ALTERING, IMPROVING, AND DEFINING THE SPECIFICITIES OF CRISPR-CAS NUCLEASES

Benjamin P. Kleinstiver, Molecular Pathology Unit ; Department of Pathology, Harvard Medical School, USA
bkleinstiver@mgh.harvard.edu
Michelle S. Prew, Nhu T. Nguyen, Moira M. Welch, Molecular Pathology Unit ; Center for Computational and Integrative Biology, Massachusetts General Hospital, USA
Vikram Pattanayak, Shengdar Q. Tsai, J. Keith Joung, Molecular Pathology Unit ; Department of Pathology, Harvard Medical School, USA
Martin J. Aryee, Molecular Pathology Unit ; Center for Computational and Integrative Biology, Massachusetts General Hospital ; Center for Cancer Research ; Department of Pathology, Harvard Medical School ; Department of Biostatistics, Harvard T.H. Chan School of Public Health, USA

Keywords: CRISPR-Cas, genome-editing, engineered nuclease, targeting range, genome-wide specificity

CRISPR-Cas9 nucleases have been widely adopted for genome editing applications to knockout genes or to introduce desired changes. While these nucleases have shown immense promise, two notable limitations of the wild-type form of the broadly used *Streptococcus pyogenes* Cas9 (SpCas9) are the restriction of targeting range to sites that contain an NGG protospacer adjacent motif (PAM), and the undesirable ability of the enzyme to cleave off-target sites that resemble the on-target site. Scarcity of PAM motifs can limit implementations that require precise targeting, whereas off-target effects can confound research applications and are important considerations for therapeutics.

To improve the targeting range of SpCas9 and an orthogonal Cas9 from *Staphylococcus aureus* (called SaCas9), we optimized a heterologous genetic selection system that enabled us to perform directed evolution of PAM specificity. With SpCas9, we evolved two separate variants that can target NGA and NGCG PAMs[1], and with SaCas9 relaxed the PAM from NNGRRT to NNNRRT[2], increasing the targetability of these enzyme 2- to 4-fold. The genome-wide specificity profiles of SpCas9 and SaCas9 variants, determine by GUIDE-seq[3], indicate that they are at least as, if not more, specific than the wild-type enzyme[1,2]. Together, these results demonstrate that the inherent PAM specificity of multiple different Cas9 orthologues can be purposefully modified to improve the accuracy of targeting.

Existing strategies for improving the genome-wide specificity of SpCas9 have thus far proven to be incompletely effective and/or have other limitations that constrain their use. To address the off-target potential of SpCas9, we engineered a high-fidelity variant of SpCas9 (called SpCas9-HF1), that contains alterations designed to reduce non-specific contacts to the target strand DNA backbone. In comparison to wild-type SpCas9, SpCas9-HF1 rendered all or nearly all off-target events imperceptible by GUIDE-seq and targeted deep-sequencing methods with standard non-repetitive target sites in human cells[4]. Even for atypical, repetitive target sites, the vast majority of off-targets induced by SpCas9-HF1 and optimized derivatives were not detected[4]. With its exceptional precision, SpCas9-HF1 provides an important and easily employed alternative to wild-type SpCas9 that can eliminate off-target effects when using CRISPR-Cas9 for research and therapeutic applications.

Finally, on-target activity and genome-wide specificity are two important properties of engineered nucleases that should be characterized prior to adoption of such technologies for research or therapeutic applications. CRISPR-Cas Cpf1 nucleases have recently been described as an alternative genome-editing platform[5], yet their activities and genome-wide specificities remain largely undefined. Based on assessment of on-target activity across more than 40 target sites, we demonstrate that two Cpf1 orthologues function robustly in human cells with efficiencies comparable to those of the widely used *Streptococcus pyogenes* Cas9. We also demonstrate that four to six bases at the 3' end of the short CRISPR RNA (crRNA) used to program Cpf1 are insensitive to single base mismatches, but that many of the other bases within the crRNA targeting region are highly sensitive to single or double substitutions[6]. Consistent with these results, GUIDE-seq performed in multiple cell types and targeted deep sequencing analyses of two Cpf1 nucleases revealed no detectable off-target cleavage for over half of 20 different crRNAs we examined. Our results suggest that the two Cpf1 nucleases we characterized generally possess robust on-target activity and high specificities in human cells, findings that should encourage broader use of these genome editing enzymes.

1. Kleinstiver, BP, et al. (2015) Nature, 523(7561):481-5
2. Kleinstiver, BP, et al. (2015) Nature Biotechnology, 33(12):1293-98
3. Tsai, SQ et al. (2015) Nature Biotechnology, 33(2):187-97
4. Kleinstiver, BP and Pattanayak, V, et al. (2016), Nature, 529(7587):490-5
5. Zetsche, B, et al. (2015) Cell, 163(3):759-71
6. Kleinstiver, BP and Tsai, SQ, et al. (2016), Nature Biotechnology, 34(8):869-74