SOFTWARE

CRISPR-TE: a web-based tool to generate single guide RNAs targeting transposable elements

Yixin Guo[1], Ziwei Xue[1,2], Meiting Gong[1], Siqian Jin[1], Xindi Wu[1] and Wanlu Liu[1,2*]

Abstract

Background The CRISPR/Cas systems have emerged as powerful tools in genome engineering. Recent studies highlighting the crucial role of transposable elements (TEs) have stimulated research interest in manipulating these elements to understand their functions. However, designing single guide RNAs (sgRNAs) that are specific and efficient for TE manipulation is a significant challenge, given their sequence repetitiveness and high copy numbers. While various sgRNA design tools have been developed for gene editing, an optimized sgRNA designer for TE manipulation has yet to be established.

Results We present CRISPR-TE, a web-based application featuring an accessible graphical user interface, available at https://www.crisprte.cn/, and currently tailored to the human and mouse genomes. CRISPR-TE identifies all potential sgRNAs for TEs and provides a comprehensive solution for efficient TE targeting at both the single copy and subfamily levels. Our analysis shows that sgRNAs targeting TEs can more effectively target evolutionarily young TEs with conserved sequences at the subfamily level.

Conclusions CRISPR-TE offers a versatile framework for designing sgRNAs for TE targeting. CRISPR-TE is publicly accessible at https://www.crisprte.cn/ as an online web service and the source code of CRISPR-TE is available at https://github.com/WanluLiuLab/CRISPRTE/.

Introduction

Since the initial discovery of the CRISPR/Cas9 system for genome editing [1, 2], the development of catalytically inactive Cas9 variants has further facilitated its application in targeted gene expression activation [3, 4], repression [3], and base editing [5]. Transposable elements (TEs) are mobile DNA sequences capable of moving within the genome [6]. Though once deemed "genomic dark matter", recent studies have suggested that TEs may act as cis-regulatory elements, contributing to gene regulation by serving as promoters, enhancers, silencers, and boundary elements [7]. For instance, in mouse early embryogenesis, the endogenous retrovirus MuERV-L serves as an alternative promoter for certain genes specific to the two-cell stage that are bound and induced by the transcription factor Dux [8]. In humans, the evolutionarily young transposable elements such as LTR7Y, LTR7, and LTR5HS harbor binding sites for several key transcription factors and are posited to regulate both human naïve pluripotency and germline lineage commitment [9–13].

The study of TE functions is challenging due to their high copy numbers and their sequence repetitiveness [7].

*Correspondence:
Wanlu Liu
wanluliu@intl.zju.edu.cn
[1] Department of Orthopedic Surgery of the Second Affiliated Hospital, and Centre of Biomedical Systems and Informatics of Zhejiang University-University of Edinburgh Institute (ZJU-UoE Institute), Zhejiang University School of Medicine, Zhejiang University, Zhejiang, Hangzhou 310003, China
[2] Future Health Laboratory, Innovation Center of Yangtze River Delta, Zhejiang University, Jiaxing 314100, China

Consequently, designing sgRNAs that efficiently recruit CRISPR/Cas9 system to TEs is key for functionally probing their biological roles. Researchers have targeted individual TE copies with CRISPR/Cas9 or CRISPR inhibition (CRISPRi) systems to delete, insert, or repress specific copies, thereby studying their biological functions [14–18]. Moreover, there have been efforts to elucidate the functions of TE subfamilies using CRISPRi or CRISPR activation (CRISPRa), involving the design of sgRNAs that target multiple copies within certain TE subfamilies [11, 17, 19, 20]. However, these attempts to manipulate TE expression via CRISPRi or CRISPRa have largely relied on sgRNAs selected using gene-centric tools and on the manual design of sgRNAs targeting consensus sequences of TE copies. Employing similar strategies, we have used CRISPRi/a to silence or activate specific TE subfamilies, assessing their potential enhancer roles in human embryonic stem cells and primordial germ cells [12, 13]. Nonetheless, the prevailing CRISPR design tools are primarily gene-centric and fail to provide adequate on- or off-target information for TEs, limiting in-depth TE functional studies.

In this study, we introduce CRISPR-TE, a web-based bioinformatics tool specifically for designing CRISPR/Cas sgRNAs targeting transposable elements. Our tool can design sgRNAs to target individual TE copies or combinations of sgRNAs to target TE subfamilies. Moreover, CRISPR-TE provides an interactive web interface with swift query capabilities, enabling convenient access and analysis of detailed sgRNA information for researchers. In summary, CRISPR-TE represents a valuable resource for researchers investigating the role of TEs in the genome, facilitating more comprehensive and precise studies of these repetitive elements.

Results

CRISPR-TE workflow

CRISPR-TE first constructs a database of sgRNAs by scanning human and mouse genomes for potential target sites containing the PAM (protospacer adjacent motif) sequence (5′-NGG-3′ for SpCas9 from *S. pyogenes*). Upon input of a genome file, the Aho-Corasick pattern matching algorithm efficiently identifies all N20NGG patterns within the reference genome [21] (Fig. 1A). A retrieval tree (trie) data structure stores all sgRNAs, their genomic locations, and 6bp downstream and upstream sequences (Fig. 1B). This data structure enables efficient computation of sequence mismatch neighborhoods. Additional data, including sgRNA on-target activity efficiency [22, 23], TE-specified off-target scores, TE subfamily, individual TE ID (if any), and overlapping genetic elements such as exons, introns, promoter-TSS, intergenic regions, are calculated and stored in the main database table managed by PostgreSQL. Queries for individual TE ID and their genomic coordinates are also available on the CRISPR-TE website. The sgRNA ID (gid) acts as a foreign key linking the database mismatch table, which contains gid, sgRNA sequence, and mismatch neighborhoods (Fig. 1C). CRISPR-TE offers two strategies for TE-specific sgRNA design: 1) targeting a single individual TE copy with minimal off-targets or 2) targeting TE subfamilies using optimized sgRNA combinations
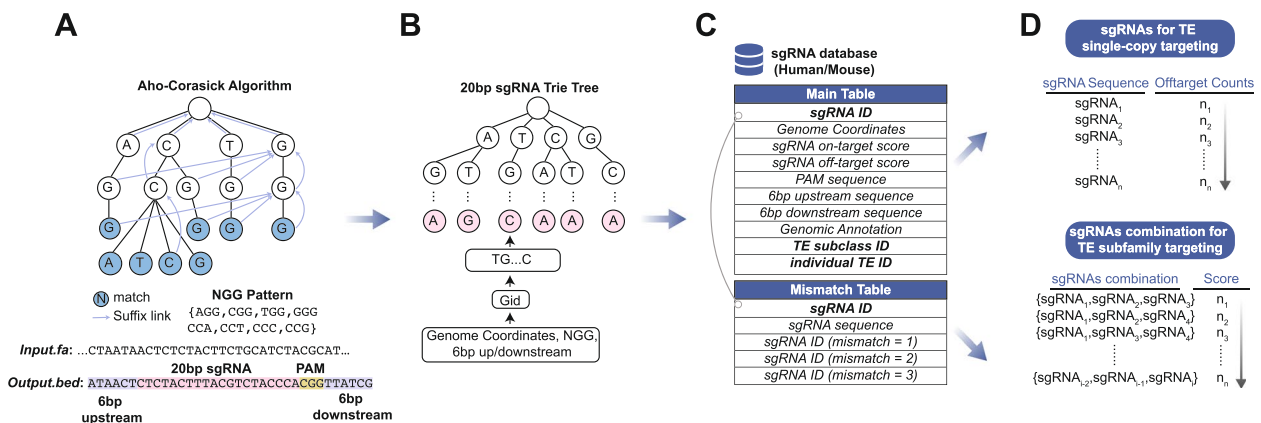
Fig. 1 CRISPR-TE Workflow for Designing sgRNAs Targeting Transposable Elements. **A** Efficiently searches for all potential sgRNA target sites (N20NGG) using the Aho-Corasick algorithm on the genome FASTA file. **B** Employs a trie data structure to efficiently store sgRNA sequences, facilitating hamming distance mismatch searches. **C** Stores comprehensive sgRNA information for human and mouse genomes in a PostgreSQL database, consisting of two tables: the main table contains sgRNA sequences and coordinates, 6bp upstream and downstream sequences, on/off-target scores, genetic element classes, and the targeted TE (if applicable). **D** CRISPR-TE provides two approaches for sgRNA design targeting TEs: (i) targeting a single copy with minimal off-targets, and (ii) targeting a TE subfamily using optimal sgRNA combinations determined by a greedy algorithm

Guo *et al. Mobile DNA*   (2024) 15:3

Page 3 of 7

ranked by a greedy algorithm (Fig. 1D). These strategies provide researchers with comprehensive options for TE studies, enabling them to select the approach best suited to their experimental goals.

Web interface

CRISPR-TE features a user-friendly web interface for designing sgRNAs targeting transposable elements. Users input the design objective (targeting TE single copies or subfamilies), the genome assembly (human or mouse), and the name of the targeted TE subfamily or individual TE ID (Fig. 2A). The annotation query function on the CRISPR-TE website allows users to search for specific genomic coordinate or individual TE IDs (Fig. 2B). After submitting the sgRNAs design, CRISPR-TE generates an interactive table displaying sgRNA sequences, coordinates, potential off-target numbers with 0, 1, 2, or 3 mismatches, and on/off-target activity scores (Fig. 2C). Detailed sgRNA information becomes accessible to users by clicking on each row in the summary table. A color-coded graphical representation of the sgRNA target site help users inspect candidate sgRNAs based on their locations (Fig. 2D). Summary pie charts depict the proportions of target sites by mismatch number and lists of off-target sgRNAs with their sequences and target sites further aid in the selection of suitable sgRNAs (Fig. 2D). For designing sgRNAs for TE subfamilies, CRISPR-TE generates combinations intended to maximize coverage of the queried TE subfamily. To balance coverage and computational complexity, CRISPR-TE currently supports designing combinations of three sgRNAs. Pie charts and bar plots visualize the proportion of on-target sites and the number of off-target sites for each sgRNA (Fig. 2E). Users can download the results in Excel, CSV, and PDF format for further analysis and documentation.

TE sgRNA analysis of human and mouse

As TEs integrate into the genome, their sequences diverge due to the accumulation of random mutations and truncations. Evolutionarily young subfamilies, often considered as currently or recently active, possess highly similar sequences across different copies. In contrast,

sequences of evolutionarily old subfamilies typically exhibit a greater degree of divergence from their consensus sequences [24]. We analyzed the percentage copies covered by three sgRNA combinations for all TE subfamilies. As anticipated, sgRNA combinations designed by CRISPR-TE target evolutionarily young families such as LTR7Y, LTR5HS, SVA-D in humans with higher coverage compared to older families (Fig. 3A and Fig. S1). Specifically, in humans, young TE subfamilies like ERVK and SVA show over 50% coverage with three sgRNAs. Similarly, in mice, B2 and ERVK rank as the top covered TE subfamilies (Fig. 3B). Furthermore, we discovered that evolutionarily young TEs in human and mouse, such as LTR5HS (coverage ranked at 15 in human TE subfamilies) and RLTR6CMm (coverage ranked at 22 in mouse TE subfamilies), can be targeted with over 70% coverage using CRISPR-TE-designed sgRNA combinations at the subfamily level, despite the possibility that some sgRNAs may also target other TE subfamilies with similar sequences (Fig. 3C). Conversely, for other relatively older TEs such as L1PA10 (coverage ranked at 223 for human TE subfamilies) and B2Mm2 (coverage ranked at 201 for mouse TE subfamilies), CRISPR-TE-designed sgRNA combinations can target only about 20% of copies, although the majority of the designed sgRNAs accurately target the intended TEs (Fig. 3C). In conclusion, the effectiveness of sgRNA targeting by CRISPR-TE is strongly correlated with the age of the TE, with younger TEs being more amenable to efficient targeting.

Discussion

This study introduces CRISPR-TE, a specialized sgRNA design tool tailored for the unique challenges associated with TE targeting in genome editing. Our novel approach for sgRNA design offers a significant advancement over traditional gene-targeting tools, addressing the high copy number and sequence repetitiveness that have long hindered effective TE manipulation. Our results indicate that CRISPR-TE can accurately target TE subfamilies, particularly those that are evolutionarily young and exhibit conserved sequences. The tool's ability to target these TEs with higher coverage suggests that CRISPR-TE

(See figure on next page.)

Fig. 2 Screenshot of The CRISPR-TE Web Tool Interface. **A** The CRISPR-TE homepage, which requires three types of input: (i) Design purpose, (ii) Genome assembly, and (iii) Target TE copy ID or genomic coordinates. **B** Tool is provided for querying individual TE copy IDs and their genomic coordinates. **C** After submitting, CRISPR-TE displays all possible sgRNAs along with detailed information, including sgRNA sequence, coordinates, GC content, mismatches, on-target score, and off-target score. **D** CRISPR-TE enables users to examine the locations of sgRNAs on the genome, alongside other genomic features, by clicking on individual sgRNAs. The pie chart on the left illustrates the proportions of target sites with various mismatch counts. A list of all off-target sgRNA, including their sequences, genomic coordinates, and associated genetic element classes, is shown on the right. **E** The results for sgRNA combinations targeting TE subfamilies are presented. This includes the sgRNA sequences, the number of on-target sites, the on-target percentage for the queried TE subfamily, the sgRNA combination coverage, and off-targets on TEs and other genetic element classes

Fig. 2 (See legend on previous page.)

Guo *et al. Mobile DNA*   (2024) 15:3

Page 5 of 7
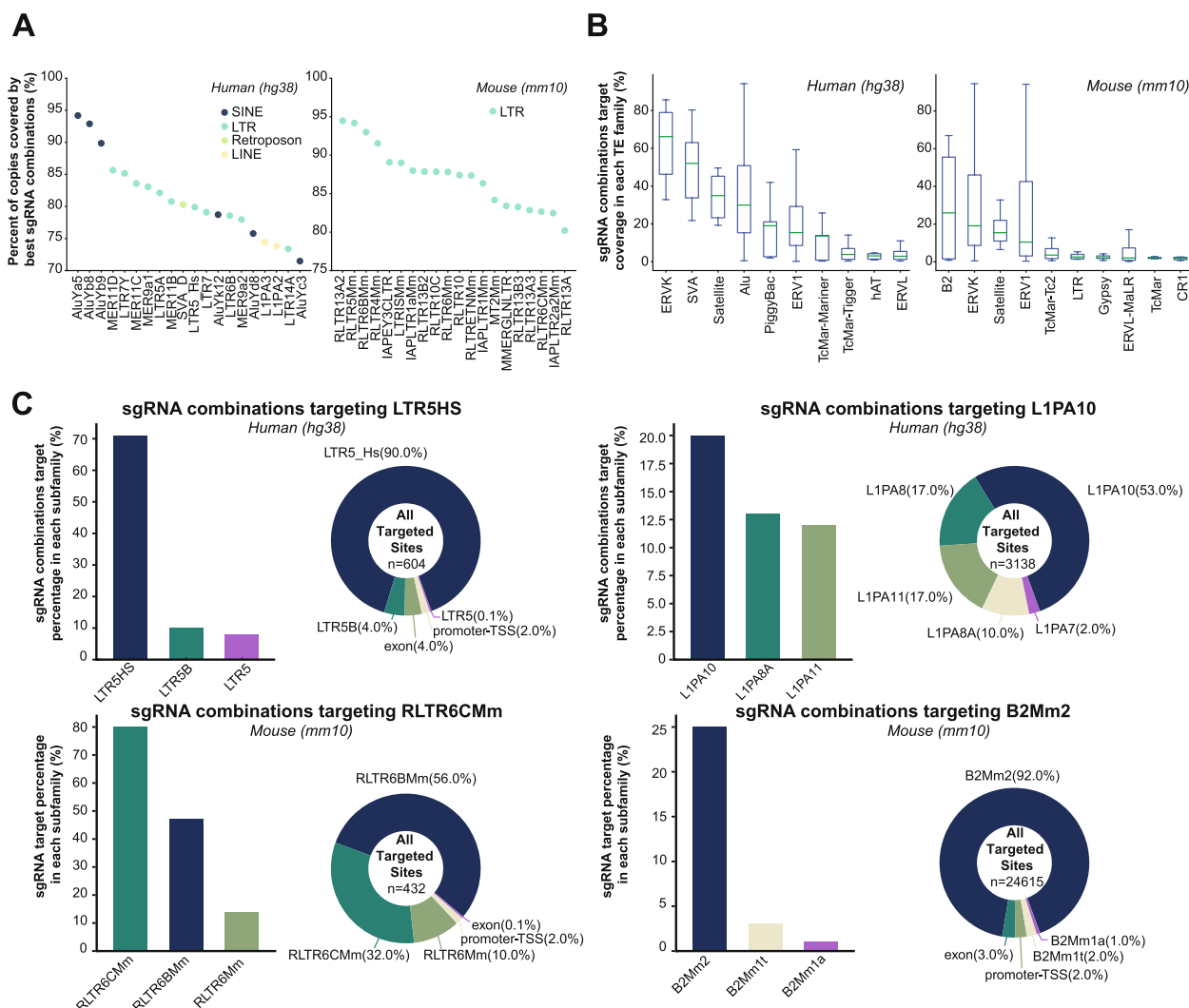
Fig. 3 Analysis of TE sgRNAs in Human and Mouse. **A** Displays the top 20 TE subfamilies with the highest coverage using the best three sgRNA combinations in human (left panel) and mouse (right panel). **B** Box plots showing the coverage achieved by the best three sgRNA combinations in each TE family for human (left panel) and mouse (right panel). **C**. Examples of sgRNAs designed by CRISPR-TE for targeting LTR5HS (upper left panel), L1PA10 (upper right panel), RLTR6BMm (bottom left panel), and B2Mm2 (bottom right panel) TEs. The bar plots indicate the targeted percentage of copies for the top three TE subfamilies using the best three sgRNA combinations. The pie charts represent the genomic distribution of all targeted sites by the corresponding best three sgRNA combinations

is adept at identifying and leveraging the less divergent sequences within these younger subfamilies. This is a crucial development, as it facilitates the functional analysis of TEs that may play significant roles in gene regulation and genome architecture.

While CRISPR-TE has shown promising results, we recognize certain limitations in its current iteration. At present, CRISPR-TE is tailored only to human and mouse genomes using SpCas9. Given the ubiquitous presence and regulatory significance of TEs in various plant species, including maize, where they play a pivotal role in phenotypic regulation [25], future enhancements will aim

to broaden the tool's species compatibility and include additional Cas enzymes like Cas12 or Cas13 orthologues [26]. We are committed to extending CRISPR-TE's functionality to encompass a wider array of species and Cas variants in our subsequent updates.

The current version supports three sgRNA combinations for targeting TE subfamilies, primarily to manage the computational complexity, which grows exponentially with additional sgRNAs. This limitation may restrict the tool's effectiveness, particularly when addressing evolutionarily older TE subfamilies that require more comprehensive sgRNA coverage. To overcome this, future

Guo *et al. Mobile DNA*   (2024) 15:3

Page 6 of 7

development will focus on refining our greedy algorithm to allow for an increased number of sgRNA combinations, which could enhance the scope of TE subfamily targeting. However, the experimental delivery of multiple sgRNAs into cellular systems or animal models poses its own set of challenges [27], particularly when investigating TE subfamily functions.

Furthermore, the potential for off-target effects is an inherent concern due to the repetitive nature of TE sequences. Although CRISPR-TE includes an on-target and off-target scoring system, these algorithms were originally developed for gene targeting and may not be fully optimized for TEs [28]. Advances in the specificity of on-target and off-target predictions for TEs remain a priority for future refinement. The incorporation of machine learning algorithms is anticipated to improve the precision of sgRNA efficacy predictions, thereby mitigating the risk of off-target effects [29, 30].

In conclusion, CRISPR-TE represents a notable step forward in the field of genome engineering, allowing researchers to explore the possible functions of TEs using genome editing tools. As we further imrpove this tool, we anticipate it to become an essential resource for TE research, providing deeper insights in understanding the repetitive elements in the genome.

Methods

sgRNA sequence search, annotation, and storage

We utilized the Aho-Corasick string matching algorithm to screen the genome sequences for all occurrences of the N20NGG pattern on both positive and negative strands. We then saved all potential sgRNA sequences into a modified Trie tree data structure. These sequences were classified based on various genetic elements, such as exons, introns, promoter-TSS, and intergenic regions. The genome assembly and annotation versions used were GRCh38.97 for human (http://www.ensembl.org/Homo_sapiens/) and GRCm38.97 for mouse (https://www.ensembl.org/Mus_musculus/). We obtained annotations for transposable element (TE) subfamilies and individual TE IDs from RepeatMasker (https://www.repeatmasker.org/). Using Trie tree structure, we performed mismatched string pattern matching to identify N20NGG sequences with fewer than 3 mismatched nucleotides in the genome for each sgRNA. We stored the resulting nucleotide sequences, their genomic coordinates, annotations, and mismatch information in a PostgreSQL (version 14.3) database for efficient indexing and rapid searching.

sgRNA combination search and off-target score

We proposed a computation time-optimized greedy search algorithm to identify all potential sgRNA combinations that can cover most copies of a TE subfamily while ensuring a minimum number of off-target sites in other genetic elements. We ranked all sgRNAs targeting any copies of each TE subfamily based on their total coverage of copies. The sgRNA combination score was computed as a weighted sum of coverage and off-target events, defined by:

$$SCORE = Coverage - \lambda_1 \times W_1 - \lambda_2$$
$$\times (\lambda_3 \times W_2 + \lambda_4 \times W_3 + \lambda_5 \times W_4 + \lambda_6 \times W_5)$$

where coverage is the percentage of TE subfamily copies covered by the current sgRNA combination, W_1 is the number of off-target TEs, W_2 is the number of off-targets to promoter-TSS, W_3 is the number of off-target exons, W_4 is the number of off-target introns, W_5 is the number of off-target intergenic regions, with weights $\lambda_1 = 1e\text{-}3$, $\lambda_2 = 1e\text{-}4$, $\lambda_3 = 0.4$, $\lambda_4 = 0.3$, $\lambda_5 = 0.4$, $\lambda_6 = 0.3$, set as default parameters.

We employed a greedy search strategy that involved selecting the top n sgRNAs with the highest combination scores and subsequently ranking the remaining sgRNAs by the increment of the combination score. We greedily added the most optimal sgRNAs to the current combination at each iteration to obtain the final combination of sgRNAs for targeting copies of a TE subfamily.

Implementation of the CRISPR-TE web server

We developed a web server that enables users to search for sgRNAs targeting TEs, using an intuitive and user-friendly data browser. The front-end interface of the web server was created with HTML5 and CSS3, and all data visualizations were produced using the D3.js framework [31]. The back-end data, containing sgRNA sequences and annotations, was managed by the PostgreSQL database system, facilitating prompt responses to user queries. Python3 (v3.9.12) and Django (v3.2.5) were used for communication between the front-end and back-end. The website is accessible at https://www.crisprte.cn/ without the need for registration or login. The CRISPR-TE website's functionality was thoroughly tested on Google Chrome and Apple Safari browsers. The site is deployed on an Nginx web server (v1.18.0) running on a Linux Ubuntu (v20.04.5 LTS) cloud server system.

Supplementary Information

> **Additional file 1: Figure S1.** sgRNA Combinations Targeting TE Subfamilies in Human and Mouse. All TE subfamilies are ranked by the targeted coverage using the best sgRNA combinations for human (left panel) and mouse (right panel).

Guo *et al. Mobile DNA*   (2024) 15:3

Page 7 of 7

Availability of data and materials
The CRISPR-TE web tool is publicly available at https://www.crisprte.cn/. The source code of CRISPR-TE is accessible at https://github.com/WanluLiuLab/CRISPRTE/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–21.
2. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819–23.
3. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-scale CRISPR-mediated control of gene repression and activation. Cell. 2014;159:647–61.
4. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. 2015;517:583–8.
5. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533:420–4.
6. Wells JN, Feschotte C. A field guide to eukaryotic transposable elements. Annu Rev Genet. 2020;54:539–61.
7. Fueyo R, Judd J, Feschotte C, Wysocka J. Roles of transposable elements in the regulation of mammalian transcription. Nat Rev Mol Cell Biol. 2022;23:481–97.
8. Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature. 2012;487:57–63.
9. Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature. 2015;522:221–5.
10. Göke J, Lu X, Chan Y-S, Ng H-H, Ly L-H, Sachs F, et al. Dynamic transcription of distinct classes of endogenous retroviral elements Marks specific populations of early human embryonic cells. Cell Stem Cell. 2015;16:135–41.
11. Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, et al. Hominoid-specific transposable elements and KZFPs facilitate human embryonic genome activation and control transcription in naive human ESCs. Cell Stem Cell. 2019;24:724–735.e5.
12. Xiang X, Tao Y, DiRusso J, Hsu F-M, Zhang J, Xue Z, et al. Human reproduction is regulated by retrotransposons derived from ancient Hominidae-specific viral infections. Nat Commun. 2022;13:463.
13. Ai Z, Xiang X, Xiang Y, Szczerbinska I, Qian Y, Xu X, et al. Krüppel-like factor 5 rewires NANOG regulatory network to activate human naive pluripotency specific LTR7Ys and promote naive pluripotency. Cell Rep. 2022;40:111240.
14. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science. 2016;351:1083–7.
15. Hummel B, Hansen EC, Yoveva A, Aprile-Garcia F, Hussong R, Sawarkar R. The evolutionary capacitor HSP90 buffers the regulatory effects of mammalian endogenous retroviruses. Nat Struct Mol Biol. 2017;24:234–42.
16. Zhang Y, Li T, Preissl S, Amaral ML, Grinstein JD, Farah EN, et al. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nat Genet. 2019;51:1380–8.
17. Todd CD, Deniz Ö, Taylor D, Branco MR. Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. eLife. 2019;8:e44344.
18. Karamitros T, Hurst T, Marchi E, Karamichali E, Georgopoulou U, Mentis A, et al. Human endogenous retrovirus-K HML-2 integration within *RASGRF2* is associated with intravenous drug abuse and modulates transcription in a cell-line model. Proc Natl Acad Sci U S A. 2018;115:10434–9.
19. Fuentes DR, Swigut T, Wysocka J. Systematic perturbation of retroviral LTRs reveals widespread long-range effects on human gene regulation. eLife. 2018;7:e35989.
20. Padmanabhan Nair V, Liu H, Ciceri G, Jungverdorben J, Frishman G, Tchieu J, et al. Activation of HERV-K (HML-2) disrupts cortical patterning and neuronal differentiation by increasing NTRK3. Cell Stem Cell. 2021;28:1566–1581.e8.
21. Aho AV, Corasick MJ. Efficient string matching: an aid to bibliographic search. Commun ACM. 1975;18:333–40.
22. Moreno-Mateos MA, Vejnar CE, Beaudoin J-D, Fernandez JP, Mis EK, Khokha MK, et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods. 2015;12:982–8.
23. Mendoza BJ, Trinh CT. Enhanced guide-RNA design and targeting analysis for precise CRISPR genome editing of single and consortia of industrially relevant and non-model organisms. Bioinform. 2018;34:16–23.
24. Lanciano S, Cristofari G. Measuring and interpreting transposable element expression. Nat Rev Genet. 2020;21:721–36.
25. Liu P, Cuerda-Gil D, Shahid S, Slotkin RK. The epigenetic control of the transposable element life cycle in plant genomes and beyond. Annu Rev Genet. 2022;56:63–87.
26. Pickar-Oliver A, Gersbach CA. The next generation of CRISPR–Cas technologies and applications. Nat Rev Mol Cell Biol. 2019;20:490–507.
27. McCarty NS, Graham AE, Studená L, Ledesma-Amaro R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. Nat Commun. 2020;11:1281.
28. Liu G, Zhang Y, Zhang T. Computational approaches for effective CRISPR guide RNA design and evaluation. Comput Struct Biotechnol J. 2020;18:35–44.
29. Abadi S, Yan WX, Amar D, Mayrose I. A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. PLoS Comput Biol. 2017;13:e1005807.
30. Chuai G, Ma H, Yan J, Chen M, Hong N, Xue D, et al. DeepCRISPR: optimized CRISPR guide RNA design by deep learning. Genome Biol. 2018;19:80.
31. Bostock M, Ogievetsky V, Heer J. D^3 Data-Driven Documents. IEEE Trans Vis Comput Graph. 2011;17:2301–9.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.