OXFORD G3 Genes | Genomes | Genetics

Modular safe-harbor transgene insertion for targeted single-copy and extrachromosomal array integration in *Caenorhabditis elegans*

Sonia El Mouridi [ID] , Faisal Alkhaldi [ID] , Christian Frøkjær-Jensen [ID] *

Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia

*Corresponding author: Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), PO Box 4700, Thuwal 84112-0840, Kingdom of Saudi Arabia. Email: cfjensen@kaust.edu.sa

Abstract

Efficient and reproducible transgenesis facilitates and accelerates research using genetic model organisms. Here, we describe a modular safe-harbor transgene insertion (MosTI) for use in *Caenorhabditis elegans* which improves targeted insertion of single-copy transgenes by homology directed repair and targeted integration of extrachromosomal arrays by nonhomologous end-joining. MosTI allows easy conversion between selection markers at insertion site and a collection of universal targeting vectors with commonly used promoters and fluorophores. Insertions are targeted at three permissive safe-harbor intergenic locations and transgenes are reproducibly expressed in somatic and germ cells. Chromosomal integration is mediated by CRISPR/Cas9, and positive selection is based on a set of split markers (*unc-119*, *hygroR*, and *gfp*) where only animals with chromosomal insertions are rescued, resistant to antibiotics, or fluorescent, respectively. Single-copy insertion is efficient using either constitutive or heat-shock inducible Cas9 expression (25–75%) and insertions can be generated from a multiplexed injection mix. Extrachromosomal array integration is also efficient (7–44%) at modular safe-harbor transgene insertion landing sites or at the endogenous *unc-119* locus. We use short-read sequencing to estimate the plasmid copy numbers for 8 integrated arrays (6–37 copies) and long-read Nanopore sequencing to determine the structure and size (5.4 Mb) of 1 array. Using universal targeting vectors, standardized insertion strains, and optimized protocols, it is possible to construct complex transgenic strains which should facilitate the study of increasingly complex biological problems in *C. elegans*.

Keywords: *Caenorhabditis elegans*; CRISPR/Cas9; transgenesis; genome engineering; long-read assembly

Introduction

The ability to modify the genome of cells and model organisms with high precision has made great advances in the past decade, primarily due to the ease of genetic engineering with CRISPR/Cas9 (Jinek *et al.* 2012; Cong *et al.* 2013; Mali *et al.* 2013). In *Caenorhabditis elegans*, early experiments demonstrated that Cas9 can be expressed from plasmids (Chen *et al.* 2013; Dickinson *et al.* 2013; Friedland *et al.* 2013; Tzur *et al.* 2013; Waaijers *et al.* 2013), mRNA (Lo *et al.* 2013; Chiu *et al.* 2013; Katic and Großhans 2013), or injected directly as a protein (Cho *et al.* 2013; Paix *et al.* 2015) and that templated repair is efficiently mediated from injected plasmids. Subsequent improvements have led to the use of short repair homology (Paix *et al.* 2014), partially single-stranded repair templates (Dokshin *et al.* 2018), improved sgRNA design (Farboud and Meyer 2015), and novel genetic (e.g. Arribere *et al.* 2014; Kim *et al.* 2014; Ward 2015; Farboud *et al.* 2019) and transgenic selection strategies (e.g. Dickinson *et al.* 2015; Schwartz and Jorgensen 2016). These advances have led to the routine use of CRISPR/Cas9 gene editing to create custom alleles and tagging genes at their endogenous loci (Frøkjær-Jensen 2013; Dickinson and Goldstein 2016; Nance and Frøkjær-Jensen 2019).

However, in some cases the goal is to create stable expression of transgenes to determine the effect of regulatory elements on gene expression (Merritt *et al.* 2008), visualize cell populations (Yemini *et al.* 2021), measure cell activity (Suzuki *et al.* 2003), or mark genomic regions (Frøkjær-Jensen *et al.* 2014). Initial transgenesis in *C. elegans* was based on transgene expression from extrachromosomal arrays (Stinchcomb *et al.* 1985; Mello *et al.* 1991; Mello and Fire 1995). Extrachromosomal arrays are formed by injecting linear or supercoiled DNA into the germline syncytium resulting in large hereditary DNA structures that contain many copies of the injected DNA (Stinchcomb *et al.* 1985) propagated as one or more ring chromosomes (Woglar *et al.* 2020). Arrays have the advantage that they are easy to generate and can drive high transgene expression in somatic cells but have the disadvantages that inheritance is unstable with loss occurring during mitosis (Lin *et al.* 2021), expression is mosaic (Okkema *et al.* 1993), and arrays are frequently silenced in germ cells (Kelly *et al.* 1997). Some of these problems can be solved by integrating arrays into the genome by X-ray or gamma irradiation but integration is inherently mutagenic, labor intensive, and does not fully solve the problem of variable expression (Mello and Fire 1995). More

recently, alternative methods for integrating arrays have been developed: reactive oxygen species generated by high-intensity blue light can trigger random array integration (Noma and Jin 2018) and Cas9 can be used for targeted array integration into the *dpy-3* or *ben-1* loci (Yoshina *et al.* 2016). However, neither methodology has become widely adopted, possible due to the requirement for a specialized light-source or the mutant phenotype of integrated strains, respectively.

In some cases, expression at near-endogenous levels or expression in the germline is desired, which can be achieved by reducing the number of plasmids. Low-copy-number plasmid integrations are primarily generated by biolistic transformation or by using a Mos1 transposon. Biolistic transformation works by bombarding animals with DNA-coated particles, a method that is scalable and can overcome germline silencing (Praitis *et al.* 2001). Biolistic transformation has the disadvantages that a variable number of plasmids are integrated, the insertion site is difficult to determine, and integration may results in chromosomal aberrations (Tyson *et al.* 2018). Targeted single-copy insertions can be generated based on the exogenous Mos1 DNA transposon from *Drosophila* which can transpose in *C. elegans* (Bessereau *et al.* 2001; Robert and Bessereau 2007; Vallin *et al.* 2012). The first widely used single-copy insertion method, Mos1-mediated single-copy insertion (MosSCI), relied on excising transposons in neutral genomic environments and transgene insertion by homology-directed repair from an extrachromosomal array (Frøkjær-Jensen *et al.* 2008; Frøkjær-Jensen 2015). Biolistic insertion and MosSCI were developed using *unc-119* selection (Maduro and Pilgrim 1995) and subsequently expanded to antibiotic resistance markers (Giordano-Santini *et al.* 2010; Semple *et al.* 2010; Radman *et al.* 2013). Later iterations of Mos1-mediated transgenesis included additional insertion sites and negative selection (Frøkjær-Jensen *et al.* 2012), and random single-copy insertion by transposition enabling a set of universal MosSCI landing sites (Frøkjær-Jensen *et al.* 2014). Single-copy transgene insertions only rely on double-strand DNA breaks and can, therefore, also be generated by CRISPR/Cas9 (Dickinson *et al.* 2013). More recent iterations of single-copy insertion techniques have incorporated specialized landing sites that contain promoters (Silva-García *et al.* 2019), self-selection based on a split antibiotic selection marker (Stevenson *et al.* 2020), or recombinase-mediated cassette exchange (Nonet 2020).

Each of the above methods have their advantages and several are continuously improved (El Mouridi *et al.* 2020; Nonet 2021) yet there is still considerable room for improvement. For example, CRISPR/Cas9-based insertion of transgenes requires different targeting vectors for every location and rely on selection protocols with substantial screening (Dickinson *et al.* 2013). The universal MosSCI sites have the disadvantage that landing sites were generated at random, and the insertion frequency is lower (Frøkjær-Jensen *et al.* 2014). Positive and negative selection markers facilitate the identification of insertions but still require a considerable amount of manual screening and rarely completely avoid false positives (McDiarmid *et al.* 2020; El Mouridi *et al.* 2021). The split antibiotic selection marker facilitates screening but is only compatible with a single antibiotic and genomic location (Stevenson *et al.* 2020). Frequently, it is also not possible to excise the selection marker used to insert transgenes which limits the complexity of transgenic strains that can be generated.

Here, we develop a method we have named <u>mo</u>dular <u>s</u>afe-harbor <u>t</u>ransgene <u>i</u>nsertion (MosTI) that improves single-copy and array integration by synthesizing many incremental improvements developed across laboratories and by developing a collection of standardized and validated reagents. MosTI is based on identifying safe-harbor insertion sites located at autosome centers in intergenic regions characterized by permissive chromatin modifications (Ho *et al.* 2014). The method was designed to be highly modular: insertion sites can be rapidly generated and existing sites are easily converted to a different selection marker. MosTI landing sites are compatible with a collection of universal targeting vectors that contain common promoters and fluorophores. The insertion sites were engineered to use a set of split selection markers that obviate the need for negative selection markers as only insertions result in rescue (*unc-119*), fluorescence (*gfp*), or antibiotic resistance (*HygroR*). Finally, MosTI is compatible with both single-copy transgene and targeted array insertion at high frequency. We have validated and characterized several landing sites and developed protocols that allow many single-copy insertions by heat-shocking animals from a single injection. Finally, we use MosTI to resolve plasmid copy number in integrated arrays and the detailed structure of one array by long-read sequencing.

Materials and methods
Strains
Strains were maintained using standard methods (Brenner 1974) and were grown at 20°C on OP50 or HB101 (for *unc-119* animals) bacteria. Strains containing MosTI landing sites have been deposited with the *Caenorhabditis* Genetics Center (CGC) (see Supplementary Table 1).

Commercial software
Some figures were created with BioRender.com. Figures were designed using Adobe Illustrator 2022 (v26.0.2). Statistical analysis was performed using GraphPad Prism 9 for macOS (v9.3.1). The manuscript was written in Microsoft Word for Mac (v.16.60) and tables were generated using Microsoft Excel for Mac (v.16.60). In silico molecular biology was performed using A plasmid Editor (ApE) (Davis and Jorgensen 2022) which is freely available at https://jorgensen.biology.utah.edu/wayned/ape/.

Molecular biology
Plasmids were generated by standard molecular techniques, including 3-fragment multisite Gateway reactions (Invitrogen cat. no. 12538200), Gibson assembly (Gibson *et al.* 2009), Golden-Gate cloning (Engler *et al.* 2009), or by gene synthesis (Twist Bioscience, CA, USA). All PCRs were performed using a high-fidelity DNA polymerase (Phusion, New England Biolabs, F530S) and constructs generated by PCR were sequence verified by Sanger sequencing. Most plasmids have been deposited with Addgene (Cambridge, MA, USA) (see Supplementary Table 1). Annotated GenBank files of all plasmids are included in Supplementary File 1.

MosTI landing sites
Please see Supplementary Table 1 for the exact insertion sites and sgRNA sequences.

unc-119 selection
We generated target plasmids containing a codon-optimized synthetic *unc-119* rescue marker [*syn-unc-119*(+)] flanked by 250 bp of homology to each side of the insertion site (pSEM203 Chr. II, pSEM224 Chr. I, pSEM226 Chr. IV). We injected these targeting plasmids (individually) into *unc-119*(ed3) animals with 25 ng/μl pCFJ2474 (P*smu-2::Cas9::gpd-2::tagRFP*), 15 ng/μl of sgRNA (specific to each site) and 10 ng/μl of fluorescent coinjection markers

pSEM233 (P*mlc-1*::*tagRFP*) or pSEM231 (P*mlc-1*::*gfp*) (El Mouridi et al. 2020) and DNA ladder (1 kb ladder, Thermo Fisher Scientific, cat. no. SM1331) added to a final concentration of 100 ng/μl. After injection, P0 plates were placed at 25°C to increase Cas9 expression. The screen for Unc rescue was done after starvation. Strains that were homozygous for the landing sites were subsequently injected with a plasmid mix containing 10 ng/μl Cre recombinase (pMDJ39), 10 ng/μl of fluorescent coinjection markers pSEM233 (P*mlc-1*::*tagRFP*) or pSEM231 (P*mlc-1*::*gfp*), and DNA ladder (1 kb ladder, Thermo Fisher Scientific, cat. no. SM1331) added to a final concentration of 100 ng/μl to remove the synthetic *unc-119* rescue marker (flanked by LoxP sites). The efficiency was high enough that no array selection was necessary and we identified excision of the integrated *cbr-unc-119* selection by screening for Unc animals. We validated all MosTI insertions by PCR and Sanger sequencing. Additionally, we whole-genome sequenced 2 MosTI insertion strains, CFJ77 (chr. I) and CFJ42 (chr. II) to identify background mutations (see Supplementary Table 1). All other insertions strains (e.g. strains with split hygroR and P*mlc-2*::*gfp* selection) were derived from the same genetic background [originally an *unc-119(ed3)* mutant strain] and are expected to share the same background SNPs.

Converting MosTI landing site to a different selection marker

We converted MosTI strains with split *unc-119* selection to split P*mlc-2*::*gfp* and hygromycin selections. MosTI strains for a given chromosomal location with *unc-119* selection were injected with a conversion plasmid (pSEM258 for hygroR and pSEM260 for P*mlc-2*::*gfp*) and standard plasmids used for MosTI single-copy insertions (Cas9, hygroR, sgRNA, *tagRFP* or *gfp* fluorescent coinjection markers). After injection, P0 animals were placed at 25°C. After 48 h, 500 μl of a 4 mg/ml hygromycin solution was added to the NGM plates to select for transgenic animals. Single-copy insertion of the conversion transgenes was identified by *unc-119* rescue after the food was exhausted. All converted MosTI sites were validated by Sanger sequencing. The conversion strategy should, in principle, work for any selection marker that can be split into 2 nonrescuing fragments.

Single-copy transgene insertion using MosTI

Constitutive Cas9 protocol. In this protocol, a codon-optimized Cas9 containing PATCs in introns to prevent silencing and a fluorophore to monitor expression (pCFJ2474, P*smu-2*::*cas9*::*gpd-2 tagRFP-T*) (Aljohani et al. 2020) was used for constitutive expression from extrachromosomal arrays. The injection mix consisted of 25 ng/μl linearized Cas9 (pCFJ2474), 25 ng/μl target vector (various), 15 ng/μl pCFJ782 (HygroR), 15 ng/μl sgRNA (pSEM318), and 10 ng/μl of fluorescent coinjection markers (P*mlc-1*::*tagRFP*, pSEM233 or P*mlc-1*::*gfp*, pSEM231) for a total DNA concentration of 100 ng/μl. The plasmids were injected into young adult Unc animals grown at 15°C or 20°C from strains with MosTI landing sites. Single, injected animals were placed on NGM plates seeded with OP50 and incubated at 25°C. Forty-eight hours after injection, 500 μl of a 4 mg/ml hygromycin solution was added to P0 plates to select for transgenic animals. Single-copy insertions were identified by screening plates for any moving animals (*unc-119* rescue) once the food was exhausted (∼7–10 days after injection).

Heat-shock protocol

In this protocol, codon-optimized Cas9 was expressed under a heat-shock promoter for inducible Cas9 expression. The injection

mix contained 25 ng/μl of pMDJ231 (P*hsp-16.41*::Cas9) linearized with ApaLI, 25 ng/μl target vector, 15 ng/μl pCFJ782 (HygroR), 15 ng/μl sgRNA (pSEM318), and 10 ng/μl of muscle coinjection markers (P*mlc-1*::*tagRFP*, pSEM233, or P*mlc-1*::*gfp*, pSEM231) for a total DNA concentration of 100 ng/μl. The plasmids were injected into young adult Unc animals grown at 15°C or 20°C from strains with MosTI landing sites. Single, injected animals were placed on NGM plates seeded with OP50 and incubated at 25°C. Forty-eight hours after injection, 500 μl of a 4 mg/ml hygromycin solution was added to P0 plates to select for transgenic animals. Two different heat-shock protocols were used to generate single-copy insertions: (a) 37°C for 1 h or (b) 30°C for 18 h (Nonet 2020). Animals with single-copy insertions were identified based on *unc-119* rescue after the food was exhausted. Note, we frequently observed transgene insertions prior to heat-shock, suggesting that the heat-shock promoter is leaky or that transgenic animals are under stress (e.g. due to temperature, antibiotic selection, or starvation). The longer heat-shock protocol (30°C for 18 h) was slightly more efficient and was more convenient as all plates placed in an incubator will be equilibrated to the surrounding temperature without the need for wrapping plates in parafilm and using a water bath (Boulin and Bessereau 2007) or splitting plates into a single layer (Frøkjær-Jensen et al. 2014).

Heat-shock induced MosTI multiplex insertions over generations

A MosTI strain containing a landing site on chr. II (CFJ42) was injected with a mix containing 10 ng/μl linearized transgene(s), 10 ng/μl pCFJ782 (HygroR), 10 ng/μl pSEM318 (sgRNA), 10 ng/μl pMDJ231 (P*hsp-16.41*::Cas9) linearized with ApaLI, and DNA ladder (1 kb ladder, Thermo Fisher Scientific, cat. no. SM1331) added to a final concentration of 100 ng/μl. Injected animals were singled to NGM plates seeded with OP50 and placed at 25°C. One day after injection, 500 μl of 4 mg/ml hygromycin was added to injection plates to select for transgenic strains. The experiment was performed in parallel with seven independent transgenic strains. For each transgenic strain, a single nonrescued transgenic Unc animal was picked to an NGM plate seeded with OP50 7 days after injection and placed at 25°C. For every generation, a single nonrescued transgenic animal was transferred to a new NGM plate with OP50 before the food was exhausted. Subsequently, the transgenic animals were heat-shocked to induce Cas9 expression 7 days after being singled to NGM plates. The plates were screened for single-copy transgene insertions (based on moving *unc-119*-rescued animals) after the food was exhausted. In some experiments, the injection mix contained several target transgenes (P*rpl-7A*::*gfp*, P*eft-3*::*mMaple*, and P*eft-3*::*mScarlet*). We determined which transgene had been inserted based on fluorescence, PCR, and Sanger sequencing.

Extrachromosomal array insertions

Extrachromosomal arrays can be inserted into MosTI sites or into the endogenous *unc-119* locus (in *ed3* mutants) using *unc-119* selection. We have also generated targeting fragments to insert arrays into MosTI sites that use hygromycin and P*mlc-2*::*gfp* but have not tested these fragments (although we deposited the reagents with Addgene). The protocols are identical except for using targeting fragments and sgRNAs that are specific to each insertion site (listed in Supplementary Table 1). Here, we describe validated array insertion protocols based on *unc-119* rescue.

One-step direct plasmid-based array integration

With this strategy, all injected plasmids are "directly" integrated at the target site, including plasmids expressing Cas9 and the sgRNA. We injected adult *unc-119* hermaphrodites with a mix composed of 10 ng/µl of pCFJ2474 (P*smu-2::cas9::gpd-2 tagRFP-T*), 5 ng/µl of array target fragment (e.g. pSEM371 for integration at the endogenous *ce-unc-119* locus), 15 ng/µl of sgRNA targeting the integration site and the array target fragment (e.g. pSEM376 for *ce-unc-119*), 10 ng/µl of transgenes for integration, and DNA ladder (1 kb plus, Invitrogen) to a final concentration of 100 ng/µl. Injected animals were placed on NGM plates and grown at 25°C. Seven to 10 days after injection, we screened plates for array integration based on *unc-119* rescue and expression of injected transgenes (typically, fluorescent markers). We treated any plate with rescued animals as a single independent integration event and picked only a single clonal strain derived from any injected animal.

One-step plasmid-based array integration in preinjected strain

With this strategy, only the transgenes, the array target fragment, and the sgRNA plasmid were integrated. First, a MosTI strain containing a landing site on chr. II (CFJ42) was injected with a mix containing: 25 ng/µl pMDJ231 (P*hsp-16.41::Cas9*) linearized with ApaLI, 15 ng/µl pCFJ594 (NeoR), 10 ng/µl pSEM235 (P*mlc-1::mCherry*), and DNA ladder (1 kb ladder, Thermo Fisher Scientific) added to a final concentration of 100 ng/µl. Injected animals were singled to NGM plates seeded with OP50 and placed at 25°C. One day after injection, 500 µl of 25 mg/ml neomycin was added to injection plates to select for transgenic strains [not phenotypically rescued for *unc-119(ed3)*]. A single strain was selected as an injection strain for a second step, where adult animals were injected with a mix containing: 15 ng/µl pSEM376 (sgRNA), 15 ng/µl pCFJ782 (hygroR), 15 ng/µl pSEM234 (P*mlc-1::mCherry::NSL*), 15 ng/µl pSEM233 (P*mlc-1::gfp*), 5 ng/µl pSEM371 (integration fragment), and 1 kb DNA ladder (Invitrogen) to a total concentration of 100 ng/µl. Single, injected animals were placed on NGM plates seeded with OP50 and incubated at 25°C. Forty-eight hours after injection, 500 µl of a 4 mg/ml hygromycin solution was added to P0 plates to select for transgenic animals. Seven to 10 days after injection (before the bacterial lawn was exhausted), strains with the 2 extrachromosomal arrays were heat-shocked (30°C for 18 h) and screened for *unc-119* rescue 1 generation later. We subsequently allowed the initial array (containing NeoR and P*mlc-1::mCherry*) to be lost on standard NGM plates with no antibiotic selection and the integrants were rendered homozygous. Cas9 should only cut the array formed in the second injection and as expected, we only observed antibiotic resistance and markers from the second injection mix in the integrated strains.

Two-step plasmid-based array integration

With this strategy, only the transgenes and the array target fragment are integrated; plasmids encoding Cas9 and the sgRNA are injected in a second round and are not part of the integrated array. As a first step, we generated transgenic lines by injecting a mix composed of 2.5 ng/µl of the array target fragment (e.g. pSEM371 for *ce-unc-119* integration), 15 ng/µl of pCFJ782 (HygroR), 10 ng/µl of transgenes, and DNA ladder (1 kb plus, Invitrogen) to a final concentration of 100 ng/µl. Injected animals were singled to NGM plates seeded with OP50 and placed at 25°C. One day after injection, 500 µl of 4 mg/ml hygromycin was added to injection plates to select for stable transgenic strains. In the F2 or F3 generation, we picked a single clonal animal to a seeded NGM plate

with hygromycin and expanded the population (note that these animals are not rescued for the *unc-119* phenotype but are hygromycin resistant). For the second step, we injected Unc transgenic animals carrying extrachromosomal arrays with a mix containing 20 ng/µl pCFJ2474 (P*smu-2::cas9::gpd-2 tagRFP-T*), 15 ng/µl of plasmid expressing an sgRNA to cut the integration site and the array fragment (e.g. pSEM376 for *ce-unc-119* integration), 10 ng/µl of fluorescent coinjection marker (pSEM233 or pSEM231), and DNA ladder (1 kb plus, Invitrogen) to a final concentration of 100 ng/µl. After injection, we placed single individual injected animals on seeded NGM plates with hygromycin and grew the animals at 25°C. We identified strains with array integrations 7–10 days after injection based on unc-119 rescue. We treated any plate with rescued animals as a single-independent integration event and picked only a single clonal strain derived from any injected animal.

Two-step protein-based array integration

With this strategy, only the transgenes and the array target fragment are integrated; Cas9 protein (IDT) and crRNA/tracrRNA (IDT) are injected in a second round. Integrations are generated faster with this protocol compared to plasmid-based protocols as array integrants are identified in the progeny of the injected animals. As for the 2-step plasmid protocol, we first generated transgenic lines by injecting a mix composed of 2.5 ng/µl of the array target fragment (e.g. pSEM371 for *ce-unc-119* integration), 15 ng/µl of pCFJ782 (HygroR), 10 ng/µl of transgenes, and DNA ladder (1 kb plus, Invitrogen) to a final concentration of 100 ng/µl. Injected animals were singled to NGM plates seeded with OP50 and placed at 25°C. One day after injection, 500 µl of 4 mg/ml hygromycin was added to injection plates to select for stable transgenic strains. In the F2 or F3 generation we picked a single clonal animal to a seeded NGM plate with hygromycin and expanded the population (note that these animals are not rescued for the *unc-119* phenotype but are hygromycin resistant). For the second step, we prepared guide RNA duplexes in vitro by incubating 3 µl of 100 uM crRNA (e.g. targeting *ce-unc-119*), 3 µl 100 uM tracrRNA, and 4 µl nuclease-free duplex buffer (IDT) at 95°C for 5 min, followed by a 5 min incubation at room temperature. We mixed 3 µl of the crRNA::tracrRNA duplex with 10 µg recombinant Cas9 protein (Alt-R S.p. Cas9 Nuclease V3, IDT), 10 ng/µl fluorescent coinjection marker (pSEM233 or pSEM231), and nuclease-free water to a final volume of 10 µl. We injected transgenic Unc animals carrying extrachromosomal arrays with this Cas9 protein mix and placed individual injected animals on NGM plates seeded with OP50 at 25°C. Rescued animals were identified in the progeny of the injected animals (the F1 generation) and isolated before the food was exhausted. We treated any plate with rescued animals as a single independent integration event and picked only a single clonal strain derived from any injected animal.

Characterization of arrays integrated by MosTI
Quantification of fluorescence expression using COPAS large-particle flow cytometer

We obtained similar mixed populations by placing four L4 worms on small NGM plates seeded with OP50 and used 4 plates per strains. The animals were grown at 25°C for 7 days and then washed off with a M9 buffer and placed on ice. After 10 min, the pellet was washed with water to remove bacteria. We quantified the fluorescence of each strain using a COPAS flow cytometer (Union Biometrica). Between each strain, we washed all the tubing by running water for at least 1 min. Settings for 1-step insertion: Gain = 2, green: 400 V; red: 450 V. Settings for 2-step

insertion: Gain = 2, green: 300 V; red: 400 V. Both green and red fluorescence were expressed along the entire length of the worm, so we based our quantification on the integral of the fluorescence signal. We selected animals with a TOF between 600 and 900 (mainly young adults) and with an extinction coefficient below 35,000 to select for animals that came straight through the flow cell. We manipulated the data in Microsoft Excel (version 16.59) and graphed the data with GraphPad Prism (version 9.3).

Background SNP analysis in injection strains

Raw sequencing reads were trimmed to remove adapters and filtered for low quality bases (Fastqc) (Andrews 2010). The filtered reads were aligned to the *C. elegans* reference genome (WBcel235) with BWA V0.7.17 (Li 2013). Single-nucleotide variants and indels were called with GATK4 Haplotypecaller (Poplin *et al.* 2018) using filters: QD < 2.0, FS > 60, MQ < 40, SOR > 4, MQRankSum < −8, ReadPosRankSum < −8. Next, we used snpEff to filter out SNPs and indels outside protein-coding genes and those with "modifier" impact (Cingolani *et al.* 2012).

Plasmid copy-number estimation

Genomic DNA from eight strains with integrated arrays was isolated using a DNeasy blood and tissue kit according to the manufacturer's instructions (Qiagen, cat. no. 69504). Whole-genome sequencing (WGS) (150 bp paired end) was performed by Novogene. Short Illumina reads were mapped to the unique regions of integrated transgenes with BWA V0.7.17 (Li 2013). We calculated copy number estimates with the formula: (mapped transgene reads * genome size)/(total genomic mapped reads/ mapped transgene length). All copy-number estimates are per haploid genome.

Integrated array assembly from Oxford Nanopore sequencing

High molecular weight genomic DNA was extracted from CFJ157 (*kstIs6* [pSEM371; pSEM233; pSEM231; pCFJ782; 1 kb ladder] III) using a MagAttract HMW DNA kit following the manufacturer's instructions (Qiagen, cat. no. 67563). Oxford Nanopore sequencing was performed by Novogene (China) using a PromethION instrument. We removed *C. elegans* reads by mapping base-called and quality filtered reads to the *C. elegans* reference genome (WBcel235) with transgene sequence homology and 40 kb of the reference sequence flanking the insertion site masked. Unmapped reads (corresponding to the integrated array, flanking genomic sequences, and *E. coli* genome) were assembled with the Canu long-read assembler using default settings except for: "corOutCoverage = 9999" "minOverlapLength = 2500" "minReadLength = 2500" "corMinCoverage = 0" "corMaxEvidenceErate = 0.15" (Koren *et al.* 2017). Array contigs were identified by aligning integrated transgene sequences to contigs with MUMmer V4 (Marçais *et al.* 2018). A single contig (~5 Mb) corresponding to the *E. coli* genome was excluded from further analysis.

Results

Engineered landing sites and split selection markers enable easy identification of insertions

We set out to develop a modular transgenesis system where insertions are easily identified, new insertion sites can be readily generated, and the selection scheme can be modified. Furthermore, we wanted the ability to integrate single-copy transgenes and extrachromosomal arrays at well-defined locations. As a first step, we developed a strategy for inserting single-copy transgenes into "safe-harbor" landing sites based on rescuing *unc-119(ed3)* animals using a split *cbr-unc-119(+)*

selection scheme (Fig. 1a), an approach that is similar to the split antibiotic selection used by Stevenson *et al.* (2020). The first part of the rescue marker is at the landing site and the second part is located on the target plasmid adjacent to the transgene; neither part can in isolation rescue the *unc-119* phenotype. Injecting plasmids expressing Cas9 and an efficient sgRNA (Moreno-Mateos *et al.* 2015) generates a DNA double-strand break (DSB) at the landing site. The target plasmid contains homology to the DSB site, and homology-directed repair reconstitutes a full rescuing *cbr-unc-119(+)* marker, while also inserting a single copy of the transgene. If desired, the reconstituted split *cbr-unc-119* selection marker can be excised after transgene insertion by expressing Cre recombinase. This approach greatly facilitates identification of single-copy insertions: we have observed no false positives and just a few rescued animals are easily visible among a population of Unc animals (see next section). Although not required, we typically enrich for nonrescued transgenic animals with extrachromosomal arrays using fluorescent markers and hygromycin selection.

The use of engineered landing sites is useful for several reasons. First, synthetic landing sites allowed us to generate 3 MosTI landing sites on different autosomes that are compatible with the same target vector and sgRNA (Fig. 1b). One landing site is near the commonly used *ttTi5605* site on chr. II (Frøkjær-Jensen *et al.* 2008), whereas the other sites were chosen for permissive chromatin marks (Ho *et al.* 2014) and minimal disruption to endogenous genes. We inserted landing site in a 2-step process: an initial landing site was inserted using a short synthetic *unc-119(+)* selectable marker flanked by LoxP sites; a second injection removed the selection marker by expressing Cre recombinase, leaving behind the final landing site with the split *cbr-unc-119* selection (Supplementary Fig. 1). This 2-step process is efficient and transgenes to generate novel landing sites are easily generated in a 1-pot Golden-Gate reaction with a synthetic gene fragment (Engler *et al.* 2008) (Supplementary Fig. 2). Second, the engineered landing sites are all compatible with a single targeting plasmid. This allowed us to generate standard target vectors compatible with restriction cloning, Golden-Gate cloning, and 3-fragment Gateway cloning (Fig. 1c). We also generated an expanded set of MosTI-compatible cloning vectors: 4 target vectors with fluorophores (*gfp* and *mScarlet*) for transcriptional and translational fusions (El Mouridi *et al.* 2017), 4 vectors with a *gpd-2*::fluorophore cassette for transgene coexpression, and 3 vectors with promoters for ubiquitous (P*eft-3*) (Wheeler *et al.* 2016), germline (P*mex-5*) (Zeiser *et al.* 2011), and neuronal (ultra-pan-neuronal, UPN promoter) (Yemini *et al.* 2021) expression (Fig. 1c; Supplementary Fig. 3). We performed WGS on 2 injection strains (chr. I: CFJ77 and chr. II: CFJ42) to identify possible passenger mutations present in the initial *unc-119(ed3)* strain or generated by landing site insertion (Supplementary Table 1). All plasmids are available from Addgene, and the universal target vector has been deposited with a gene synthesis company for direct gene synthesis instead of standard cloning, which is becoming increasingly convenient and cost-efficient.

We tested MosTI efficiency for single-copy insertions of a plasmid encoding a ubiquitously expressed GFP (P*eft-3*::*gfp*). We chose the *eft-3* promoter because single-copy insertions are bright in somatic cells whereas germline expression is sensitive to silencing in repressive chromatin domains (Frøkjær-Jensen *et al.* 2016). We generated single-copy insertions using constitutive Cas9 expression (P*smu-2*::*Cas9*) at relatively high frequencies (25–75%) (Fig. 2a). We also inserted a different transgene (P*rpl-7*::*gfp*) by inducible expression of Cas9 (P*hsp-16.21*::*Cas9*) using an improved
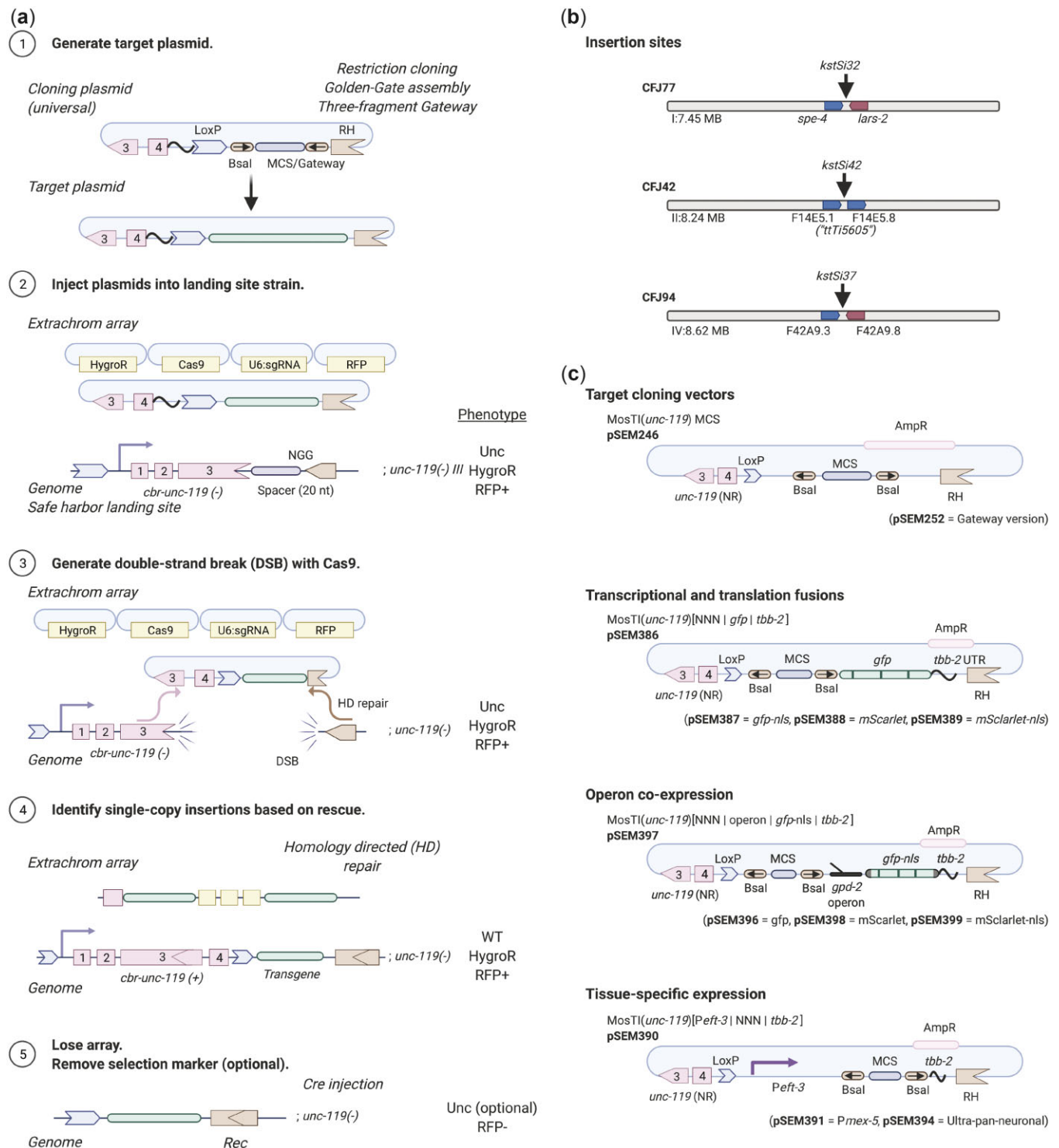
Fig. 1. Schematic of MosTI protocol. a) (1) Target transgenes are generated by cloning (standard restriction-enzyme, Gateway, Golden Gate, or gene synthesis) into a MosTI-compatible plasmid containing a nonrescuing *cbr-unc-119* rescue fragment. (2) Target plasmids are coinjected with Cas9 and sgRNA plasmids into *unc-119(ed3)* animals containing a MosTI landing site. Antibiotic or fluorescent markers can be used to enrich the population for nonrescued transgenic array animals (optional). (3) Cas9-induced DSBs are repaired by homology-directed repair from the target plasmid, leading to transgene insertion and reconstitution of a functional *cbr-unc-119(+)* gene. (4) Animals with single-copy insertions are identified by phenotypic rescue (N2-like animals on plates with Unc animals). (5) The extrachromosomal array is lost, and the *cbr-unc-119* selection can be removed by Cre expression (optional). b) MosTI-compatible insertion sites in safe-harbor landing sites were selected for permissive chromatin environment (Ho *et al.* 2014). c) A collection of MosTI-compatible cloning vectors for tissue-specific gene expression, transcriptional and translational fluorescence expression, and fluorescence coexpression with a *gpd-2* operon has been deposited with Addgene (see Supplementary Fig. 3 for all plasmids).

heat-shock protocol (Nonet 2020). These efficiencies are similar to MosSCI and other single-copy transgenesis methods (Frøkjær-Jensen *et al.* 2012; Dickinson *et al.* 2013; Nonet 2020) but do not

capture the ease of identifying transgene insertions: we readily observed *unc-119* rescued animals in the F2 or F3 generation (for both protocols) or for many generations after injection using the

(a)

| MosTI site	Cas9	Transgene	sgRNA concentration	Injected animals	F2 array transgenics	Single-copy insertions	Efficiency (inserts/transgenic)
Chr. I	P*smu-2*	P*eft-3::gfp*	15 ng/ul	10	8	2	25%
Chr. II	P*smu-2*	P*eft-3::gfp*	15 ng/ul	10	8	6	75%
Chr. II	P*smu-2*	P*eft-3::gfp*	15 ng/ul	15	11	4	36%
Chr. II	P*hsp-16.21*	P*rpl-7A::gfp*	10 ng/ul	15	10	5	50%
Chr. II	P*hsp-16.21*	P*rpl-7A::gfp*	10 ng/ul	15	10	5	50%
Chr. IV	P*smu-2*	P*eft-3::gfp*	15 ng/ul	15	10	5	50%

(b)

(c)
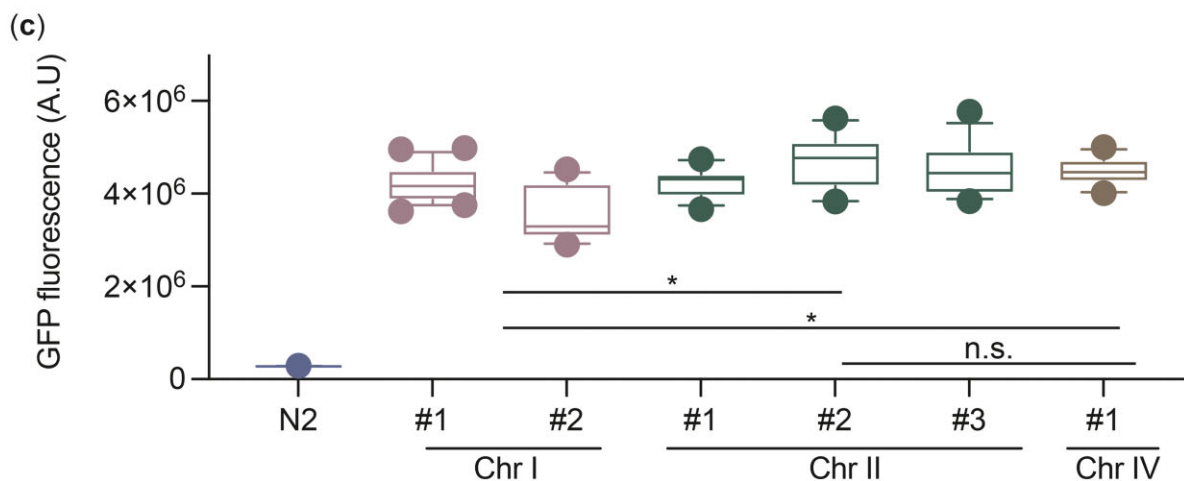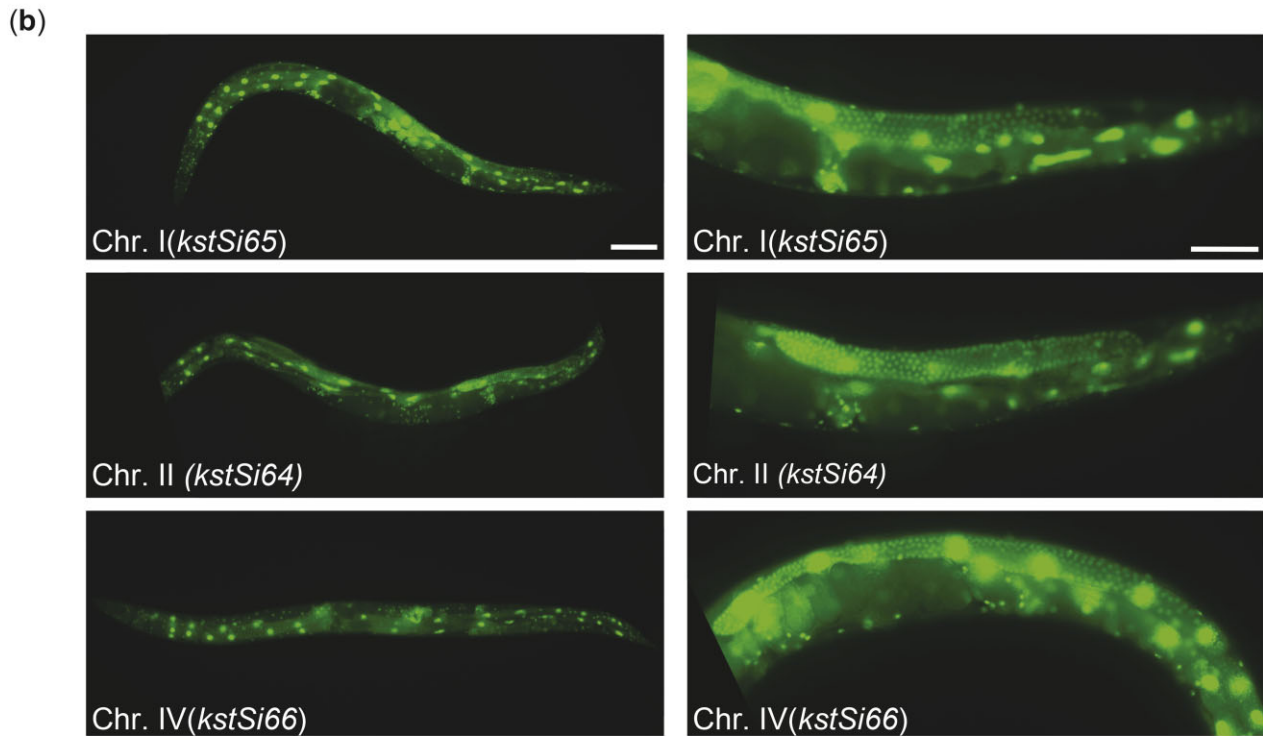
Fig. 2. MosTI insertion frequency and transgene expression. a) Table with MosTI insertion efficiencies using constitutive (P*smu-2::Cas9*) or heat-shock inducible (P*hsp-16.41::Cas9*) plasmids for Cas9 expression. b) Single-copy transgenes (P*eft-3::gfp*) are widely expressed, including the germline (13/13 inserts expressed). Left panels ×10 magnification, right panels ×40 magnification. Scale bars: 20 μm. c) GFP quantification by wide-field fluorescence microscopy (L4 stage) from single-copy P*eft-3::gfp* insertions at three MosTI landing sites. Statistics: Kruskal–Wallis's test of fluorescence intensity grouped by insertion site (chr. I: 35 images, $n = 2$; chr. II: 41 images, $n = 3$; chr. IV: 13 images, $n = 1$ insert, *$P < 0.05$).

heat-shock protocol. All rescued animals behaved as expected: the inserted fluorescent marker was expressed, transgenic animals could be rendered homozygous, the extrachromosomal array was rapidly lost when animals were grown without selection, PCR, and Sanger sequencing validated insertion at the intended MosTI landing site, and the *cbr-unc-119* selection marker

could be removed by Cre expression without losing the inserted transgene. All P*eft-3::gfp* transgenes inserted at the 3 MosTI landing sites were expressed at similarly high levels, and all transgenic strains showed GFP expression in the germline (13 of 13 strains), confirming the permissive chromatin environment (Fig. 2, b and c).

The heat-shock protocol (Nonet 2020) requires an additional heat-shock step relative to constitutive expression of Cas9. However, heat-shock induction has the advantage of being highly scalable and requires injection of fewer animals (i.e. a few transgenic animals with extrachromosomal arrays are sufficient to generate many independent single-copy insertions). The heat-shock protocol may be particularly useful for the injection of collection of different plasmids simultaneously (multiplexed MosTI) and inserting a single copy from the extrachromosomal array (Fig. 3a). Conceptually, such an approach could be used to screen many regulatory elements, for example, a library of enhancers or 3' UTRs coupled to a fluorescent reporter, as previously demonstrated by Kaymak *et al.* (2016). To test the feasibility of multiplexed MosTI, we injected a pool of three target plasmids with visually distinct expression patterns (P*rpl-7A::gfp*, P*eft-3::mMaple*, and P*eft-3::mScarlet*) together with a heat-shock inducible Cas9 plasmid (Fig. 3a). We established 4 independent transgenic strains with extrachromosomal arrays and heat-shocked a single NGM plate of each strain for 7 successive generations. In most generations, we were able to generate single-copy insertions form 3 of the 4 strains. We generated a total of 29 independent single-copy insertions from these 3 strains and observed no obvious decline in efficiency over time (Fig. 3b). In some cases, we were able to isolate animals with every type of transgene insertion from a single NGM plate (Fig. 3c), demonstrating the feasibility of scaling the approach further. Several observations were notable: first, only one of 29 transgene insertions was nonfluorescent, probably caused by an indel in the transgene, as previously observed (Frøkjær-Jensen *et al.* 2008). Second, one of the 4 transgenic strains yielded no insertions in any generation. We did not determine the exact cause but speculate that the target spacer at the MosTI site was destroyed by NHEJ in an early generation or that this array was somehow not permissive for Cas9 expression. Third, not all transgenes were inserted at similar frequencies, possibly due to differences between the injected plasmids (the *gfp* plasmid was linearized whereas the *mMaple* and *mScarlet* plasmids were supercoiled). We recently showed that linearized plasmids were incorporated into arrays more efficiently than supercoiled plasmids (Priyadarshini *et al.* 2022); the integration frequency may simply reflect the relative abundance of plasmids in the array. We note that a conceptually similar approach was developed by Kaymak *et al.* (2016) using MosSCI to assay 3' UTR gene regulation but their approach was not scalable in the same way as MosTI.

In sum, MosTI is an efficient method to insert single-copy transgenes into well-defined genomic locations using several protocols. All landing sites are compatible with robust somatic and germline expression. From a practical perspective, the split selection markers make it easier to isolate animals with single-copy transgenes and the heat-shock protocol is a simple method to generate many inserts from only a few injections.

Exchanging MosTI selection markers is simple

The use of a strong phenotypic selection marker, such as *unc-119*, has the benefit that insertions are easy to identify, and insertions are easily rendered homozygous. However, *unc-119* animals are slow growing and harder to inject than wildtype animals (Maduro and Pilgrim 1995). We designed MosTI to be modular and have developed a standardized strategy to convert selection markers at MosTI landing sites. Conversion is simple: we use MosTI to insert a transgene containing a second split selection marker and a new spacer sequence (Fig. 4a). In a second (optional) step, the *cbr-unc-119* marker is removed by Cre-mediated recombination. We validated this conversion strategy using a split hygromycin resistance gene, similar to Stevenson *et al.* (2020), and a split fluorescent marker that gives bright muscle fluorescence (P*mlc-1::gfp*) (El Mouridi *et al.* 2020). We converted MosTI sites to HygroR (chr. II) and *gfp* (chr. I and chr. II) and generated single-copy P*mex-5::gfp* insertions (Fig. 4b). Both selections were efficient and resulted in GFP expression in the germline from the *mex-5* promoter. However, we found that identifying fluorescent animals (4 inserts from 15 injected animals) (Fig. 4c) was considerably easier than isolating animals with antibiotic resistance (1 insert from 20 injected animals). MosTI insertions most frequently happen in the F2 generation when food is almost exhausted: it was relatively easy to identify moving (*unc-119* rescued) or fluorescent (*gfp* selection) L1 animals on a starved plate. In contrast, it was harder to identify healthy animals on hygromycin, which is most effective at preventing larval growth and takes several days to kill animals (Radman *et al.* 2013).

Regardless of the preferred selection, we show that it is relatively easy to change a MosTI site to fluorescence or antibiotic selection. The conversion vector, targeting vectors, and sgRNA can be cheaply generated by gene synthesis with minimal cloning. By modifying existing designs and following a few simple design guidelines, additional safe-harbor landing sites are compatible with the standard collection of targeting vectors. We also note that *cbr-unc-119* selection can be "reused" by inserting a second split *unc-119* fragment. This could, for example, be used to iteratively generate strains with multiple, tightly linked single-copy transgene insertions.

Split selection markers allow targeted integration of extrachromosomal arrays

In some cases, simultaneous expression of many transgenes is needed. For example, a method for automated neuron identification relies on expression from 41 different fluorescent promoter constructs (Yemini *et al.* 2021). In other cases, high expression of a single transgene may be necessary, for example, for genetically encoded sensors (Flytzanis *et al.* 2014; Azimi Hashemi *et al.* 2019) or for spatiotemporal control of gene expression with bi- or tripartite effectors (Wei *et al.* 2012; Wang *et al.* 2018). Transgene arrays, which contain several hundred plasmids (Stinchcomb *et al.* 1985), are useful for this purpose, especially when arrays are chromosomally integrated to stabilize expression and avoid mitotic loss in cell division (Mello and Fire 1995). We have adapted the split MosTI selection for a CRISPR/Cas9-based method that uses nonhomologous end-joining (NHEJ) to target arrays for integration (Yoshina *et al.* 2016; Yoshina and Mitani 2022) (Fig. 5a). The adaptation required a short nonrescuing *unc-119* fragment (array target fragment) to mediate array integration at MosTI landing sites by NHEJ. In this scheme, both the array target fragment and the split *cbr-unc-119* selection marker at the MosTI landing site are cut by the same sgRNA. The DSBs are generated in an intron, so NHEJ between the array and insertion site can reconstitute the full *unc-119* selection marker even if the repair generates short indels (Fig. 5a). We first tested array integration efficiency at a MosTI site (chr. II) using a 2-step protocol (2-step plasmid) which avoided cointegrating the Cas9 and sgRNA
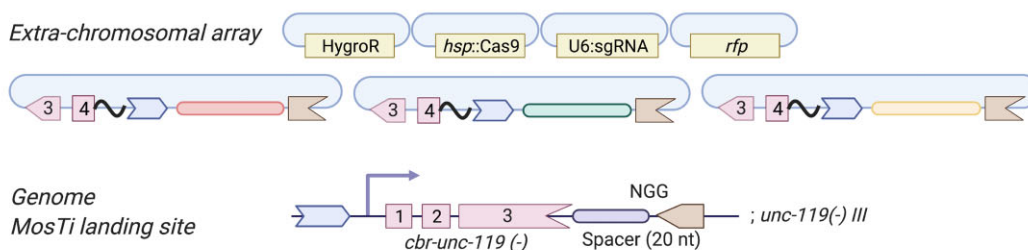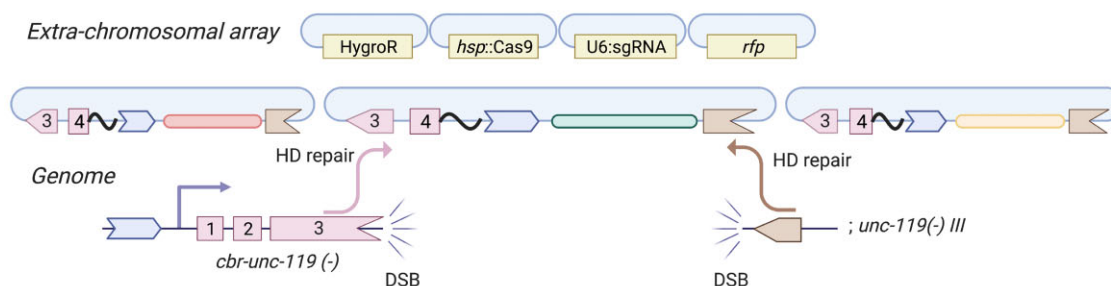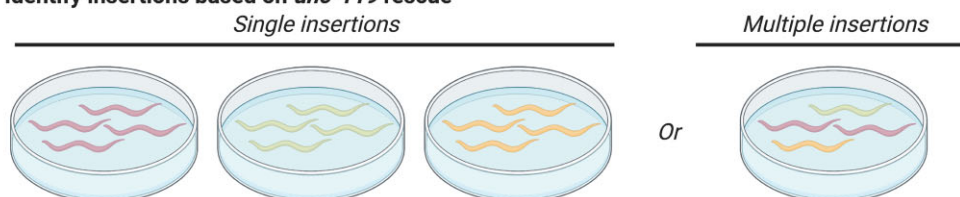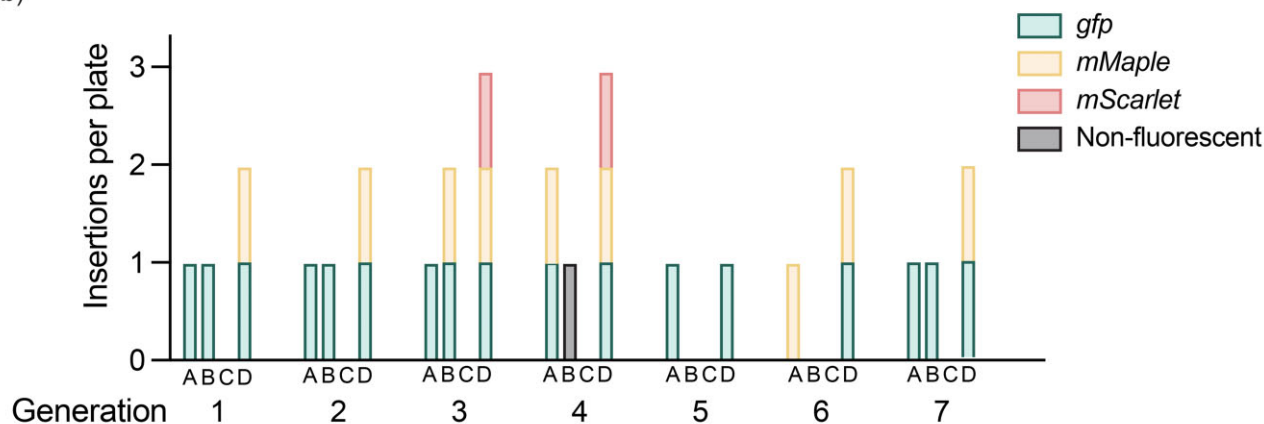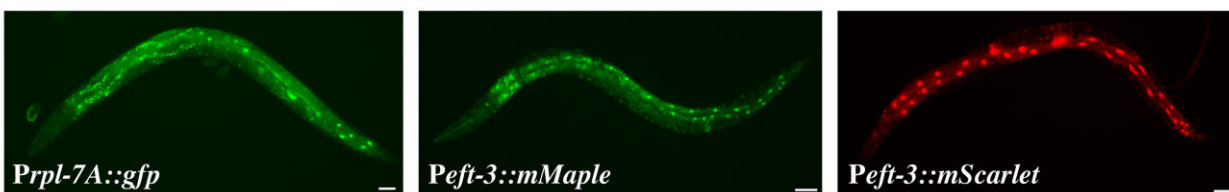
Fig. 3. Single-copy insertions from multiplex transgene injections. a) Schematic of multiplex insertions using MosTI. Multiple transgenes are pooled in one injection mix. During homology-directed repair, a single transgene from the array is inserted into the MosTI site. b) Quantification of single-copy MosTI insertions over seven generations from four independent transgenic animals with extrachromosomal arrays containing three target transgenes (P*rpl-7A::gfp*, P*eft-3::mMaple3*, and P*eft-3::mScarlet*) and a heat-shock inducible Cas9 (P*hsp-16.21::Cas9*). In each generation, a new population of animals was generated from a nonrescued (Unc) animal carrying an extrachromosomal array and these animals were heat-shocked 1 or 2 days before exhausting the bacterial lawn. c) Fluorescence microscopy showing expression from single-copy insertions of each of the 3 target transgenes. ×20 magnification, scale bar = 20 μm.
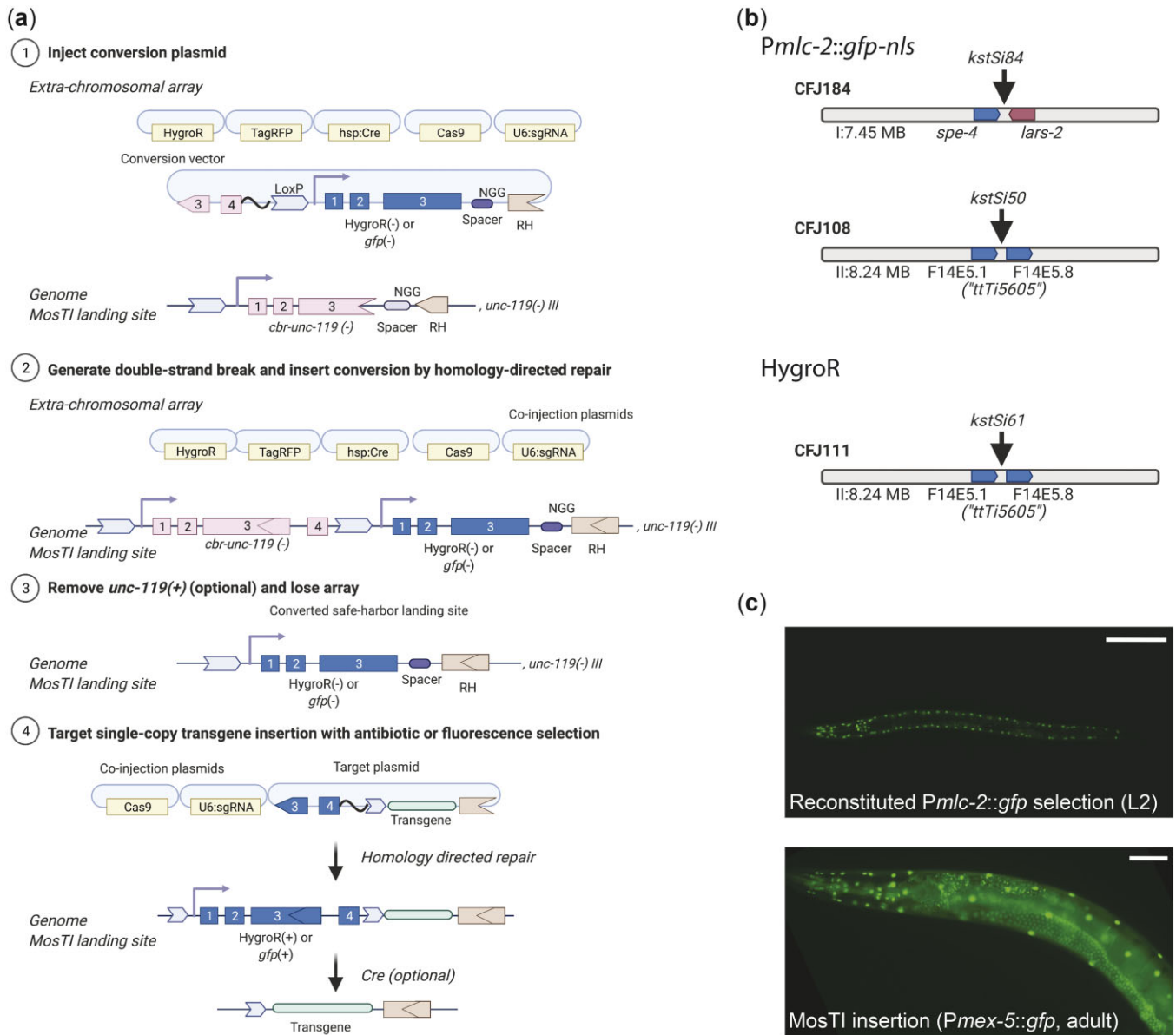
Fig. 4. MosTI sites are modular and can be converted to other split selection markers. a) Schematic of the conversion of an *unc-119* MosTI site into a MosTI site that uses a different split selection marker (here, HygroR or P*mlc-2::gfp*). (1) Injection of a conversion plasmid (e.g. generated by gene synthesis) into *unc-119(ed3)* mutants with a MosTI landing site. (2) Insertion of the conversion plasmid generates a new MosTI landing site using a different split selection marker. (3) Loss of the extrachromosomal array and (optional) excision of the reconstituted *cbr-unc-119* selection marker. (4) Injection of a modified MosTI target vector uses the novel selection marker to insert a single-copy transgene. b) Available MosTI sites for split P*mlc-2::gfp-nls* and HygroR selection. c) Example of a single-copy insertion of a P*mex-5::gfp* (germline) transgene into a converted MosTI site using a split pan-muscular *gfp* selection marker (P*mlc-2::gfp-nls*). Top panel: pan-muscular nuclear GFP expression (L2 stage). Bottom panel: nuclear GFP expression in muscles (from pan-muscular selection) and in the germline (from the P*mex-5::gfp-nls*) transgene. Top panel ×20 magnification, bottom panel ×40 magnification. Scale bar: 20 μm.

plasmids. We first generated transgenic MosTI animals with an extrachromosomal array containing 2 fluorescent markers (P*mlc-1::gfp* and P*mlc-1::tagRFP*), a plasmid encoding antibiotic resistance (HygroR), and the array target fragment. Like the protocol for single-copy insertions, these transgenic array animals expressed the injected fluorophores and were resistant to hygromycin but were not rescued for the *unc-119* phenotype. We injected these transgenic animals with a second mix containing plasmids that expressed Cas9 and an sgRNA to generate DSBs. We successfully identified integrated arrays based on *unc-119* rescue at frequencies ranging from 20% to 33% of injected animals (Fig. 5d). As expected, the rescued animals with the integration

expressed all fluorophores from the initial extrachromosomal array (Fig. 5b) and could be rendered homozygous. Most strains with array insertions were fully rescued for the Unc-119 phenotype except for 1 strain with particularly bright expression that showed variable rescue. This integration was likely an example of somatic position effect variegation since both rescued and nonrescued animals segregated a mix of rescued and nonrescued animals over several generations (Supplementary Fig. 4).

Although array insertion into a MosTI site may be preferable, the same integration method should work at any genetic locus with a strong loss-of-function phenotype caused by a late mutation. Many laboratories have the commonly used *unc-119(ed3)*
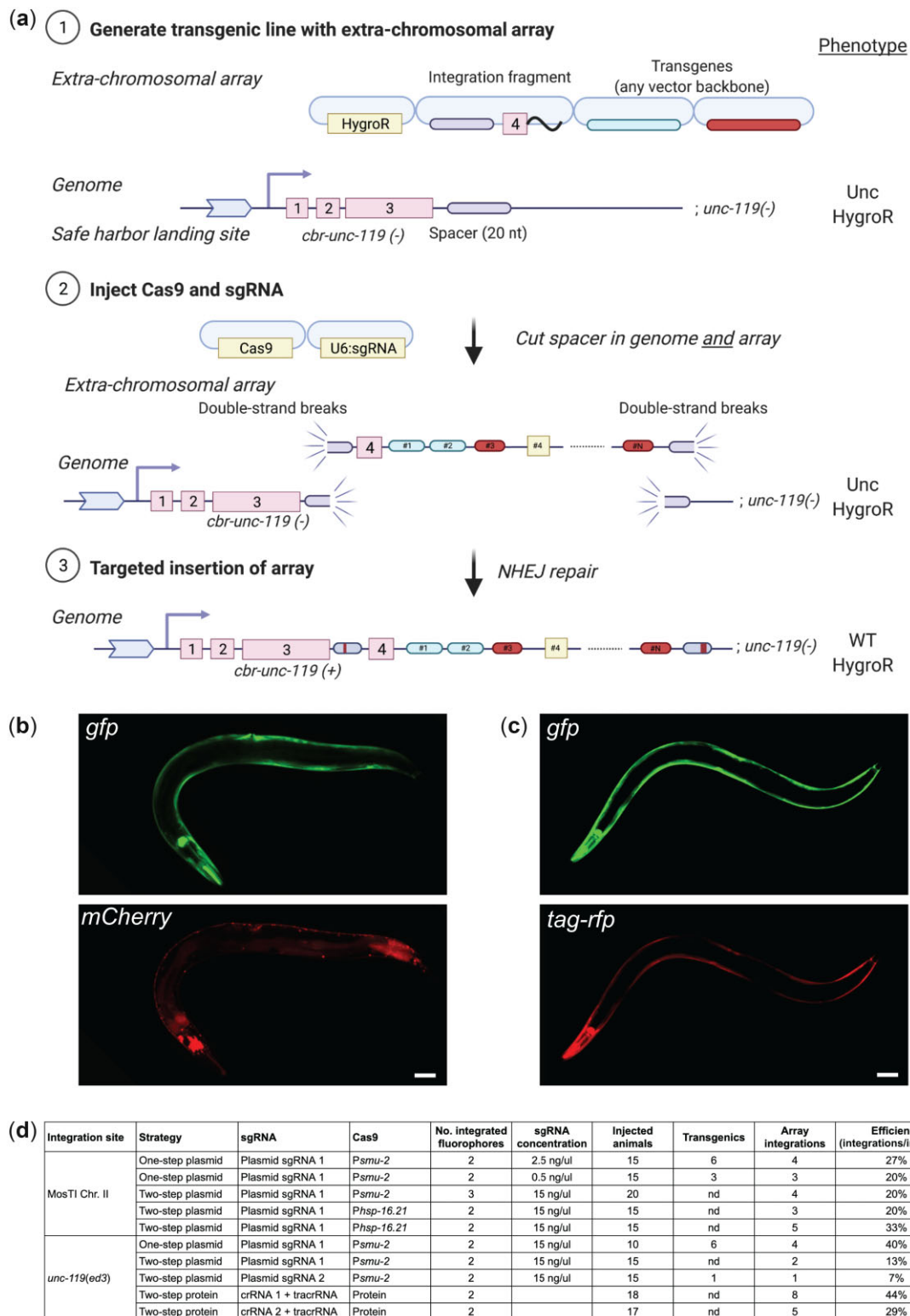
Fig. 5. Targeted integration of extrachromosomal arrays at MosTI sites and at the endogenous *ce-unc-119* locus. a) Schematic of "two-step plasmid" integration of extrachromosomal arrays at a MosTI landing site. (1) An extrachromosomal array is formed by injecting transgenes, including a plasmid containing the fourth intron of *cbr-unc-119* (integration fragment) into *unc-119(ed3)* animals containing a MosTI landing site. (2) Plasmids expressing Cas9 and an sgRNA injected into transgenic animals cause DSBs at the MosTI landing site and in the array (in integration fragments). (3) DSBs are repaired by NHEJ between the array and the MosTI site, resulting in *unc-119* rescue. The sgRNA target is in an intron to allow rescue even when NHEJ repair causes short indels. Note, arrays can also be inserted by a similar strategy at the *ce-unc-119(ed3)* locus, or by injecting Cas9 protein and crRNA/tracrRNA. b) Fluorescence microscopy of an integrated extrachromosomal array containing transgenes expressed in muscles (P*mlc-1::gfp* and P*rab-3::mCherry*) inserted into a MosTI site on chr. II. Scale bar = 20 μm. c) Fluorescence microscopy of an integrated array (P*mlc-1::gfp* and P*mlc-1::tagRFP*) inserted into the endogenous *ce-unc-119(ed3)* locus (chr. III). Scale bar = 20 μm. d) Table showing the efficiency of extrachromosomal array insertion using 1- and 2-step protocols at MosTI landing sites and at *ce-unc-119(ed3)*.

Integration site	Strategy	sgRNA	Cas9	No. integrated fluorophores	sgRNA concentration	Injected animals	Transgenics	Array integrations	Efficiency (integrations/injected)
MosTI Chr. II	One-step plasmid	Plasmid sgRNA 1	P*smu-2*	2	2.5 ng/ul	15	6	4	27%
	One-step plasmid	Plasmid sgRNA 1	P*smu-2*	2	0.5 ng/ul	15	3	3	20%
	Two-step plasmid	Plasmid sgRNA 1	P*smu-2*	3	15 ng/ul	20	nd	4	20%
	Two-step plasmid	Plasmid sgRNA 1	P*hsp-16.21*	2	15 ng/ul	15	nd	3	20%
	Two-step plasmid	Plasmid sgRNA 1	P*hsp-16.21*	2	15 ng/ul	15	nd	5	33%
unc-119(ed3)	One-step plasmid	Plasmid sgRNA 1	P*smu-2*	2	15 ng/ul	10	6	4	40%
	Two-step plasmid	Plasmid sgRNA 1	P*smu-2*	2	15 ng/ul	15	nd	2	13%
	Two-step plasmid	Plasmid sgRNA 2	P*smu-2*	2	15 ng/ul	15	1	1	7%
	Two-step protein	crRNA 1 + tracrRNA	Protein	2		18	nd	8	44%
	Two-step protein	crRNA 2 + tracrRNA	Protein	2		17	nd	5	29%

strain and we, therefore, tested insertions into the *unc-119* locus using a modified integration fragment and sgRNA. Using the same 2-step plasmid protocol, we were able to integrate arrays into the *unc-119* locus at similar frequencies to MosTI sites (7–40% efficiency) (Fig. 5, c and d). We tested 3 additional protocols for generating array insertions. In one protocol (1-step plasmid), we injected all plasmids in a single mix (fluorophores, Cas9, sgRNA, and integration fragment). Rescued integrants from this technique were obvious already in the F1 generation and were generally dimmer than integrants from the "2-step plasmid" protocol (see next section). These integrants may correspond to smaller arrays that would not have become stable inherited arrays (Mello *et al.* 1991; Lin *et al.* 2021). In a third protocol, we first generated an intermediary transgenic injection strain with plasmids encoding a heat-shock inducible Cas9 (*hsp-16.2::Cas9*) and hygromycin selection. We used this strain for a second injection with array plasmids (fluorophores, integration fragment, and sgRNA) and generated animals with two arrays. We heat-shocked these animals and were able to generate array insertions at reasonably high frequencies (20–33%) (Fig. 5d). Finally, we also tested extrachromosomal array integration with Cas9 protein and pre-complexed crRNA/tracrRNA guides. Repeated attempts at integrating plasmids using Cas9 protein in a 1-step protocol resulted in no integration events (data not shown). This result is not unexpected: arrays are formed 4–8 h after injection in embryos (Yuen *et al.* 2011) whereas DSBs are presumably generated in the germline and fertilized oocyte rapidly after injection. Instead, we developed a 2-step protein-based protocol where we injected Cas9 protein into strains carrying extrachromosomal arrays, which resulted relatively high integration efficiencies depending on the cut-site and crRNA used (29–44%) (Fig. 5d).

In sum, MosTI can be used to target extrachromosomal arrays for integration using *unc-119* selection and a variety of protocols with similar efficiencies. We have not attempted integrations at MosTI sites using other selection markers, but the approach should work for hygromycin or fluorescence selection. The choice of protocol will depend on the desired size of integrations, how many integrations are needed, and whether continued expression of Cas9 may cause a problem for subsequent experiments. We note that we have not observed arrays changing over time, despite integration of Cas9 and sgRNA plasmid (see next section).

Conceptually similar protocols can generate array integrations at the endogenous *unc-119* locus and can presumably be extended to most other endogenous genes with a strong loss-of-function phenotypes.

Molecular characterization of targeted array insertions

An early study demonstrated that arrays formed from supercoiled plasmids contain sequences organized in tandem, with some plasmids containing short duplications and deletion (Stinchcomb *et al.* 1985). The size and repetitive nature of simple arrays formed from plasmid DNA preclude full assembly using Illumina short-read sequencing. However, long-read sequencing methods, such as Oxford Nanopore and PacBio sequencing, facilitate assembly of complex and repetitive genomic regions (Logsdon *et al.* 2020). A recent study assembled a complex extrachromosomal array formed from yeast genomic DNA and linear transgene fragments by Nanopore sequencing (Lin *et al.* 2021). The array was mainly formed by NHEJ of yeast genomic DNA and spanned 50 contigs with a total estimated length of 11 Mb. In another study, Tyson *et al.* (2018) characterized a low-copy plasmid integration generated by biolistic bombardment using Nanopore

sequencing. The biolistic integrant contained 30 kb DNA but none of the transgenes (P*pie-1::gfp*) were fully intact and the integration resulted in a 2-Mb duplication of the flanking region.

To better characterize integrated arrays formed from circular plasmid DNA and to detect large-scale genomic changes, we first performed short-read Illumina WGS. We characterized arrays integrated at the *unc-119* locus by the 1-step plasmid protocol (plasmid-integrated strains) and the 2-step protein method (protein-integrated strains). All injections contained supercoiled plasmids encoding 2 fluorophores (P*mlc-1::gfp* and P*mlc-1::tagRFP*), the integration fragment, and linear stuffer DNA (1 kb plus DNA ladder, Invitrogen). Injections with Cas9 plasmid were supercoiled (P*smu-2::cas9::tagRFP*) and protein strains contained a supercoiled plasmid encoding an antibiotic resistance marker (HygroR) (Fig. 6a).

We used WGS data to estimate the copy number of all plasmids in the injection mix based on average read coverage (Fig. 6a). We detected at least 1 copy of all injected plasmid (from 2 to 23 copies) in every strain and the integrated arrays contained between 6 and 37 plasmids in total (Fig. 6a). Our plasmid copy-number estimates are likely correct within a factor of 2: our analysis estimated the copy number of 2 endogenous genes to 0.97 ± 0.14 for *tbb-2* (47% GC) and 0.48 ± 0.06 for *smu-2* (30% GC), indicating that coverage from short-read sequencing underestimates a GC poor gene by ∼50%. Plasmid-integrated strains were less bright compared with protein-integrated strains, and this correlated with lower estimated plasmid copy-number (Fig. 6a). All arrays contained multiple copies of the integration fragment with 2 strains (1 plasmid-integrated and 1 protein-integrated) containing intact spacer sequences, showing that Cas9 cutting efficiency was sometimes incomplete using either integration protocol (Supplementary Fig. 5). A couple of observations are worth noting: in the plasmid-based protocol, both sgRNA and a germline-expressed Cas9 are integrated which could possibly delete parts of the array over time. However, we propagated a plasmid-integrated strain with several intact copies of the integration plasmid for ten generations and observed no change in fluorescence (Supplementary Fig. 6). Also, all 4 protein strains were derived from a single transgenic extrachromosomal array strain. Given the variation in plasmid-copy number it is likely that only some fragments of the original array were integrated. In this model, Cas9 protein injection "catastrophically" cuts the array into fragments at most spacer sequences and the genomic location; NHEJ between array fragments and the genomic locus results in partial array integration. If this model is correct, then the relative copy-number of the integration fragment in the injection mix could be used to tune the size of the integrated array.

The *unc-119* selection scheme tolerates indels at the cut-site because the spacer is in an intron. However, large indels or duplications near the cut-site could be problematic by perturbing nearby gene expression. We analyzed the WGS data to detect small and large-scale indels near the insertion site based on read coverage (Supplementary Fig. 7). None of the integrated strains contained short- or large-scale deletions and we found only 1 instance of a putative short duplication (∼6–7 kb of *unc-119*) (Supplementary Fig. 7).

To determine the structure of a simple array (i.e. an array composed mainly of plasmid DNA and repetitive DNA ladder) we performed long-read Oxford Nanopore MinION sequencing on a strain with an integrated array containing a relatively high copy number (CFJ157). We were able to assemble the array into 40 contigs with an N50 value of 489 kb. We estimated the total array size at 5.5 Mb (Fig. 6c), which is within the range of estimates
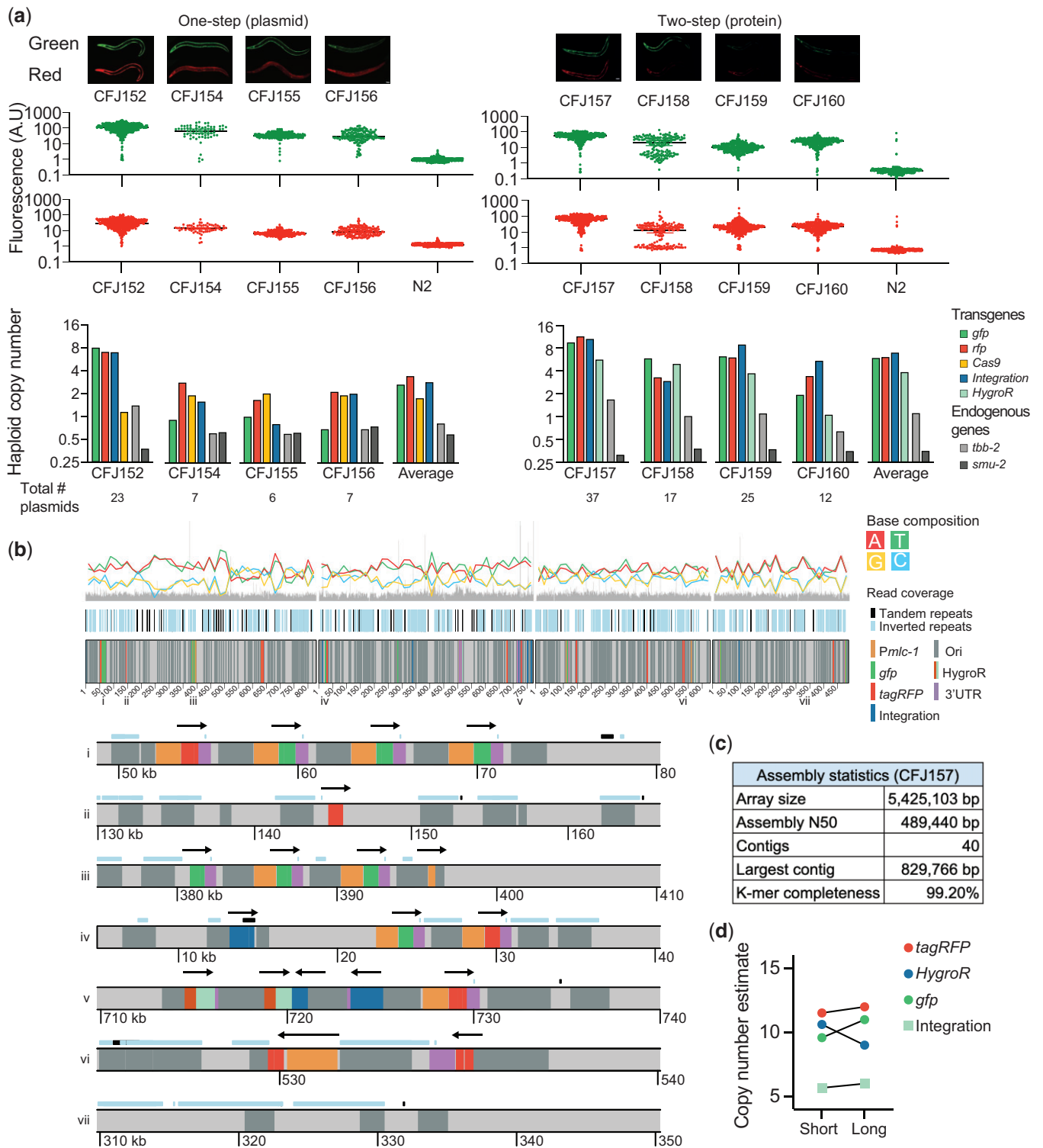
Fig. 6. Molecular characterization of targeted array integrations. a) Analysis of arrays integrated at the *ce-unc-119* locus using a "1-step plasmid" protocol (left) and a "2-step protein" protocol (right). Top: fluorescence images of strains with array integrations. Scale bar = 20 μm. Middle: quantification of total fluorescence in adult animals using a COPAS large-particle fluorescence sorter. Bottom: plasmid copy-number estimates based on WGS (Illumina short read sequencing). b) Integrated array assembly from long-read Oxford Nanopore sequencing of CFJ157 (2-step protein array integration containing *gfp*, *tagRFP*, *hygroR*, integration fragment, and 1 kb plus ladder). Top: schematic overview of 4 long contigs. Bottom: individual areas from contigs (location indicated by numbers above). Arrows indicate the transcriptional direction of transgenes (promoter::gene::UTR). c) Assembly statistics from the strain CFJ157. d) Comparison of copy-number estimates based on Illumina short-read sequencing and Oxford Nanopore long-read sequencing of CFJ157.

based on microscopy (1–2.2 Mb) (Woglar et al. 2020) and long-read assembly (11 Mb) (Lin et al. 2021). Plasmid-copy number estimates based on Illumina and nanopore sequencing were in close agreement (Fig. 6d; Supplementary Fig. 5), validating our approach for estimating plasmid copy number and suggesting that the long-range sequencing assembly for CFJ157 is near-complete.

We determined structural features from the longest contigs and observed short-tandem arrays (commonly 3–4 plasmids) interrupted by repetitive DNA sequences, presumably from the 1-kb ladder used as carrier DNA (Invitrogen) (Fig. 6b). The sequence of the DNA ladder sequences is proprietary but are derived from phage DNA and E. coli DNA, including a ColE1 replication origin and an ampicillin resistance gene with homology to the injected plasmids. We observed structures that demonstrate array assembly by a combination of homologous recombination and NHEJ. In many cases, plasmid DNA was joined to ladder DNA sequences by homologous recombination in the amp or ori sequences. In another example, a tandem array of four plasmids (3 gfp and 1 tagRFP) contained a breakpoint in the Pmlc-1 promoter. The 2 "parts" of the promoter were joined to the flanking ladder DNA with no obvious homology. We observed a similar structure where the tagRFP in a single plasmid (Pmlc-1::tagRFP) was split into 2 fragments, consistent with prior results demonstrating that some plasmids reisolated from extrachromosomal arrays contain truncations (Stinchcomb et al. 1985).

In conclusion, MosTI array insertion results in a diversity of fluorescence expression that is consistent with the plasmid-copy number in the integrated array. Integration does not generally cause wide-spread chromosomal aberrations (indels or duplications) near the insertion site, although some insertions influence the rescue marker, and possibly nearby genes, by position effect variegation. Arrays formed from a mix of supercoiled plasmids and linear carrier DNA contain short tandem plasmid arrays with some plasmids containing breakpoints in regulatory regions (promoter) or in coding regions.

Discussion

We have described protocols and a set of reagents that enable MosTI in the C. elegans genome. The insertion frequencies are high, and the isolation of transgenic strains with targeted insertions is facilitated by mutant rescue, antibiotic resistance, or fluorescence. MosTI is versatile by allowing insertion of both single-copy transgenes and extrachromosomal arrays. Finally, the reagents are modular with easy conversion between selection markers or generation of novel insertion sites.

MosTI in comparison to commonly used insertion methods

In the past decade, several novel methods for transgene insertion have been described. We were inspired by these methods to develop MosTI that incorporate and improve on several of these advances. Advances have largely been driven by CRISPR/Cas9, with Dickinson et al. (2013) first showing that Cas9 induced double-strand DNA breaks could mediate single-copy transgene insertions at similar frequencies to MosSCI. Subsequently, improvements by many laboratories have increased the versatility and efficiency of CRISPR/Cas9 by, for example, using ribonucleoproteins (Paix et al. 2015; Au et al. 2019), oligos or modified repair templates (Dokshin et al. 2018), improved cloning strategies (Schwartz and Jorgensen 2016), and efficient protospacer sequences (Farboud and Meyer 2015). These improvements have,

understandably, primarily focused on editing or tagging endogenous loci.

Some efforts have also focused on improving the ability to insert transgenes into safe-harbor locations. Aram et al. (2019) improved MosSCI by allowing the removal of selection markers. Silva-García et al. (2019) used CRISPR/Cas9 to convert a set of MosSCI sites for use with dpy-10 co-CRISPR editing to enrich for insertions (Arribere et al. 2014). This method, "SKI LODGE" is optimized for inserting PCR amplified coding regions with well-defined tissue specific promoters and protein tags present at the landing sites. The use of PCR products might be expected to allow multiplexed and high-throughput transgenesis but the low insertion efficiency (0.4–2%) of moderate size transgenes (2 kb) is currently a limitation. For this reason, we developed MosTI using plasmid templates but incorporated their use of designer landing sites with efficient protospacers.

Recently, Nonet (2020) demonstrated that recombination-mediated cassette exchange (RMCE) using Flp/FRT or Cre/LoxP could be adapted for use in worms. In this case, landing sites with a GFP were randomly inserted into the genome by transposition (Frøkjær-Jensen et al. 2014). To insert a transgene by RMCE, the GFP is exchanged for a self-excising cassette (Dickinson et al. 2015) by screening for Rol animals. Although the use of RMCE is promising, particularly for inserting large transgenes, the method has a few limitations: first, insertions are generated at relatively low frequency (1 in 3 injected animals). Second, coinsertion of the large self-excising cassette (7.7 kb) limits the size of transgene that can be inserted since most general cloning plasmids are limited to carrying 15 kb inserts (Bajpai 2014). Third, attempts by Nonet (2020) to integrate arrays were unsuccessful, possibly because constitutive expression of Flp recombinase "destroyed" arrays. Here, we have favored generating targeted landing sites by CRISPR/Cas9 and shorter selection markers. If preferred, the flexibility of modular MosTI landing sites can easily be converted to RMCE-compatible landing sites.

Finally, another research group has previously described a very similar use of a split antibiotic selection marker (HygroR) to insert single-copy transgenes (Stevenson et al. 2020). In this case, the authors generated a single landing site at the ttTi5605 site on chr. II and tested the efficiency of 6 protospacer sequences. Like our results, the authors note that single-copy insertions are easily identified based on hygromycin rescue although they report considerably lower insertion efficiencies (1–10%). Although insertion frequencies are difficult to compare across laboratories, one likely difference is our use of a recently developed germline-optimized Cas9 transgene (Aljohani et al. 2020). Stevenson et al. (2020) mainly focused on using in vivo recombineering of coinjected PCR fragments, which was first used for GFP tagging genes by Paix et al. (2016). In vivo recombineering has the significant advantage that cloning is minimized but comes at the expense of reduced insertion frequency and fidelity (Stevenson et al. 2020). We have preferred the use of clonal or synthetic transgenes but note that all MosTI landing sites and split selection markers are compatible with the in vivo recombineering strategies described by Stevenson et al. (2020).

Advantages of MosTI for single-copy insertions

MosTI has several advantages for single-copy insertions, in particular compared to MosSCI (Frøkjær-Jensen et al. 2008, 2012). First, the use of split selection markers makes it considerably easier to identify true transgene insertions, obviating the need for complex coinjection mixtures and time-consuming screens for insertions. Several negative selection markers have been

developed to facilitate the identification of single-copy insertion. The initial negative selection relied on killing array animals by expressing the toxin *peel-1* (Frøkjær-Jensen *et al.* 2012) but this method suffers from frequent escapers. A second method that relies on inducing paralysis by adding a plasmid to the injection mix which makes animals sensitive to histamine (El Mouridi *et al.* 2021) is less prone to escapers but requires the addition of drugs to every plate. In our own laboratory, we find that split selection is considerably easier, especially as it is often only necessary to generate a single extrachromosomal array line with a heat-shock inducible Cas9, which makes it possible for even inexperienced injectors to generate several independent single-copy insertions with relative ease. For the same reason, it is difficult to compare the efficiency of MosSCI and MosTI; in our experience, most new students or researchers in the lab are quickly able to generate at least one extrachromosomal array line (which enables both targeted single-copy and array insertion) whereas MosSCI insertions have typically been rather more difficult.

Each selection marker has its pros and cons. For *unc-119*, the selective pressure is very strong, it is easy to generate homozygous strains, and no specialized plates or fluorescence microscopes are required. The disadvantages include the difficulty of maintaining and injecting *unc-119* animals, no selection for the transgene in genetic crosses, and possible adverse effects of partial gene rescue. The advantages of antibiotic markers include easier strain handling, injection into wildtype animals, and strong selection in genetic crosses. The disadvantages include the cost and inconvenience of making plates with antibiotics and the difficulty of distinguishing homozygous from heterozygous inserts because nonrescued animals are "invisible" on the selective media. To our knowledge, fluorescent markers have not previously been used as the only selection marker for single-copy transgene insertions (in contrast to tagging endogenous genes), likely because fluorescence from arrays and insertions are difficult to distinguish. Split fluorescent selection markers are convenient because phenotypically normal animals are injected, homozygous and heterozygous inserts can be distinguished based on segregation, no special plates are needed, fluorescence protein expression is generally innocuous, and insertions are easily crossed to other strains by "following" the fluorescence. Minor limitations include the requirement for a fluorescence dissection microscope and possible interference between transgene and selection fluorophores. In this case, the fluorescent selection marker can be removed by Cre-mediated recombination.

Also, in contrast to MosSCI, the new MosTI insertion sites were selected based on chromosome location and local permissive chromatin marks. All the insertion sites show a high frequency of germline expression and consistently bright somatic expression. We have not attempted to do an in-depth phenotypic characterization of the strains but have not observed any obvious adverse phenotypes in animals with insertions. We have furthermore performed WGS on several of the insertion strains to detect mutations that were not outcrossed after the initial isolation of *unc-119(ed3)* or off-target effects from CRISPR/Cas9. We detected only a few mutations that cause amino-acid substitutions in genes with subtle phenotypes defined in Supplementary Table 1.

Using MosTI for targeted array insertion

Targeted array integration by MosTI may be useful when high expression or coexpression of many plasmids is necessary. One striking example is the use of 41 different fluorescence drivers with overlapping expression that allow automated cell-specific identification and monitoring neuronal activity "NeuroPAL"

(Yemini *et al.* 2021). To develop the NeuroPAL methodology, Yemini *et al.* (2021) tested many different plasmid combinations and generated integrated lines for many "landmark" fluorophores and a genetically encoded calcium sensor (GCaMP6). Each integrant required considerable effort, including integration, mapping, outcrossing, and phenotyping. Given the continuous development of improved fluorophores and optogenetic tools it is likely that tools such as NeuroPAL exemplify circumstances where rapid and reproducible array insertion would allow "prototyping" new functionalities (e.g. optogenetic inhibition or excitation of defined populations of cells). Similarly, somatic control of expression by Gal4 driver lines (Wang *et al.* 2017), FRT mediated gene activation (Davis *et al.* 2008; Voutev and Hubbard 2008), auxin-mediated degradation (Zhang *et al.* 2015), CRISPR inhibition (Long *et al.* 2015), or CRISPR activation (Gilbert *et al.* 2014) are being actively developed in *C. elegans* and often require sustained high transgene expression.

Our approach for inserting arrays using MosTI differ from the most commonly used methodology. Traditionally, arrays have been inserted by irradiation (Mello and Fire 1995; Evans 2006). Integrating arrays require specialized equipment, careful candidate screening, and is relatively slow. Furthermore, array insertion is random and inherently mutagenic. For these reasons, array integration is not standard practice despite mitotic stability and increased stability of expression in integrants (Evans 2006). In the past few years, 2 alternative array integration methods were developed. In one method, Noma and Jin (2018) adapted a genetically encoded reactive oxygen generator (miniSOG) to integrate arrays by light stimulation in just two weeks at high frequency (∼10%). Although this method is efficient, specialized equipment is still required, arrays are integrated randomly, and background mutations remain a concern. We have instead adapted a second method first developed by Yoshina *et al.* (2016) and recently expanded to additional gene loci (Yoshina and Mitani 2022) where arrays are integrated by CRISPR/Cas9. In this method, CRISPR/Cas9 is used to generate double-strand DNA breaks at a specific genomic location (*dpy-3* or *ben-1*) and in the array (typically, the AmpR marker). Arrays are integrated by NHEJ at relatively high frequencies (3–10%). The disadvantage is that arrays are inserted into coding regions causing a phenotype (Dpy) or using a somewhat complicated mix of positive (temperature selection with *vps-45*) and negative (benzimadole) leading to a relatively high frequency of false positives (Yoshina *et al.* 2016). Here, we have developed a modified CRISPR/Cas9-based integration strategy that includes a positive selection marker which is compatible with the same safe-harbor locations used for MosTI insertions or with the *unc-119(ed3)* mutant strain. MosTI allows quick isolation (∼2 weeks) of integrated arrays at high frequency (7–40%) with no requirements for extensive screening or genetic mapping. Whether or not to integrate extrachromosomal arrays depend on the exact experimental conditions: in some cases, it is useful to be able to rapidly loose an array (for example, when the array is used as a repair template for CRISPR/Cas9 genome editing). In other cases, when preventing mitotic loss of arrays is important for expression in all cells and or stable propagation without selection is required, it may be favorable to generate integrated arrays.

We characterized several array integrations using short-read sequencing and one array in detail by long-read Oxford Nanopore sequencing. From eight array integrations, we found no evidence for off-target insertions, large-scale chromosomal re-arrangements, or large indels at the integration sites, suggesting that targeted integrants are generally well-behaved. We observed unexpected effects

in only 1 array integration (out of 39 in total), where the strain showed the hallmarks of position effect variegation of the *unc-119* rescue marker. Although we have not characterized this strain in detail, it is possible that repressive chromatin may partially silence *unc-119* at the junction between the genome and the integrant. To our knowledge, there are few published examples of somatic position effect variegation in *C. elegans*; targeted insertion of a large repetitive transgene structure could perhaps be used to understand heterochromatin spreading in worms that do not encode canonical genome organizers, such as CTCF (Heger *et al.* 2009).

Our long-read sequencing confirmed the tandem array structure and partial plasmid deletions in arrays, first demonstrated by Stinchcomb *et al.* (1985), and was in agreement with size estimates based on microscopy (Woglar *et al.* 2020) and sequencing a complex array (Lin *et al.* 2021). Although we have not sequenced arrays generated by linear DNA fragments, we and others have demonstrated better transgene expression in the soma (Etchberger and Hobert 2008) and germ line (Aljohani *et al.* 2020) from PCR products. Some of this difference may be driven by increased incorporation of linear fragments in arrays (Priyadarshini *et al.* 2022) but the presence of truncated promoters and fluorophores in the array also suggests that the cell may identify unusual DNA structures that are likely to lead to active silencing (Kelly *et al.* 1997; Leyva-Díaz *et al.* 2017). We envision that the ability to routinely insert arrays into specific locations will be useful for understanding basic characteristics of chromosomes, such as how chromosome size influences large-scale genome domains (Liu *et al.* 2011), as well as enabling the initial steps in engineering synthetic *C. elegans* chromosomes.

Multiplexed transgene insertion

CRISPR/Cas9 allows multiplexed transgene insertion because a single transgenic strain with an extrachromosomal array can carry different repair templates. Kaymak *et al.* (2016) first demonstrated multiplexed repair (library MosSCI) to study the effect of 3' UTRs, which largely determine the expression within the germline (Merritt *et al.* 2008). Similar to their observations, a complex mix of repair templates did not impair the MosTI repair process. However, library MosSCI was not inducible and required a relatively large number of injections to generate unique inserts (269 injected animals resulted in 11 unique transgene insertions) (Kaymak *et al.* 2016). Here, we show that coupling improved germline expression from arrays (Aljohani *et al.* 2020), an efficient sgRNA (Moreno-Mateos *et al.* 2015), and a split selection marker has allowed us to generate many insertions from a single injection by driving Cas9 expression using heat-shock. Our results were limited to proof-of-principle experiments using three plasmids that could be distinguished based on fluorescent protein expression. However, we propose that the method can be scaled up significantly: the rescuing part of the MosTI repair template (*cbr-unc-119* and the right homology region) is relatively short (<1 kb). Thus, a ~5 Mb array can in principle carry ~2,000 different copies of a short 1.5 kb transgene, more than enough room to encode a fluorophore with regulatory elements (e.g. promoters or 3' UTRs). It should be possible to develop genetic screens that combine MosTI with large-scale gene synthesis, for example using oligo pools, and high-throughput sequencing to understand gene regulatory elements at the base-pair level (e.g. de Boer *et al.* 2020).

MosTI curation and continuous development

Finally, standardized resources enable reproducibility across a scientific field, for example, by allowing direct comparison between different transgenes inserted at the same location or sharing of compatible reagents. Reagents for MosTI are freely available from the *Caenorhabditis* Genetics Center (CGC) and Addgene. To facilitate the uptake of transgenic methodology methods, we maintain a website (www.wormbuilder.org) with detailed descriptions of reagents and updated protocols. We designed MosTI to be highly modular and plan to continuously develop the technique, including new insertion sites, strains with multiplexed insertion sites each targeted by a different sgRNA, and new split selection markers. We hope that MosTI will be a resource that is of use to the many different types of experiments done by the *C. elegans* community.

Data availability

Short-read (Illumina) and long-read (Oxford Nanopore) sequencing data were deposited with Bio-Project number: PRJNA827598, https://www.ncbi.nlm.nih.gov/sra. Full annotated plasmid sequences are included in Supplementary File 1. Plasmids are available from Addgene, https://www.addgene.org/. Strains are available from the *Caenorhabditis elegans* Genetics Center (CGC) https://cgc.umn.edu/.

Supplemental material is available at *G3* online.

Conflicts of interest

None declared.

Author contributions

CFJ: conceptualization, writing—original draft, supervision, and funding acquisition. CFJ, SEM, and FA: methodology, writing—review & editing, and visualization. FA: software. SEM and FA: formal analysis. SEM: investigation.

Literature cited

Aljohani MD, Mouridi SE, Priyadarshini M, Vargas-Velazquez AM, Frøkjær-Jensen C. Engineering rules that minimize germline silencing of transgenes in simple extrachromosomal arrays in *C. elegans*. Nat Commun. 2020;11(1):6300. doi:10.1038/s41467-020–19898-0.

Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data; 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Aram R, MacGillivray K, Li C, Saltzman A. Tools for Mos1-mediated single copy insertion (MosSCI) with excisable *unc-119*(+) or NeoR (G418) selection cassettes. MicroPubl Biol. 2019. doi:10.17912/micropub.biology.000146.

Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ. Efficient marker-free recovery of custom genetic modifications

with CRISPR/Cas9 in *Caenorhabditis elegans*. Genetics. 2014;198(3): 837–846. doi:10.1534/genetics.114.169730.

Au V, Li-Leger E, Raymant G, Flibotte S, Chen G, Martin K, Fernando L, Doell C, Rosell FI, Wang S, *et al.* CRISPR/Cas9 methodology for the generation of knockout deletions in *Caenorhabditis elegans*. G3 (Bethesda). 2019;9(1):135–144. doi:10.1534/g3.118.200778.

Bajpai B. High capacity vectors. In: I Ravi, M Baunthiyal, J Saxena, editors. Advances in Biotechnology. New Delhi (India): Springer; 2014. p. 1–10.

Bessereau JL, Wright A, Williams DC, Schuske K, Davis MW, Jorgensen EM. Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. Nature. 2001;413(6851):70–74. doi:10.1038/35092567.

Boulin T, Bessereau J-L. Mos1-mediated insertional mutagenesis in Caenorhabditis elegans. Nat Protoc. 2007;2(5):1276–1287. doi: 10.1038/nprot.2007.192.

de Boer CG, Vaishnav ED, Sadeh R, Abeyta EL, Friedman N, Regev A. Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. Nat Biotechnol. 2020;38(1):56–65. doi: 10.1038/s41587-019–0315-8.

Brenner S. The genetics of *Caenorhabditis elegans*. Genetics. 1974; 77(1):71–94.

Chen C, Fenk LA, de Bono M. Efficient genome editing in *Caenorhabditis elegans* by CRISPR-targeted homologous recombination. Nucleic Acids Res. 2013;41(20):e193. doi: 10.1093/nar/gkt805.

Chiu H, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-free genome editing in *Caenorhabditis elegans* using CRISPR-Cas. Genetics. 2013;195(3):1167–1171. doi:10.1534/genetics.113.155879.

Cho SW, Lee J, Carroll D, Kim J-S, Lee J. Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics. 2013;195(3):1177–1180. doi:10.1534/genetics.113.155853.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin). 2012;6(2):80–92. doi:10.4161/fly.19695.

Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu Jat Y, Cole MA, Liu DR, Joung JK, Bauer DE, *et al.* CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol. 2019;37(3):224–226. doi:10.1038/s41587-019–0032-3.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, *et al.* Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819–823. doi: 10.1126/science.1231143.

Davis MW, Morton JJ, Carroll D, Jorgensen EM. Gene activation using FLP recombinase in *C. elegans*. PLoS Genet. 2008;4(3):e1000028. doi:10.1371/journal.pgen.1000028.

Davis MW, Jorgensen EM. ApE, a plasmid editor: a freely available DNA manipulation and visualization program. Front Bioinformatics. 2022;2. doi:10.3389/fbinf.2022.818619.

Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat Methods. 2013;10(10):1028–1034. doi: 10.1038/nmeth.2641.

Dickinson DJ, Pani AM, Heppert JK, Higgins CD, Goldstein B. Streamlined genome engineering with a self-excising drug selection cassette. Genetics. 2015;200(4):1035–1049. doi:10.1534/genetics.115.178335.

Dickinson DJ, Goldstein B. CRISPR-based methods for *Caenorhabditis elegans* genome engineering. Genetics. 2016;202(3):885–901. doi: 10.1534/genetics.115.182162.

Dokshin GA, Ghanta KS, Piscopo KM, Mello CC. Robust genome editing with short single-stranded and long, partially single-stranded DNA donors in *Caenorhabditis elegans*. Genetics. 2018;210(3): 781–787. doi:10.1534/genetics.118.301532.

El Mouridi S, Lecroisey C, Tardy P, Mercier M, Leclercq-Blondel A, Zariohi N, Boulin T. Reliable CRISPR/Cas9 genome engineering in *Caenorhabditis elegans* using a single efficient sgRNA and an easily recognizable phenotype. G3 (Bethesda). 2017;7(5):1429–1437. doi: 10.1534/g3.117.040824.

El Mouridi S, Peng Y, Frøkjær-Jensen C. Characterizing a strong pan-muscular promoter (P*mlc-1*) as a fluorescent co-injection marker to select for single-copy insertions. MicroPubl Biol. 2020;2020. doi: 10.17912/micropub.biology.000302.

El Mouridi S, AlHarbi S, Frøkjær-Jensen C. A histamine-gated channel is an efficient negative selection marker for *C. elegans* transgenesis. MicroPubl Biol. 2021;2021. doi:10.17912/micropub.biology.000349.

Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One. 2008; 3(11):e3647. doi:10.1371/journal.pone.0003647.

Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One. 2009;4(5):e5553. doi:10.1371/journal.pone.0005553.

Etchberger JF, Hobert O. Vector-free DNA constructs improve transgene expression in *C. elegans*. Nat Methods. 2008;5(1):3. doi:10.1038/nmeth0108-3.

Evans T. Transformation and microinjection. WormBook; 2006. doi: 10.1895/wormbook.1.108.1.

Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics. 2015;199(4):959–971. doi:10.1534/genetics.115.175166.

Farboud B, Severson AF, Meyer BJ. Strategies for efficient genome editing using CRISPR-Cas9. Genetics. 2019;211(2):431–457. doi: 10.1534/genetics.118.301775.

Flytzanis NC, Bedbrook CN, Chiu H, Engqvist MKM, Xiao C, Chan KY, Sternberg PW, Arnold FH, Gradinaru V. Archaerhodopsin variants with enhanced voltage-sensitive fluorescence in mammalian and *Caenorhabditis elegans* neurons. Nat Commun. 2014;5: 4894. doi:10.1038/ncomms5894.

Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat Methods. 2013;10(8):741–743. doi:10.1038/nmeth.2532.

Frøkjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, Grunnet M, Jorgensen EM. Single-copy insertion of transgenes in *Caenorhabditis elegans*. Nat Genet. 2008;40(11): 1375–1383. doi:10.1038/ng.248.

Frøkjær-Jensen C, Davis MW, Ailion M, Jorgensen EM. Improved Mos1-mediated transgenesis in *C. elegans*. Nat Methods. 2012; 9(2):117–118. doi:10.1038/nmeth.1865.

Frøkjær-Jensen C. Exciting prospects for precise engineering of *Caenorhabditis elegans* genomes with CRISPR/Cas9. Genetics. 2013; 195(3):635–642. doi:10.1534/genetics.113.156521.

Frøkjær-Jensen C, Davis MW, Sarov M, Taylor J, Flibotte S, LaBella M, Pozniakovsky A, Moerman DG, Jorgensen EM. Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified Mos1 transposon. Nat Methods. 2014;11(5):529–534. doi: 10.1038/nmeth.2889.

Frøkjær-Jensen C. Transposon-assisted genetic engineering with Mos1-mediated single-copy insertion (MosSCI). Methods Mol Biol. 2015;1327:49–58. doi:10.1007/978-1-4939-2842-2_5.

Frøkjær-Jensen C, Jain N, Hansen L, Davis MW, Li Y, Zhao D, Rebora K, Millet JRM, Liu X, Kim SK, *et al.* An abundant class of non-coding

DNA can prevent stochastic gene silencing in the *C. elegans* germline. Cell. 2016;166(2):343–357. doi:10.1016/j.cell.2016.05.072.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6(5):343–345. doi:10.1038/nmeth.1318.

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, *et al.* Genome-scale CRISPR-mediated control of gene repression and activation. Cell. 2014;159(3):647–661. doi:10.1016/j.cell.2014.09.029.

Giordano-Santini R, Milstein S, Svrzikapa N, Tu D, Johnsen R, Baillie D, Vidal M, Dupuy D. An antibiotic selection marker for nematode transgenesis. Nat Methods. 2010;7(9):721–723. doi:10.1038/nmeth.1494.

Azimi Hashemi N, Bergs ACF, Schüler C, Scheiwe AR, Steuer Costa W, Bach M, Liewald JF, Gottschalk A. Rhodopsin-based voltage imaging tools for use in muscles and neurons of *Caenorhabditis elegans*. Proc Natl Acad Sci U S A. 2019;116(34):17051–17060. doi:10.1073/pnas.1902443116.

Heger P, Marin B, Schierenberg E. Loss of the insulator protein CTCF during nematode evolution. BMC Mol Biol. 2009;10:84. doi:10.1186/1471–2199-10–84.

Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn K-A, Minoda A, Tolstorukov MY, Appert A, *et al.* Comparative analysis of metazoan chromatin organization. Nature. 2014;512(7515):449–452. doi:10.1038/nature13415.

Katic I, Großhans H. Targeted heritable mutation and gene conversion by Cas9-CRISPR in *Caenorhabditis elegans*. Genetics. 2013;195(3):1173–1176. doi:10.1534/genetics.113.155754.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816–821. doi:10.1126/science.1225829.

Kaymak E, Farley BM, Hay SA, Li C, Ho S, Hartman DJ, Ryder SP. Efficient generation of transgenic reporter strains and analysis of expression patterns in *Caenorhabditis elegans* using library MosSCI. Dev Dyn. 2016;245(9):925–936. doi:10.1002/dvdy.24426.

Kelly WG, Xu S, Montgomery MK, Fire A. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. Genetics. 1997;146(1):227–238.

Kim H, Ishidate T, Ghanta KS, Seth M, Conte D, Shirayama MAK, Mello CC. A Co-CRISPR strategy for efficient genome editing in *Caenorhabditis elegans*. Genetics. 2014;197(4):1069–1080. doi:10.1534/genetics.114.166389.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722–736. doi:10.1101/gr.215087.116.

Leyva-Díaz E, Stefanakis N, Carrera I, Glenwinkel L, Wang G, Driscoll M, Hobert O. Silencing of repetitive DNA is controlled by a member of an unusual *Caenorhabditis elegans* gene family. Genetics. 2017;207(2):529–545. doi:10.1534/genetics.117.300134.

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio]; 2013.

Lin Z, Xie Y, Nong W, Ren X, Li R, Zhao Z, Hui JHL, Yuen KWY. Formation of artificial chromosomes in *Caenorhabditis elegans* and analyses of their segregation in mitosis, DNA sequence composition and holocentromere organization. Nucleic Acids Res. 2021;49(16):9174–9193. doi:10.1093/nar/gkab690.

Liu T, Rechtsteiner A, Egelhofer TA, Vielle A, Latorre I, Cheung M-S, Ercan S, Ikegami K, Jensen M, Kolasinska-Zwierz P, *et al.* Broad chromosomal domains of histone modification patterns in *C. elegans*. Genome Res. 2011;21(2):227–236. doi:10.1101/gr.115519.110.

Lo T-W, Pickle CS, Lin S, Ralston EJ, Gurling M, Schartner Ctl M, Bian Q, Doudna JA, Meyer BJ. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics. 2013;195(2):331–348. doi:10.1534/genetics.113.155382.

Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. Nat Rev Genet. 2020;21(10):597–614. doi:10.1038/s41576-020–0236-x.

Long L, Guo H, Yao D, Xiong K, Li Y, Liu P, Zhu Z, Liu D. Regulation of transcriptionally active genes via the catalytically inactive Cas9 in *C. elegans and D. rerio*. Cell Res. 2015;25(5):638–641. doi:10.1038/cr.2015.35.

Maduro M, Pilgrim D. Identification and cloning of *unc-119*, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics. 1995;141(3):977–988.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–826. doi:10.1126/science.1232033.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: a fast and versatile genome alignment system. PLoS Comput Biol. 2018;14(1):e1005944. doi:10.1371/journal.pcbi.1005944.

McDiarmid TA, Au V, Moerman DG, Rankin CH. *Peel-1* negative selection promotes screening-free CRISPR-Cas9 genome editing in *Caenorhabditis elegans*. PLoS One. 2020;15(9):e0238950. doi:10.1371/journal.pone.0238950.

Mello CC, Kramer JM, Stinchcomb D, Ambros V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 1991;10(12):3959–3970.

Mello C, Fire A. Chapter 19 DNA transformation. In: HF Epstein, DC Shakes, editors. Methods in Cell Biology, Caenorhabditis elegans: Modern Biological Analysis of an Organism. San Diego, California, USA: Academic Press; 1995. p. 451–482.

Merritt C, Rasoloson D, Ko D, Seydoux G. 3' UTRs are the primary regulators of gene expression in the *C. elegans* germline. Curr Biol. 2008;18(19):1476–1482. doi:10.1016/j.cub.2008.08.013.

Moreno-Mateos MA, Vejnar CE, Beaudoin J-D, Fernandez JP, Mis EK, Khokha MK, Giraldez AJ. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods. 2015;12(10):982–988. doi:10.1038/nmeth.3543.

Nance J, Frøkjær-Jensen C. The *Caenorhabditis elegans* transgenic toolbox. Genetics. 2019;212(4):959–990. doi:10.1534/genetics.119.301506.

Noma K, Jin Y. Rapid integration of multi-copy transgenes using optogenetic mutagenesis in *Caenorhabditis elegans*. G3 (Bethesda). 2018;8(6):2091–2097. doi:10.1534/g3.118.200158.

Nonet ML. Efficient transgenesis in *Caenorhabditis elegans* using Flp recombinase-mediated Cassette exchange. Genetics. 2020;215(4):903–921. doi:10.1534/genetics.120.303388.

Nonet M. Additional landing sites for recombination-mediated Cassette exchange in *C. elegans*. MicroPubl Biol. 2021;2021. doi:10.17912/micropub.biology.000503.

Okkema PG, Harrison SW, Plunger V, Aryana A, Fire A. Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. Genetics. 1993;135(2):385–404.

Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, Lu T, Smith J, Schmidt H, Krause MW, Seydoux G, *et al.* Scalable and versatile genome editing using linear DNAs with micro-homology to Cas9

sites in *Caenorhabditis elegans*. Genetics. 2014;198(4):1347–1356. doi:10.1534/genetics.114.170423.

Paix A, Folkmann A, Rasoloson D, Seydoux G. High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. Genetics. 2015; 201(1):47–54. doi:10.1534/genetics.115.179382.

Paix A, Schmidt H, Seydoux G. Cas9-assisted recombineering in *C. elegans*: genome editing using in vivo assembly of linear DNAs. Nucleic Acids Res. 2016;44(15):e128. doi:10.1093/nar/gkw502.

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, *et al.* Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv; 2018. doi:10.1101/201178 201178.

Praitis V, Casey E, Collar D, Austin J. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. Genetics. 2001;157(3): 1217–1226.

Priyadarshini M, Ni JZ, Vargas-Velazquez AM, Gu SG, Frøkjær-Jensen C. Reprogramming the piRNA pathway for multiplexed and transgenerational gene silencing in *C. elegans*. Nat Methods. 2022; 19(2):187–194. doi:10.1038/s41592-021–01369-z.

Radman I, Greiss S, Chin JW. Efficient and rapid *C. elegans* transgenesis by bombardment and hygromycin B selection. PLoS One. 2013; 8(10):e76019. doi:10.1371/journal.pone.0076019.

Robert V, Bessereau J-L. Targeted engineering of the *Caenorhabditis elegans* genome following Mos1-triggered chromosomal breaks. EMBO J. 2007;26(1):170–183. doi:10.1038/sj.emboj.7601463.

Schwartz ML, Jorgensen EM. SapTrap, a toolkit for high-throughput CRISPR/Cas9 gene modification in *Caenorhabditis elegans*. Genetics. 2016;202(4):1277–1288. doi:10.1534/genetics.115.184275.

Semple JI, Garcia-Verdugo R, Lehner B. Rapid selection of transgenic *C. elegans* using antibiotic resistance. Nat Methods. 2010;7(9): 725–727. doi:10.1038/nmeth.1495.

Silva-García CG, Lanjuin A, Heintz C, Dutta S, Clark NM, Mair WB. Single-copy knock-in loci for defined gene expression in *Caenorhabditis elegans*. G3 (Bethesda). 2019;9(7):2195–2198. doi: 10.1534/g3.119.400314.

Stevenson ZC, Moerdyk-Schauwecker MJ, Jamison B, Phillips PC. Rapid self-selecting and clone-free integration of transgenes into engineered CRISPR safe harbor locations in *Caenorhabditis elegans*. G3 (Bethesda). 2020;10(10):3775–3782. doi:10.1534/g3.120.401400.

Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. Extrachromosomal DNA transformation of *Caenorhabditis elegans*. Mol Cell Biol. 1985;5(12): 3484–3496.

Suzuki H, Kerr R, Bianchi L, Frøkjaer-Jensen C, Slone D, Xue J, Gerstbrein B, Driscoll M, Schafer WR. In vivo imaging of *C. elegans* mechanosensory neurons demonstrates a specific role for the MEC-4 channel in the process of gentle touch sensation. Neuron. 2003;39(6):1005–1017.

Tyson JR, O'Neil NJ, Jain M, Olsen HE, Hieter P, Snutch TP. MinION-based long-read sequencing and assembly extends the *Caenorhabditis elegans* reference genome. Genome Res. 2018;28(2): 266–274. doi:10.1101/gr.221184.117.

Tzur YB, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiácovo MP. Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 system. Genetics. 2013; 195(3):1181–1185. doi:10.1534/genetics.113.156075.

Vallin E, Gallagher J, Granger L, Martin E, Belougne J, Maurizio J, Duverger Y, Scaglione S, Borrel C, Cortier E, *et al.* A genome-wide

collection of Mos1 transposon insertion mutants for the *C. elegans* research community. PLoS One. 2012;7(2):e30482. doi: 10.1371/journal.pone.0030482.

Voutev R, Hubbard EJA. A "FLP-Out" system for controlled gene expression in *Caenorhabditis elegans*. Genetics. 2008;180(1):103–119. doi:10.1534/genetics.108.090274.

Waaijers S, Portegijs V, Kerver J, Lemmens BBLG, Tijsterman M, van den Heuvel S, Boxem M. CRISPR/Cas9-targeted mutagenesis in *Caenorhabditis elegans*. Genetics. 2013;195(3):1187–1191. doi: 10.1534/genetics.113.156299.

Wang H, Liu J, Gharib S, Chai CM, Schwarz EM, Pokala N, Sternberg PW. cGAL, a temperature-robust GAL4–UAS system for *Caenorhabditis elegans*. Nat Methods. 2017;14(2):145–148. doi: 10.1038/nmeth.4109.

Wang H, Liu J, Yuet KP, Hill AJ, Sternberg PW. Split cGAL, an intersectional strategy using a split intein for refined spatiotemporal transgene control in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A. 2018;115(15):3900–3905. doi:10.1073/pnas.1720063115.

Ward JD. Rapid and precise engineering of the *Caenorhabditis elegans* genome with lethal mutation co-conversion and inactivation of NHEJ repair. Genetics. 2015;199(2):363–377. doi:10.1534/genetics.114.172361.

Wei X, Potter CJ, Luo L, Shen K. Controlling gene expression with the Q repressible binary expression system in *Caenorhabditis elegans*. Nat Methods. 2012;9(4):391–395. doi:10.1038/nmeth.1929.

Wheeler BS, Anderson E, Frøkjær-Jensen C, Bian Q, Jorgensen E, Meyer BJ. Chromosome-wide mechanisms to decouple gene expression from gene dose during sex-chromosome evolution. Elife. 2016;5:e17365. doi:10.7554/eLife.17365.

Woglar A, Yamaya K, Roelens B, Boettiger A, Köhler S, Villeneuve AM. Quantitative cytogenetics reveals molecular stoichiometry and longitudinal organization of meiotic chromosome axes and loops. PLoS Biol. 2020;18(8):e3000817. doi:10.1371/journal.pbio.3000817.

Yemini E, Lin A, Nejatbakhsh A, Varol E, Sun R, Mena GE, Samuel ADT, Paninski L, Venkatachalam V, Hobert O, *et al.* NeuroPAL: a multicolor atlas for whole-brain neuronal identification in *C. elegans*. Cell. 2021;184(1):272–288.e11. doi:10.1016/j.cell.2020.12.012.

Yoshina S, Suehiro Y, Kage-Nakadai E, Mitani S. Locus-specific integration of extrachromosomal transgenes in *C. elegans* with the CRISPR/Cas9 system. Biochem Biophys Rep. 2016;5:70–76. doi: 10.1016/j.bbrep.2015.11.017.

Yoshina S, Mitani S. Integration of multicopy extrachromosomal transgenes into defined loci without phenotypes. microPubl Biol. 2022. doi:10.17912/micropub.biology.000571.

Yuen KWY, Nabeshima K, Oegema K, Desai A. Rapid de novo centromere formation occurs independently of heterochromatin protein 1 in *C. elegans* embryos. Curr Biol. 2011;21(21):1800–1807. doi: 10.1016/j.cub.2011.09.016.

Zeiser E, Frøkjær-Jensen C, Jorgensen E, Ahringer J. MosSCI and gateway compatible plasmid toolkit for constitutive and inducible expression of transgenes in the *C. elegans* germline. PLoS One. 2011; 6(5):e20082. doi:10.1371/journal.pone.0020082.

Zhang L, Ward JD, Cheng Z, Dernburg AF. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in *C. elegans*. Development. 2015;142(24):4374–4384. doi: 10.1242/dev.129635.

Communicating editor: S. Lee