Cas9 Variants Expand the Target Repertoire in *Caenorhabditis elegans*

Ryan T. Bell*, Becky X.H. Fu*, and Andrew Z. Fire†*

*Department of Genetics, Stanford University School of Medicine, Stanford, California 94305
†Department of Pathology, Stanford University School of Medicine, Stanford, California 94305
Running title: Cas9s with Alternate PAMs in C. elegans

Keywords: Alternative PAMs; C. elegans; CRISPR/Cas9; Genome editing; VQR Cas9

Ryan T. Bell
Department of Genetics
Stanford University School of Medicine
300 Pasteur Drive – L302 Stanford, CA 94305-5324
Phone: 650 723 2885
E-mail: ryanbell@stanford.edu

Andrew Z. Fire
Departments of Pathology and Genetics
Stanford University School of Medicine
300 Pasteur Drive – L235 Stanford, CA 94305-5324
Phone: 650 723 2885
E-mail: afire@stanford.edu
ABSTRACT

The proliferation of CRISPR/Cas9-based methods in *Caenorhabditis elegans* has enabled efficient genome editing and precise genomic tethering of Cas9 fusion proteins. Experimental designs using CRISPR/Cas9 are currently limited by the need for a protospacer adjacent motif (PAM) in the target with the sequence NGG. Here we report the characterization of two modified Cas9 proteins in *C. elegans* which recognize NGA and NGCG PAMs. We found that each variant could stimulate homologous recombination with a donor template at multiple loci, and that PAM specificity was comparable to wild type Cas9. In order to directly compare effectiveness, we used CRISPR/Cas9 genome editing to generate a set of assay strains with a common single-guide RNA (sgRNA) target sequence, but which differ in the juxtaposed PAM (NGG, NGA, or NGCG). In this controlled setting, we determined that the NGA PAM Cas9 variant can be as effective as wild type. We similarly edited a genomic target to study the influence of the base following the NGA PAM. Using four strains with four NGAN PAMs, differing only at the fourth position and adjacent to the same sgRNA target, we observed that efficient homologous replacement was attainable with any base in the fourth position, with an NGAG PAM being the most effective. In addition to demonstrating the utility of two Cas9 mutants in *C. elegans* and providing reagents which permit CRISPR/Cas9 experiments with fewer restrictions on potential targets, we established a means to benchmark the efficiency of different Cas9::PAM combinations that avoids variations due to differences in the sgRNA sequence.

INTRODUCTION

The characterization of the type II CRISPR-Cas bacterial immunity system has provided a powerful and efficient means to tailor the genomes of a wide variety of organisms (Cong et al. 2013; DiCarlo et al. 2013; Dickinson et al. 2013; Friedland et al. 2013; Gratzer et al. 2013; Li et al. 2013; Yang et al. 2014). The Cas9 endonuclease can be programmed by a readily modified single-guide RNA (sgRNA) to cleave a target DNA sequence (Jinek et al. 2012). In vivo, this activity can result in imperfect repair by non-homologous end joining (NHEJ), often creating a null allele of a target gene. When a homologous donor template is provided, desired modifications can be precisely introduced through homologous recombination (HR). The affinity of Cas9 for targeted sequences has also been harnessed to directly influence gene expression. The use of catalytically inactive Cas9 fused to transcriptional activators, repressors, or histone modifying enzymes has shown that the full potential of CRISPR/Cas9 techniques is still being realized (Gilbert et al. 2013; Mali et al. 2013; Hilton et al. 2015; Kearns et al. 2015; Long et al. 2015).

All of these methods share the requirement of a protospacer adjacent motif (PAM) in the target (Jinek et al. 2012). The NGG PAM required by the widely used Cas9 from *S. pyogenes* is relatively abundant in randomly composed DNA, but is less common in AT rich genomes, imposing a significant constraint. Recent studies have explored the utility of Cas9s from other bacterial species with diverse PAM specificities, as well as the directed evolution of the *S. pyogenes* Cas9 coding sequence to select mutants which recognize alternate PAMs (Esvelt et al. 2013; Kleinstiver et al. 2015; Ran et al. 2015).
Several groups have developed effective protocols utilizing CRISPR/Cas9 and additional tools for the introduction of specific mutations into the *C. elegans* genome (Dickinson et al. 2013; Arribere et al. 2014; Kim et al. 2014; Paix et al. 2014; Zhao et al. 2014; Dickinson et al. 2015; Paix et al. 2015; Ward 2015). One particularly flexible procedure involves injection of the gonad with a Cas9 transgene, an sgRNA expression vector, and an oligonucleotide repair template, resulting in gene conversion through homologous recombination with the donor template (Arribere et al. 2014; Paix et al. 2014; Zhao et al. 2014). Co-conversion with a dominant phenotypic marker mutation reduces the amount of screening necessary to isolate a desired mutant (Arribere et al. 2014; Kim et al. 2014; Ward 2015). In our application of this technique, we have observed that the recombination with the oligonucleotide donor is highly localized (Arribere et al. 2014). The portion of the donor incorporated into the genome often extends only ten or fewer bases from the site of Cas9 cleavage. In addition, a scarcity of GG dinucleotides in many regions of the genome has confounded targeting certain types of functionally critical sequences, such as 3’ UTRs. As a result, we were motivated to make more of the *C. elegans* genome accessible to efficient editing with oligonucleotide templates by expressing Cas9s with a variety of PAM requirements. In choosing an approach to extend the targetable genome, it was important to consider the substantial challenges that have been encountered in expressing foreign coding sequences in the nematode germline (Kelly et al. 1997). Expression of a “new” coding sequence in the germline can take several trials with success or failure occurring as a complex function of the corresponding DNA, RNA, and protein sequences (Fire 2005). This provides a distinct advantage to approaches for extending editing technology that rely on coding regions that derive from established vectors which are active in the germline. In this work, we describe a set of variants of an effective *S. pyogenes* Cas9 germline expression vector.

To extend the functional editing repertoire in *C. elegans*, we investigated two recently described Cas9 variants which appeared particularly promising based on their reported specificities (Kleinstiver et al. 2015). One mutant, termed the VQR variant, possesses the amino acid substitutions D1135V, R1335Q, and T1337R, which results in a Cas9 protein recognizing an NGA PAM. The second mutant, referred to as the VRER variant, has the mutated residues D1135V, G1218R, R1335E, and T1337R, which enable the recognition of an NGCG PAM (Figure 1a). In the crystal structure of wild type Cas9 in complex with an sgRNA and an NGG PAM, the R1335 residue, substituted in both the VQR and VRER variants, contacts the third PAM position (Anders et al. 2014). The T1337R substitution present in both variants is also located in this conserved PAM-interacting domain. The mutant proteins recognize the same sgRNA scaffold (secondary structure) as wild type *S. pyogenes* Cas9, a substantial advantage in working with these variants.

**MATERIALS AND METHODS**

**Strain construction and maintenance**

The VC2010 N2 strain (Thompson et al. 2013) was used as an initial starting strain for all experiments described in this paper. The strain was used directly in experiments to test the VQR and VRER Cas9 variant transgenes in *dpy-10* and *unc-58* and to assess their PAM specificity through co-injection with the wild type Cas9 transgene. When directly comparing the VQR
transgene with wild type and VRER using the same sgRNA vector, the VQR transgene was injected into VC2010 animals, while the other two were injected into mutant derivatives of VC2010 with modified PAMs (NGG or NGCG) at the dpy-10 locus. The VC2010 variants needed for the wild type and VRER assays were generated using a two-step oligonucleotide-templated CRISPR/Cas9 editing approach (ARRIBERE et al. 2014). VC2010 animals were first altered to generate dpy-10(cn64) homozygotes, which have a Dpy Rol phenotype that is easily distinguished from Rol heterozygotes. Subsequent use of an sgRNA vector targeting the dpy-10(cn64) sequence facilitated recombination of oligonucleotides carrying silent mutations to alter the endogenous PAM and to revert the dpy-10(cn64) mutation, yielding Rol animals. From this group of Rol progeny, wild type progeny were isolated and validated as the desired mutants by PCR and sequencing. All strains used to test each of the four NGAN PAMs with the VQR Cas9 were constructed in the same fashion. All strains were grown on nematode growth medium (NGM) plates seeded with E. coli strain OP50 (BRENNER 1974). All animals were grown at 16° prior to injection and at 23° after injection.

Supplementary Table S1 lists all strains created and used in this study.

**Plasmid construction**
Plasmids have been deposited with Addgene with the following accession numbers: VQR Cas9 (pRB1080), 71309; VQR dpy-10 sgRNA plasmid (pRB1081), 71479; VRER Cas9 (pRB1083), 71480; VRER dpy-10 sgRNA plasmid (pRB1084), 71516. All other plasmids used in this study are available upon request.

The VQR and VRER Cas9 transgenes (pRB1080 and pRB1083) were constructed by Gibson cloning (GIBSON et al. 2009) of gBlocks (Integrated DNA Technologies) into the vector pDD162 (Peft-3::Cas9::tbb-2 3’ UTR). Each gBlock contained a portion of the Cas9 coding sequence from pDD162 with codon substitutions to express the mutated proteins. These codons were optimized for worm expression using the C. elegans Codon Adaptor (REDEMANN et al. 2011). The sgRNA expression vectors pRB1081 (VQR, dpy-10), pRB1082 (VQR, unc-58), pRB1084 (VRER, dpy-10), pRB1085 (VRER, unc-58), and pRB1086 (wild type Cas9, dpy-10(cn64)) were built by cloning oligonucleotides into the vector pRB1017 (ARRIBERE et al. 2014). The plasmids were verified by sequencing and their concentrations were determined by the Qubit dsDNA Broad Range kit (Invitrogen).

Supplementary Table S2 lists all vectors, oligonucleotides, and gBlocks used in this study.

**Microinjection**
In all injections, Cas9 plasmid was present at 50ng/ul, such that when two different Cas9 transgenes were in the same injection mix, each was present at 25ng/ul. Each of the sgRNA plasmids was present at 25ng/ul, and each donor oligonucleotide was present at 500 nM in all injections, regardless of the number in the mix. Young adult animals were injected in both distal gonad arms in most cases. Injected animals were rehydrated in recovery buffer (1mg/ml salmon sperm DNA, 4% glucose, 2.4mM KCl, 66mM NaCl, 3mM MgCl2, 3mM CaCl2, 3mM HEPES pH 7.2) prior to being placed on an NGM plate seeded with OP50.
Screening

The F1 was screened for Rol, Dpy, and Unc phenotypes 3-4 days after injection. The progeny of animals injected with a mixture targeting *dpy-10* for conversion consistently display a range of phenotypes besides the left-handed rolling exhibited by *dpy-10(cn64)* mutants and the Dpy phenotype of *dpy-10* null homozygotes. In practice, it can be difficult to distinguish between slightly Dpy animals that can be *dpy-10(cn64)* heterozygote mosaics which hardly roll, heterozygotes for a *dpy-10* null mutation, or animals which have no heritable *dpy-10* mutation. We attribute the indistinct range of phenotypes we observe to mosaicism and the obfuscating effect of Cas9 activity in somatic tissue. In our experience observing the progeny of these individuals, we have discerned that the majority carry heritable *dpy-10* null mutations, but that a significant fraction possess the *dpy-10(cn64)* mutation resulting from HR with the donor template, while others have no heritable phenotype. For our purposes, we scored any ambiguous phenotypes with partial or complete dumpiness and any degree of body shape dysfunction as Dpy Rol. When Dpy or Rol animals were also Unc, they were scored for consistency as a single class (“Dpy Unc or Rol Unc”).

Calculation of newly targetable sequences

The *C. elegans* annotated RefSeq genes were downloaded using the UCSC Table Browser and the WS220/ce10 genome assembly. Each was downloaded with sufficient flanking sequence to determine the sgRNA sequence for Cas9 cleavage sites near the ends of each gene, but only bases within the coding sequence were checked for proximity to a potential Cas9 cleavage site. Intron sequences were included for sgRNA sequence determination, but were not considered for the proximity calculation.

*C. elegans* 3’ UTR sequences were obtained from the coordinates published in (JAN et al. 2011). The sequences were downloaded using the UCSC Table Browser and the WS190/ce6 genome assembly and included sufficient flanking sequence to determine the sgRNA sequence for potential Cas9 cleavage sites near the ends of each 3’ UTR, but only bases within the 3’ UTR sequence were considered for the proximity calculation.

The analysis was performed using custom python scripts.

RESULTS AND DISCUSSION

Two transgenes expressing Cas9 variants with alternate PAM specificities promote oligonucleotide-templated homologous replacement at multiple target loci

We introduced the sequence changes leading to the VQR and VRER Cas9 mutants into a construct that drives ubiquitous Cas9 expression in *C. elegans* with the promoter from *eft-3* (Figure 1b) (DICKINSON et al. 2013).
Figure 1: Construction of *C. elegans* transgenes expressing Cas9 mutants with alternative PAM preferences

a) Partial amino acid sequences of the *S. pyogenes* Cas9 PAM-interacting domain and mutants which have been evolved (Kleinstiver *et al.* 2015) to recognize an NGA PAM or NGCG PAM. Altered residues are numbered and highlighted in red.

b) A representation of the Cas9 transgene modified to express each variant (Dickinson *et al.* 2013). The transgene was driven ubiquitously by the *eft-3* promoter, and contains three synthetic introns and the 3’ UTR from *tbb-2*. Asterisks denote the locations of the mutations introduced to generate the variants.

In the characterization of the VQR and VRER Cas9 variants by Kleinstiver *et al.*, the assays employed to demonstrate function measured the disruption of genes by indel mutations resulting from NHEJ. As homologous recombination utilizing a defined template (HR) has been a key component in much of CRISPR/Cas9-mediated genome editing in *C. elegans*, we used a protocol that would detect both types of events. We co-injected each Cas9 variant plasmid with an sgRNA expression vector and an oligonucleotide to introduce mutations causing the easily scored Rol and Dpy phenotypes characteristic of *dpy-10(cn64)* and *dpy-10* loss-of-function, respectively (Figure 2a and 2b)(LEVY et al. 1993). We observed that each transgene was indeed capable of inducing both HR and NHEJ events at *dpy-10*. The editing process with these Cas9 variants in *C. elegans* was efficient: a majority of injected worms produced animals with modified targets among their progeny. These events were heritable, as expected for germline modification of the locus. The confirmation of germline activity was important, as somatic NHEJ events at *dpy-10* can often produce a non-heritable Dpy phenotype. In these experiments, the efficiency of the VQR variant at *dpy-10* appeared comparable to that of wild type Cas9 acting at a similar site. While the VRER variant was able to generate both HR and NHEJ events, it appeared to be somewhat less active than either wild type Cas9 or the VQR variant in the *dpy-10* assay. Additional support for the function of the VQR and VRER variants was provided by HR assays with the *unc-58* gene (Figure 2a and 2c)(ARRIBERE et al. 2014). Although the efficiency at *unc-58* was low for both variants when compared to the wild type Cas9 transgene, the appearance of worms with the unique shaking phenotype characteristic of the intended point mutant (*unc-58(e665)*) provided a confirmation of effective HR.
Figure 2: Oligonucleotide-templated homologous recombination in multiple *C. elegans* genes utilizing Cas9 variants requiring NGA and NGCG PAMs

a) A schematic of the technique used to induce readily identifiable homologous replacement in a genomic target. A Cas9 germline transgene, a vector driving sgRNA expression with a *C. elegans* U6 promoter, and a homologous oligonucleotide donor template bearing a dominant phenotypic marker mutation were injected into the gonad of young adult animals. Recombination at the endogenous locus was demonstrated by the observation of Rol or Unc phenotypes in the F1 (Arribere et al. 2014).

b) and c) A depiction of the loci, mutant alleles, PAMs, sgRNAs, and donor oligonucleotides used to demonstrate the capability of the VQR and VRER variants to perform genome editing in *C. elegans*. Initiating an editing event, a double-strand break is made by a Cas9 variant (marked with scissors). The break is repaired through recombination with a homologous donor template carrying a mutation (in green) that confers a dominant phenotype, and additional silent mutations (in yellow) that prevent repeated cleavage. Numbers of viable animals following injection and numbers of broods with successful HR events are shown; as with other microinjection studies in this system (e.g., Arribere et al. 2014), we note that raw success rates can show considerable variation between injection sets, suggesting that such frequencies be taken as illustrative rather than comparative.

**Editing of an endogenous PAM allows a direct comparison of three Cas9 transgenes using a single sgRNA target**

We used a CRISPR/Cas9 genome editing approach (Arribere et al. 2014) to build a set of target sites at an endogenous genomic location that were designed to allow parallel evaluation of efficacy for the wild type, VQR, and VRER proteins. The new alleles were engineered with
synonymous substitutions in dpy-10 which convert an endogenous NGA PAM to either NGG or NGCG (Figure 3a). The strains, in conjunction with their wild type ancestor, allowed the same sgRNA expression vector to be used with all three Cas9 transgenes (wild type, VQR, and VRER) to introduce the dpy-10(cn64) mutation by HR with the donor template (Figure 3b). Therefore, we eliminated the variables of sgRNA sequence and distance of the marker mutation from the site of Cas9 cleavage. In this controlled assay, the effectiveness of VQR Cas9 and wild type Cas9 was comparable, while the VRER Cas9 exhibited lower efficiency (Figure 3c and 3d).

Figure 3: Direct comparison of Cas9s recognizing three different PAMs using a single sgRNA by editing a C. elegans genomic target

a) The endogenous NGA PAM used in the conversion of dpy-10 by the VQR variant was edited to be an NGG or NGCG PAM, using mutations that did not alter the DPY-10 protein sequence. The resulting strains allowed a single sgRNA construct to be employed to target dpy-10 for conversion with the three Cas9 transgenes (wild type, VQR, and VRER).

b) Each Cas9 transgene was injected into a strain with its cognate PAM present in dpy-10 adjacent to the sgRNA target. The break induced by each Cas9 using this sgRNA (marked with scissors) was the same distance from the mutated base that conferred the Rol phenotype.

c) The wild type Cas9 and VQR variant were both highly competent to produce Rol progeny, with many injected animals producing “jackpots” (>10 Rol progeny). The VRER Cas9 induced Rol animals through HR at a significantly lower frequency than the other two Cas9s.
d) In addition to Rol progeny, the injected worms generally yielded as many or more Dpy and Dpy Rol progeny, which derived from largely from indels introduced by the NHEJ repair pathway, but could also result from NHEJ repair on one chromosome and HR repair on the other. The number of injected animals with >10 Dpy or Dpy Rol progeny is explicitly indicated in the plots for the wild type and VQR Cas9s, where that outcome was very frequent. The VRER Cas9, while producing fewer Rol progeny than wild type or VQR (30 of 35 injected animals had zero Rol progeny), displayed considerable activity evidenced by the number of Dpy Rol and Dpy animals observed.

Co-conversion and PAM specificity of wild type and mutant Cas9s

The VQR and VRER Cas9 mutants differ only slightly from the wild type protein, therefore we considered an evaluation of their PAM preferences to be important. In order to characterize PAM stringency, we used the \textit{dpy-10}(cn64) and \textit{unc-58}(e665) phenotypes to assay for HR events. We injected a series of six mixtures which contained donor oligonucleotides for the \textit{dpy-10}(cn64) and \textit{unc-58}(e665) marker mutations. The injection mixes also all included two sgRNA expression vectors: one targeting a sequence in \textit{unc-58} with an NGG PAM, and one targeting a sequence in \textit{dpy-10} with either an NGA or NGCG PAM (Figure 4a).

For each of the Cas9 variants, three mixes were injected. In one, both the variant (VQR or VRER) and the wild type Cas9 transgene were present, while in the other two, only one type was included (Figure 4b). If the proteins discriminated effectively between their preferred PAMs, the expectation when both transgenes were present in the injection mix was to observe an abundance of both Rol and Unc phenotypes in the F1 progeny. We also predicted the presence of Dpy Unc and Rol Unc animals with mutations at both loci (resulting from NHEJ or HR at \textit{dpy-10} and HR at \textit{unc-58}) (Figure 4c). When the wild type Cas9 was omitted from the injection mix, we expected to observe a majority of Rol animals (Figure 4c). Conversely, in the absence of either Cas9 variant, we anticipated a preponderance of Unc animals (Figure 4c).

We found that each of the classes of homologous recombinants was most efficiently recovered in experiments where the matched Cas9 was present (Figure 4d, 4e and 4f). Leaving out the cognate Cas9 resulted in a consistent and dramatic decrease in insertion activity (Figure 4e and 4f). Selectivity, although high, was not without exception. The VQR and VRER variants could occasionally promote HR with an NGG PAM (Figure 4e). Wild type Cas9 could induce HR with an NGA PAM (albeit rarely), and also exhibited a low level of activity using an NGCG PAM (Figure 4f). These results are consistent with previous observations of strong but not perfect PAM dependence for both wild type Cas9\textsc{\textit{HSU} et al. 2013; \textsc{PATTANAYAK} et al. 2013; \textsc{FU} et al. 2014; \textsc{WU} et al. 2014)} and the VQR/VRER variants\textsc{\textit{KLEINSTIVER} et al. 2015).
Figure 4: Co-conversion by each Cas9 mutant and wild type Cas9 demonstrates PAM specificity

a) A schematic of the reagents injected to assess PAM stringency. *dpy-10*(cn64) and *unc-58*(e665) donor oligonucleotides were included in all injection mixes, as well as sgRNA vectors with a target in *unc-58* with an NGG PAM. An sgRNA vector with a target in *dpy-10* with an NGA or NGCG PAM was present if the VQR or VRER Cas9 variant was to be tested, respectively.

b) One variant (VQR or VRER) Cas9 transgene and the wild type Cas9 transgene were injected together or individually, always co-injected with both donor oligonucleotides and sgRNA expression vectors.

c) The VQR and VRER variants should cleave *dpy-10* and produce Rol progeny in the F1, while the wild type Cas9 should cleave the *unc-58* locus and produce Unc progeny. An abundance of both Rol and Unc phenotypes was expected when variant and wild type Cas9s were present, in addition to Dpy Unc and Rol Unc animals with mutations at both loci. When only one Cas9 transgene was included in the injection mix, a preponderance of the corresponding phenotype (Rol for a variant and Unc for wild type) should result if the Cas9s were specific to their PAMs.

d) When either variant transgene was injected with the wild type Cas9 transgene, the progeny exhibited an abundance of Rol, Dpy Unc, Rol Unc, and Unc phenotypes, indicating robust activity at both loci.

e) When the wild type transgene was not included, Rol progeny were a substantial fraction, although some Dpy Unc, Rol Unc, and Unc animals were observed. Worms injected with the VRER Cas9 alone produced relatively few Rol animals, but additional Dpy Rol and Dpy worms were present among their progeny (Figure S1). Many injected animals had no progeny with a given mutant phenotype. For the VQR transgene, these counts were: 6 (zero Rol), 12 (zero
Unc), and 12 (zero Dpy Unc or Rol Unc). For the VRER transgene, these counts were 12 (zero Rol), 14 (zero Unc), and 14 (zero Dpy Unc or Rol Unc).

f) When the wild type Cas9 alone was injected, the result was an anticipated excess of Unc convertant progeny, but a small number of animals with Dpy Unc or Rol Unc phenotypes were also present. Many injected animals had no progeny with a given mutant phenotype. When the WT transgene was injected with an NGA PAM sgRNA expression vector, these counts were 10 (zero Rol), 7 (zero Unc), and 12 (zero Dpy Unc or Rol Unc). When the WT transgene was injected with an NGCG PAM sgRNA expression vector, these counts were 15 (zero Rol), 2 (zero Unc), and 11 (zero Dpy Unc or Rol Unc).

**Editing of an endogenous PAM elucidates the cleavage efficiency of the VQR variant with different NGAN PAMs**

Kleinstiver et al. reported differential effectiveness of the VQR variant dependent on the base following the NGA PAM. Their description of the VQR variant summarized the influence as NGAG > NGAA=NGAT > NGAC, with NGAA and NGAT having approximately 75% of the effectiveness of NGAG, while NGAC was about 50% effective. When converting dpy-10 and unc-58, we utilized NGAG and NGAA PAMs, respectively, and observed much greater efficiency with NGAG. However, the contributing factors of distinct sgRNA sequences, different distances of marker mutations from the Cas9 cleavage site, and disparate genomic locations underscored the value of a more controlled evaluation. We therefore developed a common sgRNA assessment strategy similar to the means by which we compared the three Cas9 variants. Using CRISPR/Cas9, we edited the genome near the site of the dpy-10(cn64) mutation (Arribere et al. 2014), creating four different strains. In each, an NGG PAM adjacent to an sgRNA target we had previously verified with wild type Cas9 was converted to four different NGAN PAMs with each of the four possible nucleotides at the fourth position. This approach allowed us to use a single proven sgRNA to accurately test the VQR Cas9 cleavage efficiency using each of the NGAN PAMs (Figure 5a). The modifications involved making a nonsynonymous substitution, converting an alanine codon to four different serine codons (all of the resulting Ala->Ser variants were wild type in phenotype).

We observed that injected animals derived from all four strains gave rise to Rol progeny, confirming effective HR utilizing each of the NGAN PAMs. Our results strongly support NGAG as the most effective PAM, and substantiate the report that NGAA and NGAT are both more effective than NGAC. When injected with the wild type Cas9 transgene, wild type animals with an NGG PAM adjacent to the same sgRNA target produced Rol progeny in quantities similar to worms injected with the VQR transgene with an NGAG PAM in the same position (Figure 5b).
Figure 5: Direct comparison of cleavage efficiency by the VQR Cas9 with a single sgRNA using the four NGAN PAMs by editing a *C. elegans* genomic target

a) A genomic target sequence with an NGG PAM used previously to convert *dpy-10* as a marker for Cas9 expression (Arribere et al. 2014) was edited by CRISPR/Cas9 to make an alanine to serine substitution (CCG to TCG), alter the wobble position of the upstream valine codon (GTN), and to generate all four NGAN PAMs.

b) The NGAG PAM was highly efficient in recruiting the VQR Cas9 to convert *dpy-10* using this sgRNA, with most injected worms yielding >10 Rol progeny. The NGAA and NGAT PAMs were also able to induce conversion, but to a more modest extent. The NGAC PAM was significantly less proficient, but still competent. Wild type worms with an NGG PAM injected with the wild type Cas9 transgene produced a number of Rol progeny comparable to the NGAG strain injected with the VQR variant transgene. The number of injected animals with >10 Rol progeny is explicitly indicated in the plots where that outcome was very frequent.

**Conclusions**

We have demonstrated that precise genome engineering can be extended to a broader range of target loci in the *C. elegans* genome by utilizing two Cas9 transgenes expressing
modified proteins with alternative PAM specificities. We undertook a careful comparison of the altered Cas9 proteins to the wild type by editing a PAM in a genomic target, *dpy-10*. Our engineering permitted experimental comparisons where a single sgRNA could be used in multiple tester strains in which its constant target sequence was adjacent to the PAM variants recognized by the different Cas9s. Such assays showed that the VQR Cas9, when targeted to an endogenous NGAG PAM, can promote HR at levels comparable to wild type Cas9. We similarly probed the performance of the VQR variant by editing another PAM in *dpy-10* to test the effect of the downstream base. Our results provided corroboration of a gradation in activity (KLEINSTIVER et al. 2015), with NGAC at this site being the least efficient and NGAG being the most effective, extending these results to a meiotic system and to *C. elegans*. This work demonstrated the potential of genome engineering to allow efficient and accurate testing of Cas9 transgenes and characterization of their PAM requirements. In addition, the reagents we developed can serve as a foundation to generate *C. elegans* expression vectors for Cas9 fusion proteins with alternate PAM recognition.

*C. elegans* has an AT-rich genome composition, leading to frequent challenges in identifying NGG PAMs with suitable proximity for efficient oligonucleotide-templated editing. The availability of reagents that allow the use of alternative NGA PAMs should greatly expand the regions of the genome accessible to efficient editing by such methods. Examining coding sequences (Figures S2a and S2b) and 3’ UTRs (Figures S2c and S2d), we indeed see substantial expansion of the targetable repertoires. As an example, the addition of sequences within 5 bases of NGA PAM cleavage sites to those within 5 bases of NGG cleavage sites would expand the number of efficiently targetable bases in coding sequence by 44% and more than double the number in 3’ UTR sequence.

ACKNOWLEDGEMENTS

We would like thank Nimit Jain, Christian Frøkjær-Jensen, Joshua Arribere, Elif Cenik, Karen Artiles, and Massa Shoura for their critical reading of the manuscript, as well as every member of the Fire Lab for their many helpful suggestions during the course of this work. In particular we would like to thank Joshua Arribere, who developed the oligonucleotide-templated CRISPR/Cas9 HR protocol which was the foundation for these experiments, and Christian Frøkjær-Jensen for Gibson cloning advice and suggesting the use of a single sgRNA vector to test different Cas9 variants. This work was supported by National Institutes of Health grants R01GM37706, T32GM007790, and T32HG000044.

REFERENCES


Gibson, D. G., L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison et al., 2009 Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth 6: 343-345.


