CRISPR/Cas9 Methodology for the Generation of Knockout Deletions in Caenorhabditis elegans

Vinci Au,*,1 Erica Li-Leger,*,1 Greta Raymant,*, Stephane Flibotte,*, George Chen,*, Kiana Martin,*, Lisa Fernando,*, Claudia Doell,*, Federico I. Rosell,† Su Wang,*, Mark L. Edgley,*, Ann E. Rougvie,‡ Harald Hutter,§ and Donald G. Moerman*,2

*Department of Zoology and Michael Smith Laboratories, and †Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z4, ‡Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN 55455, and §Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6

ABSTRACT

The Caenorhabditis elegans Gene Knockout Consortium is tasked with obtaining null mutations in each of the more than 20,000 open reading frames (ORFs) of this organism. To date, approximately 15,000 ORFs have associated putative null alleles. As there has been substantial success in using CRISPR/Cas9 in C. elegans, this appears to be the most promising technique to complete the task. To enhance the efficiency of using CRISPR/Cas9 to generate gene deletions in C. elegans we provide a web-based interface to access our database of guide RNAs (http://genome.sfu.ca/crispr). When coupled with previously developed selection vectors, optimization for homology arm length, and the use of purified Cas9 protein, we demonstrate a robust and effective protocol for generating deletions for this large-scale project. Debate and speculation in the larger scientific community concerning off-target effects due to non-specific Cas9 cutting has prompted us to investigate through whole genome sequencing the occurrence of single nucleotide variants and indels accompanying targeted deletions. We did not detect any off-site variants above the natural spontaneous mutation rate and therefore conclude that this modified protocol does not generate off-target events to any significant degree in C. elegans. We did, however, observe a number of non-specific alterations at the target site itself following the Cas9-induced double-strand break and offer a protocol for best practice quality control for such events.

KEYWORDS

C. elegans CRISPR/Cas9 homology dependent repair mutagenesis

CRISPR/Cas9 is the current technology of choice for genome editing (Sander and Joung 2014). This is due to its versatility, as it is an RNA-guided system where a 20-base guide RNA (crRNA) directs a Cas9 nuclease (from Streptococcus pyogenes) to the target sequence, providing high specificity and minimal off-target site effects. The endonuclease makes a double-strand cut at the target site (Gasiunas et al. 2012; Jinek et al. 2012; Jiang et al. 2013), which can then be repaired through Non-Homologous End Joining (NHEJ) or Homology-Directed Repair (HDR). CRISPR/Cas9 technology was first adapted for C. elegans in 2013 (Cho et al. 2013; Friedland et al. 2013; Chiu et al. 2013; Chen et al. 2013; Dickinson et al. 2013; Tzur et al. 2013; Lo et al. 2013; Katic and Großhans 2013; Waijers et al. 2013) and since then the community has produced increasingly sophisticated methods to mutate, delete and tag genes using this technology (for example, Dickinson et al. 2015; Norris et al. 2015; Paix et al. 2014, 2016).

Our facility, along with several international laboratories, is attempting to isolate deletion alleles for the majority of genes in C. elegans. We have tested a number of the current methodologies for the CRISPR/Cas9 system and have successfully isolated several small deletions. We found that none of the CRISPR/Cas9 protocols presently available to the worm community were suitable as is for an ambitious high-throughput approach, which is what the C. elegans Gene Knockout
Facility requires. In this paper we examine the details of several parts of the method with the aim of obtaining a more efficient experimental procedure suitable to the needs of the consortium that may in turn lead to a greater yield of deletions in a timely manner. These parts are (1) the repair mechanism and integrant selection, (2) homology arm length, (3) Cas9 delivery, and (4) guide RNA selection.

We first examined repair mechanisms and selection procedures. While one can obtain deletions utilizing NHEJ, which is reportedly polymerase theta-mediated repair in the C. elegans germline (van Schendel et al. 2015), the most versatile methodologies use HDR to introduce designer modifications at precise locations in the genome (Yang et al. 2013; Auer and Del Bene 2014; Port et al. 2014; Bottcher et al. 2014; DiCarlo et al. 2013; Kim et al. 2014; Arribere et al. 2014; Zhao et al. 2014). HDR is homologous recombination generated at the Cas9 cut site that is facilitated by exogenous DNA with homology to the regions flanking the cut site in the genome. Two recent papers coupled HDR in C. elegans with the introduction of drug selection for either hygromycin (Dickinson et al. 2015) or G418 (Norris et al. 2015) as part of the screening process. This is an important modification as drug selection improves throughput by greatly reducing the number of animals that need to be screened. For the studies described in this paper, we have opted to use the G418 protocol described by Norris et al. (2015). Their protocol requires the introduction of a repair template containing a dual-selection cassette: a G418 resistance gene (neoR) and a pharyngeal GFP reporter (Pmyo-2::GFP), flanked by long homology arms (Figure 1). As described earlier, the drug resistance marker allows F1 selection on G418 and reduces primary screening efforts by eliminating animals that do not carry the repair template. Integration of the repair template via HDR allows the simultaneous deletion of a defined genomic region and introduction of the dual-selection cassette. Furthermore, the GFP marker makes it possible to discern recombinant animals from those harboring concatemer arrays. Recombinant animals display dim and uniform expression of GFP in the pharyngeal muscle, whereas concatemer array expression is often bright and mosaic. Additional pharyngeal and body-wall RFP-bearing plasmids (Pmyo-2::RFP and Pmyo-3::RFP, respectively) aid in the identification of non-integrants, since these independent plasmids should only be expressed from concatemer arrays. As many of the genes targeted in our facility are essential and thus homozygous inviable, the pharyngeal GFP marker acts as a dominant marker and allows us to track the deletion in heterozygotes. The GFP insert also allows us to track genes with no visible phenotype. For the research community this will be particularly useful for generating mutant crosses. Another convenient feature of this vector is the ability to easily excise the selection cassette using Cre recombinase as there are LoxP sites flanking the cassette (Figure 1; Norris et al. 2015).

Once we settled on an HDR procedure, accompanied by drug selection, we then examined the homology arm length required for efficient modification via HDR at the target site. The optimal length of homology arm required for efficient modification via HDR at the target site is still a matter for debate. Paix et al. (2014) have shown that short, 30 to 60 nucleotide homologous linear repair templates can be highly

![Figure 1](image-url)  
*Figure 1* Generation of a deletion using the CRISPR/Cas9 protocol. (1) Guide RNAs direct Cas9 to create targeted DSBs in the gene of interest. (2) Through HDR, a portion of the ORF is replaced with the selection cassette containing pharyngeal GFP and G418 resistance (neoR) markers. (3) The dual-marker cassette, flanked by *loxP* sites, can be excised from the genome by injecting Cre recombinase. (Adapted from Norris et al. 2015.)
efficient for gene editing over small target intervals. In contrast, two papers (Dickinson et al. 2015; Norris et al. 2015) have shown that large inserts (>1.5 kb) require longer homology arms (>500 bp).

We next examined how best to deliver Cas9 to facilitate efficient DNA cutting. There are at least three ways to deliver Cas9 to the target DNA: via plasmid, as messenger RNA or as purified protein. The Seydoux lab has shown that the third method, direct injection of Cas9 protein as a ribonuclease (RNP) complex (Cas9 protein, the structural tracrRNA and the target-specific crRNA), is fivefold more efficient than injection of plasmids coding for these reagents in *C. elegans* (Paix et al. 2014). This advance could decrease injection time considerably and eliminate the need to generate guide RNA expression plasmids for every gene target since crRNAs could be synthesized.

We recognized that the selection of effective guide RNAs is critical to efficiently obtain deletion mutations in *C. elegans*. Although there is no clear consensus for effective guide RNA design constraints, certain metrics for efficacy have been employed by various researchers (reviewed in Mohr et al. 2016). We reasoned that our overall CRISPR/Cas9 success rate could be substantially improved by the application of a set of standard filters to the entire complement of all possible guide RNAs in the *C. elegans* genome, as a type of pre-selection to eliminate from consideration guide RNAs that could be expected to perform poorly. These filters are based on accumulated observations from many organisms. Our online *C. elegans*-specific guide RNA selection tool is available at http://genome.sfu.ca/crispr.

In this manuscript we present experimental data for each of these parts of the method, leading to an efficient standard protocol for using CRISPR/Cas9 to generate deletions in *C. elegans*, which is suitable for the requirements of our facility. We describe our new tool for the design and selection of guide RNAs specific to *C. elegans*, test the length of homology arms necessary and sufficient for engineering deletions of several hundred nucleotides while introducing a large (5.4 kb) selection cassette and compare the relative on-site efficiency of delivering Cas9 encoded in a plasmid to the direct delivery of Cas9 protein using a number of different gene targets. As our facility provides deletion strains to the Caenorhabditis Genetics Center (CGC), which then supplies the larger worm community, we want to ensure, if possible, that there are no additional mutations in these CRISPR/Cas9-generated mutants. To this end, we performed whole genome sequencing (WGS) on several of our deletion strains to determine the frequency of off-site mutations and to examine rearrangements at the target gene itself.

**MATERIALS AND METHODS**

**Designing a *C. elegans*-specific CRISPR guide RNA selection tool**

Guide RNAs were designed for the whole *C. elegans* genome and made available to the research community (see Data Availability). Briefly, a series of filters on all potential guide sequences were applied with an in-house Perl script using the reference genome sequence and gene information from WormBase version WS250 (http://www.wormbase.org). In a first step, for each Protospacer Adjacent Motif (PAM) site in the genome we kept the corresponding adjacent 20-base guide only if its GC content was between 20% and 80% and no poly-T tracts of length five or longer were present. We annotated the presence or absence of the sequence GG at the 3’ end of each guide since guides ending with GG are expected to have higher efficiency for 3’ NGG PAMs (Farboud and Meyer 2015). Guides for which the seed region, defined as 12 bases at the 3’ end plus the PAM, was not unique in the genome were eliminated. The uniqueness of those 15-mers was assessed with an in-house C code modified from Filbotte and Moerman (2008). The guide plus PAM sequences were then aligned to the whole genome with bwa aln (Li and Durbin 2009) allowing an edit distance of three and eliminating guides mapping to multiple locations in the genome. The minimum free energy in kcal/mol was calculated for each guide with the program hybrid-ss-min (Markham 2003) to help the researchers assess potential self-folding of the RNA. The location of the cut site associated with each guide was then annotated according to the gene feature being targeted, if any. All the guides were loaded into a database, which is searchable via a web interface. Users can search the database by entering a genomic interval or by searching a gene name. The database is also available as a genome browser track on WormBase (www.wormbase.org). Constraints can be applied to the GC content, folding energy, and the presence of GG at the 3’ end of the guides being returned.

**CRISPR/Cas9 deletion procedures**

CRISPR/Cas9 edits were generated using a protocol previously described by Norris et al. (2015), with some modifications. Repair templates were assembled using the NEBuilder HiFusion DNA Assembly Kit (New England Biolabs) to incorporate homology arms into a dual-marker selection cassette (loxP + Pmyo-2:GFPlacZ::unc-54 3’ UTR + prps-27:neoR::unc-54 3’ UTR + loxP vector provided by John Calarco). Homology arms ordered as DNA gBlocks from Integrated DNA Technologies (IDT), were approximately 500 bp in length including 50 bp of overhang adaptor sequences for the assembly reaction.

Depending on the experiment, CRISPR/Cas9 edits were generated using one or two guide RNAs, and Cas9 was delivered via one of two methods: a Cas9 expression plasmid (Fu et al. 2014) or purified Cas9 protein (Paix et al. 2014). For experiments using the Cas9 expression plasmid, a single guide RNA (sgRNA) vector was constructed by performing PCR on a pU6:klp-12 vector (Friedland et al. 2013) to incorporate the 20 bp guide RNA sequence via the forward primer. Injection mixes consisted of: 100 ng/µL sgRNA plasmid, 50 ng/µL repair template, 50 ng/µL Cas9 expression plasmid (Pefft-3:Cas9::NLS svac-NLS::::::::::tbb-2 3’ UTR), 5 ng/µL pCFJ104 (Pmyo-3:mCherry), and 2.5 ng/µL pCFJ90 (Pmyo-2:mCherry). For Cas9 protein injections, crRNA and TracrRNA (synthesized by IDT) were duplexed and then combined into a Cas9 RNP complex following the manufacturer’s recommendations. Cas9 protein was purified according to the protocol described by Paix et al. (2015) from plasmid nm2973 (Fu et al. 2014) and stored at -20°C at a concentration of 34 µg/µL in 20 mM Hepes buffer pH 8.0, 500 mM KCl and 20% glycerol. Injection mixes consisted of 50 ng/µL repair template, 1.5 µM RNP complex, 5 ng/µL pCFJ104, and 2.5 ng/µL pCFJ90.

For each CRISPR target, approximately 32 young adult VC3504 (Moerman lab derivative of N2) hermaphrodites were injected as previously described by Kadandale et al. (2009) and screened according to the protocol in Norris et al. (2015).

**Whole genome sequencing (WGS) and analysis to measure off-target effects after CRISPR/ Cas9 treatment**

Off-target effects were assessed using WGS for eight CRISPR/Cas9 edited strains generated for two genes from the same parental VC3504 population. Two guide RNAs were used for each of the target genes, *hgc-45* and *C34D4.2* (Table S1). To assess possible differences in off-target effects between Cas9 plasmid and Cas9 protein, we injected approximately 32 animals with either plasmid or protein for each set
of CRISPR guides. All independently generated strains were sequenced validated by PCR for correct insertion of the selection cassette. For each target, four independently generated strains (two from plasmid injections, and two from protein injections) were selected for WGS. We also sequenced the parental strain VC3504. All CRISPR strains were no more than three generations away from the parental strain when injected.

*C. elegans* strains were grown as described by Brenner (1974). Strains were allowed to grow on either 100 mm or 60 mm agar plates seeded with OP50 until just starved. Worms were washed off plates with sterile distilled water into 15 mL centrifuge tubes, pelleted, and washed with several changes of dH2O to remove residual *Escherichia coli* and other possible contaminants. The supernatant was removed after the final wash and a dense pellet of worms was transferred into a 1.5 mL freezer tube. The dense pellets of worms were frozen at -80°C and subsequently subjected to standard genomic DNA extraction. The extraction begins with Proteinase K digestion and treatment with RNase A, followed by 6M NaCl protein precipitation. The DNA is then precipitated in isopropanol, washed in 70% ethanol, and finally resuspended in distilled water. The genomic DNA was run on 1% agarose gel and high molecular weight DNA was extracted using NEB’s Monarch Gel Extraction Kit. Sequencing libraries were made using Illumina’s Nextera XT library prep kit and run on an Agilent Bioanalyzer to check average fragment size. All strains were sequenced on MiSeq 2x300 or 2x75 runs. In one case, strain VC3823, the second read was unusable, so it was treated as a single run (i.e., not paired) at the analysis stage.

Sequence reads were mapped to the *C. elegans* reference genome version WS260 (http://www.wormbase.org) using the short-read aligner BWA version 0.7.16 (Li and Durbin 2009). For each sample, this resulted in an average sequencing depth ranging from 22x to 41x with a median of 32x. Single-nucleotide variants (SNVs) and small insertions/deletions (indels) were identified and filtered with the help of the SAMtools toolbox version 1.6 (Li et al. 2009). Candidate variants at genomic locations for which the parental N2 strain VC3504 had a read depth lower than 10 reads or an agreement rate with the reference genome lower than 98% were eliminated from further consideration. Furthermore, for each of the eight CRISPR strains sequenced, only variants at locations where all of the other seven CRISPR strains agreed with the reference for at least 95% of the reads were kept. This last filter principally eliminated heterozygous candidates likely due to PCR artifacts (often in or adjacent to homopolymer stretches of A or T) or potentially present at low levels in the parental population but undetected by our sequencing of VC3504. Each variant was annotated with a custom Perl script and gene information downloaded from WormBase version WS260. The read alignments in the regions of candidate variants were visually inspected with Integrative Genomics Viewer (IGV; Robinson et al. 2011; Thorvaldsdottir et al. 2013). Copy numbers were estimated from the alignments with a procedure analogous to that of Itani et al. (2015) using 1 kb wide sliding windows with the alignments from the parental strain as the reference in order to search for potential off-target genomic rearrangements.

**Quality Control testing of CRISPR/Cas9 target deletions using PCR**

WGS is an extremely informative method for determining the structure of a CRISPR mutation, but it is more labor-intensive, time-consuming and costly than PCR. We developed a straightforward PCR protocol as an efficient screening tool, which provides basic information about the selection cassette insertion and the desired deletion in four reactions per generated mutation. Two primer pairs are used to amplify the upstream and downstream portions of the insertion, each pair with one primer in the genomic sequence adjacent to the insertion and the other within the cassette sequence. Another pair of primers is used to amplify wild type (WT) sequence in the region of the deletion from both N2 and mutant templates. The latter primer pair is sometimes designed as a pair flanking the deletion, and sometimes as a pair with one primer flanking and one internal to the deleted region. There are three possible results from these PCRs: (1) The WT product is present and of the correct size from N2 and absent from the mutant, and both insertion-site products are present and of the correct size; (2) The WT assays are correct and one or both insertion-site product are missing or of incorrect size; and (3) The WT assay on the mutant is incorrect. For the first two cases, we conclude that the gene is disrupted and that the insertion sites are correct or rearranged. In the third case, where a WT product of the size predicted for N2 is produced from the mutant template, we conclude that the gene is not disrupted.

**Data availability**

The raw sequence data from this study have been submitted to the NCBI BioProject (http://www.ncbi.nlm.nih.gov/bioproject) under accession number PRJNA 473363 and can be accessed from the Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) with accession number SRP149097. CRISPR guide RNA sequences designed for the whole genome are available using the search tool at http://genome.sfu.ca/crispr/ and can be viewed interactively within JBrowse at WormBase (http://www.wormbase.org) by activating the track called “CRISPR_Cas9 sgRNA predictions”. All the guides can also be downloaded in bulk as a bed file at http://genome.sfu.ca/crispr/. Supplemental material available at Figshare: https://doi.org/10.25387/g3.7312274.

**RESULTS**

**Guide RNA design and selection**

In the context of a production laboratory like our gene knockout facility it is crucial to use the time of the laboratory personnel effectively and eliminate unnecessary tasks. With that in mind, we attempted to streamline the design of guide RNAs by pre-calculating, pre-filtering and annotating potential guide RNAs across the whole *C. elegans* genome using current information about what defines an optimal guide RNA. Briefly, we selected guides with a balanced GC content. We annotated the presence or absence of the sequence GG at the 3’ end of each guide, which has been shown to provide higher efficiency (Farboud and Meyer 2015). Guides for which the seed region was not unique in the genome and guides mapping to multiple locations in the genome were eliminated. Since such a resource is useful to the research community, we made our database of guide RNAs accessible via a public web site (http://genome.sfu.ca/crispr/). Our website provides direct links from each guide to JBrowse at WormBase (http://www.wormbase.org). Within JBrowse one can view the location, sequence and orientation of the guide. Zooming out reveals the location of other guides and provides context for the location of the guide within the gene. The guides can also be viewed within IGV (Robinson et al. 2011; Thorvaldsdottir et al. 2013) after downloading a bed file available on our guide RNA website. Our website allows simple searches using either gene names or genomic intervals, with optional filters to limit searches to regulatory, intergenic, or non-coding regions.

As expected, the use of the guide RNA selection tool significantly reduced the amount of time spent at the design stage by laboratory personnel, but most importantly, it also improved overall laboratory throughput by promoting the selection of more efficient guides.
Our current success rate for obtaining deletions using our pre-selected guides is approximately 85%. Filtering and selecting for unique guide RNAs is most likely the major contributor to the negligible off-target background we observe after CRISPR/Cas9 use (see below).

**Homology-directed repair (HDR) and homology arm length**

HDR allows DNA fragments to be incorporated into a precise region of the *C. elegans* genome. Homology arms flanking these DNA fragments dictate the genomic region that will be simultaneously deleted (if so desired) and replaced. One of the two homologous regions must be immediately adjacent to the Cas9 double-strand cut site in the genome. To date, the homology arm length used in different protocols has varied from 35 bp (Paix et al. 2014) up to approximately 2 kb (Dickinson et al. 2015; Norris et al. 2015). Paix et al. (2016) report optimal HDR efficiency with 35 bp homology arms on a single-stranded template for making small edits. From previous studies (Dickinson et al. 2015; Norris et al. 2015) we suspected we would need longer homology arms to make larger edits. We conducted a series of experiments to test the efficiency of generating a 536 bp deletion and inserting a large selection...
cassette (5.4 kb) at a Cas9-induced double-strand break using a single guide with different lengths of homology arms (Figure 2, Table S1). The lengths of upstream and downstream homology arms differed slightly due to synthesis constraints on the sequence of gBlocks ordered from IDT.

For this experiment, we followed the plasmid-based CRISPR protocol described in Materials and Methods with one modification: injected P0s were single onto plates in order to obtain an accurate measure of editing efficiency, which was defined as the proportion of P0 plates bearing progeny with the selection cassette integrated at the desired location, as confirmed by PCR. The highest efficiency was obtained when utilizing homology arms over approximately 400 bp (Figure 2). When homology arms smaller than these were used, the editing efficiency dropped significantly \((P < 0.05)\). These results agree with previous findings for making larger edits (Dickinson et al. 2015; Norris et al. 2015), suggesting that the optimal length of homology for efficient genome editing is dependent on the particular application of HDR being used. As we observe no further gain in efficiency of HDR for larger deletions with homology arms over approximately 400 bp in length, we have adjusted our protocol accordingly. For our current production protocol, we order homology arms as 500 bp gBlocks from IDT with 50 bp of adaptor sequences included to enable Gibson assembly (Gibson et al. 2009) into our repair vector.

### Table 1 Whole-genome sequencing statistics for the parental strain and 8 mutant strains produced for the off-target mutation analysis

| Strain   | Target Gene | Cas9 Delivery | Read Length (bp) | Mean Coverage | Number of off-target SNVs | Number of off-target indels |
|----------|-------------|---------------|------------------|---------------|--------------------------|---------------------------|
| VC3823   | lgc-45      | plasmid       | 1x75             | 22            | 2                        | 0                         |
| VC3821   | lgc-45      | plasmid       | 2x300            | 34            | 0                        | 0                         |
| VC3817   | lgc-45      | protein       | 2x75             | 34            | 0                        | 0                         |
| VC3814   | lgc-45      | protein       | 2x75             | 30            | 0                        | 0                         |
| VC3843   | C34D4.2     | plasmid       | 2x75             | 27            | 1                        | 0                         |
| VC3845   | C34D4.2     | plasmid       | 2x75             | 34            | 1                        | 0                         |
| VC3834   | C34D4.2     | protein       | 2x75             | 32            | 2                        | 1                         |
| VC3840   | C34D4.2     | protein       | 2x75             | 30            | 1                        | 0                         |
| VC3504   | parent      | —             | 2x300            | 41            | —                        | —                         |

![Figure 4](image-url) Rearrangements at the CRISPR/Cas9 target site. Two examples of removing the target gene through CRISPR/Cas9 HDR visualized using IGV. In strain VC3671, there is a perfect deletion removing the first two exons of gene F01D4.9. In strain VC3674, there is evidence of a complex rearrangement. A large deletion encompasses the same two exons as well as the downstream region of F01D4.5 and a pseudogene, F01D4.3. This region likely harbors multiple copies of the repair template, since the average coverage of the homology arms is relatively high compared to the adjacent genomic sequence. At the bottom of the figure, the exons for the various genes are shown in blue, the homology arms chosen for F01D4.9 are shown in green, and the guide RNA is in red.
if they use purified Cas9 protein rather than plasmid-borne Cas9 (reviewed in Wu et al. 2014). Results confirming these observations have been reported for C. elegans (Paix et al. 2014). There is some debate on how much Cas9 protein is required for making a double-strand break (DSB) in this organism. All results reported here were achieved using Cas9 protein at a concentration between 1.5 μM and 3 μM. Our results confirm that, as in other systems and as previously reported for this nematode, purified Cas9 is more efficient at inducing DSBs than plasmid-borne Cas9. In cases where we have tested the same gene with both purified protein and plasmid-borne Cas9 (three genes), using the purified protein can be nearly four times more efficient at inducing a double-strand break at the target site (Figure 3). Note that this is not true for all genes, as we only see a twofold improvement for the gene C52B11.5. However, the overall difference for these three genes is highly significant (Chi-squared test, $P = 2 \times 10^{-5}$). On average we observe a more than twofold increase in efficiency when comparing the 112 genes we have targeted using either plasmid (56 genes) or protein (59 genes) delivery of Cas9 (Table S2). There are reports using other model systems that suggest that the use of purified protein leads to fewer non-specific effects than using the plasmid. As discussed below, that is not the case for C. elegans, where the two delivery methods exhibit an identical low level of non-specific effects.

**Off-target effects**

At present there appear to be conflicting results in the literature with regard to off-target events when using the CRISPR/Cas9 gene editing system. Whether such events are due to guide RNA homology, homology of the repair template, or unrelated spurious Cas9 cutting is unclear. As it is critical that we know of any such events in the strains we provide to the CGC and ultimately to the larger worm community, we undertook a series of tests to determine the frequency of off-target events after CRISPR/Cas9 treatment.

We performed WGS on eight CRISPR/Cas9-edited strains and their parental strain (Table 1). For this experiment we targeted two distinct genes using two different Cas9 delivery methods. We searched for small off-target mutations, both SNVs and indels, using strict filtering in order to eliminate variants already present in the parental population as well as potential technical artifacts (Materials and Methods). We found a total of eight mutations in addition to the desired edits (Table 1), six homozygous and two heterozygous calls. None of the mutations appear to be related to either the guide RNAs or the homology arms. We also searched for off-target genomic rearrangements by analyzing copy number estimates (Materials and Methods), but none were found. The frequency of extraneous SNVs and indels detected is on the order of spontaneous mutations reported for this organism (Denver et al. 2009), which suggests the observed mutations are unlikely to be due

---

**Figure 5** Complex rearrangements at the CRISPR/Cas9 target site. Whole genome sequencing of VC3743, a CRISPR-generated strain, revealed an unintended 16 kb deletion that disrupts the target gene as well as eight other genes in the surrounding area. This deletion is also accompanied by a complex rearrangement that consists of fragments of genomic sequence from the local region and the repair template. The rearrangement joins (1) the 5’ portion of srh-281 to a duplicated fragment of oac-15, (2) a duplicated inverted portion of oac-15, (3) an inverted intergenic region, and (4,5) the selection cassette flanked by homology arms. In blue are the exons for the various genes and in pink are the homology arms chosen for F11A5.3. The marker for the guide RNA is not visible at this scale.
to the CRISPR/Cas9 procedure. In our quality control studies we have used WGS to analyze an additional 30+ CRISPR/Cas9-generated deletion lines, and again, we found little or no evidence for off-target events (data not shown). We conclude that off-target mutations due to this method are rare and do not appear to be a serious concern in C. elegans when guide RNAs are carefully selected. We also conclude that in this nematode, unlike what has been reported for other systems, the delivery method (i.e., plasmid vs. protein) does not appear to influence the occurrence of off-target mutations.

**Validating deletion events at the target site**

While we detected virtually no off-target events, we now have several examples where target deletions are accompanied by local rearrangements, including local duplications, deletions of adjacent DNA, or insertions of multiple copies of the repair template. Figures 4 and 5 illustrate examples of what we observed. In Figure 4, the upper panel shows a precise deletion of the interval between the homology arms encompassing the first two exons of the gene f01D4.9. However, in the same screen we also isolated a more complex event as illustrated in the lower panel of Figure 4. Here, the region we targeted to delete is accompanied by a large downstream deletion and multiple copies of the repair template, as evidenced by the increased coverage in the homology arm regions. Figure 5 illustrates a more complex example of a rearrangement accompanying the deletion of a targeted interval in a gene. In this example, an additional eight genes are affected beyond the targeted gene. Several genes are deleted and there appears to be accompanying inversions and duplicated regions. These examples are from WGS of more than 30 CRISPR/Cas9 derived strains for several target genes.

The above observations on rearrangements led us to do PCR analysis of 330 CRISPR/Cas9-induced events for 81 genes (Figure 6 and Table S3). The PCR validation scheme described in Table 2 was used to determine the proportion of strains belonging to each mutant class for each of the gene targets. On average, when calculated per gene target, we obtained precise editing events in 35% of the integrant strains. Notably, 37% of the events contained imprecise edits, with rearrangements at one or both, of the 5' and 3' deletion sites. Curiously, we found that 28% of the strains had at least a portion of the wild-type sequence retained along with integration of the drug and GFP selectable marker. For this latter group, we suspect that a partial deletion or duplication event accompanied the marker insertion during HDR. This is suggested by PCR results that show proper integration of the selection cassette at one or both ends of the intended deletion interval yet still exhibit a wild-type band of the correct size. What is important to note here is that there is a wide range in variability from gene to gene for the ratio of precise edits vs. the other categories (Table S3). For some sites, the majority of events lead to precise insertion/deletions, but for other sites it appears to be much more difficult to generate a precise deletion. Perhaps it should not be a surprise that a double-strand break can yield precise HDR, but can also be resolved in other ways.

As this will be a recurring problem and is virtually impossible to predict for each gene (Table S3), we have developed a Quality Control (QC) deletion validation protocol. Our initial selection is based on drug resistance and Mendelian segregation of weak GFP expression in the pharynx. This is strong evidence that we have targeted and disrupted the gene of interest, but as we state above it does not guarantee that we have not disrupted flanking DNA, or even indicate whether we still have an intact copy of the target gene. Our PCR-based QC protocol tests for these possibilities by verifying the integrity of the disruption junction on both sides of the insert and confirming that the wild-type sequence is indeed absent. Primer pairs chosen and expected results are shown in Table 2.

**Figure 6** Homology-directed repair resolves CRISPR/Cas9 double-strand breaks in unpredictable ways. Using the PCR validation scheme described in Table 2, analysis of 330 CRISPR/Cas9-derived mutants (Table S3) reveals that integration of the selection cassette via HDR does not always occur as predicted. These mutant strains were generated using up to two guide RNAs and either Cas9 protein or a Cas9 expression plasmid. The proportion of integrant strains belonging to each mutant class within each gene target was determined and the average proportions for 81 gene targets were calculated.

A key test is to determine whether the wild-type sequence for the targeted deletion interval is indeed absent. For the wild-type PCR assay, we generate a genomic amplicon in wild-type and confirm its absence in the mutant. Our experience to date leads us to expect only a portion of drug-selected GFP positive animals to be correctly edited for any particular gene. It is for this reason that we inject approximately 32 worms to obtain at least five or six positive integrants per target gene.

**DISCUSSION**

We have investigated several variables that may impinge on the effectiveness of the CRISPR/Cas9 procedure to generate large deletions in C. elegans. Our C. elegans-specific guide RNA selection tool, in conjunction with the selection vector designed by the Calarco group (Norris et al. 2015), and the use of purified Cas9 protein, as pioneered by the Seydoux group for use in this organism (Paix et al. 2016), yields a very efficient and effective protocol. We feel this protocol has sufficient robustness to allow our facility to tackle the remaining genes in this organism that lack null alleles.

Among the parameters we explored in refining an HDR protocol appropriate for our purposes were homology arm length and Cas9 delivery methods, i.e., delivery via plasmid or as a purified protein in a nucleoprotein complex. These parameters were previously explored (Dickinson et al. 2015; Norris et al. 2015; Paix et al. 2014, 2016), but in somewhat different conditions than those reported in this study. The length of the homology arm required is at least partially dependent on the type of edit being done to the gene. For single nucleotide changes or small indels, as few as 35 nucleotides of homology are required when using single-stranded bridging oligonucleotides (ssODNs) (Paix et al. 2016). As our facility is looking to make larger deletions that remove most or all of an open reading frame (ORF), this approach is not feasible. For this reason, we were more attracted to protocols developed for replacing larger stretches of DNA. We also preferred protocols that replaced direct PCR screening of DNA from individual strains with a drug and/or visible marker for selection as we envision performing hundreds of screens using the protocol. As stated earlier, we examined the Dickinson et al. (2015) and Norris et al. (2015) vectors and settled on using the G418-resistant, pharyngeal GFP selection system of Norris...
Table 2 Quality Control PCR Interpretation

| Mutant Class          | Left Insertion PCR | Right Insertion PCR | WT PCR on Mutant | WT PCR on N2 | Conclusion                      |
|-----------------------|--------------------|---------------------|------------------|--------------|---------------------------------|
| Precise Edit          | Correct            | Correct             | No product       | Correct      | Mutant, probably as expected    |
| Imprecise Edit        | Correct            | Incorrect           | No product       | Incorrect    | Mutant, probably with rearrangement |
| Imprecise Edit        | Incorrect          | Correct             | No product       | Correct      | Mutant, probably with rearrangement |
| Imprecise Edit        | Incorrect          | Incorrect           | No product       | Correct      | Mutant, probably with rearrangement |
| Gene Partially Intact | Any result         | Any result          | WT product       |              | Not a null mutant                |

et al. (2015). We chose the latter system because it has a smaller selection cassette. The GFP marker is advantageous to the work in our facility as many of the genes we target are essential and thus a deletion allele will lead to homozygous lethality. The GFP acts as a dominant selectable marker and is quite useful for tracking the deletion and constructing balanced lines.

Both the Dickinson et al. (2015) and Norris et al. (2015) groups agree that longer homology arms are required for making large insertions/deletions via HDR. They both experimented with homology arms of 1-2 kb but also pointed out that 500 bp arms would work. We concur with these findings and extend these studies to show there is no advantage to making arms longer than 500 bp. With homology length of only 450 bp, we were able to generate up to 20 kb deletions (using two guide RNAs – data not shown). From a recent paper on using CRISPR/Cas9 to generate inversions in C. elegans it appears there is no limit to the spacing of simultaneous cut sites (Dejima et al. 2018).

Similar to Paix et al. (2016), we find that purified protein is more efficient at generating on-target DSBs. In our hands, the purified protein is at least two times, and often four times more effective than plasmid-borne Cas9 when tested against numerous genes. There are a number of possible reasons for this difference but the most likely one is that there is substantial lag between injecting a plasmid and producing a protein, which then has to be combined with the guide RNA. This lag time is removed when injecting an RNP complex directly. Overall protein concentration between the two methods may also differ.

We detect virtually no off-target events after using our CRISPR/Cas9 protocol. We suspect the major contributor to eliminating off-target events is the use of our C. elegans-specific guide RNA selection tool. Several reviews on the issue of off-target effects consider poor guide RNA design to be a major contributor to these effects (see, for example Wu et al. 2014; Hendel et al. 2015; Jamal et al. 2017). Additionally, while we saw no difference in the frequency of off-target events between Cas9 delivery systems, others have shown that in mammalian cells, direct delivery of Cas9 protein and guide RNA as a nucleoprotein complex reduces off-target events (Kim et al. 2014; Ramakrishna et al. 2014). Considering that the deletion lines we generate will go to the CGC for distribution to the nematode community, identifying and eliminating the sources of off-target events is imperative. It seems the protocol described here satisfies this criterion. Perhaps a more important issue to address at this time is on-target rearrangements that may occur during HDR. We have documented that these events happen at most target sites and therefore proper analysis has to be done when assaying any event generated using the CRISPR/Cas9 method.

ACKNOWLEDGMENTS

Dr. John Calarco (University of Toronto) and Dr. Geraldine Seydoux (Johns Hopkins University) kindly provided us with reagents and much valued advice as we refined the CRISPR/Cas9 system in our laboratory. We also thank Mei Zhen and Julie Claycomb, for their critical support during the early stages of these studies. This work was funded by CIHR grant PJT-148549 to DGM, CIHR grant MOP 93719 and NSERC grant 312498 to HH, NIH grant 3P40OD010440 to AR and NIH grant 5R24OD023041 to AR, Paul Sternberg, Geraldine Seydoux and DGM.

LITERATURE CITED

Arribere, J. A., R. T. Bell, B. X. Fu, K. L. Artiles, P. S. Hartman et al., 2014 Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in C. elegans. Genetics 198: 837–846.
Auer, T. O., and F. Del Bene, 2014 CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. Methods 69: 142–150.
Bottcher, R., M. Hollmann, K. Merk, V. Nitschko, C. Obermaier et al., 2014 Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells. Nucleic Acids Res. 42: e89.
Brenner, S., 1974 The genetics of Caenorhabditis elegans. Genetics 77: 71–94.
Chen, C., I. A. Fenk, and M. de Bon, 2013 Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic Acids Res. 41: e193. https://doi.org/10.1093/nar/gkt805
Chiu, H., H. T. Schwartz, I. Antoshechkin, and P. W. Sternberg, 2013 Transgene-free genome editing in C. elegans using CRISPR/Cas9. Genetics 195: 1167–1171. https://doi.org/10.1534/genetics.113.155879
Cho, S. W., J. Lee, D. Carroll, J. S. Kim, and J. Lee, 2013 Heritable gene knockout in C. elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics 195: 1177–1180. https://doi.org/10.1534/genetics.113.155853
Dejima, K., S. Hori, S. Iwata, Y. Suehiro, S. Yoshida et al., 2018 An aneuploidy-free and structurally defined balancer chromosome toolkit for Caenorhabditis elegans. Cell Reports 22: 232–241. https://doi.org/10.1016/j.celrep.2017.12.024
Denver, D. R., P. C. Dolan, L. J. Wilhelm, W. Sung, J. I. Lucas-Lledo et al., 2009 A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc. Natl. Acad. Sci. USA 106: 16310–16314. https://doi.org/10.1073/pnas.0904895106
DiCarlo, J. E., J. E. Norville, P. Mali, X. Rios, J. Aach et al., 2013 Genome engineering n S. cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41: 4336–4343.
Dickinson, D. J., J. D. Ward, D. J. Reiner, and B. Goldstein, 2013 Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nat. Methods 10: 1028–1034. https://doi.org/10.1038/nmeth.2641
Dickinson, D. J., A. M. Panì, J. K. Heppert, C. D. Higgins, and B. Goldstein, 2015 Streamlined genome engineering with a self-excision drug selection cassette. Genetics 200: 1035–1049. https://doi.org/10.1534/genetics.115.178335
Farboud, B., and B. J. Meyer, 2015 Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics 199: 959–971. https://doi.org/10.1534/genetics.115.175166
Filbotte, S., and D. G. Moerman, 2008 Experimental analysis of oligonucleotide microarray design criteria to detect deletions by comparative genomic hybridization. BMC Genomics 9: 497. https://doi.org/10.1186/1471-2164-9-497
Fu, B. X., L. L. Hansen, K. L. Artiles, M. L. Nonet, and A. Z. Fire, 2014 Landscape of target/guide homology effects on Cas9-mediated cleavage. Nucleic Acids Res. 42: 13778–13787. https://doi.org/10.1093/nar/gku1102
Friedland, A. E., Y. B. Tzur, K. M. Evelt, M. P. Colaiacovo, G. M. Church et al., 2013 Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat. Methods 10: 741–743. https://doi.org/10.1038/nmeth.2532

Gasiunas, G., R. Barrangou, P. Horvath, and V. Siksnys, 2012 Cas9-cRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA 109: E2579–E2586. https://doi.org/10.1073/pnas.1208507109

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchinson, 3rd et al., 2009 Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6: 343–345. https://doi.org/10.1038/nmeth.1318

Hendel, A., E. J. Fine, G. Bao, and M. H. Porteus, 2015 Quantifying on- and off-target genome editing. Trends Biotechnol. 33: 132–140. https://doi.org/10.1016/j.tibtech.2014.12.001

Itani, O. A., S. Flibotte, K. J. Dumas, D. G. Moerman, and P. J. Hu, 2003 Hybrid: A software system for nucleic acid folding, base hybridizing and melting predictions. Masters thesis. Rensselaer Polytechnic Institute, Troy, NY.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, 2013 RNA-guided DNA endonucleases as programmable genome engineers. Cell 157: 86–90. https://doi.org/10.1016/j.cell.2014.04.012

Kadandale, P., I. Chatterjee, and A. Singson, 2009 Germline transformation of Caenorhabditis elegans by injection. Methods Mol. Biol. 518: 123–133. https://doi.org/10.1007/978-1-59745-202-1_10

Katic, I., and H. Großhans, 2013 Targeted heritable mutation and gene conversion by Cas9-CRISPR in C. elegans. Genetics 195: 1173–1176. https://doi.org/10.1534/genetics.113.155754

Kim, S., D. Kim, S. W. Cho, J. Kim, and J. S. Kim, 2014 Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 24: 1012–1019. https://doi.org/10.1101/gr.171322.113

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan et al., 2009 The sequence alignment/map format and SAMtools. Bioinformatics 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lo, T. W., C. S. Pickle, S. Lin, E. J. Ralston, M. Gurling et al., 2013 Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics 195: 331–348. https://doi.org/10.1534/genetics.113.155382

Markham, N. R., 2003 Hybrid: A software system for nucleic acid folding, hybridizing and melting predictions. Masters thesis. Rensselaer Polytechnic Institute, Troy, NY.

Mohan, S. E., Y. Hu, B. Ewen-Campen, B. E. Housden, R. Viswanatha et al., 2016 CRISPR guide RNA design for research applications. FEBS J. 283: 3232–3238. https://doi.org/10.1111/febs.13777

Norris, A. D., H. M. Kim, M. P. Colaiácovo, and J. A. Calarco, 2015 Efficient genome editing in C. elegans with a toolkit of dual-marker selection cassettes. Genetics 201: 449–458. https://doi.org/10.1534/genetics.115.180679

Paix, A., Y. Wang, H. E. Smith, C. Y. Lee, D. Calidas et al., 2014 Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 sites in Caenorhabditis elegans. Genetics 198: 1347–1356. https://doi.org/10.1534/genetics.114.170423

Paix, A., H. Schmidt, and G. Seydoux, 2015 Cas9-assisted recombineering in C. elegans: genome engineering using in vivo assembly of linear DNAs. Nucleic Acids Res. 44: e128. https://doi.org/10.1093/nar/gkw302

Port, F., H. M. Chen, T. Lee, and S. L. Bullock, 2014 Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc. Natl. Acad. Sci. 111: E2967–E2976.

Ramakrishna, S., A. B. Kwaku Dad, J. Beloor, R. Gopalappa, S. K. Lee et al., 2014 Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res. 24: 1020–1027. https://doi.org/10.1101/gr.171264.113

Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander et al., 2011 Integrative genomics viewer. Nat. Biotechnol. 29: 24–26. https://doi.org/10.1038/nbt.1754

Sander, J. D., and K. Joung, 2014 CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32: 347–355. https://doi.org/10.1038/nbt.2842

Thorvaldsdottir, H., J. T. Robinson, and J. P. Mesirov, 2013 Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14: 178–192. https://doi.org/10.1093/bib/bbs017

Tzur, Y. B., A. E. Friedland, S. Naderanaj, G. M. Church, J. A. Calarco et al., 2013 Heritable custom genomic modifications in C. elegans via a CRISPR-Cas9 system. Genetics 195: 1181–1185. https://doi.org/10.1534/genetics.113.156075

van Schendel, R., S. F. Roerink, V. Portegijs, S. van den Heuvel, and M. Tijsterman, 2015 Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. Nat. Commun. 6: 7394. https://doi.org/10.1038/ncomms8394

Waalviers, S., V. Portegijs, J. Kerver, B. B. Lemmens, M. Tijsterman et al., 2013 CRISPR/Cas9-targeted mutagenesis in C. elegans. Genetics 195: 1187–1191. https://doi.org/10.1534/genetics.113.156299

Wu, X., A. J. Kriz, and P. A. Sharp, 2014 Target specificity of the CRISPR-Cas9 system. Quant. Biol. 2: 59–70. https://doi.org/10.1007/s40484-014-0030-x

Yang, H., H. Wang, C. S. Shivvalla, A. W. Cheng, L. Shi et al., 2013 One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 154: 1370–1379.

Zhao, P., Z. Zhang, H. Ke, Y. Yue, and D. Xue, 2014 Oligonucleotide- based targeted gene editing in C. elegans via the CRISPR/Cas9 system. Cell Res. 24: 247–250.