Check for updates

RESEARCH ARTICLE SUMMARY

GENOME ENGINEERING

Megabase-scale human genome rearrangement with programmable bridge recombinases

Nicholas T. Perry, Liam J. Bartie, Dhruva Katrekar, Gabriel A. Gonzalez, Matthew G. Durrant, James J. Pai, Alison Fanton, Juliana Q. Martins, Masahiro Hiraizumi, Chiara Ricci-Tam, Hiroshi Nishimasu, Silvana Konermann*, Patrick D. Hsu*

Full article and list of author affiliations: https://doi.org/10.1126/science.adz0276

INTRODUCTION: Editing the human genome has revolutionized biology and medicine, but current tools remain constrained by the length scale of edits they can make. CRISPR systems enable precise, programmable changes, yet they are limited to relatively small DNA modifications. By contrast, recombinases can mobilize large DNA fragments but at fixed or preinstalled recognition sites, often leaving behind unwanted sequence scars. A method that combines programmable precision with the ability to insert, excise, or invert megabase-sized stretches of DNA would greatly expand our ability to study and treat human disease.

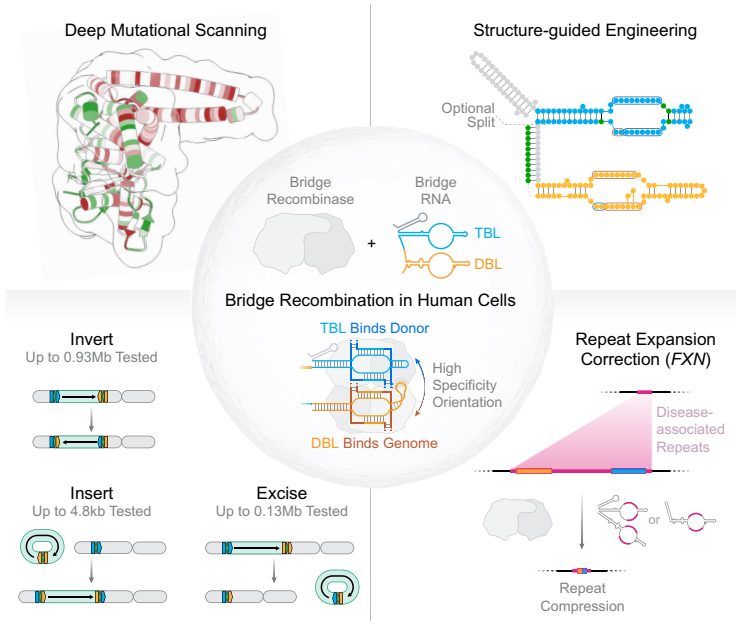
RATIONALE: Bridge recombinases are a recently discovered family of enzymes that use a short "bridge RNA" (bRNA) to direct sequence-specific recombination between a pair of DNA substrates. This bispecific guidance is conceptually and mechanistically distinct from that of CRISPR nucleases and classical recombinases. We previously showed they could be used to insert DNA into bacterial genomes, suggesting the possibility of much broader applications. In this study, we asked whether bridge recombinases could be adapted to human cells, systematically optimized for activity and specificity, and used to perform programmable rearrangements across unprecedented genomic distances.

RESULTS: By screening dozens of natural variants, we identified ISCro4, a bridge recombinase with robust activity in human cells. Through structural modeling, RNA engineering, and deep mutational scanning of the recombinase protein, we built an optimized system that achieved up to 20% efficiency and 82% specificity for inserting new DNA at chosen loci. Notably, an enhanced ISCro4 also performed programmable, precise, and scarless inversions approaching 1 Mb and excision events >100 kb, with no apparent distance dependency.

To demomstrate therapeutic potential, we programmed ISCro4 to excise the *BCL11A* enhancer, a clinically validated target for sickle cell anemia therapy, and remove expanded GAA (G, guanine; A, adenine) repeats that cause Friedreich's ataxia. These proof-of-concept applications show that bridge recombinases can tackle disease-relevant genomic lesions beyond the reach of current methods.

CONCLUSION: We present ISCro4 as a bridge recombinase system that establishes the first single-effector, RNA-guided technology capable of programmable megabase-scale editing of the human genome. By uniting the targeting flexibility of CRISPR with the payload mobility of recombinases, this technology enables flexible and large-scale DNA rearrangements that were previously inaccessible. Bridge recombinases can modify the genome at arbitrary new scales, ranging from single-gene insertions to megabase-sized rearrangements, which unlocks substantial potential for understanding cellular function and human disease pathology. □

*Corresponding author. Email: patrick@arcinstitute.org (P.D.H.); silvana@arcinstitute.org (S.K.) Cite this article as N. T. Perry *et al.*, *Science* **391**, eadz0276 (2026). DOI: 10.1126/science.adz0276



Developing bridge recombinases for megabase-scale genome rearrangements in human cells. (Center) Schematic of bridge recombinase, bRNA, and the recombination mechanism. (Top) Enhancing the bridge recombinase protein by deep mutational scanning (left) and bRNA by rational engineering (right). (Bottom) General applications of bridge recombinases in human cells (left) and potential therapeutic uses of bridge recombinase–mediated excision (right). TBL, target-binding loop; DBL, donor-binding loop.

GENOME ENGINEERING

Megabase-scale human genome rearrangement with programmable bridge recombinases

Nicholas T. Perry[1,2,3], Liam J. Bartie[1,2,3], Dhruva Katrekar[1], Gabriel A. Gonzalez[1], Matthew G. Durrant[1], James J. Pai[1], Alison Fanton[1,2,3], Juliana Q. Martins[1,4], Masahiro Hiraizumi[5], Chiara Ricci-Tam[1], Hiroshi Nishimasu[5,6,7], Silvana Konermann[1,8]*, Patrick D. Hsu[1,2,9]*

Bridge recombinases are naturally occurring RNA-guided DNA recombinases that we previously demonstrated can programmably insert, excise, and invert DNA in vitro and in *Escherichia coli*. In this study, we report the discovery and engineering of the bridge recombinase ortholog ISCro4 for universal rearrangements of the human genome. We defined strategies for the optimal application of bridge systems, leveraging mechanistic insights to improve their targeting specificity. Through rational engineering of the ISCro4 bridge RNA and deep mutational scanning of its recombinase, we achieved up to 20% insertion efficiency into the human genome and genome-wide specificity as high as 82%. We further demonstrated intrachromosomal inversion and excision, mobilizing up to 0.93 megabases of DNA. Lastly, we provided proof of concept for plasmid-based excision of disease-relevant gene regulatory regions or repeat expansions.

Genomes are the foundational information layer that encode biological complexity, from the activity of individual enzymes to coordinated cellular networks that orchestrate organism-level behavior (*1*, *2*). These genotype-to-phenotype relationships are determined by combinations of mutations across megabase-scale genomic regions, limiting our understanding of even the simplest genomes. The development of RNA-guided technologies, such as CRISPR genome engineering, has revolutionized our ability to interrogate biological function through programmable DNA modification (*3*, *4*). However, these approaches rely on complex multistep mechanisms, such as homology-directed repair or prime editing, that restrict the length scale of possible genome edits (*5*, *6*). In contrast to CRISPR, DNA recombinases, such as Cre, can mobilize many thousands of DNA base pairs but require preinstalled recognition sites to enable mammalian genome recombination and leave residual scars (*7*–*10*). Although fusion to RNA-guided systems or directed evolution can modify the sequence specificity of recombinase enzymes, such tools are largely limited to gene integration and leave long insertion site scars (*11*–*14*).

As the first described single-effector class of RNA-guided DNA recombinases, bridge recombinases have the potential to combine the flexible specificity of CRISPR with the payload scale of recombinases (*15*). Derived from IS110 insertion sequence elements, bridge recombinases recognize their target and donor DNA using a bridge RNA (bRNA) that confers programmable specificity for both substrates, enabling a universal mechanism for DNA insertion, inversion, and excision (Fig. 1A). By contrast, CRISPR-associated transposases specify only the target DNA through a guide RNA while requiring a fixed DNA donor sequence that limits them to gene insertion (*16*–*19*). Structural studies of the bridge recombinase IS621 revealed that two recombinase dimers bind both bRNA loops separately and then bind the cognate DNA to form a tetrameric synaptic complex (*20*) (Fig. 1B). Strand exchange occurs after the formation of covalent bonds between the DNA and the recombinase, resulting in a Holliday junction intermediate that is resolved to yield the final recombination product (Fig. 1C).

Our initial study provided proof of concept for using bridge recombinases in programmable modification of prokaryotic genomes (*15*). In this study, we describe the discovery of a bridge recombinase with high activity in human cells, ISCro4. We enhanced its activity by rational engineering of its bRNA through point mutations and scaffold modifications. Using these enhanced bRNAs, we discovered design principles for maximizing the specificity of insertion into the human genome, achieving as high as 82% specificity. To further augment activity of ISCro4, we performed a deep mutational scan of the recombinase directly in human cells. We combined a rationally engineered high-activity recombinase mutant with our enhanced bRNAs, achieving up to 20% insertion efficiency into the human genome. Beyond gene insertion, we used ISCro4 for programmable, precise, and scarless genome rearrangements, inverting up to 0.92 Mb and excising up to 0.13 Mb with no apparent distance dependency. Lastly, we demonstrated therapeutic proof of concept with excision of the *BCL11A* enhancer for sickle cell anemia and of repeat sequences found in Friedreich's ataxia.

Results

Diverse bridge recombinases are active in human cells

We previously characterized reprogrammable bridge recombination in vitro and in *Escherichia coli*. In our previous efforts to effectively use prokaryotic enzymes in eukaryotic systems, broad screening of diverse orthologs with a wide range of activity was necessary before an optimal ortholog was selected for further development (*21*, *22*). Leveraging our database of IS110 elements (*15*), we predicted the recombinase coding sequences and cognate bRNAs for 72 diverse IS110 insertion sequence elements (fig. S1, A to C, and table S1). Using the IS110 element boundaries determined by comparative genomics, we reconstituted the natural target and donor sequences and tested them in an inversion reporter assay in human cells (Fig. 1D). The inverted target-donor junction was detected for 18/72 (25%) orthologs, confirming that at least some bridge recombinases do not require prokaryotic host factors to function in human embryonic kidney (HEK) 293FT cells (Fig. 1E).

Next, we quantified the efficiency of these 18 functional orthologs by measuring inversion of *Gaussia* luciferase (Fig. 1F). In this assay, only a single ortholog exhibited activity above the limit of detection despite all orthologs having observable activity through nonquantitative amplification of the inversion junction (Fig. 1G). The active ortholog, previously labeled 23122 (*15*) and with a fourfold change in relative luciferase over the background, is from *Citrobacter rodentium* and identified as ISCro4 in ISFinder (*23*). This ortholog is highly similar to the previously described IS621 ortholog (88% recombinase and 86.6% bRNA identity) (fig. S2, A to B). A comparison of an AlphaFold 3–predicted model of the ISCro4 recombinase with our cryo–electron microscopy (cryo-EM) structure of the IS621 recombinase (*20*) confirmed the high structural similarity between the recombinases, with root mean square deviation (RMSD) of 2.5 Å for equivalent Cα atoms (fig. S3, A to D), allowing modeling of the likely bRNA structure (Fig. 1H).

[1]Arc Institute, Palo Alto, CA, USA. [2]Department of Bioengineering, University of California, Berkeley, Berkeley, CA, USA. [3]University of California, Berkeley - University of California, San Francisco, Graduate Program in Bioengineering, Berkeley, CA, USA. [4]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA. [5]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan. [6]Structural Biology Division, Research Center for Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo, Japan. [7]Inamori Research Institute for Science, Shimogyo-ku, Kyoto, Japan. [8]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA. [9]Center for Computational Biology, University of California, Berkeley, Berkeley, CA, USA. ***Corresponding author. Email: patrick@arcinstitute.org (P.D.H.); silvana@arcinstitute.org (S.K.)**

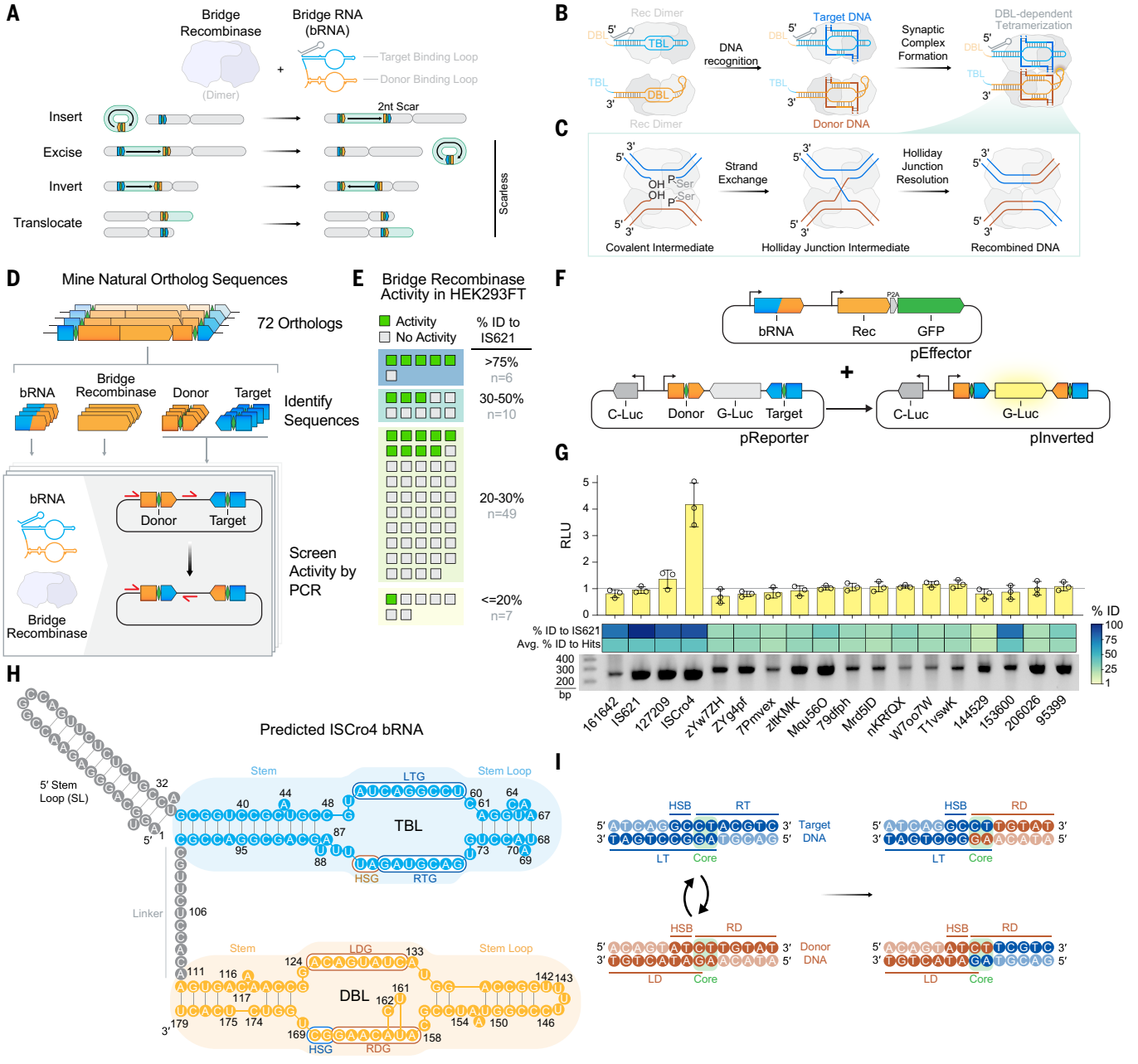


Fig. 1. Screening activity of bridge recombinases in human cells. (**A**) Schematic of chromosome-scale genome modifications with bridge recombination. (**B**) Overview of bridge recombinase synaptic complex assembly based on IS621. (**C**) Overview of strand exchange, Holliday junction formation, and resolution within a bridge recombinase transpososome based on IS621. (**D**) Schematic of ortholog screening approach. Key components of bridge recombinase systems from IS110 elements were encoded on reporter plasmids, where inversion between target and donor is measured to determine activity. (**E**) Activity of bridge recombinases in HEK293FT cells based on PCR detection of the inversion product junction. Orthologs are categorized by percent identity to IS621. (**F**) Schematic of C- and G-luciferase inversion reporter assay. p, plasmid. (**G**) Activity of orthologs in luciferase assay. Amplification of inversion junction shown below. Percent identity to IS621 and percent identity to other active orthologs is shown. RLU, relative luminescence units. Mean ± s.d. of three biological replicates. (**H**) Schematic of ISCro4 bRNA, based on IS621 bRNA structure (PDB: 8WT6). LTG, left target guide; RTG, right target guide; LDG, left donor guide; RDG, right donor guide; HSG, handshake guide. (**I**) Target and donor DNA recognized by the ISCro4 bRNA. Recombination occurs between CT cores. HSB, handshake bases.

The model suggests that the ISCro4-bRNA complex recognizes 14 nucleobases in both its target and donor DNA substrates, using its target and donor binding loops (TBL and DBL) in the bRNA, respectively. These two bRNA loops recognize both strands of the target and donor sequences, resulting in conservative recombination around a dinucleotide core sequence (Fig. 1I). Additionally, both loops encode handshake guides (HSGs) that base pair upstream of the core sequences with the handshake bases (HSBs) after top strand exchange, a mechanism that drives the recombination reaction forward (*20*) (fig. S4).

Structure-guided bRNA engineering

We next reasoned that rational engineering of the bRNA may yield increased activity of ISCro4 in human cells, as guide RNA engineering has been a longstanding strategy for increasing the activity of CRISPR

systems in human cells (*24*–*26*). Despite encoding specificity for both DNA substrates on a single RNA molecule, our structural study of IS621 suggested that two distinct bRNA molecules may be used during recombination by the tetrameric complex (*20*). Building on this observation, we asked whether separate expression of TBL and DBL may enhance transpososome assembly by preventing the linkage of multiple synaptic complexes through single bRNAs (*20*) (Fig. 2A). Using an mCherry inversion reporter, we compared the full-length bRNA with two versions of a split bRNA with or without the 5′ stem-loop. We observed that split bRNAs were not only functional but also resulted in ~2.1-fold increase in recombination efficiency by mCherry mean fluorescence intensity (MFI) (Fig. 2, B and C).

In the wild-type (WT) bRNA, the TBL and DBL are connected by a linker region that connects the stem regions of the two loops to one

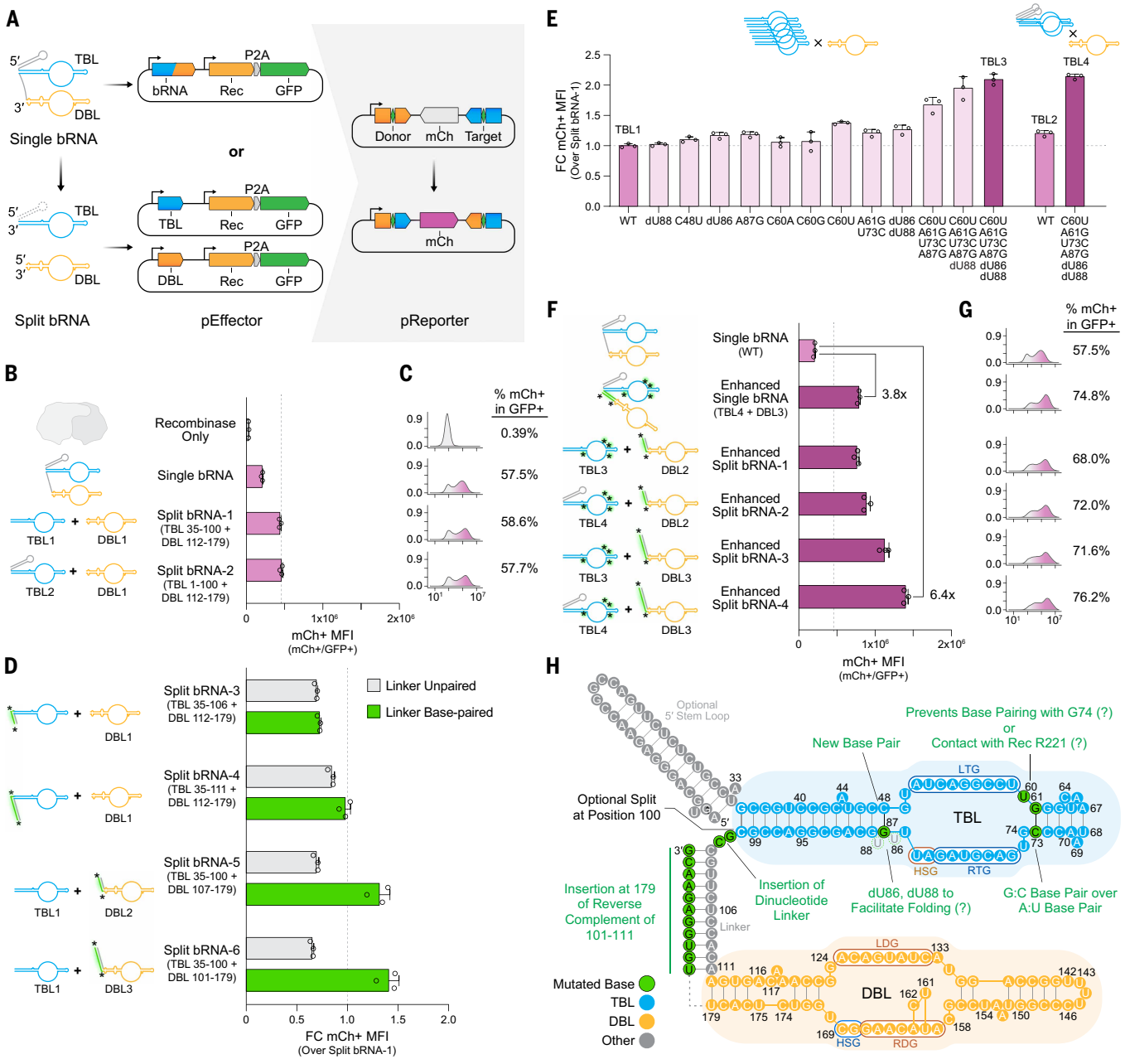


Fig. 2. bRNA engineering for efficient recombination. (**A**) Overview of recombinase and bRNA expression plasmids with mCherry (mCh) inversion reporter. (**B**) Inversion efficiency with single and split bRNAs. TBL and DBL intervals based on 179-nt single bRNA. Data are mean ± SD of three biological replicates. (**C**) Inversion efficiency as measured by percentage of mCherry+ cells within GFP+ (recombinase expressing) cells. (**D**) Inversion efficiency (MFI) of TBL and DBL with extended stems. All or part of the bRNA linker sequence (positions 101 to 111) was included on a TBL or DBL with or without insertion of a complementary sequence (green line with asterisks). Represented as fold change (FC) in MFI compared with that of split bRNA-1 in (B). Data are mean ± SD of three biological replicates. (**E**) Inversion efficiency of mutated TBLs when paired with DBL1. TBL2 and TBL4 include the 5′ stem-loop of bRNA (positions 1 to 34). Data are mean ± SD of three biological replicates. (**F**) Inversion efficiency (MFI) of single and split bRNAs featuring point mutations and extended stems. Data are mean ± SD of three biological replicates. (**G**) Inversion efficiency (percentage mCherry+ within GFP+) of single and split bRNAs featuring point mutations and extended stems. (**H**) Schematic of enhanced bRNA. Green text indicates enhancing mutations relative to the WT bRNA.

another. When determining the optimal region to split the bRNA into its two loops, we found that inclusion of part or all of the linker region (positions 101 to 111) onto either loop greatly reduced activity relative to eliminating the linker (Fig. 2D, gray boxes). Whereas the TBL contains a 12 bp continuous stem with a single base (A44) flipped out, the DBL stem contains three unpaired positions in the middle flanked by three and four paired bases on each side (Fig. 1H). We therefore asked whether increased stem length would modulate activity. By inserting bases to pair with the natural unpaired linker region, we increased the TBL stem length by 6 or 11 bp. However, we did not observe an increase of activity compared with that of Split bRNA-1, which terminates the TBL at the end of its natural stem. In comparison, lengthening the shorter DBL stem by 6 bp (DBL2) and 11 bp (DBL3) resulted in 1.3- and 1.4-fold greater activity over DBL1, respectively (Fig. 2D, green boxes). Taken together, we found that limiting unpaired linker bases and modulating stem length are effective at increasing recombination activity in a loop-specific manner.

We next tested the effects of rational substitutions and deletions in both loops, focusing on maximizing the number of base-paired nucleotides to help stabilize the bRNA structure. We initially assessed these changes on each loop separately using the split bRNA architecture. Whereas DBL mutations yielded limited improvement over the unmodified split system (0.2- to 1.1-fold change) (fig. S5A), substitutions and deletions in the TBL were more successful (1.0- to 1.4-fold change) (Fig. 2E). We next combined these three substitutions and two deletions with an A:U→G:C base pair modification (A61G and U73C) and observed cumulative increases in activity, culminating in TBL3 (2.1-fold greater activity over that of TBL1). Inclusion of the natural 5′ stem-loop (TBL4) additionally slightly increased activity (2.1-fold greater activity over that of TBL1), similar to our previous study of the IS621 complex (*20*).

We next modified the single bRNA scaffold to contain TBL4 and DBL3, which each performed optimally in the split system. We observed a 3.8-fold increase in activity in the context of the single bRNA (Fig. 2, F to G) and named this the enhanced bRNA. To determine the optimal split bRNA configuration, we tested all four combinations of the top two TBLs and top two DBLs. The best split loop pairing, TBL4 with DBL3, exhibited a 6.4-fold increase in activity over that of the WT bRNA, highlighting synergistic effects of individual loop modifications.

Modeling of our enhanced bRNA supports potential mechanistic explanations for the improved activity (Fig. 2H). Multiple mutations (A87G, A61G, and U73C) in the TBL that were observed to increase activity likely contribute to stabilizing the TBL stem by replacing two existing base pairs with a stronger G:C pair. C60U (the strongest individual mutation) is expected to either help delineate the TBL stem boundary by preventing undesirable base pairing with G74 or by making a base-specific contact with the recombinase at R221 (U60:R221), similar to the corresponding interaction in the IS621 structure (U59:R221) (PDB: 8WT6; Fig. 2E). We found that the unpaired U86 and U88 bases on the other hand were dispensable and that their removal in fact enhanced activity (dU86 and dU88) (Fig. 2E). In the context of the split bRNA, unpaired bases corresponding to the bRNA linker between the TBL and DBL reduced efficiency (Fig. 2D), suggesting base pairing to this region, which extends the stem region, contributing to additional stabilization. Lastly, although lengthening the linker between the loops in a single bRNA slightly enhanced efficiency, split bRNAs remained more efficient (fig. S5B).

DBL-DBL recombination with ISCro4

The 14–nucleotide (nt) target sequence naturally recognized by the ISCro4 bRNA appears only once in the human genome, making it a straightforward test case for sequence-specific insertion. We cotransfected plasmids bearing the recombinase and WT bRNA into human cells with a plasmid encoding the WT donor sequence and a constitutive puromycin marker (Fig. 3A). After 18 days of selection to enrich for both on- and off-target insertions, we measured the relative efficiency of insertion genom-wide through deep sequencing (materials and methods). Although we observed insertion reads mapping to the expected target, these represented only 0.05% of insertions. Upon further inspection, we found that the consensus insertion sequence across all insertion sites closely matched the donor sequence rather than the target sequence (Fig. 3B), and 91.3% of insertions were more similar to the plasmid donor sequence than the expected genomic target (Fig. 3C). Out of all insertions, 75.5% were within 4-nt Levenshtein distance from the WT donor sequence, including 1.59% into human genome occurrences of the WT donor sequence. By contrast, only 0.25% of insertions were within 4 nt of the WT target sequence (fig. S6, A to B). The higher similarity of off targets to the sequence specified by the DBL suggested a distinct sequence recognition mechanism for ISCro4.

In our previous study of insertion specificity of IS621 in *E. coli*, we observed low-frequency recombination between the plasmid-encoded donor sequence and donor-like sequences matching the DBL in the genome (*15*). Structural and biochemical studies indicated that this "donor-donor" reaction facilitated by IS621 is likely a result of transpososomes containing two DBLs from separate bRNAs rather than one TBL and one DBL (Fig. 3D) (*20*). DBL-DBL but not TBL-TBL transpososomes are possible owing to the distinctive stem-loop of the DBL, which is involved in tetramer stabilization (*20*) (figs. S2B and S6, C to D).

Given that donor-like insertion sites (matching the DBL instead of the TBL) were much more abundant, we used an mCherry inversion reporter to compare recombination between target-donor and donor-donor with the WT bRNA, confirming that both recombination types can occur robustly (Fig. 3, E to F). We then delivered only a DBL from the split design, demonstrating that it is sufficient for recombination and in fact leads to higher donor-donor recombination than the full-length single bRNA containing both DBL and TBL. Altogether, these results illuminate factors influencing donor-donor DNA rearrangement with bridge recombinases.

Principles of improved human genome insertion specificity

To favor on-target insertion and mitigate DBL-only mediated recombination, we made three adjustments: (i) We encoded a unique bRNA recognition sequence not found in the genome (i.e., genome orthogonal) on the delivered plasmid donor, (ii) used the enhanced version of the single bRNA (which may limit donor-donor recombination in comparison to the more active split bRNA system), and (iii) programmed the DBL to recognize the genome and the TBL to recognize the donor plasmid. To test this strategy, we selected 10 genomic loci that contain a unique 14-nt sequence and paired these sequences with a genome-orthogonal donor sequence, encoded within a puromycin-resistant plasmid (Fig. 3G). We designed two enhanced single bRNAs for each sequence pair: one where the TBL recognizes the genome and the DBL recognizes the plasmid and vice versa. We first evaluated the efficiency of the two orientations on day three after transfection using droplet digital polymerase chain reaction (ddPCR), observing a range of 0.18 to 3.38% for the condition in which the TBL binds the genome and 0.10 to 9.82% for that in which the DBL binds the genome (Fig. 3H).

We next measured the specificity of these insertions after 18 days of puromycin selection. For 8 of 10 loci, we noticed that DBL targeting the genome was substantially more specific than TBL targeting the genome, reducing the number of off targets by an average of 90.4% (range 87.0 to 94.1% reduction) (Fig. 3I and fig. S7A). Despite confirming on-target insertion at each of the 10 loci 3 days after transfection (Fig. 3H), we observed on-target insertion for only four loci after 18 days of selection (Fig. 3, J to M, and table S2), including for each of the top three most-efficient loci (*SBF2*, *ANP32A*, and *IFNGR1*). We did not observe any correlation between insertion efficiency and specificity, indicating that the specificity gain with DBL is not an artifact of lower insertion efficiency (fig. S7B). Further analysis demonstrated that encoding a genome-orthogonal sequence on the plasmid donor additionally

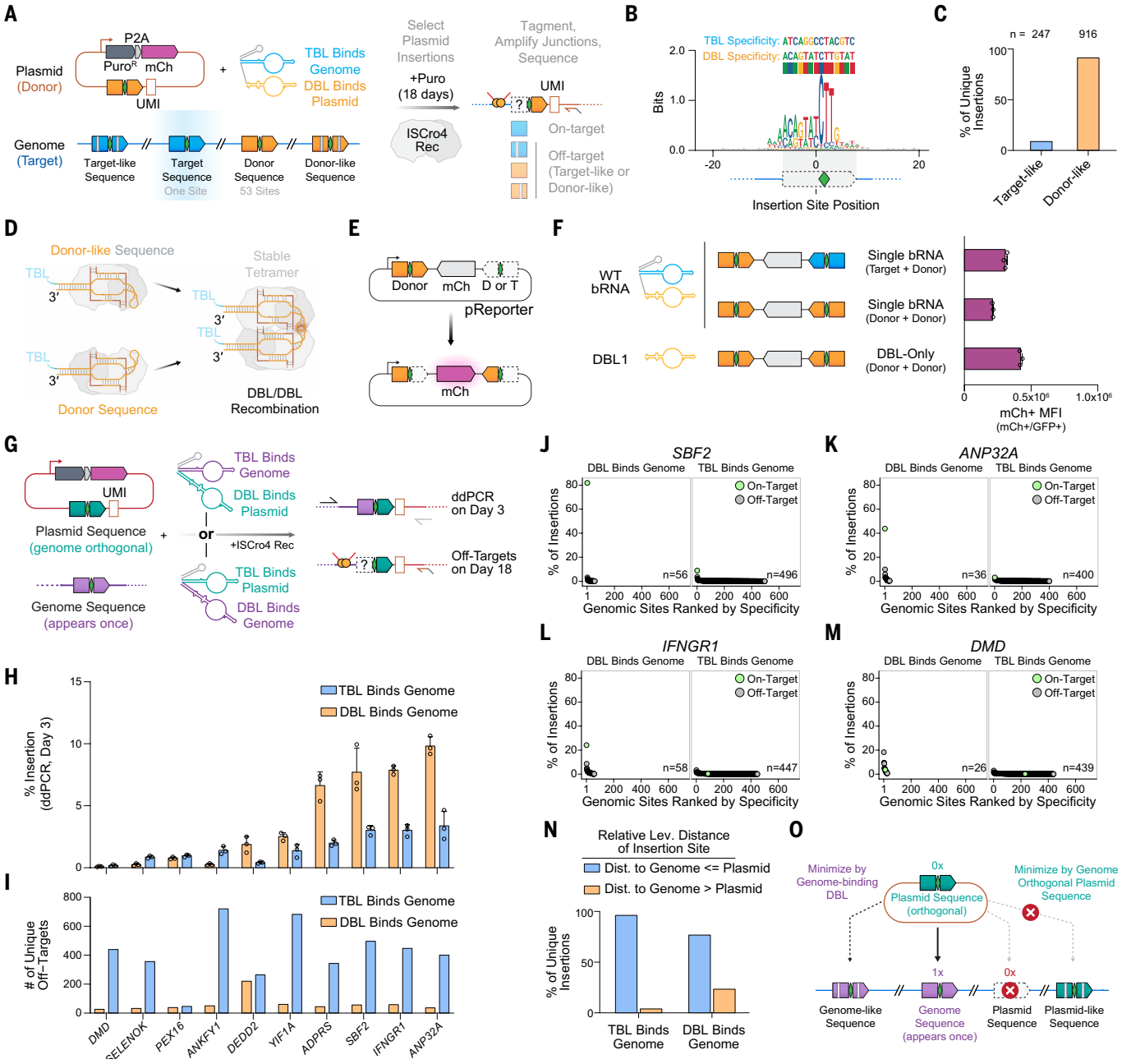


Fig. 3. DBL binding to the genome enables high-specificity insertion. (**A**) Schematic of genome-wide insertion specificity assay, with WT bRNA shown as an example. On and off targets are measured after insertion of puromycin resistance gene and antibiotic selection. UMIs are counted genome-wide to determine specificity across all insertions. (**B**) Sequence logo of all insertion sites using the WT bRNA, weighted by observed UMIs. TBL and DBL recognition sequences are shown. (**C**) Percentage of all insertions binned by similarity to ISCro4 WT target sequence or donor sequence. (**D**) Schematic of DBL-DBL–mediated recombination. DBL ribonucleoproteins bind DBL-like DNA and tetramerize, yielding recombination. (**E**) Schematic of mCherry inversion reporter. (**F**) Inversion efficiency of single and split bRNAs for target + donor or donor + donor recombination. Data are mean ± SD of three biological replicates. (**G**) Schematic of approach for measuring efficiency and specificity of insertion into the genome sites using both enhanced bRNA orientations. (**H**) Insertion efficiency by ddPCR 3 days after transfection (no selection) at 10 genomic loci using both bRNA orientations. Data are mean ± SD of three biological replicates. (**I**) Number of unique off targets observed for each locus with each bRNA orientation. (**J** to **M**) Insertion specificity into *SBF2* (J), *ANP32A* (K), *IFNGR1* (L), and *DMD* (M) using both bRNA orientations after 18 days of selection. (**N**) Percentage of unique insertions binned by the relative Levenshtein distance of the insertion site to the intended genome site or the plasmid-encoded sequence. Results are cumulative of all 10 loci in (I). (**O**) Model for minimizing off-target insertions with ISCro4 bridge recombinases.

helped to eliminate plasmid-like off-targets for either bRNA orientation (Fig. 3N and fig. S7, C to D).

Given the recovery of on-target insertions for the least-efficient target (*DMD*), factors influencing specificity beyond loop orientation and initial insertion efficiency, such as differential fitness, should be considered. The interruption of genomic loci by plasmid insertion, whether on or off target, may positively or negatively influence fitness on a locus-specific basis, an effect that may be magnified with each additional day the population is under selection for the inserted antibiotic resistance gene. In addition to fitness changes due to interruption of specific loci, certain loci may induce differential silencing of the inserted payload, which could lead to their depletion and skew the off-target

profile. However, neither of these effects would be expected to bias cells specifically toward the on-target insertion site across diverse loci. Although we did not observe overt toxicity resulting from bRNA or recombinase expression, further studies would be necessary to determine whether they could affect the off-target profile. We used a genome-wide off-target assay that measures insertions at day 18, which we preferred over targeted amplification of genomic sites with sequence similarity, as that would have likely underestimated off-target insertions. Measurement of off-targets at earlier time points (e.g., day three) will likely be closer to the ab initio off-target profile but would require technical innovations to deplete uninserted donor plasmids that contaminate genome-derived reads carrying off-target information (*22*, *27*).

Our results support our strategy for maximizing efficiency while balancing specificity (Fig. 3O). We observed that recombination efficiency between a given pair of sequences can vary depending on which loop is programmed to which substrate, suggesting that both orientations should be tested for any gene of interest when optimizing for efficiency. Off targets were consistently reduced by targeting the DBL instead of the TBL to the genome, an observation that is agnostic to the relative on-target insertion efficiency of the two loop orientations (fig. S7B). With DBL targeting of the genome, we achieved up to 82% on-target specificity (*SBF2*) and up to 9.8% efficiency (*ANP32A*), metrics that are comparable to nonprogrammable, site-specific recombinases (*22*).

To confirm the generalizability of bridge recombination in multiple cell types, we performed episomal recombination in U2OS, K562, and H1 human embryonic stem cells (hESCs), observing robust recombination rates across all lines (fig. S8, A to D). To further assess bridge recombination in dose-limited settings beyond high-copy plasmid delivery, we lentivirally integrated the recombinase at low multiplicity of infection (MOI) and demonstrated low-efficiency genome insertion in U2OS, HEK293FT, and H1 hESC cells (fig. S9, A to D). Taken together, we propose that using a genome-orthogonal sequence on the donor plasmid and programming the DBL to target the genome is the optimal approach for performing programmable genome insertion with ISCro4. These insights will facilitate future applications of bridge recombination in additional cell types.

Deep mutational scan of the ISCro4 recombinase

To further augment the effectiveness of ISCro4, we performed a deep mutational scan of the recombinase protein directly in human cells with the intention of bypassing variants that only improve activity in bacteria. To ensure compatibility with short-read sequencing, we generated five libraries, each representing one-fifth of the recombinase (65 amino acids), which included all possible mutations at each position (materials and methods). To incorporate these libraries into the genome, we first created a large serine recombinase (LSR) landing pad K562 cell line, which enables site-specific and single-copy integration in contrast to random lentiviral integrations (*28*, *29*). This landing pad cell line was designed to capture ISCro4 recombinase variants at the AAVS1 safe harbor locus using our previously reported Dn29 LSR (*22*), enabling high library coverage for robust screening (Fig. 4A and fig. S10, A to G). We transfected the resulting ISCro4 variant cell libraries with the enhanced bRNA and an inversion reporter, sorted them to isolate higher activity mutants, and then sequenced the recombinase in the pre- and postsort populations (Fig. 4A). We detected 99.6% of all possible single–amino acid substitutions in the presort cell population, and 51% of variants were enriched in the sorted population (fig. S11A and table S3). Recombinase mutants with 65 consecutive amino acids mutated (serving as negative controls) and mutations of the recombinase catalytic residues were highly depleted in the sorted population, and less than 8% of mutations, spread across 30% of residues, were enriched above WT in the pooled screen (Fig. 4B and fig. S11B).

To interpret the mutational patterns, we developed a mutability score, where we ranked each position of the recombinase according to how many amino acid substitutions per position fell within one standard deviation of the WT sequence (Fig. 4C and fig. S11C). Based on this score, the RuvC domain is more mutable than the CC or Tnp domains, with the CC domain being the least mutable overall. This finding is consistent with the critical roles of the CC domain in dimerization and the Tnp domain in bRNA binding (*20*). Based on the structural comparison between IS621 (PDB: 8WT6) and ISCro4 (AlphaFold 3 model) (fig. S3, A to D), we decided to use the IS621 structure to explore the likely solvent accessibility of specific amino acids in the context of the synaptic complex (materials and methods). We found that solvent-exposed residues were enriched in the screen (fig. S11D). Additionally, most mutations in the wedge residues of the Tnp domain were also depleted, consistent with their importance for DNA unwinding for base pairing with the bRNA (*20*). Based on both their mutability scores and fold-change activity, we selected variants and tested their activities for insertion into *SBF2* (Fig. 4D and fig. S11C). Among these mutations, the S30T, P54Q, and S243H mutations consistently increased activity across genes (Fig. 4E). We combined these mutations to generate a triple mutant, enhanced ISCro4, which exhibited consistent activity of 1.5-fold across genes and no apparent change in on-target insertion relative to the wild-type ISCro4 recombinase (Fig. 4E and fig. S12, A to C).

To interpret the potential mechanism of these activity-enhancing mutations, we mapped the S30T, P54Q, and S243H mutations in ISCro4 onto the IS621 transpososome structure (PDB: 8WT6) (Fig. 4F), as IS621 and ISCro4 share 88% sequence identity with no gaps and have the same number of residues (fig. S2A). P54 in ISCro4, equivalent to H54 in IS621, is likely located at a solvent-exposed loop, indicating that the P54Q mutation may improve solubility or aid in protein folding. In the IS621 structure, T30 hydrogen bonds with U142 of the DBL, stabilizing the tetrameric complex (*20*) (fig. S13A). It seems likely that S30 of ISCro4 forms a similar hydrogen bond with the DBL and that the $Ser^{30}Thr$ mutation enhances activity owing to a favorable additional contact between the methyl group of T30 and the adjacent D20 residue (fig. S13, B and C). The catalytic site of IS621 is created only upon tetramerization and consists of the RuvC domain from one dimer and the Tnp domain from the opposing dimer (*20*). The catalytic S241 forms a covalent bond with DNA, whereas the main chains of K14 and S243 interact with the DNA phosphate backbone at the catalytic site (Fig. 4G). Modeling of S243H suggests that the H243 side chain forms a new hydrogen bond with the DNA phosphate backbone (Fig. 4H), thereby increasing activity.

In the validations of our deep mutational scan, we evaluated mutations using the enhanced bRNA, where the DBL was programmed to bind the genome. To assess whether the enhanced ISCro4 recombinase was generalizable to other bRNA formats, we evaluated either bRNA orientation with various bRNA scaffolds (Fig. 4, I to K). Across all conditions, increased activity was consistently observed with the enhanced bRNA scaffolds in comparison to the unmodified bRNAs. Pairing these with the engineered recombinase yielded up to 20.2% insertion at *SBF2*. Across three targets, we show that the original bridge recombination system (using WT bRNA where the TBL binds the genome) can be outperformed by up to 9.8-fold by utilizing the optimal loop configuration, our enhanced split bRNA scaffold, and our engineered recombinase.

Megabase-scale genome rearrangements with bridge recombination

Genomic insertion has been achieved with many natural and engineered molecular systems (*30*, *31*). Some of these tools have been co-opted for genomic rearrangement, but none of them are naturally capable of this function. Bridge recombinases are distinctive in that they can be programmed to recognize almost any pair of genomic sequences without the preinstallation of recombinase recognition sites common among other approaches (*11*, *12*).

To demonstrate this capability in human cells, we took a similar approach to our genome insertion assays. We first nominated pairs of

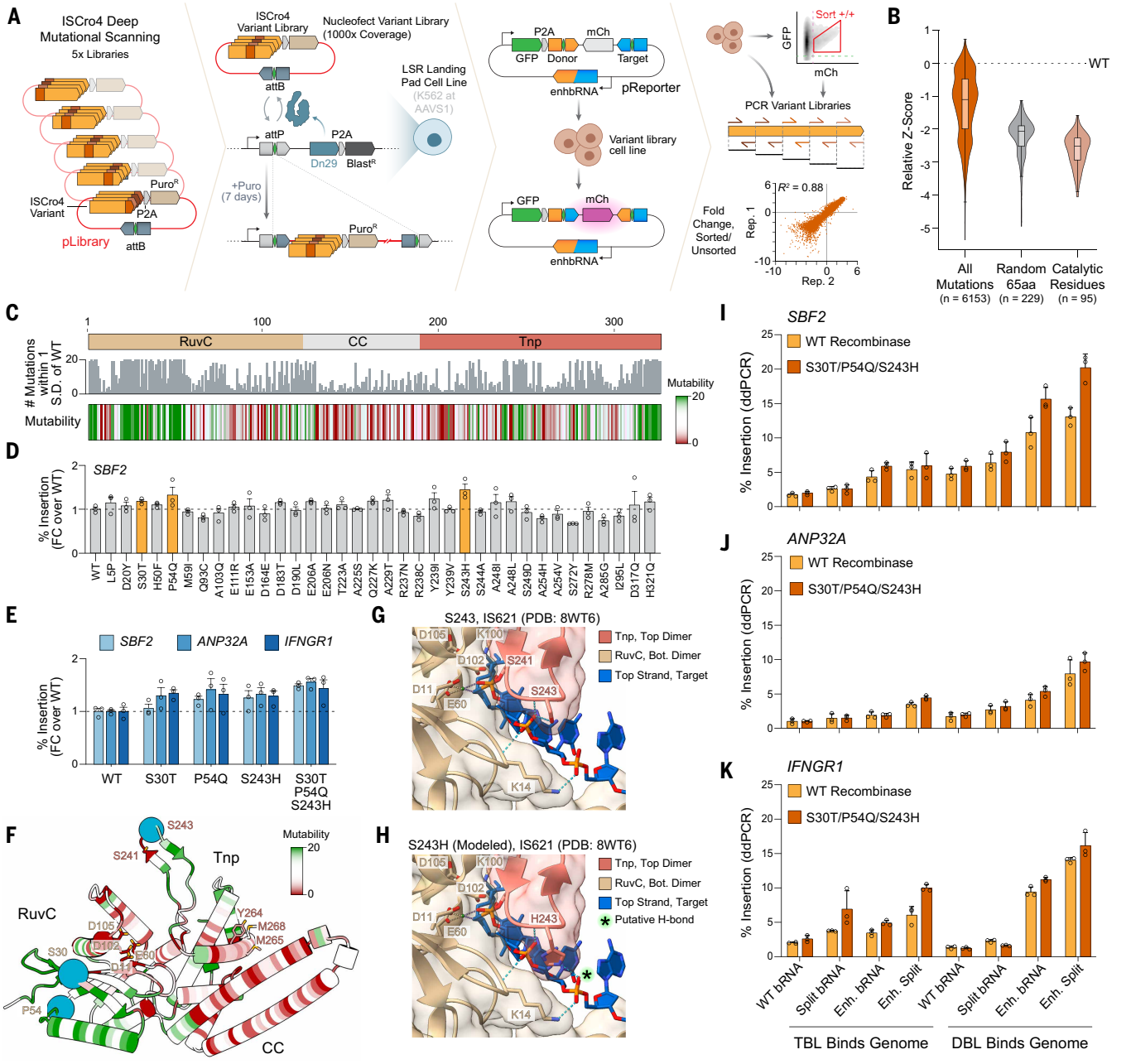


Fig. 4. Improving ISCro4 activity by deep mutational scanning in human cells. (**A**) Schematic of approach for a deep mutational scan of ISCro4 recombinase in human cells. Five libraries of single mutants are inserted into AAVS1 using Dn29 large serine recombinase. The genomically expressed recombinase uses a transfected enhanced bRNA to recombine an inversion reporter. Cells are sorted based on mCherry/GFP ratio prior to sequencing. (**B**) *Z*-score relative to that of the WT for all mutations in deep mutational scan compared with random sequences serving as negative controls and the catalytic residue positions (11, 60, 102, 105, and 241). (**C**) Mutability of the recombinase represented as the number of mutations within one SD of the WT at each position. (**D**) Relative insertion efficiency into *SBF2* with selected single mutants. Insertion efficiency was measured using the enhanced bRNA with DBL binding the genome. Data are mean ± SD of three biological replicates. (**E**) Relative insertion efficiency across genes with selected single mutants and a triple mutant. Data are mean ± SD of three biological replicates. (**F**) AlphaFold 3 model of ISCro4 recombinase monomer. Color represents the per-position mutability score. Catalytic residues and wedge residues are highlighted as stick models, and residues mutated in the triple mutant are highlighted as sphere models. (**G** and **H**) DNA recognition by S243 in the IS621 structure (PDB: 8WT6) (G) and S243H in the IS621 model (H). Hydrogen bonds are depicted with blue dashed lines. (**I** to **K**) Genome insertion efficiency with WT and S30T/P54Q/S243H recombinase using various bRNA formats into *SBF2* (I), *ANP32A* (J), and *IFNGR1* (K). Enh., enhanced. Data are mean ± SD of three biological replicates. Single-letter abbreviations for the amino acid residues referenced throughout the paper and figures are as follows: S, Ser; T, Thr; P, Pro; Q, Gln; S, Ser; H, His; D, Asp; K, Lys; Y, Try; L, Leu; F, Phe; M, Met; I, Ile; A, Ala; E, Glu; R, Arg; M, Asn; C, Cys; V, Val.

unique sequences in the genome, filtering for matching core sequences and nonmatching positions 6 and 7 (materials and methods). Depending on the relative strands of these two sequences, recombination between the two sites yields intragenomic inversion or excision, which can be measured by tagmentation and amplicon sequencing (Fig. 5A). The effect of loop orientation on insertion efficiency prompted us to assess both configurations of the enhanced bRNA for these rearrangements (site 1 bound by TBL, site 2 bound by DBL, and vice versa). Using WT ISCro4 recombinase, we measured recombination efficiency for seven pairs of sequences resulting in inversion (Fig. 5B) and seven

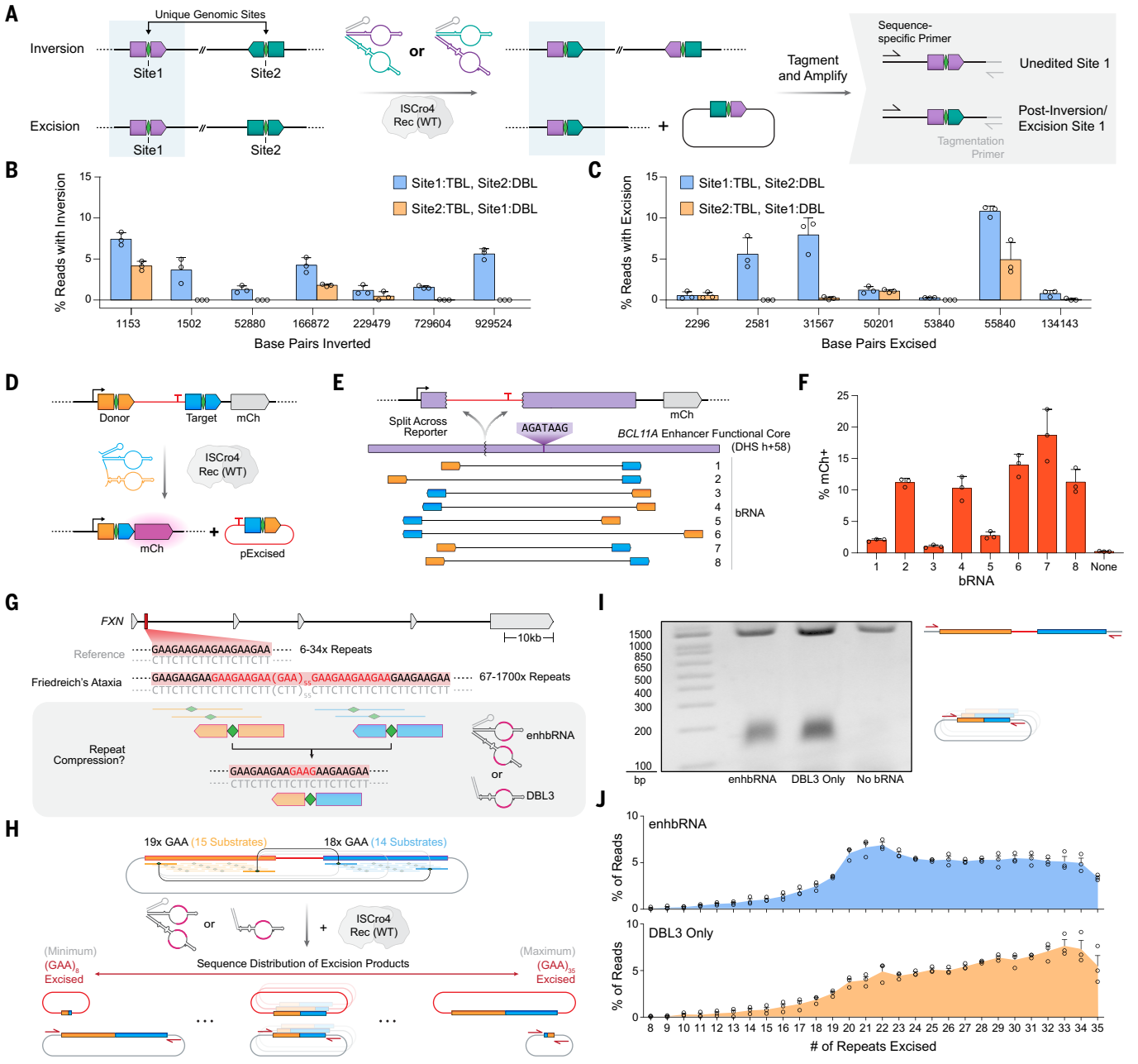


Fig. 5. Applications of programmable genomic rearrangement with ISCro4. (**A**) Schematic of performing intrachromosomal recombination. Enhanced bRNAs in both orientations are used to yield inversion or excision, which is measured by tagmentation and next-generation sequencing. (**B**) Efficiency of genome inversion across seven sequence pairs. Each pair of bars represents one pair of sequences recombined with either orientation of bRNA. Data are mean ± SD of three biological replicates. (**C**) Efficiency of genome excision across seven sequence pairs. Each pair of bars represents one pair of sequences recombined with either orientation of bRNA. Data are mean ± SD of three biological replicates. (**D**) Schematic of excision reporter assay. Removal of a terminator ahead of the promoter yields mCherry expression. (**E**) Schematic of bRNA recognition sites used to recombine the *BCL11A* enhancer region on the plasmid excision reporter. (**F**) Efficiency of *BCL11A* enhancer excision with eight unique bRNAs. Data are mean ± SD of three biological replicates. (**G**) Overview of *FXN* repeat expansion and compression through bridge recombination. (**H**) Schematic of plasmid reporter for recombination of repeat regions. Possible recognition sites are represented by colored lines, with each green diamond representing the core of a recognition site. Recombination results in a distribution of product lengths. (**I**) Agarose gel showing efficiency of repeat recombination with (i) an enhanced bRNA and (ii) a DBL (DBL3) in comparison to (iii) no bRNA. The lower band represents the range of remaining repeats. (**J**) Distribution in the number of repeats excised, measured through next-generation sequencing. The data are mean ± SD of three biological replicates.

pairs resulting in excision (Fig. 5C). We observed up to 7.42% efficiency of inversion (1153 bp) and 10.8% efficiency of excision (55,840 bp). The efficiency of recombination had no clear correlation with the distance between the recombined sites, and we were able to invert as long as 929,524 bp (0.92 Mbp, 5.62% efficiency) and excise as long as 134,143 bp (0.13 Mbp, 0.80% efficiency).

Across both rearrangement types, we observed that the bRNA orientation again influences the outcome. For 9 of the 14 sequence pairs we tested, only one orientation mediated recombination, with the other yielding little to no rearrangement (0 to 0.21%). For the remaining five sequence pairs, we observed recombination with both orientations at similar efficiency (relative fold change of 1.1 to 2.5). Furthermore, we

observed that the enhanced split bRNA increased the efficiency of recombination for 11 out of 12 bRNAs tested (0.89- to 5.65-fold change; mean, 2.1) (fig. S14, A to C). These results suggest that the main driver of rearrangement efficiency is intrinsic to the recombinase and bRNA rather than to the specific pair of sequences and their respective genomic contexts.

As with any genome editing tool, it is important to consider the potential for bridge recombination to cause unwanted changes to the genome. To begin to address this, we measured the rate of indel formation when performing inversion or excision, observing low rates of indels both prior to (mean, 0.71%; median, 0.21%) and after (mean, 0.354%; median, 0.221%) recombination. We compared these rates against background indel formation in the absence of bRNA (mean, 0.61%; median, 0.13%) (fig. S14, D to E). Furthermore, bRNAs may mediate recombination between one of the intended sites and off-target sites or between pairs of off-target sites within the human genome. These potential off-target genomic rearrangements, which could be mediated through TBL-DBL or DBL-DBL activity, will be an important consideration for their application in therapeutically relevant contexts, whether performing insertion, excision, or inversion. Future work to systematically understand the off-target profile of bridge recombinases will be needed to advance the clinical translation of this genome rearrangement technology.

BCL11A enhancer excision with bridge recombination

The ability to modify the genome in a sequence-specific manner is a proven method for interrogating genotype-phenotype relationships (*32*–*35*). In particular, CRISPR functional genomics has unveiled a number of new disease targets (*36*), some of which have led to the development of effective therapies. Prominently, an enhancer (BCL11A DHS58+) of the transcription factor gene *BCL11A* was found to be essential to its expression in a CRISPR screen (*37*), and targeted knockout of this region in erythroid cells is now used as a therapy for sickle cell anemia and β-thalassemia (*38*).

The ability of bridge recombination to programmably excise specific genomic intervals represents a substantial opportunity for high-resolution functional genomics. As a proof of concept, we used a reporter plasmid containing the *BCL11A* enhancer region with a disrupting intervening sequence that must be excised in order to express mCherry, allowing us to measure excision efficiency (Fig. 5, D to E). This region encodes a GATA motif, elimination of which attenuates *BCL11A* expression in erythroid progenitors (*37*). Across a variety of binding sites for both the TBL and DBL, we were able to achieve robust excision of up to 18.7%, indicating that bridge-mediated excision can be applied at user-defined intervals (Fig. 5F).

Excision of Friedreich's ataxia repeat expansion

Repeat expansion disorders are a class of ~50 genetic diseases arising from the accumulation of repetitive sequences in the genome. These repeats are typically within coding genes and cause disease through loss of function, dysregulation of gene expression, generation of toxic RNAs, or expression of gain-of-function proteins (*39*, *40*). Among these diseases is Friedreich's ataxia, a neuromuscular disorder resulting from the expansion of a GAA trinucleotide repeat in intron 1 of the gene *FXN* (Fig. 5G). In healthy individuals, there are as few as six repeats, whereas patients exhibiting the disease encode as many as 1700 repeats (*41*).

Using bridge recombinases to excise excess repeats should be scarless and in frame, enabling the restoration of a corrected genotype. However, this type of recombination would require both the TBL and DBL to bind identical substrates, potentially causing competitive inhibition of one loop by the other. To this end, we wondered whether DBL-only recombination (Fig. 3F) could be exploited for the removal of undesired repeat regions. To test this, we programmed both loops of the enhanced bRNA or an enhanced DBL (DBL3) to recognize a 14-nt sequence spread across five GAA repeats. Within just 40 GAA repeats, there are 36 stretches of this 14-nt recognition sequence, with each CT core on the bottom strand representing a unique recombination breakpoint. Recombination can occur between any two of these potential binding sites, all of which result in the excision of repeats, leading to variable outcomes across different molecules (Fig. 5G).

To test this in the context of a reporter plasmid, we encoded two GAA repeat regions on either side of a stuffer sequence to compare excision efficiency between a full bRNA and a DBL alone (Fig. 5H). We amplified across this region to quantify the proportion of intact and excised reporters (Fig. 5I). Because the total number of repeats is related to the severity of disease (*40*), we sought to understand how many repeats were removed from the plasmid substrates. As expected, a range of outcomes was observed for both bRNA and DBL only, with as few as 8 and as many as 35 repeats excised from the plasmids (Fig. 5J). Notably, the DBL appeared to be more effective at excising a larger number of repeats, with 40% of plasmids exhibiting excision having 80% (>29 repeats) or more of the repeats removed. These results show that excision with bridge recombination simply requires the 326–amino acid recombinase (976-nt) and the 90-nt DBL to mediate excision. Future translational efforts may be enabled by the transient delivery of these components as RNA molecules, facilitating genomic modification without any DNA delivery.

Discussion

In this study, we present ISCro4, the first bridge recombinase for programmable insertion, inversion, and excision of the human genome. We interrogate the mechanism and application of ISCro4 through systematic engineering, enhancing activity of both its bRNA and recombinase. We highlight key strategies for maximizing the effectiveness of applying bridge recombination in human cells, particularly the use of the donor-binding loop to base-pair with the genome, revealing functional and mechanistic differences between the two bRNA binding loops. Our optimized approach for universal chromosomal DNA rearrangements achieves as high as 20% insertion efficiency and mobilizes genomic DNA at the megabase scale.

As the applications for DNA recombinases have grown increasingly ambitious over the past few decades, the need to overcome the limitations of existing tools has also grown. Recombinase platforms have advanced beyond early implementations of Cre-LoxP knockouts (*42*, *43*) to sophisticated applications in studying development (*44*), building biological circuitry (*45*, *46*), and developing gene therapies (*47*, *48*). These approaches have been developed through many systematic engineering and discovery efforts to tackle challenges such as recognizing user-defined sequences, minimizing scars, or targeting of safe harbor loci (*14*, *22*, *49*–*51*). The bispecific and programmable mechanism of bridge recombinases naturally circumvents many of the fundamental engineering challenges in the recombinase field. We expect future efforts for improving bridge recombinases to focus on further improvement of targeting efficiency and specificity, delivery of megabase-scale DNA payloads, and development of therapeutic delivery formulations, including RNA-only delivery for genomic rearrangements such as repeat excisions. Additionally, recombinase-bRNA electroporation may prove effective, similar to previous efforts for delivery of CRISPR effectors and guide RNAs (*52*).

Bridge recombinases can modify the genome at arbitrary new scales, ranging from single gene insertions to megabase-sized rearrangements, which unlocks substantial potential for understanding cellular function and human disease pathology. Locus- and chromosome-scale excisions or inversions will enable higher-resolution understanding of sequence-to-function relationships across organisms. In combination with chromosomal translocation, these recombinations can enable precise creation of cell lines and animal models that mimic the genotypes of cancers and other chromosomal abnormalities. Lastly, bridge recombination is poised to be combined with artificial intelligence–generated DNA sequences of high complexity (*53*, *54*), enabling programmable genome design at unprecedented length scales.

Materials and methods

Cell lines and culture

Experiments were conducted in HEK293FT, K562, U2OS, and H1 hESCs. HEK293FT cells were cultured at 37°C with 5% CO_2 in DMEM + GlutaMax supplemented with 10% FBS and 100 U/mL PenStrep. K562 cells were cultured at 37°C with 5% CO_2 in RPMI + GlutaMax supplemented with 10% FBS and 100 U/mL PenStrep. U2OS cells were maintained in DMEM High Glucose with Glutamax (Gibco, Cat: 10566016) supplemented with 10% FBS and 100 U/mL PenStrep. For routine passaging of HEK293FT, cells were dissociated with TrypLE (Gibco) and seeded at 1-2 million cells per 10 cm dish. K562 cells were routinely diluted every 2-3 days upon reaching 1E6 cells/mL to a concentration of 1-2E5 cells/mL. U2OS cells were routinely subcultured upon reaching 80% confluency by washing twice with PBS, incubating for 5-10 min with TrypLE (Gibco), and seeding at 1.5 million cells per 10 cm dish.

H1 hESCs (WA01; WiCell Research Institute, Madison, WI; catalog WA01; RRID: CVCL_9771) were maintained feeder-free on Matrigel-coated plates in mTeSR Plus (STEMCELL Technologies) per the manufacturer's instructions. For routine passaging, colonies were dissociated with ReLeSR (STEMCELL Technologies). For experimental seeding prior to plasmid transfection or lentiviral transduction, single cells were generated using Accutase Cell Detachment Solution (Innovative Cell Technologies), and cultures were supplemented with 10 μM Y-27632 (STEMCELL Technologies) for 24 hours post-dissociation. Following transfection or transduction, the culture medium was supplemented with 1x penicillin–streptomycin (Gibco).

Bridge recombinase ortholog mining

Bridge recombinase candidates were identified from our previously generated database of assembled genomic and metagenomic sequences, focusing on those that were within the broader IS621-like clade (*15*). We identified bRNA sequences in IS110 loci using cmsearch (INFERNAL v1.1.4) and an RNA covariance model that was constructed using the IS621 bRNA as a seed sequence. The search was performed using the–toponly parameter as all loci were oriented with respect to the forward strand of the bridge recombinase coding sequence. bRNA sequences were filtered by an E-value threshold of 1e-6 and by requiring at least 80% coverage of the covariance model length.

Ortholog screening in human cells

The recombinase, bRNA, target, and donor sequences for each nominated ortholog were cloned into plasmids for evaluating activity. For orthologs with >75% amino acid identity to IS621, positions 12-14 of the target were mutated to match the target binding loop RTG. Bridge recombinases were human codon optimized and then cloned into the same construct as the bRNA. The bRNA is expressed via a U6 promoter, and the recombinase is expressed via an EF1α promoter. The recombinases were cloned with a C-terminal NLS as Rec-3xSV40NLS-P2A–enhanced green fluorescent protein (EGFP). The target and donor were cloned on a plasmid separated by 1 kb. 100 ng of recombinase and bRNA plasmid was cotransfected with 288 ng of target + donor plasmid into 1.4×10^4 HEK293FT cells using 0.5 μL Lipofectamine 2000 in 96-well plates. After 60 hours at 37°C, plasmid DNA was recovered from cells using 50 μL of QuickExtract (Biosearch Technologies) per well. Extracted DNA was cleaned up using 0.9x AmpureXP magnetic beads (Beckman Coulter). PCR was performed with 2x Platinum SuperFi II across the newly formed inversion junction. Activity was determined via agarose gel electrophoresis of the PCR and subsequent visual inspection.

Luciferase reporter assay in human cells

100 ng of recombinase and bRNA plasmid was cotransfected with 288 ng of C- and G-luciferase inversion reporter plasmid into 1.4×10^4 HEK293FT cells using 0.5 μL Lipofectamine 2000 in 96-well plates. C-luciferase was expressed constitutively from the CMV promoter. G-luciferase was flanked by the bRNAs natural target and donor sequence, or, a target and donor sequence cognate for the bRNA binding loops. Inversion of G-luciferase results in in-frame expression of the luciferase from an EF1α promoter. After 60 hours at 37°C, 20 μL of cell media from each well was transferred to two different 96-well black flat bottom plates. Luciferase expression was measured using the luciferase activity kit from Targeting Systems. G-luciferase expression was measured by adding 50 μL of GAR diluted 1:100 in GAR buffer to cell media, incubating in the dark for 10 min, then measuring with luminescence on Tecan with 10 ms exposure time. C-luciferase expression was measured by adding 50 μL of VLAR diluted 1:100 in VLAR buffer to cell media, incubating in the dark for 10 min, then measuring with luminescence on Tecan with10 ms exposure time. Inversion activity was measured as the ratio of G-luciferase luminescence divided by C-luciferase luminescence, yielding a measurement in relative fluorescence units (RLU). Inversion was confirmed via amplification of the inversion junction on the luciferase reporter plasmid. For 161642, IS621, 127209, and ISCro4, the natural target sequence was modified such that positions 12-14 matched the target binding loop RTG.

AlphaFold modeling of ISCro4 recombinase and alignments

The three-dimensional structure of ISCro4 recombinase was predicted using the AlphaFold 3 (*55*) Server from the amino acid sequence alone and modeled as a single monomer. From the five models outputted, the resulting model was selected for analysis based on the minimal RMSD to one of two monomers with resolved catalytic loops from the experimentally determined structure of IS621 (PDB ID: 8WT6) using UCSF ChimeraX 1.9 (*56*). Alignments were performed using the matchmaker function with backbone atoms for the three functional domains alongside the whole protein.

ISCro4 bRNA engineering

The activity of ISCro4 engineered bRNAs was measured in HEK293FT cells. In the case of a single bRNA, 100 ng of a plasmid encoding U6-bRNA and EF1α-Rec-3xSV40NLS-P2A-EGFP was cotransfected with 288 ng of a plasmid encoding EF1α-Target-revcomp(mCherry)-Donor, such that inversion yields mCherry expression. In the case of a split bRNA, the same construct was prepared, except with only one loop of the bRNA encoded. 50 ng each of two plasmids encoding one loop and the recombinase were cotransfected with the 288 ng of the same mCherry inversion reporter. In all cases, 1.4×10^4 cells were transfected in 96-well plates with 0.5 μL well Lipofectamine 2000 per well. After 60 hours at 37°C, media was aspirated and cells were detached using 50μL of TrypLE Express (Gibco). TrypLE was neutralized with 50 μL Stain Buffer (BD), then, cells were resuspended then spun down in 96-well U-bottom plates at 300xg for 5 min. The supernatant was aspirated and the cells were resuspended in 120 μL Stain Buffer (BD). Flow cytometry was performed using the Quanteon Novocyte and resultant data was analyzed using NovoExpress software. Efficiency was recorded as the MFI of mCherry+ cells within the GFP+ gate or as the percentage of mCherry+ cells within the GFP+ gate.

Genome insertion of plasmids

The activity of ISCro4 for genome insertion was measured in HEK293FT cells. For screening genomic sites, 137 ng of a plasmid encoding EF1α-Rec-3xSV40NLS-P2A-EGFP was cotransfected with 60 ng of a plasmid encoding U6-bRNA and 534 ng of a plasmid encoding EF1α-Puro^R-P2A-mCherry and a recognition sequence for the bRNA. For assessing the efficiency and specificity of insertion with TBL or DBL binding the genome, 137 ng of a plasmid encoding U6-bRNA and EF1α-Rec-3xSV40NLS-P2A-EGFP was cotransfected with 534 ng of a plasmid encoding EF1α-Puro^R-P2A-mCherry and a recognition sequence for the bRNA. For assessing the efficiency of insertion with engineered recombinases with various bRNAs, 137 ng of a plasmid encoding EF1α-Rec-3xSV40NLS-P2A-EGFP was cotransfected with 60 ng of a plasmid encoding U6-bRNA and 534 ng of a plasmid encoding EF1α-Puro^R-P2A-mCherry and a recognition

sequence for the bRNA. In the case of a split bRNA, 30 ng each of two plasmids, each one expressing one loop, were cotransfected. In all cases, 1.4×10^4 cells were transfected in 96-well plates with 0.59 μL well Lipofectamine 2000 per well.

Cell culture for selection of on- and off-target insertions

After 96 hours, all cells were passaged from 96-well to 24-well plates. After two additional days, 20% of cells were passaged in 10 μg/mL puromycin. Cells were passaged under puromycin selection for an additional 12 days, for a total of 18 days post-transfection.

Mapping insertion sites genome-wide

After selection, genomic DNA was extracted from cells using *Quick*-DNA Miniprep Plus Kit (Zymo). Genomic DNA was measured with Qubit HS dsDNA Assay (Thermo). 150 ng of gDNA was then tagmented with Tn5 then amplified using a plasmid donor specific bait primer. Sample preparation, sequencing, and sequencing analysis for determination of insertion sites was performed as previously described (*22*), with some modifications to account for a unique molecular identifier (UMI) of length 20 rather than of length 12. Prefiltering of the reads was conducted to remove technical artifacts and contaminants. This included removing reads that mapped to three genomic "blacklist" sites, including an EF1α promoter near chr6:73519430-73521921, a U6 promoter near chr15:67840054-67840294 and a G-rich site at chr2:32915322-32917351, which were found to be technical artifacts. Contaminants and off-target amplicons were filtered by excluding all reads that didn't contain at least one 17-mer that was specific to the donor-plasmid between the primer sequence and the donor sequence dinucleotide core, allowing for up to 2 mismatches with that 17-mer. Samples were sequenced on an Illumina NextSeq with a 600 cycle P1 or P2 kit. For insertion site mapping of the WT bRNA, unique amplicons were used in lieu of unique molecular identifiers to map unique insertions.

Selection of unique genomic sites for genome insertion

Genomic sites were selected from amongst 14 nucleotide sequences found only once within the hg38 reference genome. Of these, we filtered for sequences with either a CT or GT core sequence at position 8 and 9 of the 14-nt sequence. Amongst these, we selected sequences that are between the transcriptional start site and 5 kbp downstream of the transcriptional start site. Sequences were further filtered by essentiality, using a published list of essential genes (*57*). Most sequences were then selected from amongst the top 30% of expression level in HEK293FT cells, based on RNA-seq data of unmodified cells (GEO: GSE164956) (*58*). Genome sequences were paired with plasmid sequences based on matching core sequences and a lack of a matching nucleotide at position 7, which interferes with the handshake base pairing mechanism (*20*). For intragenomic inversion and excision, pairs of unique 14 nucleotide sequences were selected for relative distance to one another. Next, the pairs of sequences were filtered to require both to encode an identical core sequence. Finally, the sequences were filtered again to ensure position 6 and 7 of the sequences did not match.

Selection of genome-orthogonal sequences

Sequences of length 14 that were orthogonal to the human genome were identified. Sequences were subsequently filtered for those containing a CT or GT at position 8 and 9 of the 14 nucleotide sequences. Next, sequences were compared against the genome to identify all genome locations matching positions 1-11 of the 14-nucleotide sequence. Genome orthogonal sequences were then rank-ordered according to the number of 11-mer off-targets in the genome, with zero being optimal.

Lentivirus production

Third-generation lentiviral particles encoding EF1α-ISCro4Rec-P2A-EGFP-Puro^R were produced with the Gibco LV-MAX Lentiviral Production System following the manufacturer's optimized suspension protocol. Briefly, LV-MAX Viral Production Cells (HEK293F-derived) were expanded in LV-MAX Production Medium and maintained at 37°C and 8% CO_2 on an orbital shaker (125 ± 5 rpm; 19-mm orbit) in vented-cap, plain-bottom PETG Erlenmeyer flasks. Cultures were subcultured to ensure transfection at 4.0×10^6 viable cells/mL as specified by the kit. Transfections used a total of 2.5 μg plasmid DNA per mL of final production volume with a 3:2 mass ratio of LV-MAX Lentiviral Packaging Mix to transfer vector. Plasmid DNAs were complexed with LV-MAX Transfection Reagent (6 μL/mL) in Opti-MEM, with LV-MAX Supplement (5% v/v) added at setup and LV-MAX Enhancer (4% v/v) added 12 hours post-transfection. For the 125-mL scale, we used a 10-mL final production volume per flask. Supernatants were collected 48–55 hours post-transfection, clarified by centrifugation, and passed through a 0.45-μm low-protein-binding filter before concentration.

Lentivirus was concentrated 20x using the Lenti-X Concentrator (Takara/Clontech) by mixing 1 volume of Concentrator with 3 volumes of clarified supernatant, incubating at 4°C for 60 min, and centrifuging at 1,500xg for 45 min at 4°C. Pellets were gently resuspended in PBS to 1:20 of the starting volume, aliquoted, and stored at −80°C until use. All work with lentiviral materials was performed under BSL-2 containment using institutional biosafety-approved procedures and PPE.

Generation of stable hESC lines by lentiviral transduction

For lentiviral delivery, H1 hESCs were exposed to a dilution series of bridge-recombinase lentivirus to empirically establish conditions yielding ~20–30% EGFP+ cells at 48–72 hours, consistent with low multiplicity of infection and favoring single-copy integration. Using this calibration, bulk transductions were performed at low MOI with 1:500 and 1:1,000 dilutions of the concentrated virus. Puromycin (Gibco; 1 μg/mL) selection was initiated 48 hours post-transduction and maintained for 7 days, after which flow cytometry demonstrated near-uniform EGFP expression (~100% EGFP+).

hESC transfections

For plasmid transfection, H1 human embryonic stem cells were plated as single cells using Accutase Cell Detachment Solution (Innovative Cell Technologies) one day prior to transfection at a density of 22,000 cells per well of a 96-well plate on Matrigel-coated plates in mTeSR Plus (STEMCELL Technologies) supplemented with 10 μM Y-27632 (STEMCELL Technologies). On the day of transfection, the medium was replaced with mTeSR Plus supplemented with 10 μM Y-27632 and 1x penicillin–streptomycin (Gibco). Transfection master mixes were prepared in mTeSR Plus with plasmid DNA and FuGENE® HD Transfection Reagent (Promega) at a 1:3 ratio of total plasmid DNA to FuGENE® HD, using 2 μg total DNA with 6 μL FuGENE® HD per 100 μL of transfection master mix.

For hESC episomal mCherry reporter recombination experiments, wild-type H1 stem cells were co-transfected per well of a 96-well plate with 127.03 ng of each pEffector, together with 98.54 ng of the pReporter for recombination. For hESC experiments for genome integration in lentivirally-transduced H1 stem cell lines with stable expression of the WT ISCro4 recombinase, the H1 stem cell lines were co-transfected with 47.50 ng of bRNA plasmid targeting each of the three different loci in the human genome, along with 63.1 ng of pDonor for genome integration. One day after transfection, the medium was replaced with mTeSR Plus supplemented with 10 μM Y-27632 and 1× penicillin–streptomycin (Gibco). Two days after transfection, the medium was replaced with mTeSR Plus supplemented with 1x penicillin–streptomycin. Three days after transfection, cells were harvested using 50 μL of Accutase per well of a 96-well plate for flow-cytometry analysis and genomic DNA extraction for ddPCR.

Generation of stable U2OS and HEK293FT lines by lentiviral transduction

Cells were exposed to a dilution series of bridge-recombinase lentivirus to empirically establish conditions yielding ~30% EGFP+ cells at 48 hours.

Bulk transductions were performed at low MOI with 1:500 and 1:1000 dilutions of the concentrated virus. 48 hours post transductions, cells were confirmed to be ~30% EGFP+ after which the culture medium was refreshed and supplemented with Puromycin (Gibco; 2 μg/mL). Cells from both dilutions were pooled together and selected for 7 days before proceeding with transfections or nucleofections.

U2OS and K562 nucleofections

For U2OS nucleofection experiments, cells were passaged one day prior at a density of 4 million cells per 15 cm dish. After 24 hours, cells were dissociated and counted to ensure an approximate doubling time of one day, then 200,000 cells were nucleofected with 1.5 μg DNA using the SE Cell Line 96-well Nucleofector Kit (Lonza, Cat: V4SC-1096) and the DN-100 pulse code. For K562 nucleofection experiments, cells were diluted one day prior to a density of 2E5 cells/mL. After 24 hours, 200,000 cells were nucleofected with 1.5 μg DNA using the SF Cell Line 96-well Nucleofector Kit (Lonza, Cat: V4SC-2096) and the FF-120 pulse code.

For episomal plasmid recombination experiments, cells received 1113 ng of pReporter along with either 387 ng of single bridgeRNA pEffector or 193 ng each of split bridgeRNA pEffector, maintaining a 1:4 molar ratio of pEffector to pReporter plasmids. For U2OS experiments using the single copy lentivirus recombinase cell line, 119 ng pbRNA and 1381 ng of pDonor were delivered at a 1:3 molar ratio of bRNA to donor. Following nucleofection, U2OS cells were incubated for 15 min in cuvettes at room temperature before adding 80 μL of complete media and transferring the total volume to 12-well plates containing 1 mL of pre-warmed media. For K562s, cells were incubated for 10 min in cuvettes at room temperature before adding 80 μL of complete media and dividing the total volume into two wells of a 96-well plate containing 100 μL of pre-warmed media. Three days post-nucleofection, cells were split equally for flow cytometry analysis and genomic DNA extraction, followed by ddPCR.

Lentivirus recombinase HEK293FT cell line transfections

14,000 lentivirus modified HEK293FT cells were seeded into each well of a 96 well plate one day prior to transfections with puromycin removed. 119 ng bRNA plasmid and 1381 ng of pDonor were delivered at a 1:3 molar ratio of bRNA to donor. Briefly, a 25 μl mix of DNA + Optimem was pipette mixed with a 25 μl mix containing 0.6 μl Lipofectamine 2000 + OptiMEM. This was incubated for 10 min before adding it dropwise onto cells. Three days post-transfection, cells were split equally for flow cytometry analysis and genomic DNA extraction, followed by ddPCR.

Generation of LSR landing pad cell line

A landing pad HDR template plasmid was designed to express the LSR Dn29 protein, with the attP sequence positioned between the EF1α promoter and the Dn29 coding sequence. This design enables a promoter swapping mechanism where delivery of a plasmid library containing the attB sequence interrupts Dn29 expression and activates expression of the library member. For selection purposes, a blasticidin resistance gene was linked to the Dn29 coding sequence via a P2A peptide, allowing for selection of successfully edited cells that maintain expression.

1×10^6 K562 cells were nucleofected with the HDR template plasmid (5 μg) along with TALEN expression vectors targeting the AAVS1 locus: 1 μg each of hAAVS1 1L TALEN (Addgene #35431) and hAAVS1 1R TALEN (Addgene #35432). Nucleofection was performed using the Lonza SF Cell Line 4D-Nucleofector X Kit L according to the manufacturer's protocol. Beginning three days post-nucleofection, the media was supplemented with 20 μg/mL blasticidin for 10 days, followed by maintenance at 10 μg/mL blasticidin. To establish monoclonal populations, single cells were sorted into 96-well plates using a BD FACSAria Fusion and expanded over a two-week culture period. Integration of the landing pad at the AAVS1 locus was validated by PCR. To evaluate Dn29 integration efficiency, cell lines were nucleofected with 1.2 μg of a promoterless plasmid containing the attB sequence upstream of an mCherry coding sequence. Eleven days after nucleofection, mCherry expression was quantified using an Attune NxT Flow Cytometer (Thermo Fisher), comparing results to wildtype K562 cells nucleofected with the same promoterless mCherry donor as a control for background expression.

Design of libraries for ISCro4 recombinase deep mutational scan

To generate a comprehensive library of single amino acid substitutions in the ISCro4 recombinase, we ordered a 250 bp oligonucleotide pool from Twist Bioscience. This pool was divided into five sub-pools, each covering a distinct 65 amino acid stretch of the ISCro4 sequence: residues 2–66, 67–131, 132–196, 197–261, and 262–326. The library was designed such that each amino acid substitution was encoded by 2 different codons (when available) and was designed to include completely random 65 amino acid stretches that resulted in the formation of a full-length protein.

Each sub-pool was PCR-amplified in a 50 μL reaction using Kapa HiFi HotStart PCR Mix (Kapa Biosystems), with 10 ng of synthesized oligonucleotide as template and sub-pool-specific primers. The thermocycling parameters were: 98°C for 120 s; 10 cycles of 98°C for 10 s, 65°C for 15 s, and 72°C for 15 s; and 72°C for 1 min. PCRs were cleaned up using a 0.8x ratio of Agencourt AMPure XP beads (Beckman Coulter). Plasmid backbones for cloning each sub-pool were generated by replacing the corresponding sequence of the ISCro4 with BsaI recognition sites, to enable Golden Gate assembly of the libraries. Each backbone (3 μg) was digested with BsaI in a 50 μL reaction for 3 hours at 37°C, followed by heat inactivation at 65°C for 20 min. Digested products were purified using DNA Clean & Concentrator-5 columns (Zymo). Golden Gate assemblies were performed using 500 ng of backbone and 100 ng of insert split across multiple 50 μL reactions, incubated at 37°C for 3 hours and heat-inactivated at 65°C for 20 min. Reaction products were purified and eluted in 10 μL of water.

Electrocompetent Endura DUOs (Lucigen) were transformed with the assembled libraries following the manufacturer's protocol. A small aliquot (1–5 μL) was plated on carbenicillin LB agar to estimate library coverage, while the remaining culture was expanded overnight in 70 mL LB with carbenicillin. Library coverage of at least 1000x was confirmed before downstream applications. Plasmid libraries were sequenced using the NextSeq 2000 (150 bp PE run).

Generation of recombinase library cell lines

Each sub-library contained approximately 2,500 elements. Our K562-Dn29 landing pad cell line exhibited an integration efficiency of around 15%. To ensure a minimum of 1,000x coverage, we started with 2.5×10^7 K562-Dn29 landing pad cells per replicate of the library and used 50 μg of library plasmid. Cells were divided into multiple 100 μL nucleocuvettes and nucleofected according to the manufacturer's protocol (Lonza). Starting three days after nucleofection, the culture media was supplemented with 2 μg/mL puromycin for a minimum of seven days.

Transfection and sorting of recombinase library cell lines

For each replicate, 3×10^7 recombinase library cells were nucleofected with 60 μg of plasmid containing the bRNA and inverted mCherry reporter. Cells were divided into multiple 100 μL nucleocuvettes and nucleofected according to the manufacturer's protocol (Lonza).Three days after nucleofection, 20% of the cells were collected as the pre-sort fraction. The remaining cells were sorted using a FACSAria Fusion Cell Sorter based on their mCherry to GFP fluorescence ratio. Post-sort, a minimum of 1,000x library coverage was maintained.

Extraction and preparation of recombinase libraries

Genomic DNA (gDNA) was extracted using the Quick-DNA miniprep plus kit (Zymo), following the manufacturer's protocol. To achieve a

minimum of 1,000x coverage, at least 15 μg of gDNA was used for the first PCR (PCR1). For each replicate, 16 × 50 μL PCR1 reactions were prepared using Kapa HiFi HotStart PCR Mix (Kapa Biosystems) and primers that specifically amplify recombinases integrated into the genomic landing pad. PCR1 products were purified with a 0.7x volume of Agencourt AMPure XP beads (Beckman Coulter). The thermocycling parameters were: 98°C for 120 s; 24–27 cycles of 98°C for 10 s, 65°C for 15 s, and 72°C for 40 s; and 72°C for 2 min. For the second PCR (PCR2), 100 ng of purified PCR1 product was used to amplify recombinase regions corresponding to each library in a 50 μL reaction, again using Kapa HiFi HotStart PCR Mix. The thermocycling parameters were: 98°C for 120 s; 9-12 cycles of 98°C for 10 s, 65°C for 15 s, and 72°C for 15 s; and 72°C for 1 min. This PCR2 was cleaned up with 0.7x ratio of Agencourt AMPure XP beads. The third PCR (PCR3) was performed to add sample indices, using 100 ng of purified PCR2 product as input in a 50 μL reaction, using Kapa HiFi HotStart PCR Mix. The thermocycling parameters were: 98°C for 120 s; 5-7 cycles of 98°C for 10 s, 65°C for 15 s, and 72°C for 20 s; and 72°C for 1 min. PCR3 products were gel extracted using the Monarch DNA gel extraction kit (New England Biolabs) and then cleaned up with 0.7x ratio of Agencourt AMPure XP beads. In all the above PCRs, the numbers of cycles were tested to ensure that they fell within the linear phase of amplification. The libraries were quantified using the Qubit dsDNA HS assay kit (Thermo Fisher) and pooled together and sequenced on a 150 bp PE run on the NextSeq 2000 (Illumina).

Analysis of deep mutational scan

Paired-end reads with overlapping regions were merged using BBMerge (*59*). To analyze the sequencing data, we developed a Python-based pipeline to count variant frequencies within specific regions. The script begins by initializing a dictionary from a reference library file containing all expected variant sequences. For each read in the FASTQ file, it searches for two constant flanking sequences within the merged read. If both flanking sequences are identified, the sequence between them is extracted and counted only if it matches exactly with an entry in the reference library.

For each sample, variant counts are normalized using the formula (count + 1) / total counts across all library members. Enrichment scores are calculated by dividing the normalized post-sort frequency by the pre-sort frequency for each variant. These enrichment scores are specific to each library. A z-score is then computed as $z = (x - \mu)/\sigma$ where x is the enrichment score of the variant, μ is the mean enrichment score of all members of a library, and σ is the standard deviation of the enrichment scores. The relative z-score is determined by subtracting the z-score of the library-specific WT variant from that of each individual variant.

Finally, mutability scores were calculated for each position by counting the number of variants whose relative z-score falls within one standard deviation of the WT.

Prediction of structural contacts

To analyze the structural interactions of bridge recombinases with DNA and RNA substrates, we developed a computational approach for the identification of protein residues involved in critical contacts. We utilized the cryo-EM structure of IS621 (PDB: 8WT6) to predict and compare key interaction interfaces. Protein-nucleic acid and protein-protein contacts were identified using a distance-based approach implemented in Python (3.12.2) with Biopython (v1.84). Residues were classified as contacting if any atom in the residue was within 4.0 Å of any atom in a DNA, RNA, ion, or protein molecule. For each residue, we calculated the solvent-accessible surface area (SASA) using the Shrake-Rupley algorithm and classified residues as solvent-exposed (SASA ≥ 10.0 $Å^2$) or internal. We further categorized protein contacts as monomer (intra-chain), dimer, or tetramer contacts based on the identity of the contacting chains.

Modeling of recombinase mutations

Recombinase mutations were modeled using the IS621 transpososome structure (PDB: 8WT6) and modeled in ChimeraX. Amino acids were mutated using the swapaa command with Dunbrack rotamers. Hydrogen bonds for rotamers were predicted using the default settings, with a distance tolerance of 0.40 Å and an angle tolerance of 20°.

ddPCR of genome insertion

Three days post-transfection, cells were trypsinized with 50 μL TrypLE (Gibco) for 10 min and then quenched with 50 μL Stain Buffer (BD). The 100 μL cell suspension was split into two 50 μL aliquots in U-bottom 96-well plates, centrifuged (300 × g, 5 min), and supernatant aspirated. One plate was resuspended in 120 μL Stain Buffer (BD) and analyzed with Novocyte Quanteon Flow Cytometer with Autosampler (Agilent). The other plate was resuspended in 50 μL QuickExtract DNA Solution (Biosearch Technologies), vortexed for 15 s, and thermocycled: 65°C for 15 min, 68°C for 15 min, 98°C for 10 min. DNA was cleaned with 0.9x AmpureXP (Beckman Coulter) beads. To assess insertion efficiency, qPCR/ddPCR primers and probes were designed to span the left insertion junction of various loci in the human genome, using a constant primer that binds to the donor plasmid sequence (PR_NPXX) a genome binding primer near the locus (PR_NPXX-XX) and a FAM probe within the amplicon (pbNPXX) (Table S4). Genomic reference primers and probes located nearby each insertion locus were designed to measure locus copy number for efficiency percentage calculations.ddPCR reaction mix (22 μL total): 11 μL ddPCR Supermix for Probes (no dUTP) (Bio-Rad), 1.98 μL of each primer (10 μM), 0.55 μL of each probe (10 μM), 1.65 μL cleaned gDNA, 0.22 μL SacI-HF (NEB), water to volume. Each reaction contained primers and probes for the target site (FAM probe) and a nearby reference locus (HEX probe). Reactions were run on QX200 AutoDG Droplet Digital PCR System (Bio-Rad). For off-target detection or low concentration samples, primers were increased to 20 μM and volume halved, and gDNA volume was increased to 4.95 μL.

Genomic inversion and excision and sample preparation

750 ng of a plasmid encoding U6-bRNA and EF1α-Rec-P2A-EGFP was transfected into $1.4 \times 10^4$4 HEK293FT cells using 0.59 μL Lipofectamine 2000 in 96-well plates. After 96 hours at 37°C, genomic DNA was extracted from cells using 50 μL of QuickExtract (Biosearch Technologies) per well. Extracted DNA was cleaned up using 0.9x AmpureXP magnetic beads (Beckman Coulter). 150 ng of genomic DNA was then tagmented with i5 handles using Tn5 transposase as previously described (*22*). Tagmented gDNA was cleaned up using 0.9x AmpureXP magnetic beads. Then, a PCR was performed with a i5-specific primer and a second primer designed to bind upstream of the rearrangement junction. PCR was performed with 2x Platinum SuperFi with a protocol of 1 Cycle, 98°C for 2 min; 12 cycles, 98°C for 10 s, 68°C for 10 s, 72°C for 90 s; 1 cycle 72°C for 5 min. After cleanup with magnetic beads, a second PCR was performed with an outer nested i5 primer and a full-length i7 primer with a binding region specific for a region downstream of the first primer but upstream of the rearrangement junction. PCR was performed with 2x Platinum SuperFi with a protocol of 1 Cycle, 98°C for 2 min; 22 cycles, 98°C for 10 s, 68°C for 10 s, 72°C for 90 s; 1 cycle 72°C for 5 min. Genome specific primers were designed using Primer3. For both inversion and excision, genome specific primers were designed such that they bound upstream of a junction but not within the region being inverted or excised. Samples were run on a 2% agarose gel and then gel extracted from 200-600 bp using Monarch DNA Gel Extraction Kit (NEB). Samples were then sequenced on an Illumina NextSeq with 600 cycle P1 or P2 kit.

Measurement of inversion and excision efficiency

Fastq files of sequenced amplicons were trimmed using FLASH and Cutadapt. Human genome reference sequences were prepared for each measured junction; one fasta file was prepared with the wild-type

genome sequence, and a second fasta file was prepared with the expected rearrangement product based on the sequence pair targeted by the bRNA. Trimmed, single-end reads (read 2 with known starting sequence upstream of the junction) were separately mapped to both files using bwa. Alignments were then filtered to segregate reads originating from wild-type and rearranged genomes and to eliminate reads that map only upstream of the recombination junction, which cannot be assigned. To do this, we required reads to overlap with 10 nucleotides upstream of the core sequence and 20 nucleotides downstream of the core sequence, while allowing for indels between these regions. The resultant alignments contain only reads that align to the expected rearrangement junction. The number of reads in both files was counted, and the rearrangement efficiency was calculated as the number of reads in the rearranged alignment divided by the number of reads in both the wild-type and rearranged alignments.

Measurement of indel formation during genomic inversion and excision

Filtered alignments used for measuring efficiency of excision and inversion were further analyzed for the presence of indels. Briefly, CIGAR strings for each read in the alignment were parsed, analyzing a window 20nt upstream and downstream of the recombination junction. Reads containing an "I" or "D" of any length in the CIGAR string at any position within this window were classified as containing indels. Indels were not counted individually; a read with one indel and a read with multiple indels are each counted only once. After classifying reads, the total number of reads with or without indels were summed across biological replicates. The percentage of indel formation was then calculated as the all indel-bearing reads divided by the total number of reads.

Excision of *BCL11A* enhancer sequences

The *BCL11A* DHS58+ sequence was cloned into a plasmid such that the regions upstream and downstream of the GATA motif were separated by a 1 kbp sequence containing a terminator. This cassette was between a EF1α promoter and mCherry gene, such that recombination between the two regions of *BCL11A* would result in excision and mCherry expression. 100 ng of a plasmid expressing a U6-bRNA and EF1α-Rec-P2A-EGFP was cotransfected with 372 ng of this reporter plasmid into 1.4×10^4 HEK293FT cells using 0.5 μL Lipofectamine 2000 in 96-well plates. After 60 hours at 37°C, media was aspirated and cells were detached using 50 μL of TrypLE Express (Gibco). TrypLE was neutralized with 50 μL Stain Buffer (BD), then, cells were resuspended then spun down in 96-well U-bottom plates at 300xg for 5 min. The supernatant was aspirated and the cells were resuspended in 120 μL Stain Buffer (BD). Flow cytometry was performed using the Quanteon Novocyte and resultant data was analyzed using NovoExpress software. Efficiency was recorded as the MFI of mCherry+ cells within the GFP+ gate or as the percentage of mCherry+ cells within the GFP+ gate.

Excision of GAA repeats

GAA repeats sequences were cloned into a plasmid such that two GAA regions were separated by 1 kbp of intervening sequence, such that recombination between the two GAA regions results in excision. 100 ng of a plasmid expressing a U6-bRNA or U6-DBL and EF1α-Rec-P2A-EGFP was cotransfected with 372 ng of this reporter plasmid into 1.4×10^4 HEK293FT cells using 0.5 μL Lipofectamine 2000 in 96-well plates. After 72 hours at 37°C, genomic DNA was extracted from cells using 50 μL of QuickExtract (Biosearch Technologies) per well. Extracted DNA was cleaned up using 0.9x AmpureXP magnetic beads (Beckman Coulter). Primers were designed such that amplification would result in equal amplification of the unmodified and modified plasmids. PCR was performed with 2x Platinum SuperFi with a protocol of 1 Cycle, 98°C for 1 min; 20 cycles, 98°C for 10 s, 68°C for 10 s, 72°C for 10 s; 1 cycle 72°C for 2 min. PCR products were run on a 2% agarose gel, stained with SYBR Gold and visualized with a Bio-Rad ChemiDoc Imaging System. The lower band was then gel extracted and sequenced using the Premium PCR service from Plasmidsaurus. The number of repeats between the primer binding regions was counted using custom python scripts. Then, the count table was filtered for repeat counts possible via intra-plasmid excision rather than via inter-plasmid recombination. The number of repeats excised was calculated as the number of total repeats in the initial plasmid minus the number of repeats counted.

REFERENCES AND NOTES

1. F. Jacob, J. Monod, On the Regulation of Gene Activity. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 193–211 (1961). doi: 10.1101/SQB.1961.026.01.024
2. L. H. Hartwell, J. J. Hopfield, S. Leibler, A. W. Murray, From molecular to modular cell biology. *Nature* **402** (Suppl), C47–C52 (1999). doi: 10.1038/35011540; pmid: 10591225
3. P. D. Hsu, E. S. Lander, F. Zhang, Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014). doi: 10.1016/j.cell.2014.05.010; pmid: 24906146
4. M. Pacesa, O. Pelea, M. Jinek, Past, present, and future of CRISPR genome editing technologies. *Cell* **187**, 1076–1100 (2024). doi: 10.1016/j.cell.2024.01.042; pmid: 38428389
5. H. Liao, J. Wu, N. J. VanDusen, Y. Li, Y. Zheng, CRISPR-Cas9-mediated homology-directed repair for precise gene editing. *Mol. Ther. Nucleic Acids* **35**, 102344 (2024). doi: 10.1016/j.omtn.2024.102344; pmid: 39494147
6. A. V. Anzalone *et al.*, Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019). doi: 10.1038/s41586-019-1711-4; pmid: 31634902
7. S. Pinglay *et al.*, Multiplex generation and single-cell analysis of structural variants in mammalian genomes. *Science* **387**, eado5978 (2025). doi: 10.1126/science.ado5978; pmid: 39883753
8. J. Koeppel, J. Weller, T. Vanderstichele, L. Parts, Engineering structural variants to interrogate genome function. *Nat. Genet.* **56**, 2623–2635 (2024). doi: 10.1038/s41588-024-01981-7; pmid: 39533047
9. B. Sauer, Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol.* **7**, 2087–2096 (1987). pmid: 3037344
10. B. Sauer, N. Henderson, Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5166–5170 (1988). doi: 10.1073/pnas.85.14.5166; pmid: 2839833
11. A. V. Anzalone *et al.*, Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **40**, 731–740 (2021). doi: 10.1038/s41587-021-01133-w; pmid: 34887556
12. M. T. N. Yarnall *et al.*, Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat. Biotechnol.* **41**, 500–512 (2022). doi: 10.1038/s41587-022-01527-4; pmid: 36424489
13. M. Pallarès-Masmitjà *et al.*, Find and cut-and-transfer (FiCAT) mammalian genome engineering. *Nat. Commun.* **12**, 7071 (2021). doi: 10.1038/s41467-021-27183-x; pmid: 34862378
14. A. Fanton *et al.*, Site-specific DNA insertion into the human genome with engineered recombinases. bioRxiv, 2024.11.01.621560 (2024).
15. M. G. Durrant *et al.*, Bridge RNAs direct programmable recombination of target and donor DNA. *Nature* **630**, 984–993 (2024). doi: 10.1038/s41586-024-07552-4; pmid: 38926615
16. S. E. Klompe, P. L. H. Vo, T. S. Halpin-Healy, S. H. Sternberg, Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219–225 (2019). doi: 10.1038/s41586-019-1323-z; pmid: 31189177
17. G. D. Lampe *et al.*, Targeted DNA integration in human cells without double-strand breaks using CRISPR-associated transposases. *Nat. Biotechnol.* **42**, 87–98 (2023). doi: 10.1038/s41587-023-01748-1; pmid: 36991112
18. C. J. Tou, B. Orr, B. P. Kleinstiver, Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. *Nat. Biotechnol.* **41**, 968–979 (2023). doi: 10.1038/s41587-022-01574-x; pmid: 36593413
19. J. Strecker *et al.*, RNA-guided DNA insertion with CRISPR-associated transposases. *Science* **365**, 48–53 (2019). doi: 10.1126/science.aax9181; pmid: 31171706
20. M. Hiraizumi *et al.*, Structural mechanism of bridge RNA-guided recombination. *Nature* **630**, 994–1002 (2024). doi: 10.1038/s41586-024-07570-2; pmid: 38926616
21. S. Konermann *et al.*, Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **173**, 665–676.e14 (2018). doi: 10.1016/j.cell.2018.02.033; pmid: 29551272
22. M. G. Durrant *et al.*, Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat. Biotechnol.* **41**, 488–499 (2022). doi: 10.1038/s41587-022-01494-w; pmid: 36217031
23. P. Siguier, J. Perochon, L. Lestrade, J. Mahillon, M. Chandler, ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **34**, D32–D36 (2006). doi: 10.1093/nar/gkj014; pmid: 16381877
24. L. Cong *et al.*, Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi: 10.1126/science.1231143; pmid: 23287718

25. D. Y. Kim *et al.*, Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat. Biotechnol.* **40**, 94–102 (2021). doi: 10.1038/s41587-021-01009-z; pmid: 34475560
26. P. D. Hsu *et al.*, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013). doi: 10.1038/nbt.2647; pmid: 23873081
27. I. P. Witte *et al.*, Programmable gene insertion in human cells with a laboratory-evolved CRISPR-associated transposase. *Science* **388**, eadt5199 (2025). doi: 10.1126/science.adt5199; pmid: 40373119
28. K. A. Matreyek, J. J. Stephany, D. M. Fowler, A platform for functional assessment of large variant libraries in mammalian cells. *Nucleic Acids Res.* **45**, e102–e102 (2017). doi: 10.1093/nar/gkx183; pmid: 28335006
29. J. Cao *et al.*, High-throughput 5′ UTR engineering for enhanced protein production in non-viral gene therapies. *Nat. Commun.* **12**, 4138 (2021). doi: 10.1038/s41467-021-24436-7; pmid: 34230498
30. L. Villiger *et al.*, CRISPR technologies for genome, epigenome and transcriptome editing. *Nat. Rev. Mol. Cell Biol.* **25**, 1–24 (2024). doi: 10.1038/s41580-023-00697-6
31. C. J. Tou, B. P. Kleinstiver, Recent Advances in Double-Strand Break-Free Kilobase-Scale Genome Editing Technologies. *Biochemistry* **62**, 3493–3499 (2023). doi: 10.1021/acs.biochem.2c00311; pmid: 36049184
32. O. Shalem, N. E. Sanjana, F. Zhang, High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015). doi: 10.1038/nrg3899; pmid: 25854182
33. A. Dixit *et al.*, Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853–1866.e17 (2016). doi: 10.1016/j.cell.2016.11.038; pmid: 27984732
34. B. Adamson *et al.*, A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867–1882.e21 (2016). doi: 10.1016/j.cell.2016.11.048; pmid: 27984733
35. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014). doi: 10.1126/science.1246981; pmid: 24336569
36. L. Przybyla, L. A. Gilbert, A new era in functional genomics screens. *Nat. Rev. Genet.* **23**, 89–103 (2022). doi: 10.1038/s41576-021-00409-w; pmid: 34545248
37. M. C. Canver *et al.*, BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192–197 (2015). doi: 10.1038/nature15521; pmid: 26375006
38. H. Frangoul *et al.*, CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *N. Engl. J. Med.* **384**, 252–260 (2021). doi: 10.1056/NEJMoa2031054; pmid: 33283989
39. I. Malik, C. P. Kelley, E. T. Wang, P. K. Todd, Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **22**, 589–607 (2021). doi: 10.1038/s41580-021-00382-6; pmid: 34140671
40. A. N. Khristich, S. M. Mirkin, On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170 (2020). doi: 10.1074/jbc.REV119.007678; pmid: 32060097
41. K. Bürk, Friedreich Ataxia: Current status and future prospects. *Cerebellum Ataxias* **4**, 4 (2017). doi: 10.1186/s40673-017-0062-x; pmid: 28405347
42. P. C. Orban, D. Chui, J. D. Marth, Tissue- and site-specific DNA recombination in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6861–6865 (1992). doi: 10.1073/pnas.89.15.6861; pmid: 1495975
43. P. Soriano, Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70–71 (1999). doi: 10.1038/5007; pmid: 9916792
44. K. K. Chow *et al.*, Imaging cell lineage with a synthetic digital recording system. *Science* **372**, eabb3099 (2021). doi: 10.1126/science.abb3099; pmid: 33833095
45. C. A. Merrick, J. Zhao, S. J. Rosser, Serine Integrases: Advancing Synthetic Biology. *ACS Synth. Biol.* **7**, 299–310 (2018). doi: 10.1021/acssynbio.7b00308; pmid: 29316791
46. B. H. Weinberg *et al.*, Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nat. Biotechnol.* **35**, 453–462 (2017). doi: 10.1038/nbt.3805; pmid: 28346402
47. S. Pandey *et al.*, Efficient site-specific integration of large genes in mammalian cells via continuously evolved recombinases and prime editing. *Nat. Biomed. Eng.* **9**, 22–39 (2025). doi: 10.1038/s41551-024-01227-1; pmid: 38858586
48. J. Xie *et al.*, Curative levels of endogenous gene replacement achieved in non-human primate liver using programmable genomic integration. bioRxiv, 2024.10.12.617700 (2024).
49. F. Fauser *et al.*, Systematic Development of Reprogrammed Modular Integrases Enables Precise Genomic Integration of Large DNA Sequences. bioRxiv, 2024.05.09.593242 (2024).
50. M. Jelicic *et al.*, Discovery and characterization of novel Cre-type tyrosine site-specific recombinases for advanced genome engineering. *Nucleic Acids Res.* **51**, 5285–5297 (2023). doi: 10.1093/nar/gkad366; pmid: 37158248
51. A. D. Schmitt, M. Hu, B. Ren, Genome-wide mapping and analysis of chromosome architecture. *Nat. Rev. Mol. Cell Biol.* **17**, 743–755 (2016). doi: 10.1038/nrm.2016.104; pmid: 27580841
52. P. Hoellerbauer, M. Kufeld, P. J. Paddison, Efficient Multi-Allelic Genome Editing of Primary Cell Cultures via CRISPR-Cas9 Ribonucleoprotein Nucleofection. *Curr. Protoc. Stem Cell Biol.* **54**, e126 (2020). doi: 10.1002/cpsc.126; pmid: 32833346
53. E. Nguyen *et al.*, Sequence modeling and design from molecular to genome scale with Evo. *Science* **386**, eado9336 (2024). doi: 10.1126/science.ado9336; pmid: 39541441
54. G. Brixi *et al.*, Genome modeling and design across all domains of life with Evo 2. bioRxiv 2025.02.18.638918 [Preprint] (2025); https://doi.org/10.1101/2025.02.18.638918.
55. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). doi: 10.1038/s41586-024-07487-w; pmid: 38718835
56. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021). doi: 10.1002/pro.3943; pmid: 32881101
57. L. Funk *et al.*, The phenotypic landscape of essential human genes. *Cell* **185**, 4634–4653.e22 (2022). doi: 10.1016/j.cell.2022.10.017; pmid: 36347254
58. D. Katrekar *et al.*, Comprehensive interrogation of the ADAR2 deaminase domain for engineering enhanced RNA editing activity and specificity. *eLife* **11**, e75555 (2022). doi: 10.7554/eLife.75555; pmid: 35044296
59. B. Bushnell, J. Rood, E. Singer, BBMerge - Accurate paired shotgun read merging via overlap. *PLOS ONE* **12**, e0185056 (2017). doi: 10.1371/journal.pone.0185056; pmid: 29073143

ACKNOWLEDGMENTS

We thank all members of the Hsu lab for helpful input throughout the course of the project. We are grateful to J. Athukoralage for insights on the bridge recombination mechanism and V. Tran for discussions about deep mutational scanning. We also thank B. Plosky, R. Senturia, and J. Caputo of the Arc Institute. **Funding:** N.T.P., L.J.B., D.K., G.A.G., M.G.D., J.J.P., A.F., and S.K. are supported by funding from the Arc Institute. J.Q.M. is supported by funding from the Arc Institute and the Curci Foundation. M.H. is supported by JSPS KAKENHI grant no. 23K14133; Takeda Medical Research Foundation; and JST, ACT-X grant no. JPMJAX232F. H.N. is supported by JSPS KAKENHI grant nos. 21H05281, 22H00403, and 25H00436; Takeda Medical Research Foundation; the Inamori Research Institute for Science; and JST CREST grant no. JPMJCR23B6. P.D.H. is supported by funding from the Arc Institute, Rainwater Foundation, Curci Foundation, Rose Hill Innovators Program, Yosemite, S. Altman, V. and N. Khosla, and anonymous gifts to the Hsu Lab. **Author contributions:** N.T.P., M.G.D., S.K., and P.D.H. conceived the study. N.T.P., L.J.B., D.K., S.K., and P.D.H. designed experiments. N.T.P., L.J.B., D.K., G.A.G., A.F., J.Q.M., and J.J.P. performed experiments. N.T.P., L.J.B., D.K., G.A.G., M.G.D., A.F., J.Q.M., M.H., H.N., S.K., and P.D.H. analyzed and interpreted data. D.K. designed deep mutational scan strategy. A.F. designed and prepared engineered cell lines. J.Q.M. performed stem cell experiments. N.T.P. and M.G.D. performed off-target analyses. M.G.D. nominated orthologs for testing. G.A.G. analyzed AlphaFold structures of engineered variants. M.H. and H.N. advised on bRNA and protein engineering. N.T.P. and C.R.-T. designed figures with assistance from D.K. N.T.P. and P.D.H. wrote the manuscript with important contributions from H.N. and S.K. and input from all other authors. **Competing interests:** P.D.H. acknowledges outside interest as a cofounder of Monet AI and Stylus Medicine, a member of the board of directors at Stylus Medicine, a scientific advisory board member at Amgen and Veda Bio, and a venture partner at Thrive Capital. M.G.D. and A.F. acknowledge outside interest in Stylus Medicine. D.K. acknowledges outside interest in Shape Therapeutics. A.F., M.G.D., and P.D.H. are inventors of PCT/US2022/079227 held by the Regents of the University of California, the Board of Trustees of the Leland Stanford Junior University, and the Salk Institute for Biological Studies that covers use of large serine recombinases. N.T.P, M.G.D., and P.D.H. are inventors on patent application number PCT/US23/82192 submitted by the Arc Institute that covers bridge recombinase in human cells. N.T.P., M.G.D., P.D.H., M.H., and H.N. are inventors on patent application number PCT/US25/12632 submitted by the Arc Institute that covers bRNA engineering. N.T.P., D.K., G.A.G., and P.D.H. are inventors on patent application number 63/804,543 submitted by the Arc Institute that covers bridge recombinase engineering. **Data, code, and materials availability:** Data supporting the results are in the main text, figures, and supplementary tables. Next-generation sequencing datasets generated as part of this study are available on the NCBI Sequence Read Archive at Bioproject PRJNA1313122. ISCro4, enhanced ISCro4, and enhanced single- and split-bRNA plasmids can be requested from Addgene under the Uniform Biological Material Transfer Agreement with Arc Research Institute. **License information:** Copyright © 2026 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adz0276
Figs. S1 to S14; Tables S1 to S4; MDAR Reproducibility Checklist

Submitted 15 May 2025; accepted 13 September 2025; published online 25 September 2025
10.1126/science.adz0276

Science
AAAS

Megabase-scale human genome rearrangement with programmable bridge recombinases

Nicholas T. Perry, Liam J. Bartie, Dhruva Katrekar, Gabriel A. Gonzalez, Matthew G. Durrant, James J. Pai, Alison Fanton, Juliana Q. Martins, Masahiro Hiraizumi, Chiara Ricci-Tam, Hiroshi Nishimasu, Silvana Konermann, and Patrick D. Hsu

Science **391** (6790), eadz0276. DOI: 10.1126/science.adz0276

Editor's summary

Precisely inserting large pieces of DNA into specific locations in the genomes of cells remains a major challenge in biotechnology and medicine. Perry *et al*. and Pelea *et al*. now introduce a genome-editing approach using a bacterial recombinase enzyme called ISCro4, guided by programmable guide RNA molecules that direct the enzyme to recognize target and donor DNA sites. These studies demonstrate that ISCro4 works efficiently in human cells, where it can remove, flip, or insert DNA segments thousands of nucleotides in length, albeit with some specificity constraints. RNA-guided recombinases could thus provide a powerful genome-editing platform to complement and extend beyond current technologies. —Di Jiang

View the article online
https://www.science.org/doi/10.1126/science.adz0276
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works