Molecular identification of a peroxidase gene controlling body size in the entomopathogenic nematode *Steinernema hermaphroditum*

Hillel T. Schwartz[1,4], Chieh-Hsiang Tan[1], Jackeline Peraza[2], Krystal Louise T. Raymundo[3], and Paul W. Sternberg[1,4]

[1] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[2] Department of Biology, Barnard College of Columbia University, New York, NY 10027, USA

[3] Dept. of Neuroscience, Princeton University, Princeton, NJ 08544, USA

[4] To whom correspondence should be addressed: hillels@caltech.edu and pws@caltech.edu

Running head: Cloning a *dpy* gene in an entomopathogenic nematode

Keywords: *Steinernema*, entomopathogenic nematode, CRISPR, body size, mutagenic spectrum

Abstract

The entomopathogenic nematode *Steinernema hermaphroditum* was recently rediscovered and is being developed as a genetically tractable experimental system for the study of previously unexplored biology, including parasitism of its insect hosts and mutualism with its bacterial endosymbiont *Xenorhabdus griffiniae*. Through whole-genome re-sequencing and genetic mapping we have for the first time molecularly identified the gene responsible for a mutationally defined phenotypic locus in an entomopathogenic nematode. In the process we observed an unexpected mutational spectrum following EMS mutagenesis in this species. We find that the

© The Author(s) 2023. Published by Oxford University Press on behalf of The Genetics Society of America. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com This article is published and distributed under the terms of the Oxford University Press, Standard Journals Publication Model (https://academic.oup.com/pages/standard-publication-reuse-rights)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

ortholog of the essential *C. elegans* peroxidase gene *skpo-2* controls body size and shape in *S. hermaphroditum*. We confirmed this identification by generating additional loss-of-function mutations in the gene using CRISPR-Cas9. We propose that the identification of *skpo-2* will accelerate gene targeting in other *Steinernema* entomopathogenic nematodes used commercially in pest control, as *skpo-2* is X-linked and males hemizygous for loss of its function can mate, making *skpo-2* an easily recognized and maintained marker for use in co-CRISPR.

Introduction

Entomopathogenic nematodes of the genera *Steinernema* and *Heterorhabditis* reside in the soil as developmentally arrested dispersal-stage infective juvenile (IJ) larvae (Dillman and Sternberg 2012; Schwartz 2015). Upon encountering a suitable insect host, an entomopathogenic nematode invades its body and resumes development, releasing endosymbiotic pathogenic bacteria from its intestine into its host (Dziedziech *et al.* 2020).The nematode and its bacterial symbiote rapidly kill the insect and convert the carcass into an incubator for the nematode-bacterial pair. When the carcass is exhausted of nutrients, a subsequent generation of IJs, each carrying pathogenic bacteria, disperse to begin the process anew. The entomopathogenic nematode lifecycle offers an opportunity to study the development and behavior of parasitic nematodes and their interactions with their bacterial symbiotes and their insect prey, along with other aspects of their biology shared with or differing from those described in other nematodes.

The extensively described biology of the free-living soil nematode *Caenorhabditis elegans* offers a model for establishing entomopathogenic nematodes as a tool for laboratory research. Work on *C. elegans* has provided major contributions to our understanding of development and disease (Horvitz 2003; Sulston 2003; Brenner 2003; Fire 2007; Mello 2007) in part because *C. elegans* is a small animal with a rapid generation time and reproduces by selfing

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

hermaphroditism (Apfeld and Alper 2018; Singh 2021). More recently, CRISPR-Cas9 genome editing has opened new possibilities for exploring gene function (Frøkjær-Jensen 2013).

We are developing the entomopathogenic nematode *Steinernema hermaphroditum* into a similarly tractable and powerful platform for laboratory research. This would enable research into aspects of the entomopathogenic nematode life cycle not amenable to study in previously available nematode species, such as interactions between the nematodes and their bacterial symbiotes, or specific to this nematode, such as its unusual mode of reproduction. First reported in 2000 from studies in the Moluccan islands of Indonesia, *S. hermaphroditum* was subsequently lost until its rediscovery outside New Delhi was reported in 2019 (Griffin *et al.* 2000; Stock *et al.* 2004; Bhat *et al.* 2019). We recently reported that *S. hermaphroditum* consistently reproduces as a selfing hermaphrodite, established an inbred wild-type strain and protocols for its propagation in the laboratory, and used chemical mutagenesis screens to recover mutants that we complementation tested, genetically mapped, and cryopreserved (Cao *et al.* 2022). No other entomopathogenic nematode species is known to reproduce as hermaphrodites in every generation.

The first entomopathogenic nematode mutants described had a short body size (Dumpy, or Dpy) phenotype (Zioni (Cohen-Nissan) *et al.* 1992; Rahimi *et al.* 1993; Tomalak, M. 1994). Continuing our development of *S. hermaphroditum* as a platform for laboratory exploration, we sought proof-of-principle for molecular identification of a mutationally defined locus. Through whole-genome sequencing of three independent alleles of an X-linked gene with a strong Dpy phenotype we identified *Sthm-skpo-2*, the *S. hermaphroditum* ortholog of the *C. elegans* peroxidase gene *Cel-skpo-2*, as the only mutated gene likely to be responsible for this Dpy phenotype. *Sthm-skpo-2* mutants generated using CRISPR-Cas9 phenocopied and failed to complement the existing Dpy mutants.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Materials and Methods

Nematode genetics

Steinernema hermaphroditum strains were derived from the inbred wild-type strain PS9179 and cultured with the bacterial strains *Xenorhabdus griffiniae* HGB2511 and *Comamonas sp*. DA1877 as food sources (Cao *et al.* 2022). Individual HGB2511 or DA1877 colonies were grown overnight at room temperature in 20 g/L Proteose Peptone No. 3 containing 0.1% sodium pyruvate and dispensed onto agar media in Petri plates to grow bacterial lawns as a food source for *S. hermaphroditum*. HGB2511 lawns were grown on NGM agar media as described (Cao *et al.* 2022). DA1877 lawns were grown on Enriched Peptone Plates except with 1.8% (weight/volume) agar (Evans 2006).

Caenorhabditis elegans were derived from the wild-type strain N2 and cultured on *E. coli* OP50 (Brenner 1974). Existing *C. elegans* mutants obtained for use in this study included *skpo-1(ok1640) II* and *mlt-7(tm1794) II*, along with the balancer chromosome *tmC6[dpy-2(tmIs1189)] II* (Thein *et al.* 2009; Tiller and Garsin 2014; Dejima *et al.* 2018). Existing *S. hermaphroditum* mutants used included *unc(sy1647)*, *dpy(sy1639) X*, *dpy(sy1644) X*, *dpy(sy1646) X*, *dpy(sy1662) X*, and *unc(sy1636) X* (Cao *et al.* 2022).

A genetic screen for visible phenotypic mutants of *S. hermaphroditum* was performed using ethyl methansulfonate (EMS) mutagenesis as described (Cao *et al.* 2022). A single phenotypic mutant, PS9839 *dpy(sy1926) X*, was recovered. Complementation tests were performed using *dpy(sy1926)* and the X-linked Dpy mutants *sy1646* and *sy1662*, marked with *unc(sy1636) X* to identify cross progeny.

DNA sequencing and analysis

Genomic DNA was prepared essentially as described, except without grinding of frozen animals (Emmons *et al.* 1979). Animals were grown on 10 cm Petri plates containing NGM agar with a lawn of HGB2511 bacteria. Animals were washed repeatedly in M9 buffer and digested using

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

proteinase K in the presence of SDS and beta-mercaptoethanol. Lysate was extracted with phenol/chloroform/isoamyl alcohol followed by chloroform. Nucleic acids were precipitated from the aqueous fraction using ethanol and recovered by spooling. RNA was removed by digestion with RNase A, after which DNA was recovered by ethanol precipitation. Purified DNA was sent to Novogene (Sacramento, CA) for Illumina sequencing with a target of 26.6 million paired-end 150 nt reads for each sample.

Analysis of high-throughput sequencing data was adapted from a published pipeline for *C. elegans* (Smith and Yun 2017). Sequencing reads were filtered using BBTools bbduk (http://sourceforge.net/projects/bbmap/) to remove reads matching an assembly of *X. griffiniae* HGB2511 genome sequence (Jennifer Heppert and Heidi Goodrich-Blair, personal communication). Reads were mapped to a draft annotated *S. hermaphroditum* PS9179 genome (Erich Schwarz, personal communication), reads were sorted, duplicate reads were removed, and reads were indexed using Samtools (Danecek *et al.* 2021). Mutations were detected using Freebayes (Garrison and Marth 2012) and were mapped onto gene models and categorized for coding changes using ANNOVAR (Wang *et al.* 2010). Annotated changes were sorted, compared, and counted using Excel (Microsoft, Redmond, WA).

Individual animals or small groups of animals were lysed and sequences were amplified from them using PCR as described for *C. elegans* (Wicks *et al.* 2001) using oligonucleotide primers whose sequences are listed in Table S1. Restriction enzymes were obtained from New England Biolabs (Beverly, MA). For Sanger sequencing, at least two PCR products were combined for each sample; nucleic acid was purified using QiaQuick (QIAGEN, Germantown, Maryland) and sent to Laragen for Sanger sequencing (Laragen, Culver City, CA).

Homology searches of additional *Steinernema* nematodes were performed using BLAST 2.2.24+ on a Debian GNU server (Altschul *et al.* 1990) using genome and transcriptome assemblies downloaded from the NCBI or from WormBase ParaSite (Howe *et al.* 2017); accession numbers were *Steinernema carpocapsae* GCA_000757645.3 (DNA), *Steinernema*

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

carpocapsae WBPS16 (mRNA), *Steinernema diaprepesi* GCA_013436035.1, *Steinernema feltiae* GCA_000757705.1, *Steinernema glaseri* GCA_000757755.1, *Steinernema hermaphroditum* GCA_030435675.1 (DNA and mRNA), *Steinernema khuongi* GCA_016648015.1, *Steinernema monticolum* GCA_000505645.1, and *Steinernema scapterisci* GCA_000757745.1 (Dillman *et al.* 2015; Serra *et al.* 2019; Baniya *et al.* 2019; Baniya and DiGennaro 2021). MEGA11 software (Tamura *et al.* 2021) was used to generate a neighbor-joining phylogeny of predicted peroxidases identified by a BLAST search of the *C. elegans* proteome as having significant homology to *Sthm*-SKPO-2, and the predicted *S. hermaphroditum* proteins closely related to them, using protein sequences from *C. elegans* version WS290 (Davis *et al.* 2022) and from *S. hermaphroditum* GCA_030435675.1.

Identification of candidate genes from whole-genome sequencing

To identify the mutations responsible for the Dpy phenotypes of our three sequenced strains, we first searched for genes on the X chromosome that had coding mutations in all three strains, ideally distinct mutations (each of the four alleles was descended from an independently mutagenized P_0 animal, and *sy1926* was recovered in a separate screen from the first three alleles). Mutations were expected to be single-nucleotide C-to-T changes consistent with EMS mutagenesis (Anderson 1995; Volkova *et al.* 2020). These criteria resulted in four candidates, encoding the hypothetical proteins QR680_001060, QR680_001389, QR680_001390, and QR680_002483. Further inspection suggested the latter three candidates were likely the result of sequencing and software issues: the mutations associated with these three candidates were defined by low read counts that had low quality scores. Proteins 001389 and 001390 are encoded by neighboring genes and include nearly identical sequence; these two genes have 14 different mutations annotated between them among the three strains, which did not seem consistent with the mutations having arisen after mutagenesis and being causative for the Dpy phenotype. Protein 002483 has 14 mutations annotated, of which three were annotated in more

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

than one strain; this also is not consistent with the gene having been mutated to cause the Dpy phenotype. By contrast, the gene encoding protein 001060 has only three mutations annotated among the three strains, one in each strain; all three annotations have high read counts and quality scores.

Homology analysis of candidate genes

We examined the four multiply mutated X-linked genes' homology to assess them as candidates. Predicted proteins 001389 and 001890 lack identifiable homologs by BLAST searches, with none found even in the other available *Steinernema* genomes or the *Steinernema carpocapsae* transcriptome, and lack conserved domains identifiable by SMART or by Pfam (Letunic *et al.* 2021; Mistry *et al.* 2021).

The closest characterized homolog of protein 002483 is in *C. elegans*, Cel-HGRS-1; protein 002483 is also the predicted *S. hermaphroditum* protein most closely related to Cel-HGRS-1. RNAi-mediated inactivation of *Cel-hgrs-1* causes a Dpy phenotype and other defects (Kamath *et al.* 2003). This reported Dpy phenotype nominated it as a viable candidate for the Dpy phenotypes of our *S. hermaphroditum* mutants, despite the low read count and the poor quality scores of the sequence data implicating this gene.

The last of the four candidates is the gene encoding protein 001060, orthologous to *C. elegans* Cel-SKPO-2, predicted to encode a peroxidase (see Figure 1B). *Cel-skpo-2* does not have a reported abnormal mutant phenotype, but it is closely related to *Cel-mlt-7*, loss of which causes defects in cuticle formation and molting along with nearly fully penetrant lethality and a Dpy phenotype in the survivors (Figure 1C). Protein 001060 is more distantly related to the product of *Cel-bli-3*, which mutates to cause a blistered cuticle defect; *bli-3* functions with *mlt-7* to regulate cuticle structure, and other blister mutants genetically interact with cuticular Dumpy phenotypes (Higgins and Hirsh 1977; Cox *et al.* 1980; Simmer *et al.* 2003; Thein *et al.* 2009). This homology implicated the gene encoding protein 001060 as a strong candidate.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Genetic linkage of *dpy(sy1926)* to *Sthm-skpo-2*

One of the three molecularly identified mutations in *Sthm-skpo-2*, the mutation in PS9839, disrupts a locally unique endogenous restriction site (FokI: GGATG). We used this restriction site to assess linkage between the *Sthm-skpo-2* locus and the Dpy phenotype of PS9839: 0/118 Dpy self-progeny of *dpy(sy1926)*/+ heterozygotes contained wild-type sequence at *Sthm-skpo-2*, indicating extremely tight linkage, within 2 map units ($p<0.001$). The *Cel-hgrs-1*-homologous gene encoding protein 002483 is seven million base pairs from *Sthm-skpo-2*, on a chromosome of approximately 18.4 million base pairs; tight linkage of the Dpy phenotype with *Sthm-skpo-2* is inconsistent with the causative mutations being in the gene encoding protein 002483, leaving *Sthm-skpo-2* as the only strong candidate unanimously identified by sequencing, homology, and genetic linkage.

CRISPR-Cas9

CRISPR-Cas9 targeting *skpo-2* in *C. elegans* was performed as described (Arribere *et al.* 2014) using *dpy-10* and *unc-58* co-conversion markers to obtain the two alleles *sy2121* and *sy2122*, respectively. In both cases the co-conversion marker was lost and the mutation was balanced using *tmC6[dpy-2(tmIs1189)]*. The guide RNA used contained the *C. elegans* genomic sequence CCCCAACATCGACCCATCTG, targeting cleavage at codon 480, in exon 11. An oligonucleotide with the sequence CATCGGCGCCTACCCAGGCTATGACCCCAACATCGACCCATgggaagtttgtccagagcagaggtgact aagtgataagctagcCTGTGGCCAACGAGTTCACATCGTGCGCGTTCCGTTTTGG was included as a template for homology-directed repair of the double-strand break in *Cel-skpo-2* repair, including a STOP-IN cassette, in lowercase (Wang *et al.* 2018). Homology-directed repair was confirmed by Sanger sequencing. The *C. elegans* CRISPR protocol including its injection mixture was adapted for use in *Steinernema* with the exceptions that there was no

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

co-conversion marker used and the injection mix was 1/10 Lipofectamine RNAiMAX by volume (Invitrogen, Waltham, MA) according to a protocol modification reported to be helpful in *Auanema* (Adams *et al.* 2019): 1.35 µL each of 100 µM tracrRNA and 100 µM crRNA were combined and heated at 94°C for two minutes and allowed to cool at room temperature; 1 µL 1M KCl and 2 µL 10µg/mL Cas9 protein were added and the mixture was incubated for five minutes at room temperature; 0.6 µL 10 µM repair oligo and 0.7 µL Lipofectamine were added and the mixture was incubated for twenty minutes at room temperature; immediately after this incubation the mixture was used to load injection needles and treat animals. The guide RNA used included the *S. hermaphroditum* sequence GCACCCGAGGAAGGTACTCG, targeting cleavage at codon 447 in exon 8, and a repair template oligonucleotide with the sequence CATCGGCGCCTACCCAGGCTATGACCCCAACATCGACCCATgggaagtttgtccagagcagaggtgact aagtgataagctagcCTGTGGCCAACGAGTTCACATCGTGCGCGTTCCGTTTTGG, containing a STOP-IN cassette shown in lowercase, was included in the injection mix. Animals used for CRISPR-Cas9 genome editing were grown on *Comamonas* DA1877 in preference to *Xenorhabdus* HGB2511. CRISPR reagents were injected into the gonad syncytia of day-old adult hermaphrodites.

In the first CRISPR experiment targeting *Sthm-skpo-2*, three of eleven injected P_0 animals survived to give progeny; survival and recovery of *S. hermaphroditum* after injection is thus far considerably worse than is seen for injection of *C. elegans*. 78 F_1 progeny of these animals were moved to new Petri plates, from one to four F_1s per plate. Phenotypically Dpy F_1 and F_2 progeny of P_0 animals injected with CRISPR reagents were recovered and used to establish clonal lines; these clonal lines were composed entirely of healthy, fertile animals with a strong and consistent Dpy phenotype that was stable for at least ten generations. The clonal lines were genotyped by PCR using the oligonucleotides GACGTGTGTTTCCTCCCGT and GCATCTTAGCCGGGAGACT followed by restriction digest with RsaI to detect changes at the CRISPR cleavage site and with NheI seeking evidence that the oligonucleotide template had

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

been used as a template for homology-directed repair. Two F_1 animals were Dpy hermaphrodites; F_2 animals from each of these were placed singly on Petri plates to establish subclones that could segregate CRISPR-induced mutant alleles that might be present if the Dpy F_1 were a *trans*-heterozygote of two different *Sthm-skpo-2* alleles. One of the Dpy F_1s contained two molecularly distinguishable mutant alleles of *Sthm-skpo-2* (*sy2106* and *sy2107*), the other was apparently homozygous for the allele *sy2105* (8/8 progeny contained only the *sy2015* allele by PCR and sequencing). Three more alleles (*sy2107*, *sy2108*, and *sy2120*) were identified as the Dpy self-progeny of nonDpy F_1 progeny of P_0 animals injected with CRISPR reagents. The sixth CRISPR allele of *Sthm-skpo-2* was recovered as the Dpy self-progeny of nonDpy F_1 progeny of P_0 animals injected with CRISPR reagents in a second round of injections.

A complementation test was performed using males of a representative CRISPR-generated Dpy mutant, *Sthm-skpo-2(sy2108)*, mated to hermaphrodites homozygous for *dpy(sy1644) X* and the autosomal mutation *unc(sy1647)*, which was used to distinguished self-progeny from cross-progeny.

RNAi

HT115 bacteria containing plasmids derived from L4440 for the expression of dsRNA corresponding to the *C. elegans* genes *mlt-7*, *skpo-1*, and *skpo-2* were obtained from a library initially generated by the laboratory of Dr. Julie Ahringer and distributed by Source BioScience (San Diego, CA), and used to perform RNAi experiments as described (Kamath *et al.* 2003). Plasmid inserts were confirmed by Sanger DNA sequencing. Individual colonies were grown in LB media containing carbenicillin and tetracycline and used to seed lawns on NGM agar containing carbenicillin and IPTG. Fourth-stage (L4) larval hermaphrodites were placed on these lawns and their progeny were examined for abnormal phenotypes for two generations. As has been previously reported (Kamath *et al.* 2003), no effect was seen from feeding dsRNA targeting *skpo-2* or *skpo-1*, while animals that grew up feeding on *mlt-7* dsRNA often displayed

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

a molting defect, with blistering or an unshed cuticle, and often were defective in locomotion, possibly as an effect of the molting defect.

Microscopy

Images were acquired with a Zeiss Imager Z2 microscope equipped with an Apotome 2 and Axiocam 506 mono using Zen 2 Blue software (Zeiss, White Plains, NY). Animals were immobilized with 1 mM levamisole in M9 buffer or with 30 mM sodium azide in S basal and mounted on 2% or 4% agarose pads (for *S. hermaphroditum* and *C. elegans*, respectively) on microscope slides for imaging.

Chemical analysis

Gas chromatography-mass spectrometry of an EMS sample was performed by Dr. Mona Shagholi of the Caltech Mass Spectrometry service center to confirm its molecular identity. Samples were analyzed by Field Ionization using a JEOL AccuTOF GC-Alpha (JMS-T2000GC) mass spectrometer (JEOL USA, Peabody, MA) interfaced with an Agilent 8890 gas chromatograph (Agilent, Santa Clara, CA). The gas chromatograph was fitted with a Restek Rxi-5ms column (30 m x 0.25 mm ID, 0.25 micron df) (Restek Corporation, Bellefonte, PA). The temperature gradient was started at 50°C, held for 1 minute, then ramped at 32°C/minute to 300°C with another 1-minute hold. The desired species eluted at 3.7 minutes and was detected as a radical cation.

Nomenclature

Genes and proteins in the species *Caenorhabditis elegans* (ToLID: NrCaeEleg) are identified with the prefix "Cel-". Genes and proteins in the species *Steinernema hermaphroditum* (ToLID: NxSteHerm) are identified with the prefix "Sthm-". See http://id.tol.sanger.ac.uk for more on ToLID identifiers.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Results

Molecular identification of a *Steinernema hermaphroditum dpy* gene by screens and sequencing

The first chemical mutagenesis screens in *S. hermaphroditum* used ethyl methanesulfonate (EMS) to recover 32 independent mutant strains with visible phenotypes such as uncoordinated (Unc) or dumpy (Dpy) (Cao *et al.* 2022). Three X-linked mutations – PS9260 *dpy(sy1639),* PS9265 *dpy(sy1644)*, and PS9267 *dpy(sy1646)* – caused an identical Dumpy (Dpy) phenotype (Figure 1A) and failed to complement each other. Another X-linked mutation with a similar phenotype, *dpy(sy1662)*, complemented these mutations, indicating that *sy1639, sy1644*, and *sy1646* are in one complementation group and *sy1662* is in another.

Additional EMS mutagenesis screens recovered one mutant, PS9839 *dpy(sy1926)*, with an indistinguishable Dpy phenotype. *dpy(sy1926)* was also X-linked and failed to complement *dpy(sy1646)* but did complement *dpy(sy1662)*, indicating *sy1926* was a fourth member of the complementation group containing *sy1639*, *sy1644*, and *sy1646*. We sequenced the genomes of three of these four allelic mutants: PS9260 *dpy(sy1639)*, PS9267 *dpy(sy1646)*, and PS9839 *dpy(sy1926)*. After filtering reads for bacterial contamination, mapping reads to a draft *S. hermaphroditum* genome, and removing duplicate reads, we had 18.8x, 34.3x, and 46.0x genome coverage of these mutants, respectively. As detailed in the Materials and Methods we identified candidate genes, examined sequence quality, identified homologous genes in *C. elegans*, and demonstrated genetic linkage to determine that the Dpy phenotypes of these mutants were likely caused by mutations in a gene we named *Sthm-skpo-2*, the ortholog of *C. elegans Cel-skpo-2* (see Figure 1B).

Sanger sequencing confirmed the three mutations in *Sthm-skpo-2*: PS9260 has a three-nucleotide deletion removing amino acid R503, PS9267 has a single-nucleotide G-to-T change causing the predicted coding change E469ochre; and PS9839 has a single-nucleotide

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

G-to-T change causing the predicted change C178F (Figures 1D and S1A). Attempts to identify a coding change in the fourth allelic mutant, PS9265 *dpy(sy1644)*, which was not selected for whole-genome sequencing, demonstrated that the ninth exon of *Sthm-skpo-2* could not be amplified using PCR primers that reliably amplified this sequence from the wild type, indicating a large deletion, insertion, or other rearrangement in this region of *Sthm-skpo-2*.

Targeted inactivation of *Sthm-skpo-2* using CRISPR-Cas9 causes a Dpy phenotype

To confirm that loss of *Sthm-skpo-2* function causes a Dpy phenotype, we used CRISPR-Cas9 to knock out *Sthm-skpo-2*. A guide RNA was chosen to induce double-strand breaks within 65 nucleotides of the ochre stop mutation *sy1646* in PS9267. Six mutations were identified following CRISPR-Cas9 injection; each caused a stable, fully penetrant, healthy Dpy phenotype. We confirmed our gene identification by complementation testing between the CRISPR-induced Dpy mutant *Sthm-skpo-2(sy2108)* and *dpy(sy1644)*.

All six CRISPR alleles caused changes at the targeted site likely to disrupt gene function: a genomic abnormality that prevented PCR of the *skpo-2* locus (*sy2107*, *sy2108*, and *sy2120*) or an alteration identified using Sanger sequencing (*sy2015*, *sy2106*, and *sy2123*; Figure S1B). Although an oligonucleotide donor was included as a template for homology-directed repair, the induced lesions were consistent with non-homologous end joining (NHEJ) (4/6 lesions) or microhomology-mediated end joining (MMEJ) (Figure S1B). Insertion of a STOP-IN cassette (Wang *et al.* 2018) at the corresponding site of *Cel-skpo-2* caused fully penetrant recessive embryonic lethality in *C. elegans*; *trans*-heterozygotes between this lethal null mutation in *Cel-skpo-2* and the nearly lethal mutation *tm1794* in the closely related gene *Cel-mlt-7* were grossly wild type. Growth on bacteria expressing dsRNA targeting *skpo-2* for RNAi caused no apparent phenotypic defects.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

EMS mutagenesis of *S. hermaphroditum* induced a mutational spectrum consistent with double-strand breaks

Of the four EMS-induced mutations in *Sthm-skpo-2*, none were consistent with the expected mutational spectrum of EMS, which causes 95% single-nucleotide C-to-T substitutions in organisms ranging from *C. elegans* to the flowering plant *Arabidopsis thaliana* (Pastink *et al.* 1991; Greene *et al.* 2003; Volkova *et al.* 2020). Examination of all annotated sequence changes in each strain, focusing on changes that were unique to each strain and should therefore have arisen subsequent to mutagenesis, showed that their single-nucleotide substitutions comprised roughly equal numbers of each nucleotide change (Figure 2A). Single-nucleotide substitutions accounted for over half of annotated mutations; approximately 15% were multi-nucleotide changes, and nearly all of the remainder were single-nucleotide insertions. The single-nucleotide insertions were almost exclusively found in noncoding sequences; as noncoding sequence is enriched for mononucleotide repeats, annotated single-nucleotide insertions might include spurious reports from the software used to detect mutations. The observed sequence changes were consistent with mutagenesis using neither EMS nor any other chemical mutagen causing single-nucleotide changes well characterized for mutational spectrum in nematodes, but were consistent with the mutations observed following the induction of double-strand breaks (Volkova *et al.* 2020). The mutations tended to appear in clusters rather than distributed evenly across the chromosome (Figure 2B). Mass spectrometry confirmed the chemical used to mutagenize was EMS, but as discussed below there are possible explanations for how EMS treatment could induce double-strand breaks instead of causing C-to-T single-nucleotide transitions.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Discussion

Molecular identification of a mutationally defined locus in an entomopathogenic nematode

Caenorhabditis elegans has been a uniquely powerful species for the use of experimental genetics to explore an animal's biology (Horvitz and Sulston 1990; Meneely *et al.* 2019), providing a model for other nematode species with interestingly different biology that are similarly amenable to laboratory experimentation. Entomopathogenic nematodes offer experimental access to biology not previously extensively explored in the laboratory. The relationships of entomopathogenic nematodes with their insect prey and with their pathogenic bacterial endosymbionts lack endogenous parallels in *C. elegans* or other established nematode experimental systems, such as *Pristionchus* nematodes, which have a necromenic instead of pathogenic relationship with their insect hosts, and plant-pathogenic *Bursaphelenchus* nematodes that have a commensal relationship with insects that act as their vector (Sommer and McGaughran 2013; Félix *et al.* 2018; Kirino *et al.* 2023). Our finding that the peroxidase gene *Sthm-skpo-2* is required for normal body size and shape in *S. hermaphroditum* is the first molecular identification of the gene responsible for a mutant phenotype in an entomopathogenic nematode, and only the second time this has been reported for any clade IV nematode, following the recent molecular identification of a *tra-1* homolog mutated in sex determination mutants of *Bursaphelenchus okinawaensis* (Shinya *et al.* 2022).

Sthm-skpo-2 encodes a predicted peroxidase orthologous to *C. elegans* SKPO-2. *skpo-2* is among several nematode genes closely related to the human peroxidasin PXDN (Thein *et al.* 2009), shown to crosslink collagen and regulate the structure of the endothelial basement membrane (Cheng *et al.* 2008; Bhave *et al.* 2012); peroxidase genes of *C. elegans* modify cuticle collagen structure and permeability, and many genes that impact the Dpy phenotype in *C. elegans* encode collagens or proteins known to modify collagens (Edens *et al.* 2001; Myllyharju and Kivirikko 2004; Thein *et al.* 2009). It is likely that the Dpy phenotype of our

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Sthm-skpo-2 mutants is the result of altered collagen structure, arising from an evolutionarily conserved role of peroxidases in modifying collagens.

Loss of function of *Cel-mlt-7*, a paralog of *Cel-skpo-2*, causes a Dpy phenotype resembling *Sthm-skpo-2* mutants – except that *Cel-mlt-7* mutants nearly all die during development, with rare survivors being Dpy, sick, and uncoordinated (Thein *et al.* 2009). The *S. hermaphroditum* ortholog of *Cel-mlt-7* is distinct from *Sthm-skpo-2* (Figure 1B). The related but differing effects of mutating different peroxidasin homologs in different nematodes reflects an established theme: although all nematodes share a highly similar body plan, the functions of orthologous genes can differ significantly even within a genus (Wang and Chamberlin 2002; Félix 2007; Mahalak *et al.* 2017). Even if a particular phenotype has been extensively studied in one nematode, studies in a distantly related species may identify genes that cannot easily mutate to cause the phenotype in *C. elegans*, for example because of redundancy, because pleiotropies prevent recovery of such mutants, or because different genes are involved in the two species. In *Steinernema* we found viable, healthy Dpy mutants from a gene class known to be capable of modifying the cuticle, but that in *C. elegans* lacks similarly healthy Dpy mutants; unlike related *C. elegans* peroxidases, *Sthm-skpo-2* mutants could be studied to understand how these peroxidases affect cuticle structure and animal shape.

An unexpected mutational spectrum following ethyl methanesulfonate mutagenesis

Our *Sthm-skpo-2* mutations were not the single-nucleotide C-to-T changes expected from EMS mutagenesis. Examining thousands of unique mutations in each sequenced strain revealed a mutational spectrum resembling those seen when double-strand breaks are induced rather than the spectrum normally expected following EMS mutagenesis; the mutational spectrum also did not resemble those of other chemicals causing single-nucleotide changes, nor the observed effects of EMS mutagenesis on animals mutant for selected genes involved in DNA repair (Volkova *et al.* 2020). The entomopathogenic nematode *Heterorhabditis bacteriophora*, similar

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

in ecological niche but not closely related to *S. hermaphroditum*, also has a mildly divergent spectrum of EMS-associated mutations: 80% C-to-T, versus 95% in *C. elegans* (Wang *et al.* 2023).

Our EMS mutagenesis screens rapidly produced dozens of stably phenotypic mutants, but phenotypic mutants were not found from observing vastly more animals in the absence of a chemical mutagen in the course of mapping, complementation testing, and cryopreserving our mutant collection. Treatment with EMS must therefore have induced the genetic changes we detected, even though those changes did not conform to the mutational spectrum expected from EMS mutagenesis.

While EMS is noted for its ability to cause mutations distributed across the genome evenly (Figure 2C and Thompson *et al.* 2013), we recovered 33 phenotypic mutations of which four were alleles of one gene and two were alleles of another (Cao *et al.* 2022); selected loci were apparently highly susceptible to our mutagenesis. The vast majority of identified mutations were tightly clustered at a few positions on each chromosome; these were the same positions in all three sequenced strains, and were also the sites of mutations found in common among the three strains, that must have predated mutagenesis. Between this clustering of mutations and the mutagenic spectrum we observed, we hypothesize that our mutagenesis induced double-strand breaks, often at sensitive loci, rather than evenly distributed single-nucleotide C-to-T transitions. An increased incidence of EMS induction of deletion mutations, presumed to be secondary to double-strand breaks, has previously been reported when an increased concentration of EMS was used to mutagenize *C. elegans* (Lesa 2006), suggesting that different dose-responses to EMS might explain our results; alternatively, EMS mutagenesis of cells in a state of cell-cycle arrest could cause double-strand breaks instead of C-to-T transitions. EMS mutagenesis normally converts cytosine to thymine when a guanine residue modified by EMS to become O_6-ethylguanine is misread during DNA replication and paired with thymine instead of cytosine (Sega 1984). If DNA replication were halted – if for example DNA checkpoint activity

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

were different in *S. hermaphroditum* from in *C. elegans*, or if the two species respond differently to incubation in M9 buffer in the absence of food during mutagenesis – then the DNA base modifications caused by exposure to EMS would not rapidly be resolved to induce C-to-T single-nucleotide conversions, and might instead be repaired by error-prone nucleotide excision repair. Alternatively, an accumulation of modified residues might trigger stalling of DNA replication, followed by error-prone translesion repair of the clustered changes, or single-strand or double-strand breaks, whose resolution could result in a mutagenic spectrum characteristic of double-strand break repair (Kondo *et al.* 2010; Schärer 2013; Khatib *et al.* 2023). Further investigation of this difference between EMS mutagenesis of *S. hermaphroditum* and other nematodes should improve our ability to perform genetic screens probing the unique biology of this entomopathogenic nematode and may provide an opportunity to examine the basis of the differing effects of EMS mutagenesis on these different species.

Prospects for CRISPR-Cas9 gene targeting in *Steinernema*

We confirmed our identification of *Sthm-skpo-2* using CRISPR-Cas9, generating new *Sthm-skpo-2* mutants with an identical phenotype: healthy animals with a dumpy body shape. The resulting alleles were consistent with NHEJ and MMEJ double-strand break repair (Xue and Greene 2021); we saw no evidence of homology-directed repair using the oligonucleotide template we included. CRISPR-Cas9 is used extensively in established laboratory research animals and is becoming a powerful tool in a growing variety of nematodes new to intensive laboratory research (Mendez *et al.* 2022; Cadd *et al.* 2022; Hellekes *et al.* 2023; Dutta *et al.* 2023); with our work and that of Cao (Cao 2023) CRISPR-Cas9 has been extended to entomopathogenic nematodes.

The identification of *skpo-2* could facilitate CRISPR in other *Steinernema* entomopathogens. All other *Steinernema* species whose reproduction has been described are diecious (Hunt and Nguyen 2016). Any mutant phenotype used in these other *Steinernema* species must not

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

interfere with mating ability. Visible phenotypic markers are important in CRISPR-Cas9 genome editing: individual *C. elegans* that show the phenotypic consequences of CRISPR-mediated genome modification at one locus are highly enriched for additional genome modifications at other sites simultaneously targeted using CRISPR (Kim *et al.* 2014). Variations of this method, called co-CRISPR and co-conversion, have been transformative for the efficiency of CRISPR-mediated genome modification in *C. elegans* and other nematodes (Arribere *et al.* 2014; Cohen and Sternberg 2019; Choi and Villeneuve 2023). The *skpo-2* locus is well suited to serve as a marker in divergent *Steinernema* species: we readily recovered *Sthm-skpo-2* mutants using their easily recognized phenotype, and hemizygous *Sthm-skpo-2* Dumpy males are healthy and mate well. There is one *skpo-2* gene in each of the seven sequenced *Steinernema* species (Dillman *et al.* 2015; Serra *et al.* 2019; Baniya *et al.* 2019; Baniya and DiGennaro 2021). *skpo-2* is X-linked in *S. hermaphroditum* and *S. carpocapsae*, and likely in all *Steinernema* species. If the mutant phenotype is conserved, newly arising *skpo-2* mutant males in the first generation after CRISPR treatment should be Dumpy and fully capable of mating, and will transmit any mutations they simultaneously carry in other loci targeted using CRISPR. The *skpo-2* locus should therefore make it efficient to modify the genomes of *Steinernema* species used commercially in agriculture to control insect pests (Karabörklü *et al.* 2017; Poinar 2018) and to probe the genomes of *Steinernema* species that possess novel biological abilities not yet observed or described in *S. hermaphroditum*, such as the abilities of *S. carpocapsae* to leap into the air and to secrete venom proteins (Campbell and Kaya 1999; Lu *et al.* 2017; Dillman *et al.* 2021).

Strain and data availability

Strains are available upon request. Sequence data have been deposited in the NCBI Short Read Archive as part of NCBI BioProject PRJNA1037740.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Acknowledgments

We thank Erich Schwarz for generously providing early access to unpublished versions of the *S. hermaphroditum* genome and its annotation; Jennifer Heppert and Heidi Goodrich-Blair for unpublished HGB2511 sequence; Heenam Park and Tsui-Fen Chou for CRISPR reagents and advice; Mengyi Cao for information about *Steinernema* CRISPR; Barbara Perry, Wilber Palma, and Stephanie Nava for technical assistance; WormBase and WormBase ParaSite for *C. elegans* and *Steinernema* genome information; Mona Shagholi of the Caltech Mass Spectrometry service center; and Daniel Semlow and Anton Gartner for advice about the effects of EMS mutagenesis. Some strains were provided by the CGC, funded by P40 OD010440.

Funding

This work was supported by NSF-EDGE grant 2128267 (to PWS) and Caltech's Center for Evolutionary Science (CES) and Center for Environmental Microbial Interactions (CEMI). This research benefited from the use of instrumentation made available by the Caltech CCE Multiuser Mass Spectrometry Laboratory, enabled by funds from DOW Next Generation Instrumentation.

References

Adams S., P. Pathak, H. Shao, J. B. Lok, and A. Pires-daSilva, 2019 Liposome-based transfection enhances RNAi and CRISPR-mediated mutagenesis in non-model nematode systems. Sci Rep 9: 483. https://doi.org/10.1038/s41598-018-37036-1

Altschul S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990 Basic local alignment search tool. Journal of Molecular Biology 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anderson P., 1995 Chapter 2 Mutagenesis, pp. 31–58 in *Methods in Cell Biology*, *Caenorhabditis elegans*: Modern Biological Analysis of an Organism. edited by Epstein H. F., Shakes D. C. Academic Press.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Apfeld J., and S. Alper, 2018 What Can We Learn About Human Disease from the Nematode *C. elegans*?, pp. 53–75 in *Disease Gene Identification: Methods and Protocols*, Methods in Molecular Biology. edited by DiStefano J. K. Springer, New York, NY.

Arribere J. A., R. T. Bell, B. X. H. Fu, K. L. Artiles, P. S. Hartman, *et al.*, 2014 Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. Genetics 198: 837–846. https://doi.org/10.1534/genetics.114.169730

Baniya A., J. C. Huguet-Tapia, and P. DiGennaro, 2019 A draft genome of *Steinernema diaprepesi*. Journal of Nematology 52: 1–4. https://doi.org/10.21307/jofnem-2020-069

Baniya A., and P. DiGennaro, 2021 Genome announcement of *Steinernema khuongi* and its associated symbiont from Florida. G3 Genes|Genomes|Genetics 11: jkab053. https://doi.org/10.1093/g3journal/jkab053

Bhat A. H., A. K. Chaubey, E. Shokoohi, and P. William Mashela, 2019 Study of *Steinernema hermaphroditum* (Nematoda, Rhabditida), from the West Uttar Pradesh, India. Acta Parasit. 64: 720–737. https://doi.org/10.2478/s11686-019-00061-9

Bhave G., C. F. Cummings, R. M. Vanacore, C. Kumagai-Cresse, I. A. Ero-Tolliver, *et al.*, 2012 Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. Nat Chem Biol 8: 784–790. https://doi.org/10.1038/nchembio.1038

Brenner S., 1974 The genetics of *Caenorhabditis elegans*. Genetics 77: 71–94. https://doi.org/10.1093/genetics/77.1.71

Brenner S., 2003 Nobel Lecture: Nature's Gift to Science. Bioscience Reports 23: 225–237. https://doi.org/10.1023/B:BIRE.0000019186.48208.f3

Cadd L. C., B. Crooks, N. J. Marks, A. G. Maule, A. Mousley, *et al.*, 2022 The *Strongyloides* bioassay toolbox: A unique opportunity to accelerate functional biology for nematode parasites. Molecular and Biochemical Parasitology 252: 111526. https://doi.org/10.1016/j.molbiopara.2022.111526

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Campbell J. F., and H. K. Kaya, 1999 Mechanism, kinematic performance, and fitness consequences of jumping behavior in entomopathogenic nematodes (*Steinernema* spp.). Can. J. Zool. 77: 1947–1955. https://doi.org/10.1139/z99-178

Cao M., H. T. Schwartz, C.-H. Tan, and P. W. Sternberg, 2022 The entomopathogenic nematode *Steinernema hermaphroditum* is a self-fertilizing hermaphrodite and a genetically tractable system for the study of parasitic and mutualistic symbiosis. Genetics 220: iyab170. https://doi.org/10.1093/genetics/iyab170

Cao M., 2023 CRISPR-Cas9 genome editing in *Steinernema* entomopathogenic nematodes. bioRxiv https://doi.org/10.1101/2023.11.24.568619

Cheng G., J. C. Salerno, Z. Cao, P. J. Pagano, and J. D. Lambeth, 2008 Identification and characterization of VPO1, a new animal heme-containing peroxidase. Free Radic Biol Med 45: 1682–1694. https://doi.org/10.1016/j.freeradbiomed.2008.09.009

Choi C. P., and A. M. Villeneuve, 2023 CRISPR/Cas9 mediated genome editing of *Caenorhabditis nigoni* using the conserved *dpy-10* co-conversion marker. microPublication Biology. https://doi.org/10.17912/micropub.biology.000937

Cohen S. M., and P. W. Sternberg, 2019 Genome editing of *Caenorhabditis briggsae* using CRISPR/Cas9 co-conversion marker *dpy-10*. microPublication Biology. https://doi.org/10.17912/micropub.biology.000171

Cox G. N., J. S. Laufer, M. Kusch, and R. S. Edgar, 1980 Genetic and phenotypic characterization of roller mutants of *Caenorhabditis elegans*. Genetics 95: 317–339. https://doi.org/10.1093/genetics/95.2.317

Danecek P., J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, *et al.*, 2021 Twelve years of SAMtools and BCFtools. GigaScience 10: giab008. https://doi.org/10.1093/gigascience/giab008

Davis P., M. Zarowiecki, V. Arnaboldi, A. Becerra, S. Cain, *et al.*, 2022 WormBase in 2022—data, processes, and tools for analyzing *Caenorhabditis elegans*. Genetics 220: iyac003. https://doi.org/10.1093/genetics/iyac003

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Dejima K., S. Hori, S. Iwata, Y. Suehiro, S. Yoshina, *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

Dillman A. R., and P. W. Sternberg, 2012 Entomopathogenic Nematodes. Curr Biol 22: R430–R431. https://doi.org/10.1016/j.cub.2012.03.047

Dillman A. R., M. Macchietto, C. F. Porter, A. Rogers, B. Williams, *et al.*, 2015 Comparative genomics of *Steinernema* reveals deeply conserved gene regulatory networks. Genome Biology 16: 200. https://doi.org/10.1186/s13059-015-0746-6

Dillman A. R., W. Korff, M. H. Dickinson, and P. W. Sternberg, 2021 *Steinernema carpocapsae* jumps with greater velocity and acceleration than previously reported. microPublication Biology. https://doi.org/10.17912/micropub.biology.000435

Dutta T. K., S. Ray, and V. Phani, 2023 The status of the CRISPR/Cas9 research in plant–nematode interactions. Planta 258: 103. https://doi.org/10.1007/s00425-023-04259-0

Dziedziech A., S. Shivankar, and U. Theopold, 2020 High-Resolution Infection Kinetics of Entomopathogenic Nematodes Entering *Drosophila melanogaster*. Insects 11: 60. https://doi.org/10.3390/insects11010060

Edens W. A., L. Sharling, G. Cheng, R. Shapira, J. M. Kinkade, *et al.*, 2001 Tyrosine cross-linking of extracellular matrix is catalyzed by Duox, a multidomain oxidase/peroxidase with homology to the phagocyte oxidase subunit gp91*phox*. The Journal of Cell Biology 154: 879–892. https://doi.org/10.1083/jcb.200103132

Emmons S. W., M. R. Klass, and D. Hirsh, 1979 Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans*. Proc Natl Acad Sci U S A 76: 1333–1337.

Evans T., 2006 Transformation and microinjection. WormBook. https://doi.org/10.1895/wormbook.1.108.1

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Félix M.-A., 2007 Cryptic Quantitative Evolution of the Vulva Intercellular Signaling Network in *Caenorhabditis*. Current Biology 17: 103–114. https://doi.org/10.1016/j.cub.2006.12.024

Félix M.-A., M. Ailion, J.-C. Hsu, A. Richaud, and J. Wang, 2018 *Pristionchus* nematodes occur frequently in diverse rotting vegetal substrates and are not exclusively necromenic, while *Panagrellus redivivoides* is found specifically in rotting fruits. PLOS ONE 13: e0200851. https://doi.org/10.1371/journal.pone.0200851

Fire A. Z., 2007 Gene Silencing by Double-Stranded RNA (Nobel Lecture). Angewandte Chemie International Edition 46: 6966–6984. https://doi.org/10.1002/anie.200701979

Frøkjær-Jensen C., 2013 Exciting Prospects for Precise Engineering of *Caenorhabditis elegans* Genomes with CRISPR/Cas9. Genetics 195: 635–642. https://doi.org/10.1534/genetics.113.156521

Garrison E., and G. Marth, 2012 Haplotype-based variant detection from short-read sequencing

Greene E. A., C. A. Codomo, N. E. Taylor, J. G. Henikoff, B. J. Till, *et al.*, 2003 Spectrum of Chemically Induced Mutations From a Large-Scale Reverse-Genetic Screen in *Arabidopsis*. Genetics 164: 731–740. https://doi.org/10.1093/genetics/164.2.731

Griffin C. T., R. Chaerani, D. Fallon, A. P. Reid, and M. J. Downes, 2000 Occurrence and distribution of the entomopathogenic nematodes *Steinernema* spp. and *Heterorhabditis indica* in Indonesia. J. Helminthol. 74: 143–150. https://doi.org/10.1017/S0022149X00000196

Hellekes V., D. Claus, J. Seiler, F. Illner, P. H. Schiffer, *et al.*, 2023 CRISPR/Cas9 mediated gene editing in non-model nematode *Panagrolaimus* sp. PS1159. Frontiers in Genome Editing 5. https://doi.org/10.3389/fgeed.2023.1078359

Higgins B. J., and D. Hirsh, 1977 Roller mutants of the nematode *Caenorhabditis elegans*. Molec. Gen. Genet. 150: 63–72. https://doi.org/10.1007/BF02425326

Horvitz H. R., and J. E. Sulston, 1990 “Joy of the worm”. Genetics 126: 287–292. https://doi.org/10.1093/genetics/126.2.287

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Horvitz H. R., 2003 Worms, Life, and Death (Nobel Lecture). ChemBioChem 4: 697–711. https://doi.org/10.1002/cbic.200300614

Howe K. L., B. J. Bolt, M. Shafie, P. Kersey, and M. Berriman, 2017 WormBase ParaSite − a comprehensive resource for helminth genomics. Molecular and Biochemical Parasitology 215: 2–10. https://doi.org/10.1016/j.molbiopara.2016.11.005

Hunt D. J., and K. B. Nguyen, 2016 *Advances in Entomopathogenic Nematode Taxonomy and Phylogeny*. Brill.

Kamath R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, *et al.*, 2003 Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421: 231–237. https://doi.org/10.1038/nature01278

Karabörklü S., U. Azizoglu, and Z. B. Azizoglu, 2017 Recombinant entomopathogenic agents: a review of biotechnological approaches to pest insect control. World J Microbiol Biotechnol 34: 14. https://doi.org/10.1007/s11274-017-2397-0

Khatib J. B., C. M. Nicolae, and G.-L. Moldovan, 2023 Role of Translesion DNA Synthesis in the Metabolism of Replication-associated Nascent Strand Gaps. Journal of Molecular Biology 168275. https://doi.org/10.1016/j.jmb.2023.168275

Kim H., T. Ishidate, K. S. Ghanta, M. Seth, D. Conte Jr., *et al.*, 2014 A Co-CRISPR Strategy for Efficient Genome Editing in *Caenorhabditis elegans*. Genetics 197: 1069–1080. https://doi.org/10.1534/genetics.114.166389

Kirino H., N. Maehara, and R. Shinya, 2023 How did *Bursaphelenchus* nematodes acquire a specific relationship with their beetle vectors, *Monochamus*? Frontiers in Physiology 14.

Kondo N., A. Takahashi, K. Ono, and T. Ohnishi, 2010 DNA Damage Induced by Alkylating Agents and Repair Pathways. J Nucleic Acids 2010: 543531. https://doi.org/10.4061/2010/543531

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Lesa G. M., 2006 Isolation of *Caenorhabditis elegans* gene knockouts by PCR screening of chemically mutagenized libraries. Nat Protoc 1: 2231–2240. https://doi.org/10.1038/nprot.2006.345

Letunic I., S. Khedkar, and P. Bork, 2021 SMART: recent updates, new developments and status in 2020. Nucleic Acids Research 49: D458–D460. https://doi.org/10.1093/nar/gkaa937

Lu D., M. Macchietto, D. Chang, M. M. Barros, J. Baldwin, *et al.*, 2017 Activated entomopathogenic nematode infective juveniles release lethal venom proteins. PLOS Pathogens 13: e1006302. https://doi.org/10.1371/journal.ppat.1006302

Mahalak K., A. Jama, S. Billups, A. Dawes, and H. Chamberlin, 2017 Differing roles for *sur-2*/*MED23* in *C. elegans* and *C. briggsae* vulval development. Development Genes & Evolution 227: 213–218. https://doi.org/10.1007/s00427-017-0577-4

Mello C. C., 2007 Return to the RNAi world: rethinking gene expression and evolution. Cell Death Differ 14: 2013–2020. https://doi.org/10.1038/sj.cdd.4402252

Mendez P., B. Walsh, and E. A. Hallem, 2022 Using newly optimized genetic tools to probe *Strongyloides* sensory behaviors. Molecular and Biochemical Parasitology 250: 111491. https://doi.org/10.1016/j.molbiopara.2022.111491

Meneely P. M., C. L. Dahlberg, and J. K. Rose, 2019 Working with Worms: *Caenorhabditis elegans* as a Model Organism. Current Protocols Essential Laboratory Techniques 19: e35. https://doi.org/10.1002/cpet.35

Mistry J., S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, *et al.*, 2021 Pfam: The protein families database in 2021. Nucleic Acids Research 49: D412–D419. https://doi.org/10.1093/nar/gkaa913

Myllyharju J., and K. I. Kivirikko, 2004 Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends in Genetics 20: 33–43. https://doi.org/10.1016/j.tig.2003.11.004

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Pastink A., E. Heemskerk, M. J. Nivard, C. J. van Vliet, and E. W. Vogel, 1991 Mutational specificity of ethyl methanesulfonate in excision-repair-proficient and -deficient strains of *Drosophila melanogaster*. Mol Gen Genet 229: 213–218. https://doi.org/10.1007/BF00272158

Poinar G. O., 2018 *Nematodes for Biological Control of Insects*. CRC Press.

Rahimi F. R., T. R. McGuire, and R. Gaugler, 1993 Morphological Mutant in the Entomopathogenic Nematode, *Heterorhabditis bacteriophora*. Journal of Heredity 84: 475–478. https://doi.org/10.1093/oxfordjournals.jhered.a111374

Schärer O. D., 2013 Nucleotide Excision Repair in Eukaryotes. Cold Spring Harb Perspect Biol 5: a012609. https://doi.org/10.1101/cshperspect.a012609

Schwartz H. T., 2015 *Bob The Worm: The life cycle of an entomopathogenic Heterorhabditis nematode*. CreateSpace Independent Publishing Platform.

Sega G. A., 1984 A review of the genetic effects of ethyl methanesulfonate. Mutation Research/Reviews in Genetic Toxicology 134: 113–142. https://doi.org/10.1016/0165-1110(84)90007-1

Serra L., M. Macchietto, A. Macias-Muñoz, C. J. McGill, I. M. Rodriguez, *et al.*, 2019 Hybrid Assembly of the Genome of the Entomopathogenic Nematode *Steinernema carpocapsae* Identifies the X-Chromosome. G3 Genes|Genomes|Genetics 9: 2687–2697. https://doi.org/10.1534/g3.119.400180

Shinya R., S. Sun, M. Dayi, I. J. Tsai, A. Miyama, *et al.*, 2022 Possible stochastic sex determination in *Bursaphelenchus* nematodes. Nat Commun 13: 2574. https://doi.org/10.1038/s41467-022-30173-2

Simmer F., C. Moorman, A. M. van der Linden, E. Kuijk, P. V. E. van den Berghe, *et al.*, 2003 Genome-Wide RNAi of *C. elegans* Using the Hypersensitive *rrf-3* Strain Reveals Novel Gene Functions. PLoS Biol 1: e12. https://doi.org/10.1371/journal.pbio.0000012

Singh J., 2021 Harnessing the power of genetics: fast forward genetics in *Caenorhabditis elegans*. Mol Genet Genomics 296: 1–20. https://doi.org/10.1007/s00438-020-01721-6

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Smith H. E., and S. Yun, 2017 Evaluating alignment and variant-calling software for mutation identification in *C. elegans* by whole-genome sequencing. PLoS One 12: e0174446. https://doi.org/10.1371/journal.pone.0174446

Sommer R. J., and A. McGaughran, 2013 The nematode *Pristionchus pacificus* as a model system for integrative studies in evolutionary biology. Molecular Ecology 22: 2380–2393. https://doi.org/10.1111/mec.12286

Stock S. P., C. T. Griffin, and R. Chaerani, 2004 Morphological and molecular characterisation of *Steinernema hermaphroditum* n. sp. (Nematoda: Steinernematidae), an entomopathogenic nematode from Indonesia, and its phylogenetic relationships with other members of the genus. Nematology 6: 401–412. https://doi.org/10.1163/1568541042360555

Sulston J. E., 2003 *Caenorhabditis elegans*: The Cell Lineage and Beyond (Nobel Lecture). ChemBioChem 4: 688–696. https://doi.org/10.1002/cbic.200300577

Tamura K., G. Stecher, and S. Kumar, 2021 MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution 38: 3022–3027. https://doi.org/10.1093/molbev/msab120

Thein M. C., A. D. Winter, G. Stepek, G. McCormack, G. Stapleton, *et al.*, 2009 Combined Extracellular Matrix Cross-linking Activity of the Peroxidase MLT-7 and the Dual Oxidase BLI-3 Is Critical for Post-embryonic Viability in *Caenorhabditis elegans*. Journal of Biological Chemistry 284: 17549–17563. https://doi.org/10.1074/jbc.M900831200

Thompson O., M. Edgley, P. Strasbourger, S. Flibotte, B. Ewing, *et al.*, 2013 The million mutation project: A new approach to genetics in *Caenorhabditis elegans*. Genome Res. 23: 1749–1762. https://doi.org/10.1101/gr.157651.113

Tiller G. R., and D. A. Garsin, 2014 The SKPO-1 Peroxidase Functions in the Hypodermis to Protect Caenorhabditis elegans From Bacterial Infection. Genetics 197: 515–526. https://doi.org/10.1534/genetics.113.160606

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Tomalak, M., 1994 Phenotypic and genetic characterization of dumpy infective juvenile mutant in *Steinernema feltiae* (Rhabditida : Steinernematidae). Fundamental and Applied Nematology 17: 485–495.

Volkova N. V., B. Meier, V. González-Huici, S. Bertolini, S. Gonzalez, *et al.*, 2020 Mutational signatures are jointly shaped by DNA damage and repair. Nat Commun 11: 2169. https://doi.org/10.1038/s41467-020-15912-7

Wang X., and H. M. Chamberlin, 2002 Multiple regulatory changes contribute to the evolution of the *Caenorhabditis lin-48 ovo* gene. Genes Dev 16: 2345–2349. https://doi.org/10.1101/gad.996302

Wang K., M. Li, and H. Hakonarson, 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research 38: e164. https://doi.org/10.1093/nar/gkq603

Wang H., H. Park, J. Liu, and P. W. Sternberg, 2018 An Efficient Genome Editing Strategy To Generate Putative Null Mutants in *Caenorhabditis elegans* Using CRISPR/Cas9. G3 Genes|Genomes|Genetics 8: 3607–3616. https://doi.org/10.1534/g3.118.200662

Wang Z., C. Ogaya, V. Dörfler, M. Barg, R.-U. Ehlers, *et al.*, 2023 Pheno- and genotyping in vitro dauer juvenile recovery in the nematode *Heterorhabditis bacteriophora*. Appl Microbiol Biotechnol 107: 7181–7196. https://doi.org/10.1007/s00253-023-12775-y

Wicks S. R., R. T. Yeh, W. R. Gish, R. H. Waterston, and R. H. A. Plasterk, 2001 Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. Nat Genet 28: 160–164. https://doi.org/10.1038/88878

Xue C., and E. C. Greene, 2021 DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends in Genetics 37: 639–656. https://doi.org/10.1016/j.tig.2021.02.008

Zioni (Cohen-Nissan) S., I. Glazer, and D. Segal, 1992 Phenotypic and Genetic Analysis of a Mutant of *Heterorhabditis bacteriophora* Strain HP88. J Nematol 24: 359–364.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

Figure legends

Figure 1 The Dpy phenotype of *S. hermaphroditum* and *C. elegans* peroxidase mutants. **A** Young adult hermaphrodites of the wild-type strain of *S. hermaphroditum* (PS9179) and of two *S. hermaphroditum skpo-2* mutants: *sy1644*, recovered in an EMS mutagenesis screen, and *sy2108*, generated by CRISPR-Cas9. Scale bar, 200 µM. **B** A phylogeny of *C. elegans* SKPO-2 and its closely related peroxidase proteins in *C. elegans* and *S. hermaphroditum*. *C. elegans* SKPO-2 and its *S. hermaphroditum* ortholog are indicated with a gray box. Branch strength bootstrap scores were generated using neighbor-joining with 1000 repetitions. *S. hermaphroditum* protein accession numbers are listed in Table S2. **C** The Dpy phenotype of rare surviving *C. elegans mlt-7(tm1794)* homozygotes, with the wild type (N2) for comparison; young adults of each are shown. Although these rare Dpy survivors are generally healthy when their mother carried a wild-type allele, in subsequent generations the rare survivors are increasingly sickly and display pleiotropic phenotypes, as can be seen in the irregular body shape and blistered cuticle of this animal. Scale bar, 200 µM. **D** The *skpo-2* genomic locus in *S. hermaphroditum*. The positions of the CRISPR cleavage target and of the mutations identified in PS9260, PS9267, and PS9839 are indicated. Primers that amplified exon 9 from the wild type and from other mutants did not amplify sequence from PS9265. Scale bar, 250 bp.

Figure 2 Mutational spectrum resulting from EMS mutagenesis of *S. hermaphroditum*. **A** The homozygous mutations found in the three sequenced strains were divided according to whether they were unique to the strain in question, or were found in multiple strains and so must have existed prior to mutagenesis. They were then sorted by the nature of the mutation: the single-nucleotide changes indicated, a single-nucleotide insertion, a single-nucleotide deletion, or a multi-nucleotide variation (MNV). The single-nucleotide insertion category was more common in pre-existing mutations, and almost all were intronic (not shown). The total numbers of unique mutations in each strain and the total number of shared mutations are indicated as

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

"n=". **B** Distribution of unique and common annotated mutations on the *S. hermaphroditum* X chromosome. Each row consists of from 227 to 295 mutations, each indicated with a colored, partially transparent circle; the intensely colored spots corresponding to hotspots for mutation detection indicate dozens of overlapping dots. The position of *skpo-2* on the chromosome is indicated. **C** Distribution of 868 EMS-induced mutations detected on the *C. elegans* X chromosome, from a collection of whole-genome sequencing data (HTS and PWS, manuscript in preparation). Each mutation is indicated with a colored, partially transparent circle. The distribution of EMS-induced mutations in *C. elegans* is noticeably more even than is seen in *S. hermaphroditum*.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

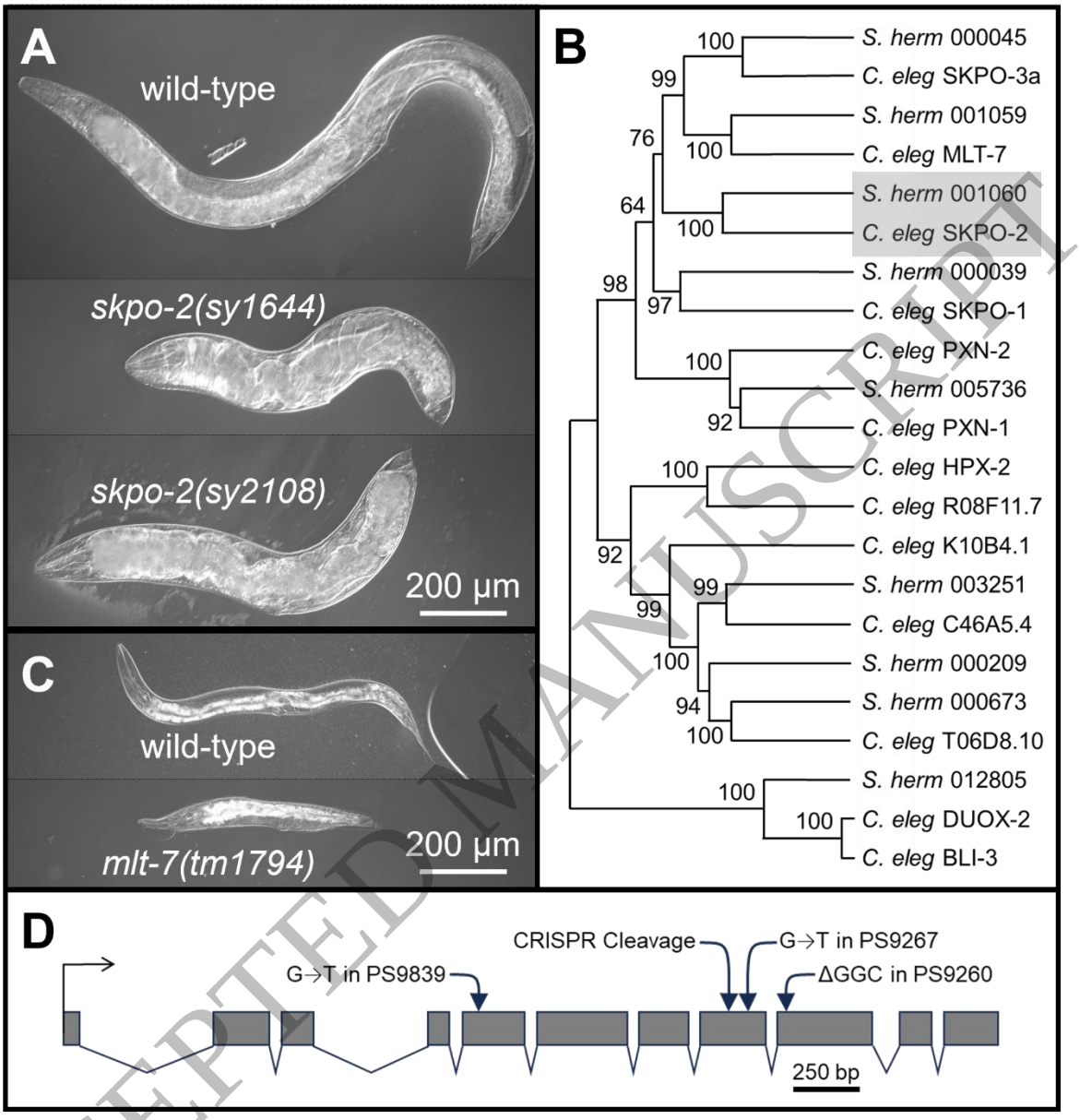

Figure 1
153x159 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023

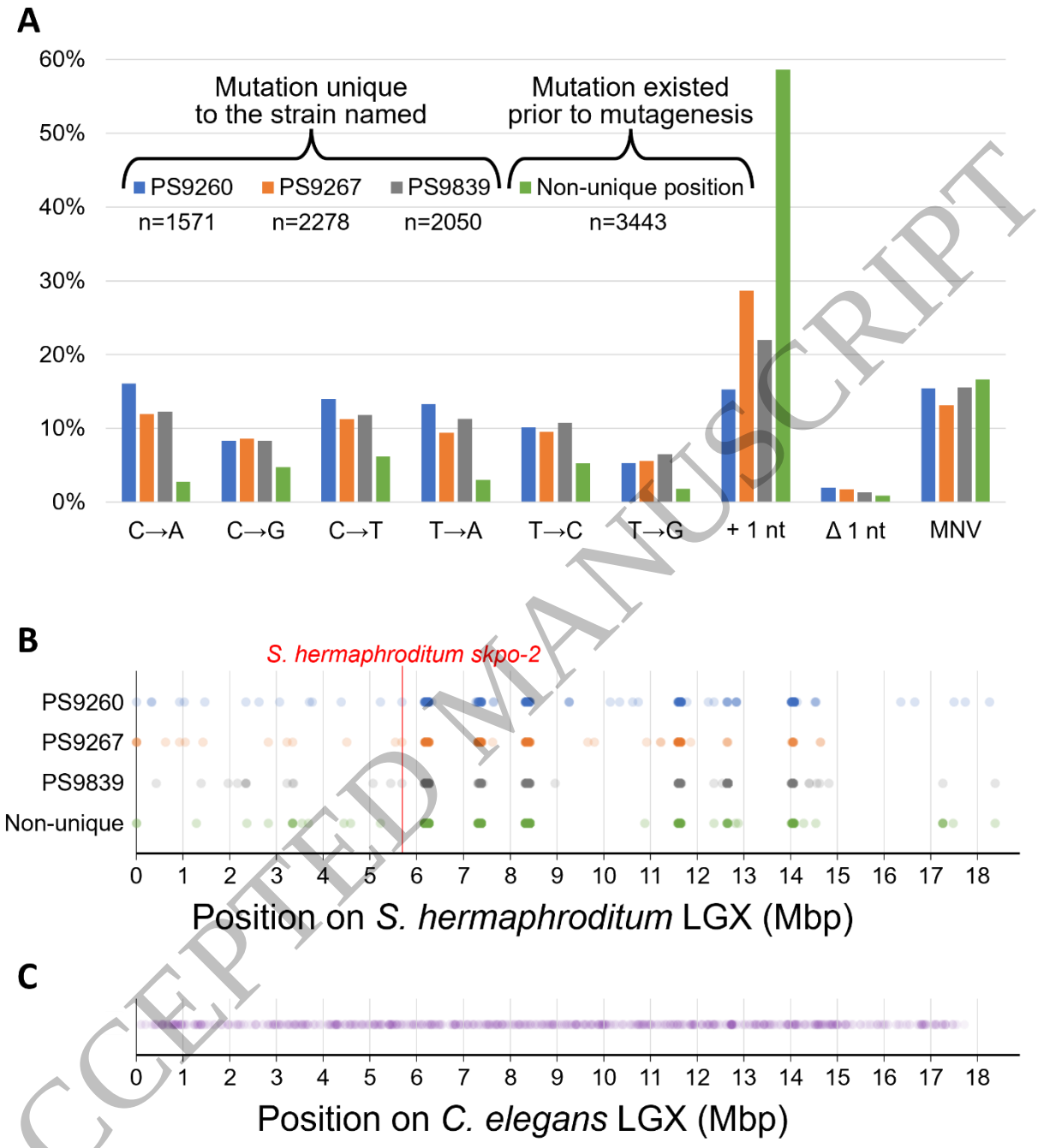

Figure 2
194x213 mm (x DPI)

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyad209/7469079 by California Institute of Technology user on 21 December 2023