dCas9 targeting, enriched sgRNAs corresponded to DHSs, which harbor binding sites for many transcription factors. More interestingly, dCas9 targeting of *GATA1* or histone deacetylase 6 (*HDAC6*) enhancers reduced HDAC6 expression, suggesting competition between genes for common enhancers. Identified *MYC* enhancers corresponded to alternative transcription start sites and CCCTC-binding factor (CTCF)–mediated chromatin loops, all of which likely affect cellular proliferation.

Whereas both CRISPR-spCas9 and CRISPRi screening methods were successful in identifying noncoding regulators, CRISPRi is limited to transcriptional repression, which often varies between genes. The studies of Sanjana *et al.* and Fulco *et al.* serve as specific cases that contribute to the larger goal of identifying all noncoding regulatory regions by laying the groundwork for generating genome-wide screens tiling all noncoding regions. In addition, specific CRISPR screens can be generalized to other disease-based phenotypic assays, as studies have confirmed that noncoding mutations causing small changes in gene expression can have large phenotypic effects (*12*). Although germline variants have been identified in genome-wide association studies, systematic studies that provide functional annotation of all noncoding regions will be exceedingly important for identifying disease-causing somatic variants. For example, specific somatic variants that include gain of transcription factor binding sites, fusion events due to genomic rearrangements, and variants caused by ncRNAs and pseudogenes have been identified in focused studies on the disease gene (*12*). ■

REFERENCES

1. M. Bulger, M. Groudine, *Cell* **144**, 327 (2011).
2. W. Akhtar *et al.*, *Cell* **154**, 914 (2013).
3. D. E. Dickel *et al.*, *Nat. Methods* **11**, 566 (2014).
4. N. Sanjana *et al.*, *Science* **353**, 1545 (2016).
5. C. P. Fulco *et al.*, *Science* **354**, 769 (2016).
6. P. Mali *et al.*, *Science* **339**, 823 (2013).
7. L. Cong *et al.*, *Science* **339**, 819 (2013).
8. O. Shalem *et al.*, *Science* **343**, 84 (2014).
9. T. Wang *et al.*, *Science* **343**, 80 (2014).
10. M. H. Larson *et al.*, *Nat. Protoc.* **8**, 2180 (2013).
11. L. A. Gilbert *et al.*, *Cell* **159**, 647 (2014).
12. E. Khurana *et al.*, *Nat. Rev. Genet.* **17**, 93 (2016).

10.1126/science.aak9849

EPIDEMIOLOGY

First flu is forever

A change in the properties of influenza virus in 1968 has left a profound mark on population immunity

By **Cécile Viboud**[1] *and* **Suzanne L. Epstein**[2]

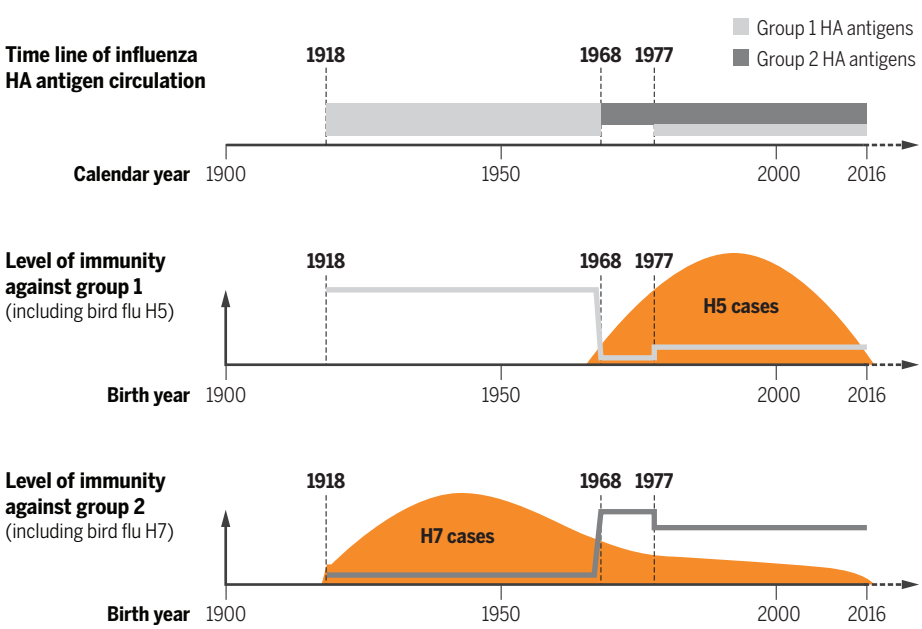
Influenza is a threat that has been with humans throughout history, fueled by a constant race between host immunity and viral evolution. Control strategies rely on annual immunizations and require frequent updates of the vaccine, an expensive, cumbersome, and not always foolproof process. Efforts are therefore under way to develop vaccines that confer broadly cross-protective immunity to diverse influenza strains. Cross-immunity is pervasive in nature; in multistrain viral diseases such as influenza or dengue, response to a primary infection can profoundly influence response to the next strain encountered. Even unrelated viruses can be recognized by the same cross-reactive T cells. On page 722 of this issue (*1*), Gostic *et al.* show that severe infection with a bird flu virus depends on the individual's first encounter with influenza in childhood.

Gostic *et al.* dissected the age patterns of human infections with avian influenza A/H5N1 and A/H7N9 reported globally since 1997. These viruses represent two distinct types of hemagglutinin (HA), type 1 for A/H5N1 and type 2 for A/H7N9, which differ in genetic sequence. This HA difference parallels intriguing epidemiological differences: A/H5N1 cases are found mainly in children and young adults, whereas A/H7N9 cases are concentrated in older individuals. These contrasting age profiles have sparked several hypotheses, including the effect of age-specific contacts with infected birds for A/H7N9 (*2*) and previous immunity to the neuraminidase surface protein for A/H5N1 (*3*).

Gostic *et al.* offer a single explanation for the contrasting age profiles of A/H5N1 and A/H7N9 cases and for the abrupt change in infection risk around birth year 1968. The latter coincides with the emergence of a new influenza virus in human populations and a shift in circulating antigens from HA group 1 to group 2. A previously little-noted consequence of this event was an altered immune status of the population (see the figure).

Population immunity to bird flu depends on birth year

In 1968, there was a change in a major protective antigen of influenza, hemagglutinin (HA). This altered the type of flu virus that new birth cohorts first encountered in life. Gostic *et al.* show that resulting levels of broadly protective immunity differ by birth year and that these differences can predict the risk of severe infection with different types of bird flu.



GRAPHIC: C. BICKEL/*SCIENCE*

Making the cut in the dark genome
Jaclyn M. Einstein and Gene W. Yeo (November 10, 2016)
Science **354** (6313), 705-706. [doi: 10.1126/science.aak9849]

Editor's Summary

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.