Targeted Integration to Endogenous Sites in the Human Genome Using CRISPR-Associated Transposases Discovered from Natural Environments

Metagenomi

Jason Liu*, Daniela S. Aliaga Goltsman*, Lisa M. Alexander, Khak Khak Khayi, Christine A. Romano, Morayma Temoche-Diaz, Shailaja Chadha, Rodrigo Fregoso-Ocampo, Jennifer Hong, Owen P. Janson, Liliana Gonzalez-Osorio, Sarah Laperriere, Keirstinne Turcios, Audra E. Devoto, Cindy J. Castelle, Cristina N. Butterfield, Brian C. Thomas, Gregory J. Cost+, Christopher T. Brown+

Abstract

Integration of DNA into targeted genomic sites has historically been a challenge for therapeutic gene editing. Established transposase and lentiviral systems are efficient at inserting large DNA cargos into the human genome, but result in non-specific and sometimes hyperactive integration. These integrations have resulted in severe adverse events during clinical trials in the form of neoplasia. CRISPR associated transposases (CAST) are a potential solution to this problem, as they could provide user-directed and programmable DNA integration. Although relatively rare in nature, the handful of known CAST systems are efficient at delivering large DNA payloads into bacterial genomes; however, translation to mammalian cells has not been possible. We hypothesized that novel CAST identified from metagenomic sequences of environmental samples would enable discovery of systems more amenable to use in human cells. Mining of millions of assembly-driven, metagenomic sequences from diverse environments uncovered active CAST capable of efficient transposition *in vitro* and into the *E. coli* genome. When delivered to mammalian cells, these CAST components are expressed in an active form and localized to the nucleus. When tested in cells, we reproducibly achieved programmable transposition into multiple endogenous sites in the human genome. Our results augur well for the development of CAST into tools for treatment of genetic disease.

Overview

Our approach to CAST discovery

Proprietary sampling

Our scientists collect samples from diverse climates and geographies, building a database that spans biodiversity from high-altitude and high-temperature environments to hydrothermal vents below the ocean

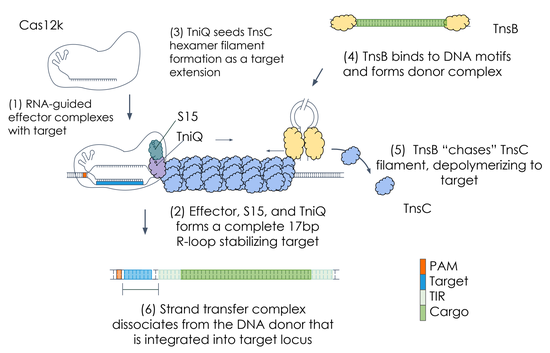
High-throughput screening

Leveraging high-throughput screening, AI-based cloud computing, and proprietary algorithms, we have identified and filed on over 20,000 novel systems from over 180 novel enzyme families

Toolbox engineering

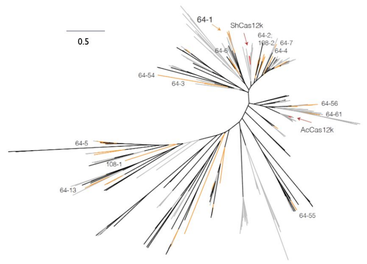
We have engineered a broad toolbox of next-generation gene editing systems - including nucleases, base editors, prime editors and CRISPR transposases (CASTs) to address a wide variety of genetic diseases

CAST mechanism

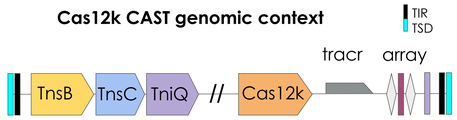


Novel Cas12k CAST are active in vitro

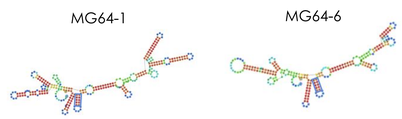
Cas12k effector diversity



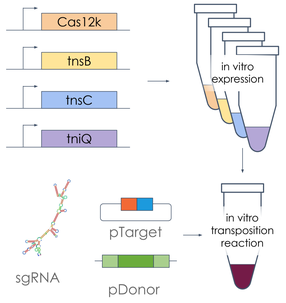
Cas12k CAST genomic context



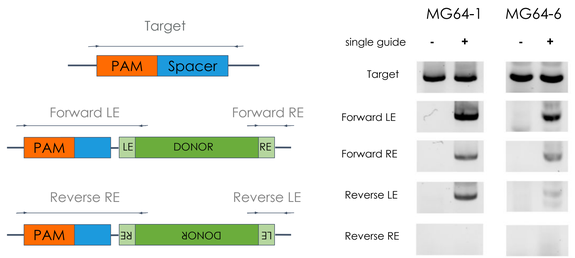
Active single guides



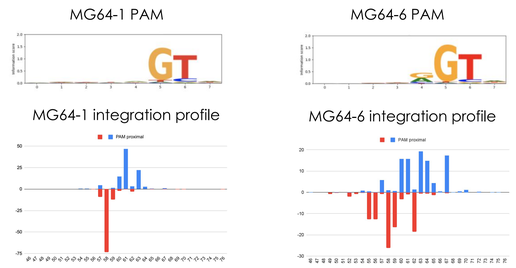
Schematic of CAST testing



Activity determination

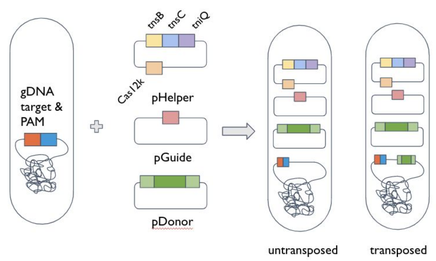


CAST characterization

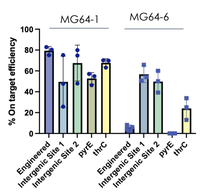


Efficient, targeted transposition in *E. coli*

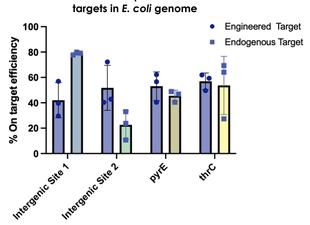
Schematic of *E. coli* transposition



Efficiency of transposition at engineered and endogenous targets

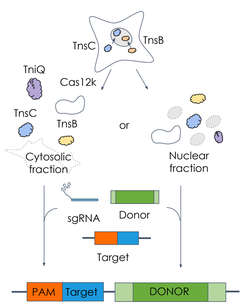


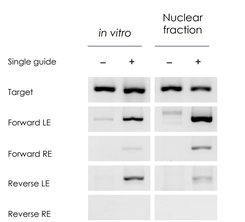
MG64-1 multiplexes across targets in *E. coli* genome



Translation of CASTs to human cells

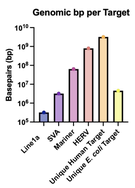
CAST components retain activity after engineering for nuclear localization

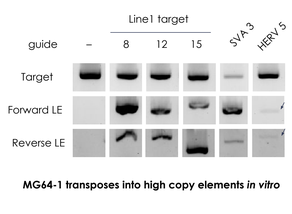




Nuclear fraction of cells stably expressing CAST components are sufficient for transposition *in vitro*

CAST can navigate human genome *in vitro*

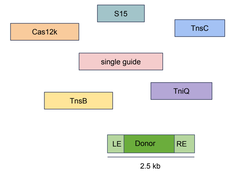


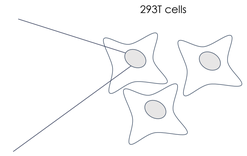


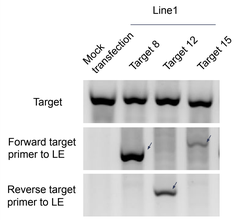
MG64-1 transposes into high copy elements *in vitro*

Targeted transposition in human cells

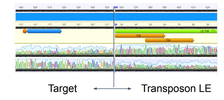
Targeting high copy number sites





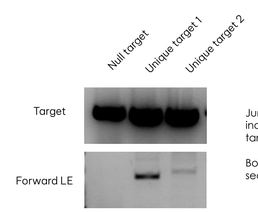


Sanger sequencing confirms transposition



Two plasmid MG64-1 system transposes donor into high copy targets of human genomic DNA in HEK293T cells. Directionality is determined by forward orientation to the PAM.

Integration at single copy sites in the human genome



Junction PCR of donor-target forward LE orientation indicates that MG64-1 is capable of transposing to unique target loci in HEK293T cells.

Both unique sites are confirmed for transposition by Sanger sequencing and NGS

Summary

- Identified novel Cas12k CAST systems from natural environments capable of RNA-directed integration of donor cargo
- Characterized systems with a nGTN/rGTN 5' PAM that complete transposition in the forward direction, and are capable of partial product formation in the reverse direction. CAST systems integrate donor in a tight ~10 bp window centered at 61 bp away from the PAM.
- Minimized versions of the LE/RE and single guide components are capable of transposition and will improve delivery
- In *E. coli,* CASTs show robust targeted integration at multiple genomic loci, achieving efficiencies up to 80%
- Systematic targeting to high copy sites enabled integration of CAST payloads into the genome of mammalian cells
- MG64-1 is capable of integration to single copy loci in the human genome

References

1. Strecker J, Ladha A, Gardner Z, Schmid-Burgk JL, Makarova KS, Koonin EV, Zhang F. RNA-guided DNA insertion with CRISPR-associated transposases. Science. 2019 Jul 5;365(6448):48-53. doi: 10.1126/science.aax9181. Epub 2019 Jun 6. PMID: 31171706; PMCID: PMC6659118.
2. Park JU, Tsai AW, Mehrotra E, Petassi MT, Hsieh SC, Ke A, Peters JE, Kellogg EH. Structural basis for target site selection in RNA-guided DNA transposition systems. Science. 2021 Aug 13;373(6556):768-774. doi: 10.1126/science.abi8976. Epub 2021 Jul 15. PMID: 34385391; PMCID: PMC9080059.
3. Querques, I., Schmitz, M., Oberli, S. et al. Target site selection and remodelling by type V CRISPR-transposon systems. Nature 599, 497–502 (2021). https://doi.org/10.1038/s41586-021-04030-z
4. Xiao R, Wang S, Han R, Li Z, Gabel C, Mukherjee IA, Chang L. Structural basis of target DNA recognition by CRISPR-Cas12k for RNA-guided DNA transposition. Mol Cell. 2021 Nov 4;81(21):4457-4466.e5. doi: 10.1016/j.molcel.2021.07.043. Epub 2021 Aug 26. PMID: 34450043; PMCID: PMC8571069.
5. Schmitz M, Querques I, Oberli S, Chanez C, Jinek M. Structural basis for the assembly of the type V CRISPR-associated transposon complex. Cell. 2022 Dec 22;185(26):4999-5010.e17. doi: 10.1016/j.cell.2022.11.009. Epub 2022 Nov 25. PMID: 36435179; PMCID: PMC9798831.
6. Park JU, Tsai AW, Chen TH, Peters JE, Kellogg EH. Mechanistic details of CRISPR-associated transposon recruitment and integration revealed by cryo-EM. Proc Natl Acad Sci U S A. 2022 Aug 9;119(32):e2202590119. doi: 10.1073/pnas.2202590119. Epub 2022 Aug 1. PMID: 35914146; PMCID: PMC9371665.
7. Park, JU., Tsai, A.W., Rizo, A.N. et al. Structures of the holo CRISPR RNA-guided transposon integration complex. Nature 613, 775–782 (2023). https://doi.org/10.1038/s41586-022-05573-5