Rapid Mapping and Identification of Mutations in *Caenorhabditis elegans* by Restriction Site-Associated DNA Mapping and Genomic Interval Pull-Down Sequencing

Sean M. O'Rourke, John Yochem, Amy A. Connolly, Meredith H. Price, Luke Carter, Joshua B. Lowry, Douglas W. Turnbull, Nick Kamps-Hughes, Nicholas Stiffler, Michael R. Miller, Eric A. Johnson, and Bruce Bowerman¹

Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403

ABSTRACT Forward genetic screens provide a powerful approach for inferring gene function on the basis of the phenotypes associated with mutated genes. However, determining the causal mutation by traditional mapping and candidate gene sequencing is often the rate-limiting step, especially when analyzing many mutants. We report two genomic approaches for more rapidly determining the identity of the affected genes in *Caenorhabditis elegans* mutants. First, we report our use of restriction site-associated DNA (RAD) polymorphism markers for rapidly mapping mutations after chemical mutagenesis and mutant isolation. Second, we describe our use of genomic interval pull-down sequencing (GIPS) to selectively capture and sequence megabase-sized portions of a mutant genome. Together, these two methods provide a rapid and cost-effective approach for positional cloning of *C. elegans* mutant loci, and are also applicable to other genetic model systems.

DETERMINING mutant gene identity is a key step for understanding gene function in forward genetic screens following mutagenesis and phenotype-based mutant isolation. In some organisms such as fungi and bacteria, a recessive mutant allele can be complemented with a plasmid-borne wild-type gene to establish gene identification. In organisms that lack robust DNA transformation methods, mapping with visible or selected single nucleotide polymorphism (SNP) markers to progressively finer genomic intervals is the traditional route to ascertain identity of the mutant gene. Now whole genome sequencing (WGS) methods can significantly reduce the time required to identify the causal mutation. For example, WGS can simply be used to determine all of the sequence alterations present in a mutant strain (Sarin *et al.* 2008; Smith *et al.* 2008; Srivatsan *et al.* 2008; Blumenstiel

et al. 2009; Irvine *et al.* 2009). However, some mapping data are still required to differentiate the background mutational load from the causal mutation. More recently, WGS has been performed on outcrossed mutant progeny to combine mapping and sequencing for pinpointing the position of the causal mutation (Doitsidou *et al.* 2010; Zuryin *et al.* 2010).

While resequencing a genome to identify mutant alleles is being used more frequently, in some cases it is more efficient to sequence only a portion of a genome. For example, sequencing of a single chromosome, a defined genomic interval, exonic sequences, or a single locus can be more cost effective when there is evidence that a mutation resides within a specific genome feature. There have been several throughput-enhancing advances in capturing targeted regions of a genome using DNA annealing since the first reported use of this methodology whereby individual microarray spots were physically scraped from the substrate (Ksiazek *et al.* 2003; Rota *et al.* 2003; Wang *et al.* 2003). For example, genomic DNA can be annealed to microarrays printed with oligonucleotides covering the region to be targeted, washed, and then eluted for sequencing (Albert *et al.* 2007; Hodges *et al.* 2007; Okou *et al.* 2007). Alternatively, oligonucleotides can be used to capture homologous genomic

Copyright © 2011 by the Genetics Society of America
doi: 10.1534/genetics.111.134031

Manuscript received June 29, 2011; accepted for publication August 25, 2011

Supporting information is available online at <http://www.genetics.org/content/suppl/2011/09/06/genetics.111.134031.DC1>.

¹Corresponding author: Institute of Molecular Biology, University of Oregon, 1370 Franklin Boulevard, Eugene, OR 97403. E-mail: bbowerman@molbio.uoregon.edu

DNA in solution (Gnirke *et al.* 2009). While these approaches are extremely high throughput, they also can be prohibitively expensive.

We have developed two Illumina-based sequencing methods in *Caenorhabditis elegans* that offer an alternative pipeline for mutation detection. First, we have performed restriction site associated DNA (RAD) polymorphism mapping to position the causal mutation to a relatively small region of the genome. Second, we have used genome interval pull-down sequencing (GIPS) to sequence a defined genomic interval. Genome intervals are captured by annealing sheared genomic DNA to sheared fosmids containing wild-type *C. elegans* DNA, eliminating the need for customized microarray or oligonucleotide production. Because multiple RAD mapping and genome interval sequencing samples can be combined in a single Illumina lane, it is possible to positionally clone and identify the mutant loci rapidly and cost effectively without performing WGS.

Materials and Methods

C. elegans strains and culture

Strains were grown under standard laboratory conditions (Brenner 1974). The temperature-sensitive (ts) mutants were maintained in a 15° incubator and shifted to a 26° incubator to perform temperature upshifts for determining embryonic lethality. Mutants were isolated in a *lin-2(e1309)* background, as previously described (Encalada *et al.* 2000).

Genetic crosses for RAD mapping

To map *or1167ts*, we crossed the polymorphic *C. elegans* strain CB4856 into the original mutagenized background [*or1167ts/or1167ts*; *lin-2(e1309)/lin-2(e1309)*]. After self-fertilization of the heterozygous F₁ outcross, we pooled 200 of the 1/16 of the F₂ progeny that were again *or1167ts/or1167ts*; *lin-2(e1309)/lin-2(e1309)*, taking advantage of the *lin-2* egg-laying defect to identify with a stereomicroscope within F₁ self-progeny *or1167ts/or1167ts*; *lin-2(e1309)/lin-2(e1309)* F₂'s filling up with dead embryos [avoiding laboriously singling out hundreds of F₂'s to look for production of dead embryos by egg-laying *lin-2(+/+ or +/e1309)* F₂ progeny]. Similarly, for mapping *unc-13*, we also performed a cross to CB4856 but selected ~200 *Unc* F₂ progeny for the genomic DNA preparations. For mapping *or1089ts*, we crossed the original mutant to CB4856 males and isolated 800 F₂ hermaphrodites that were tested for embryonic lethality. Approximately 200 homozygous *or1089ts* animals were recovered and used for the RAD mapping procedure.

Illumina sequencing for RAD mapping

Genomic DNA was isolated from pools of ~200 homozygous *unc-13* F₂'s, and ~200 *or1054ts*; *lin-2(e1309)* F₂'s as well as the N2 and CB4856 parental strains using the Qiagen DNeasy kit. A total of 150 ng of each sample was digested

with *Eco*RI and processed into barcoded RAD libraries as previously described (Baird *et al.* 2008) with the minor modification of using the paired end P2 adapter (Hohenlohe *et al.* 2010). Briefly, each sample was individually digested with *Eco*RI, and a P1 adapter (with a 4-bp barcode; see below) was ligated to the overhangs. After this step multiple samples were multiplexed. Next, the DNA was sheared and gel extracted to obtain ~400-bp fragments and the Illumina P2 adapter was ligated. Samples were then run on the Illumina flow cell. For RAD mapping, there is no need to use an Illumina kit (for full protocol see Baird *et al.* 2008). The RAD library from the mutant pool was sequenced at ≥30× coverage in an Illumina Genome Analyzer Ix machine. With SNPs present at about every 1000 bp in the polymorphic CB4856 strain, and sequencing reads of ~75 bp from each *Eco*RI site, we anticipated detecting a SNP near 1 in 10 *Eco*RI sites, or about one every 50,000 bp, which was close to the observed value of one SNP in 64,000 bp achieved, on average. The RAD sequences were aligned to the reference Bristol N2 genome using the Bowtie software package (Langmead *et al.* 2009). The Bowtie output was then exported to SAMtools (Li *et al.* 2009) and converted into BAM format. We then produced a pileup file, to which we applied the samtools.pl script “varFilter” command (using default options) to identify SNPs. The varFilter results were then saved as a tab-delimited file for use with graphing software (Microsoft Excel and Adobe Illustrator). As an alternative method for identifying N2/CB4856 SNPs, one could use the MAQGene program (Bigelow *et al.* 2009), which may be more accessible to nonbioinformaticians (as used by Doitsidou *et al.* 2010).

Illumina sequencing of genomic intervals

To pull down intervals of genomic DNA to which *or195ts*, semidominant *or600(sd)*,ts, and *or683ts* were mapped, we used magnetic bead pull-downs. A total of 5 µg of genomic DNA was purified from each mutant strain using a DNeasy Blood and Tissue kit (Qiagen) and sheared to an average size of 500 bp by sonication in a Bioruptor (Diagenode). The ends of the sheared DNA fragments were blunted using a QuickBlunt kit (New England Biolabs) and the fragments purified with a PCR purification kit (Qiagen). A-overhangs were added to the genomic fragments by incubation of the purified, blunted DNA with 150 units of Klenow DNA polymerase exo- (New England Biolabs) and dATP at 37° for 30 min. The modified fragments were purified with a mini-elute PCR purification kit (Qiagen). A total of 7 µl of 1-µM modified Illumina sequencing adapters were ligated to the sheared genomic fragments at 16° for 2 hr using 2000 units of T4 DNA ligase (New England Biolabs). [Top strand: 5' ACACCTCTTTCCCTACACGACGCTCTTCCGATCxxxx*T 3'; bottom strand: 5' phosphate xxxxGATCGGAAGAGCGGTT-CAGCAGGAATGCCGAG 3' (where x indicates the barcode bases and *, phosphorotioate bond.) The ligation reaction was size separated by agarose gel electrophoresis, and fragments between 150 and 500 bp in size were purified from the gel using a Gel Extraction kit (Qiagen). The purified

ligation products were PCR amplified using Phusion high fidelity DNA polymerase (New England Biolabs) and the Illumina amplification primers 5' AATGATACGGCGACCACCGAG ATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT 3' and 5' CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCT GCTGAACCGCTCTTCCGATCT 3'. The following cycling conditions were used for PCR: 98° for 2 min, 15 cycles of 98° for 10 sec, 65° for 30 sec, and 72° for 15 sec. Following amplification, samples were size separated by agarose gel electrophoresis, and fragments between 150 and 500 bp were purified with a Gel Extraction kit (Qiagen).

Biotinylated probe preparation: DNA preps homologous to the targeted genomic interval were prepared from genomic fosmids, using a nearly genome-wide fosmid library for *C. elegans* that was developed by the Genome Sciences Centre in Vancouver, BC, Canada. A total of 100 ng of the fosmid DNA mixtures were combined with 20 μ l of 2.5 $\times$ random octamer solution (Life Technologies) and heated to 100° for 5 min. The mixture was rapidly cooled in an ice/water bath, following which 5 μ l of biotin dNTP mixture [1 mM biotin-14-dCTP, 1 mM dCTP, 2 mM dATP, 2 mM dGTP, and 2 mM dGTP, in 10 mM Tris-HCl (pH 7.5), 1 mM Na₂ EDTA] (Life Technologies) was added, along with 1 μ l of Klenow fragment DNA polymerase (Life Technologies) and ultrapure water to bring the reaction volume to 50 μ l. The reaction was then incubated at 37° for 1 hr, following which, the products were size separated on an agarose gel and the predominant 100-bp product was purified with a Gel Extraction kit (Qiagen).

Streptavidin bead preparation: A total of 50 μ l of M270 streptavidin Dynabeads (Life Technologies) were washed three times with 100 μ l of 6 $\times$ SSC and resuspended in 100 μ l of bead block buffer [2% I-Block (Tropix), 0.5% SDS, 1 $\times$ PBS]. Beads were incubated at room temperature for 30 min with occasional mixing and were then magnetically captured and washed three times with 6 $\times$ SSC.

Hybridization, immobilization, elution, and sequencing:

A total of 5 μ g of adapted, purified genomic DNA was combined with 150 ng of purified biotinylated probe in 300 μ l of hybridization buffer [54% formamide, 1 $\times$ SSC, 1% SDS, 5.4 $\times$ Denhardt's solution (Sigma), 1 mg/ml Salmon sperm DNA (Life Technologies)]. The mixture was heated to 100° for 2 min and then transferred to a 42° incubator, where it was incubated with mixing overnight. Following overnight incubation, biotinylated probe/genomic DNA fragment hybrids were immobilized by binding to prepared blocked and washed streptavidin beads by combining the hybridization mixture (300 μ l) with the bead/SSC mixture (100 μ l), and incubating at room temperature for 15 min with occasional mixing. Beads were then magnetically captured, and washed three times with wash solution 1 (1 $\times$ SSC, 0.15% SDS), three times with wash solution 2 (0.2 $\times$ SSC), and three times with wash solution 3 (0.05 $\times$ SSC). After the final wash step, the beads were resuspended in 200 μ l of ultra-

pure water, heated to 100° for 2 min, and quickly magnetically captured. The supernatant was carefully collected and concentrated to a volume of 20 μ l in a Speedvac concentrator (Savant). A total of 10 μ l of the concentrated supernatant was then used as template for a PCR reaction utilizing Illumina amplification primers and Phusion high fidelity DNA polymerase (New England Biolabs) (2 min 98°, 24 cycles of 98° for 10 sec, 65° for 30 sec, and 72° for 15 sec). The PCR products were purified with a PCR cleanup kit (Qiagen), quantified, and submitted for Illumina sequencing on an Illumina Genome Analyzer II.

Results

RAD mapping of C. elegans mutations

To rapidly map *C. elegans* mutations, we have used an Illumina sequencing-based genome-wide single nucleotide polymorphism mapping procedure called RAD polymorphism mapping (Lewis *et al.* 2007; Miller *et al.* 2007; Baird *et al.* 2008). RAD markers are SNPs adjacent to restriction enzyme recognition sequences in the genomes of divergent strains. In our case, we used the N2 background (isolated in Bristol, UK) to isolate mutants and subsequently crossed them to the polymorphic Hawaiian CB4856 strain for mapping. The N2 and CB4856 genome sequences have diverged substantially but their hybrid progeny are fertile. On average, there is a SNP approximately every 1 kb, allowing physical mapping using a large number of markers.

To experimentally identify RAD tags, we crossed wild-type N2 hermaphrodites to CB4856 males. F₁ hybrid progeny were isolated and genomic DNA was digested with *Eco*RI. After ligation of Illumina adapters and selective amplification of the RAD tags (Figure 1), Illumina sequencing was performed with an Illumina Genome Analyzer IIx system. Selective amplification was carried out by using a "Y" adapter for the P2 adapter, which prevents fragments that lack a P1 adapter from being amplified after first round synthesis initiated from the P1 site (as described in Coyne *et al.* 2004; Baird *et al.* 2008). We detected 3,462 SNPs with an average distance between them being 29 kb. Most SNPs (95%) were separated from an adjacent SNP by <100 kb (Figure 2). The largest distance separating adjacent SNPs occurred near the center of chromosome V (515 kb). Most of the SNPs we identified could be predicted *in silico* from the sequence of the CB4856 strain (data not shown). Because the sequencing is done with purified RAD tags instead of total genomic DNA, multiple samples can be multiplexed on a single lane in an Illumina sequencer, with each sample containing unique barcodes for subsequent sequence data deconvolution. The barcodes used for RAD mapping are 6-bp sequences added to the P1 adapter primer. In one test, we used one Illumina Genome Analyzer IIx lane to process 13 RAD mapping crosses. We used single-end sequencing with 80-bp reads, to achieve 50 million reads yielding ~40 $\times$ coverage. We tested the applicability of RAD mapping coupled with Illumina sequencing using three different approaches.

Figure 1 Restriction site-associated DNA (RAD) mapping schematic. Genomic DNA was isolated from 200 pooled F_2 progeny in crosses between the N2 mutants and the polymorphic CB4856 (Hawaiian) strain. Genomic DNA was digested with *EcoRI* and P1 Illumina adaptors were ligated to the fragments. This DNA was mechanically sheared and Illumina P2 adaptors were ligated to the fragment ends. Next, RAD tags were selectively amplified and sequenced from the Illumina sequencing primer site encoded on the P2 adaptor on an Illumina Genome Analyzer IIx machine.

First, we mapped a known mutant, *unc-13(e450)*. We crossed the *unc-13* mutant to the polymorphic *C. elegans* strain CB4856 and pooled 200 F_2 progeny that were homozygous for the *unc-13* mutation. We chose 200 recombinants as a goal because it afforded a relatively large number of independent recombination events, although it is possible that using fewer recombinant F_2 's would also yield sufficient resolution. After producing a RAD library from the F_2 genomic DNA sample, we performed sequencing on the Illumina machine to detect SNPs across the genome (see *Materials and Methods*). We used graphing software (Microsoft Excel and Adobe Illustrator) to plot the ratio of CB4856/Bristol SNPs across the *C. elegans* genomic sequence, indicating the fraction of samples for any one SNP that correspond to the polymorphic CB4856 sequence (Figure 3A). In this test, we identified 683 RAD tags throughout the genome. The ratio of CB4856 to N2 SNPs was ~ 0.5 across the genome, except for chromosome I, where a large trough was present. The center of the trough on chromosome I is within ~ 800 kb of the known location of *unc-13* (Figure 3B). We conclude that

Illumina-based RAD mapping can quickly provide the approximate physical position of mutant loci.

We similarly applied RAD mapping to *or1089ts*, a mutant of unknown molecular identity with severe defects in mitotic spindle assembly in one-cell-stage embryos (Figure 4A). In this case, we identified 3134 RAD markers in the F_2 RAD library. Chromosome I was highly enriched for N2 DNA (Figure 4B), with the trough of N2 DNA centered at 9.25 Mb on chromosome I (Figure 4C), positioning the *or1089ts* mutation near the center of chromosome I. In a 1-Mb region centered on the trough, we found one candidate gene in online databases, *spd-2*, that when reduced in function using RNAi, results in defects that closely resemble the *or1089ts* mutant phenotype (Sonnichsen *et al.* 2005; Harris *et al.* 2010). The center of the reduced CB4856 ratio is 276 kb to the left of the known location of *spd-2* (Figure 4C). We Sanger sequenced the *spd-2* gene in genomic DNA from *or1129ts* mutants after amplifying the locus using gene-specific primers. We found a single nucleotide mutation that changes a phenylalanine codon to an isoleucine codon at

Figure 2 *EcoRI*-associated RAD tag locations. (A) RAD mapping results from an N2/CB4856 cross. Vertical lines represent *EcoRI*-associated RAD markers on each of the *C. elegans* chromosomes. The total chromosome sizes are listed on the right. For a list of the RAD markers, sequences, and positions, see File S1. (B) Distance between adjacent RAD markers. The plot represents the number of RAD marker pairs at the indicated distances, in kilobase pairs.

Figure 3 RAD mapping results for *unc-13*. (A) Genome-wide RAD mapping results for *unc-13* after crossing it to CB4856. A total of 683 SNPs across the genome in F_2 progeny were detected and the ratio of the CB4856 SNPs plotted along the chromosomes. The vertical axis represents the percentage of CB4856 SNPs in the F_2 population. (B) Magnification of chromosome I. The trough on chromosome I correlates with the known location of *unc-13*.

amino acid 544 (Figure 4D). We also found that *or1089ts* failed to complement *spd-2(or293ts)* (data not shown). We conclude that *or1089ts* is a new *spd-2* allele.

We have screened for temperature-sensitive, embryonic-lethal *C. elegans* mutants using an egg laying-defective (*Egl*) *lin-2(e1309)* mutant background that enables one to screen mutagenized populations for animals that produce inviable embryos without singling out individual worms (see also Kempfues *et al.* 1988; Encalada *et al.* 2000; Jorgensen and Mango 2002; O'Rourke *et al.* 2011). To take further advantage of this *Egl* background for RAD mapping, we crossed *or1167ts*; *lin-2(e1309)* hermaphrodites to CB4856 males and obtained F_2 animals that were mutant for both *lin-2(e1309)* and *or1167ts*. Instead of testing embryonic lethality of ~800 F_2 animals individually on plates, as was done for *spd-2(or1089ts)*, we were able to isolate ~200 *or1167ts*; *lin-2(e1309)* animals that accumulated mostly dead embryos from a mixed F_2 population after shifting them to the restrictive temperature as L4 larvae. After preparing and sequencing the RAD library, we identified >3400 RAD tags. The ratios of CB4856 to N2 DNA were plotted and we found two troughs that correspond to an enrichment of N2 DNA (Figure 5). The trough on the X chromosome is 550 kb from the known location of *lin-2*, while the trough on the right arm of chromosome IV presumably correlates with the

location of *or1167ts*. The fact that the *lin-2* trough does not reach zero likely relates to inadvertently picking some worms that were in fact not *lin-2*, as even wild-type worms sometimes hold their embryos and can appear *Egl*. We found one candidate gene, *sas-6*, positioned 250 kb left of the center of the chromosome IV trough. Like animals depleted for *sas-6*, *or1167ts* animals shifted to the restrictive temperature as L4 larvae produce embryos that appear to assemble monopolar mitotic spindles in early embryonic cells (not shown). We sequenced the *sas-6* locus with the Sanger method after amplification of the region by PCR. We found a single missense mutation that changes an aspartic acid to a valine in the ninth amino acid of SAS-6 (Figure 5D). As *or1167ts* also failed to complement a known *sas-6* mutant (not shown), we conclude that *or1167ts* is a *sas-6* allele. This example demonstrates that RAD mapping can be easily applied in the context of the *lin-2(e1309)* marker to minimize the effort required to isolate F_2 animals that are homozygous for embryonic lethal mutations. We continue to explore the use of RAD methodologies to rapidly map temperature-sensitive embryonic lethal *C. elegans* mutants and note that this approach can be used to map virtually any locus that can be assayed in N2/CB4856 F_2 animals.

Illumina-based GIPS

To quickly and more cost-effectively identify causal mutations in mutant strains, we have also applied Illumina DNA sequencing to defined genomic intervals, rather than sequencing entire mutant genomes (as has been done to identify some mutant loci in *C. elegans* (Sarin *et al.* 2008; Zuryn *et al.* 2010)). Sequencing an entire genome involves more cost than our procedure because we multiplexed multiple barcoded sequencing experiments on a single Illumina Genome Analyzer IIx flow cell. Briefly, we used wild-type genomic DNA from defined genomic intervals, linked to magnetic beads, to partially purify regions of mutant genomic DNA (see Figure 6A and *Materials and Methods*). We first tested the feasibility of using interval pull-downs and Illumina sequencing by resequencing a previously identified mutation present in the *dhc-1* locus of *dhc-1(or195ts)* mutants worms (O'Rourke *et al.* 2007). We then identified the mutations responsible for conditional lethality in two previously reported mutants [*tbb-2(or600sd,ts)* and *plk-1(or683ts)*; O'Rourke *et al.* 2011], after sequencing genomic intervals of 1.8 and 1.3 Mb, respectively (Figure 7).

To test this methodology, we first selected a fosmid that includes the entire *dhc-1* locus, available from Source BioScience (<http://lifesciences.sourcebioscience.com>). We purified the fosmid, sheared it, and linked the fragments to biotinylated beads as described in *Materials and Methods*. After using these beads to isolate sheared *dhc-1(or195ts)* genomic DNA, the mutant genomic DNA was eluted and prepared for Illumina sequencing (see *Materials and Methods*). We aligned 66,853 30-base reads to the 30.1 kb *dhc-1*-containing fosmid on chromosome I and the average

Figure 4 Phenotype and RAD mapping of *spd-2(or1089ts)*. (A) Defective mitotic spindle formation and cytokinesis failure in an early *or1089ts* embryo produced from a worm shifted to 26° for 6 hr (lower) compared to a wild-type embryo (upper). (B) Genome-wide RAD mapping of *or1089ts*. A reduction of CB4856 DNA on chromosome I results from the selection of the mutant homozygotes. (C) The center of the trough is near 9 Mb and is positioned near the causal mutation (see text). (D) DNA sequence analysis of *or1089ts* identifies a mutation in *spd-2* open reading frame (GenBank: AY340594.1). The mutation, causing a phenylalanine-to-isoleucine change in codon 544, is shown relative to the changes of *or183* and *or188*, two known temperature-sensitive alleles of *spd-2*.

coverage for each position in the fosmid was 66×. For comparison, we also identified 570,311 reads that could be aligned to the rest of the genome, yielding an average read coverage of 0.17× for each nucleotide position. Therefore, we achieved a 388-fold enrichment for reads in the targeted region using our interval pull-down sequencing method. We identified the previously sequenced C-to-T *dhc-1(or195ts)* mutation in a total of 66 reads, and no other mutations were detected in the *dhc-1* locus (Figure 7A). We conclude that

GIPS can readily identify the mutations present in relatively large regions of the genome.

Next, we used GIPS to identify the mutation responsible for the early embryonic cell division defects caused by the semidominant, temperature-sensitive mutation *or600sd,ts*. We defined the *or600sd,ts* interval using standard mapping crosses with visible morphological and behavioral markers. We localized the mutation to chromosome III between positions 3,618,381 and 5,447,436. We used the WormBase

Figure 5 RAD mapping results for a *sas-6(or1167ts); lin-2* double mutant. *or1167ts; lin-2* mutants were crossed to CB4856 males and F₂ progeny, homozygous for both the *or1167ts* embryonic lethal mutation and the *lin-2* mutation, were isolated. (A) Genome-wide RAD mapping results show enrichment for N2 DNA on chromosomes IV and X. (B) The *sas-6* gene is located near the chromosome IV trough. (C) The trough on the X chromosome lies near the known position of the *lin-2* locus. (D) The *sas-6* locus contains a missense mutation that changes an aspartic acid codon to a valine (NCBI: NP_502660.1). An alignment was performed with the *C. elegans* (Cel) wild-type *SAS-6*, *C. elegans* *SAS-6(or1167ts)*, *C. remanei* (Cre) *SAS-6*, and *C. briggsae* (Cbr) *SAS-6*.

Figure 6 Genome interval pull-down sequencing (GIPS) using the Illumina platform. (A) Schematic overview of the interval pull-down sequencing method. First, fosmids of wild-type DNA covering a region of the genome are purified, sheared, and ligated to biotinylated adapters. Next, mutant genomic DNA is sheared and annealed to the biotinylated fosmids. After purification and release of the mutant DNA using magnetic beads, the fragments are subjected to sequencing on an Illumina machine. Finally, the reads are assembled onto the genome scaffold and polymor-

phisms are identified. (B) Example output of mutant genome assembly. Shown is a small region of the *tbb-2* locus with portions of reads aligned beneath the reference sequence. In this case, each read shows that a cytosine in wild type has been changed to a thymidine in *tbb-2(or600sd,ts)*. Also shown is one apparent sequencing error where an A > G change was called in one of the reads. Nucleotides are color coded, uppercase letters, and periods in the sequence reads represent identity with the reference sequence, while reads containing lowercase letters and commas represent sequence data obtained from the reverse complement strand.

genome browser to identify a minimal tiling path using genomic DNA from fosmid clones available from Source BioScience (<http://lifesciences.sourcebioscience.com>). We identified 65 fosmids that spanned the region with 7 gaps that totaled <45 kb (~2.5% of the region). We purified, sheared, and linked the fosmid fragments to biotinylated beads. After using these beads to isolate sheared *or600sd,ts* genomic DNA, the mutant genomic DNA was eluted and prepared for Illumina sequencing (see *Materials and Methods*). We found 1,596,403 48-base reads that could be aligned to the region, giving an average coverage for each nucleotide in the interval of 42x. The total number of reads corresponding to the *C. elegans* genome was 6,034,221.

We used software (SAMtools) to output a text file that lists the mutations identified in the interval. One can also view the sequence reads aligned to the reference wild-type genome sequence with SAMtools (Figure 6B). There were 45 mutations in the 1.8-Mb *or600sd,ts* interval (Figure 7B). We found 14 extragenic changes, 23 intronic changes, one mutation in a pseudogene, two mutations that cause the transcript to go out of frame, three potential annotation errors, and two missense mutations (Table 1). As >90% of our identified temperature-sensitive mutations are caused by missense mutations (the remaining are due to premature stop codons, small deletions, and mutations in splice-site boundaries) (O'Rourke *et al.* 2011), we narrowed our

Figure 7 Analysis of the genome interval pull-down sequencing. The protocol shown was applied to three mutants. The positions of nucleotide alterations vs. the reference sequence is shown in graphical form. Each line represents a single change, diagonal lines attached to the top of the vertical lines represent multiple changes located very close together. Lines are color coded to show various types of mutations. The ^ in the black bars point to the causative mutations. (A) Sequencing results for the 15 kb *dhc-1(or195ts)* locus purified using a single genomic fosmid clone. The previously identified missense mutation was found. (B) Sequencing results for a 1.82-Mb region on chromosome III in the *or600sd,ts* genome. We identified 45 total mutations in the region including the missense mutation in *tbb-2* (see Table 1 for details). (C) Sequencing results for a 1.3-Mb region on chromosome III in the *or683ts* genome. We identified 29 total mutations in the region including the missense mutation in *plk-1* (see Table 2 for details).

Table 1 Mutations present in the *or600sd,ts* interval

Position ^a	Gene	WT ^b	Alteration ^c	Notes ^d	Reads ^e
3,645,769	<i>unc-93</i>	C	A	Intron	20
3,645,821	<i>unc-93</i>	T	G	Intron	58
3,716,562	C14B1.6	*	C insertion	Intron	22
3,805,251	<i>toh-1</i>	C	Single base deletion	Intron	591
3,830,219	Intergenic	*	G insertion		23
3,885,443	<i>tir-1</i>	*	G insertion	Intron	30
3,971,283	<i>msi-1</i>	T	A	Intron	45
3,997,832	Intergenic	A	T		142
4,007,014	<i>pdx-1</i>	*	C insertion	Intron	22
4,017,108	<i>tbb-2</i>	C	T	Missense G > E, E	614
4,019,053	Intergenic	*	C insertion		920
4,019,157	Intergenic	*	G insertion		1104
4,032,835	<i>hum-5</i>	*	C insertion	Intron	72
4,056,056	<i>rcq-5</i>	A	T	Intron	28
4,056,057	<i>rcq-5</i>	T	A	Intron	28
4,056,058	<i>rcq-5</i>	A	C	Intron	29
4,059,721	<i>his-70</i>	T	Single base deletion	Altered C terminus, NE	112
4,074,141	<i>unc-79</i>	T	Single base deletion	Intron	341
4,091,848	C03C10.2	*	C insertion	Intron	122
4,209,222	Intergenic	C	A		1959
4,209,223	Intergenic	A	C		2105
4,209,244	Intergenic	C	A		2694
4,209,245	Intergenic	G	C		2546
4,210,819	F43C1.1	*	G insertion	Intron	716
4,286,570	R10E4.3	C	T	Intron	30
4,445,015	Intergenic	*	G insertion		24
4,447,068	<i>ceh-43</i>	*	G insertion	Intron	81
4,447,641	<i>ceh-43</i>	*	C insertion	Intron	83
4,502,805	<i>cct-5</i>	*	G insertion	Intron	1140
4,530,454	<i>cdh-4</i>	A	Single base deletion	Intron	55
4,605,001	<i>sel-2</i>	T	Single base deletion	Intron	25
4,612,783	Intergenic	*	C insertion		411
4,640,496	<i>clec-154</i>	*	G insertion	Altered C terminus, NE	195
4,684,393	T04A8.3	G	Single base deletion	Pseudogene	45
4,687,673	Intergenic	*	C insertion		82
4,689,556	T04A8.5	*	C insertion	Intron	64
4,708,949	Intergenic	*	C insertion		780
4,712,953	T04A8.15	C	G ^f	Annotation error?	626
4,738,437	<i>ras-2</i>	C	T	Missense G > R, NE	53
4,808,804	Intergenic	*	C insertion		91
4,947,903	Intergenic	*	C insertion		59
4,962,073	C27F2.5.1	T	Single base deletion	Intron	44
5,105,425	F54D8.6	*	C insertion	Intron	280
5,105,939	F54D8.6	*	G insertion ^f	Annotation error?	133
5,333,260	F52C9.1	*	C insertion ^f	Annotation error?	977

^a Chromosome III position from the WS200 reference version of WormBase.^b The wild-type sequence.^c The changes found in the *or600* genome. For insertions, the alteration takes place immediately following the indicated position.^d For mutations that cause protein coding changes, we have classified the loci as NE for not essential and E for essential.^e The total number of Illumina reads for the position. Only mutations with ≥20 reads were considered and furthermore, >75% of the reads had to show the alteration presented in column 4.^f These apparent mutations may represent the wild-type sequence in some backgrounds because expressed sequence tags also show these alterations.

analysis of mutations in exon and intron splice sites. Three mutations that appeared to cause exonic changes are likely not the causal mutation because they were present in wild-type DNA; expressed sequence tags show the same mutations, perhaps indicating errors in the reference sequence (we have labeled these as “annotation error?” in Table 1). None of the intronic changes occurred at intron boundaries and thus are unlikely to interfere with RNA splicing. The

four remaining exonic mutations occur in the *tbb-2*, *ras-2*, *his-70*, and *clec-154* genes. Single nucleotide deletion and insertion mutations in the *his-70* and *clec-154* loci, respectively, encode proteins with altered C termini, while the *ras-2* and *tbb-2* mutations are missense. Of these four genes, only RNAi that targets the *tbb-2* locus phenocopies the *or600sd,ts* early embryonic phenotype [note that *tbb-2(RNAi)* also depletes the paralogous redundant *tbb-1* gene product].

Table 2 Mutations present in the *or683ts* interval

Position ^a	Gene	WT ^b	Alteration ^c	Notes ^d	Reads ^e
7,029,617	Intergenic	*	A insertion		11
7,047,283	Intergenic	*	C insertion		20
7,063,069	R01H2.4	*	C insertion	Altered C terminus, NE	15
7,214,105	R151.2	A	C ^f	Annotation error?	27
7,228,289	R151.8	AAC	Deletion	Intron	11
7,272,182	<i>psd-1</i>	*	G insertion	Intron	26
7,297,779	B0361.2	*	AA insertion	Intron	10
7,318,365	F56C9.8	*	TT insertion	Intron	11
7,587,395	<i>cup-5</i>	G	Deletion	Intron	12
7,628,358	<i>fbn-1</i>	TG	Deletion	Altered C terminus, E	11
7,686,828	C18H2.3	*	TA insertion	3'-UTR	10
7,696,424	C18H2.4	TA	Deletion	Altered C terminus, NE	12
7,712,933	Intergenic	*	G insertion		17
7,781,836	<i>mab-5</i>	AG	Deletion	Intron	11
7,864,173	C02C2.4	C	Deletion	Intron	23
7,919,577	<i>tag-250</i>	*	GA insertion	Altered C terminus, NE	12
7,937,298	<i>ran-2</i>	TC	Deletion	3'-UTR	11
7,958,664	F54H12.3	C	T	Transposon	15
8,042,705	K12H4.6	*	G insertion	Altered C terminus, NE	32
8,053,797	Intergenic	C	Deletion		10
8,059,766	Intergenic	C	T		21
8,064,605	Intergenic	C	T		17
8,068,402	<i>ceh-26</i>	*	G insertion ^f	Annotation error?	11
8,070,697	<i>ceh-26</i>	*	G insertion	Intron	25
8,075,189	<i>dcr-1</i>	*	GG insertion	Altered C terminus, E	19
8,103,295	<i>plk-1</i>	T	A	Missense M > K, E	12
8,127,908	Intergenic	*	T insertion		12
8,128,979	Intergenic	*	C insertion		30
8,154,518	Intergenic	T	A		10

^a Chromosome III position from the WS200 reference version of Wormbase.

^b The wild-type sequence.

^c The changes found in the *or683* genome. For insertions, the alteration takes place immediately following the indicated position.

^d For mutations that cause protein coding changes, we have classified the loci as NE for not essential and E for essential.

^e The total number of Illumina reads for the position. Only mutations with ≥ 10 reads were considered and furthermore, >90% of the reads had to show the alteration presented in column 4.

^f These apparent mutations may represent the wild-type sequence in some backgrounds because expressed sequence tags available from WormBase also show these alterations.

Depletion of the other three genes by RNAi does not result in any lethal phenotypes (WormBase). On the basis of sequence data, the embryonic phenotype and genetic interactions with a previously isolated *tbb-2* allele, *or362sd,ts* (O'Rourke *et al.* 2011), we conclude that *or600sd,ts* is a *tbb-2* allele.

We also identified a recessive temperature-sensitive embryonic lethal mutant, *or683ts*, that mapped between 6,862,157 and 8,214,712 on chromosome III, on the basis of mapping crosses with visible markers and some PCR-based SNP mapping (not shown). To identify all of the mutations in the interval, we again performed GIPS, as detailed for *dhc-1(or195ts)* and *tbb-2(or600sd,ts)*. There were 197,752 48-base reads (from a total of 5,616,109 reads) that aligned to the region, giving an average coverage of 7× for the interval. Even though the sequencing coverage was greatly reduced compared to the *dhc-1(or195ts)* and *or600sd,ts* results, we were still able to identify 29 mutations in the interval after performing Illumina sequencing (Figure 7C). We found nine intergenic changes, eight mutations in introns, two mutations in 3'-untranslated regions, one mutation in a transposon, six insertions or deletions that

caused frame shifting, two potential annotation errors, and one missense allele (Table 2). The annotation errors may represent the wild-type sequence because expressed sequence tags show this same alteration. The single missense mutation changes a methionine codon to a lysine codon in the *plk-1* gene. Since *or683ts* contains a mutation in the *plk-1* gene and fails to complement a known allele of *plk-1* (O'Rourke *et al.* 2011), we conclude that *or683ts* is a novel *plk-1* allele.

Discussion

The utility of *C. elegans* as an animal model, in which one can readily isolate temperature-sensitive mutations in essential genes, and the power of next generation DNA sequencing for greatly reducing the time required to positionally clone mutant loci, now make it possible to much more rapidly isolate experimentally useful conditional mutations in essential genes. Our two new Illumina-based sequencing methods should allow for increased throughput when analyzing large numbers of mutants.

RAD polymorphism mapping has been used successfully to map genes in *Drosophila* (Miller *et al.* 2007), threespine stickleback fish (Miller *et al.* 2007), *Neurospora* (Lewis *et al.* 2007), diamondback moths (Baxter *et al.* 2011), barley (Chutimanitsakun *et al.* 2011) and now *C. elegans*. So long as a hybrid strain is available to generate F₂ progeny, the methodology should be feasible in any organism. RAD mapping makes it possible to simultaneously and rapidly determine the approximate location of large numbers of mutations isolated after mutagenesis. We have used *Eco*RI to cut the genomic DNA because it provides relatively good resolution between RAD markers (Figure 1). However, any other restriction enzyme could also be used, or multiple enzymes could be used to gain increased mapping resolution. RAD mapping provides an ideal procedure to identify mutant loci when the number of gene candidates is limited.

We explored the use of RAD mapping in three different contexts. First we have performed a proof-of-principle experiment where we mapped the position of *unc-13*, a known mutant (Figure 3). While the numbers of RAD markers obtained was relatively low in this example (683, presumably because we experienced some sample loss), the center of the trough on chromosome I is still within 800 kb of the known position of *unc-13*. In the second approach, we used RAD mapping to clone the *or1089ts* mutant. After picking 800 F₂ animals from an *or1089ts*/CB4856 cross, we identified those that were homozygous for the *or1089ts* mutation. We found a substantial enrichment of the N2 DNA on the center of chromosome I that was positioned within 276 kb from the known position of the *spd-2* locus. We performed Sanger DNA sequencing on PCR products derived from the *spd-2* locus and identified one sequence alteration that causes a missense mutation (Figure 4D). Thus, RAD mapping can rapidly identify candidate genes that can be further investigated by sequencing candidate genes, GIPS, or complementation tests with existing mutants. We also tested the feasibility of performing RAD mapping on F₂ animals that were doubly mutant for an embryonic lethal mutation and an egg-laying defective mutant, *lin-2* (present in the original mutagenized strain). Since homozygous *lin-2* animals hold their embryos, it was possible to more easily and rapidly select F₂ progeny homozygous for the mutation being mapped from an *or1167ts*; *lin-2* and CB4856 cross (Figure 5). As expected, we found two regions of the genome that were enriched for N2 DNA: one corresponds to the *lin-2* locus, while the second corresponds to the *or1167ts* mutation in the *sas-6* locus. Finally, we are currently exploring the use of RAD mapping by crossing temperature-sensitive mutants to CB4856 males and allowing the progeny to reproduce at the nonpermissive temperature for many generations. This method may significantly reduce the labor in isolating the homozygous F₂ progeny and would show an exclusion of N2 DNA corresponding to the lethal locus.

In a second use of Illumina sequence technology, we have developed a genome interval pull-down method to sequence

defined portions of the genome. By sequencing only intervals that contain the causal mutation, one can reduce the expenses associated with whole genome sequencing. We have successfully applied this technology to positionally clone two new mutants so far (Figure 7), as well as the control *dhc-1(or195ts)* mutant. We identified 45 sequence alterations in the 1.83 Mb *or600sd,ts* interval and 29 mutations in the 1.3 Mb *or683ts* interval vs. the WormBase reference sequence. Thus, if we had sequenced the entire genome of these mutants, we would have found many mutations. Therefore, it is clearly important to have some mapping data to narrow the search for the causal mutation, and RAD mapping fills this role well. As costs continue to come down, whole genome sequencing using either recombinant F₂ animals (Doitsidou *et al.* 2010) or backcrossed mutants (Zuryn *et al.* 2010), two techniques that simultaneously map and sequence mutations, may become more cost effective than the relatively more labor-intensive GIPS. Nevertheless, RAD mapping may continue to prove useful for analyzing large numbers of mutants, as many mutants can be identified by sequencing only candidate genes in the vicinity or by complementation tests with previously identified alleles in the region. In fact, a large number of nonconditional mutants exist that can be used for performing complementation tests (for example, Johnsen and Baillie 1991; Clark and Baillie 1992; Stewart *et al.* 1998; Johnsen *et al.* 2000). GIPS should also remain useful for sequencing candidate genes that are too large to easily amplify with PCR for Sanger sequencing.

As of August 2011, the cost to sequence the entire *C. elegans* genome at 30× coverage was about U.S.\$600 on the HiSeq2000 platform. With this many reads, one could perform 50 RAD mapping experiments (at 30× coverage) or ~50 GIPS procedures using 2-Mb pull-down regions. Both of these sequencing techniques can also be run on Illumina runs with other samples (with the use of barcoded adapters). Depending on the type of mutant being sequenced, using WGS strategies will be more straightforward (and more time- and cost effective; M. Doitsidou and O. Hobert, unpublished results.) For example, if the mutant locus is resistant to RNAi, or if large-scale RNAi screens have not assayed the phenotype represented by the mutant, then WGS is likely the best approach. However, if the mutant phenotype is likely to be recapitulated by RNAi, such as early embryonic lethality (as we are studying), then RAD mapping will reveal a limited number of candidate genes. Sanger sequencing, complementation tests with existing mutants, or GIPS could then identify the causal mutation, although the time required to perform this two-step approach is longer than using a mapping/WGS approach. Thus, RAD mapping may be a viable alternative to WGS when large numbers of mutants are being investigated. Similarly, GIPS can be useful for sequencing loci that are known to be defective in many different mutants. For example, suppressor screens often identify many intragenic suppressor alleles (for example, see Greenwald and Horvitz 1980)

which, depending on size, can be very time consuming to Sanger sequence; yet performing WGS may be too expensive with many alleles to sequence. GIPS fills this gap and allows the simultaneous sequencing of many different mutant loci on the same Illumina lane with the use of barcoded samples. In conclusion, we offer two new strategies for mutant identification in *C. elegans* that can fill roles not currently provided by WGS for certain applications.

The following additional data are available with the online version of this paper: [File S1](#) is a table listing the genomic positions and sequence of the EcoRI-associated RAD tags present in the N2 and CB4856 strains using the WS190 referential version of WormBase.

Acknowledgments

We thank members of the Bowerman and Johnson labs for help in isolating conditional mutants and developing mapping and sequencing protocols. This work was supported by a Leukemia and Lymphoma Society of America fellowship to S.M.O. and National Institutes of Health grants GM050817 and GM049869 to B.B.

Literature Cited

- Albert, T. J., M. N. Molla, D. M. Muzny, L. Nazareth, D. Wheeler *et al.*, 2007 Direct selection of human genomic loci by microarray hybridization. *Nat. Methods* 4: 903–905.
- Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver *et al.*, 2008 Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3: e3376.
- Baxter, S. W., J. W. Davey, J. S. Johnston, A. M. Shelton, D. G. Heckel *et al.*, 2011 Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE* 6: e19315.
- Bigelow, H., M. Doitsidou, S. Sarin, and O. Hobert, 2009 MAQGene: software to facilitate *C. elegans* mutant genome sequence analysis. *Nat. Methods* 6: 549.
- Blumenstiel, J. P., A. C. Noll, J. A. Griffiths, A. G. Perera, K. N. Walton *et al.*, 2009 Identification of EMS-induced mutations in *Drosophila melanogaster* by whole-genome sequencing. *Genetics* 182: 25–32.
- Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. *Genetics* 77: 71–94.
- Chutimanitsakun, Y., R. W. Nipper, A. Cuesta-Marcos, L. Cistue, A. Corey *et al.*, 2011 Construction and application for QTL analysis of a restriction site associated DNA (RAD) linkage map in barley. *BMC Genomics* 12: 4.
- Clark, D. V., and D. L. Baillie, 1992 Genetic analysis and complementation by germ-line transformation of lethal mutations in the unc-22 IV region of *Caenorhabditis elegans*. *Mol. Gen. Genet.* 232: 97–105.
- Coyne, K. J., J. M. Burkholder, R. A. Feldman, D. A. Hutchins, and S. C. Cary, 2004 Modified serial analysis of gene expression method for construction of gene expression profiles of microbial eukaryotic species. *Appl. Environ. Microbiol.* 70: 5298–5304.
- Doitsidou, M., R. J. Poole, S. Sarin, H. Bigelow, and O. Hobert, 2010 *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. *PLoS ONE* 5: e15435.
- Encalada, S. E., P. R. Martin, J. B. Phillips, R. Lyczak, D. R. Hamill *et al.*, 2000 DNA replication defects delay cell division and disrupt cell polarity in early *Caenorhabditis elegans* embryos. *Dev. Biol.* 228: 225–238.
- Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E. M. LeProust *et al.*, 2009 Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* 27: 182–189.
- Greenwald, I. S., and H. R. Horvitz, 1980 unc-93(e1500): a behavioral mutant of *Caenorhabditis elegans* that defines a gene with a wild-type null phenotype. *Genetics* 96: 147–164.
- Harris, T. W., I. Antoshechkin, T. Bieri, D. Blasiar, J. Chan *et al.*, 2010 WormBase: a comprehensive resource for nematode research. *Nucleic Acids Res.* 38: D463–D467.
- Hodges, E., Z. Xuan, V. Balija, M. Kramer, M. N. Molla *et al.*, 2007 Genome-wide in situ exon capture for selective resequencing. *Nat. Genet.* 39: 1522–1527.
- Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson *et al.*, 2010 Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6: e1000862.
- Irvine, D. V., D. B. Goto, M. W. Vaughn, Y. Nakaseko, W. R. McCombie *et al.*, 2009 Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing. *Genome Res.* 19: 1077–1083.
- Johnsen, R. C., and D. L. Baillie, 1991 Genetic analysis of a major segment [LGV(left)] of the genome of *Caenorhabditis elegans*. *Genetics* 129: 735–752.
- Johnsen, R. C., S. J. Jones, and A. M. Rose, 2000 Mutational accessibility of essential genes on chromosome I(left) in *Caenorhabditis elegans*. *Mol. Gen. Genet.* 263: 239–252.
- Jorgensen, E. M., and S. E. Mango, 2002 The art and design of genetic screens: *Caenorhabditis elegans*. *Nat. Rev. Genet.* 3: 356–369.
- Kemphues, K. J., M. Kusch, and N. Wolf, 1988 Maternal-effect lethal mutations on linkage group II of *Caenorhabditis elegans*. *Genetics* 120: 977–986.
- Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret *et al.*, 2003 A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 348: 1953–1966.
- Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.
- Lewis, Z. A., A. L. Shiver, N. Stiffler, M. R. Miller, E. A. Johnson *et al.*, 2007 High-density detection of restriction-site-associated DNA markers for rapid mapping of mutated loci in *Neurospora*. *Genetics* 177: 1163–1171.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko, and E. A. Johnson, 2007 Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17: 240–248.
- Okou, D. T., K. M. Steinberg, C. Middle, D. J. Cutler, T. J. Albert *et al.*, 2007 Microarray-based genomic selection for high-throughput resequencing. *Nat. Methods* 4: 907–909.
- O'Rourke, S. M., M. D. Dorfman, J. C. Carter, and B. Bowerman, 2007 Dynein modifiers in *C. elegans*: light chains suppress conditional heavy chain mutants. *PLoS Genet.* 3: e128.
- O'Rourke, S. M., C. Carter, L. Carter, S. N. Christensen, M. P. Jones *et al.*, 2011 A survey of new temperature-sensitive, embryonic-lethal mutations in *C. elegans*: 24 alleles of thirteen genes. *PLoS ONE* 6: e16644.
- Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli *et al.*, 2003 Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 300: 1394–1399.

- Sarin, S., S. Prabhu, M. M. O'Meara, I. Pe'er, and O. Hobert, 2008 *Caenorhabditis elegans* mutant allele identification by whole-genome sequencing. *Nat. Methods* 5: 865–867.
- Smith, D. R., A. R. Quinlan, H. E. Peckham, K. Makowsky, W. Tao *et al.*, 2008 Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res.* 18: 1638–1642.
- Sonnichsen, B., L. B. Koski, A. Walsh, P. Marschall, B. Neumann *et al.*, 2005 Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* 434: 462–469.
- Srivatsan, A., Y. Han, J. Peng, A. K. Tehranchi, R. Gibbs *et al.*, 2008 High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. *PLoS Genet.* 4: e1000139.
- Stewart, H. I., N. J. O'Neil, D. L. Janke, N. W. Franz, H. M. Chamberlin *et al.*, 1998 Lethal mutations defining 112 complementation groups in a 4.5 Mb sequenced region of *Caenorhabditis elegans* chromosome III. *Mol. Gen. Genet.* 260: 280–288.
- Wang, D., A. Urisman, Y. T. Liu, M. Springer, T. G. Ksiazek *et al.*, 2003 Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol.* 1: E2.
- Zuryn, S., S. Le Gras, K. Jamet, and S. Jarriault, 2010 A strategy for direct mapping and identification of mutations by whole-genome sequencing. *Genetics* 186: 427–430.

Communicating editor: O. Hobert

GENETICS

Supporting Information

<http://www.genetics.org/content/suppl/2011/09/06/genetics.111.134031.DC1>

Rapid Mapping and Identification of Mutations in *Caenorhabditis elegans* by Restriction Site-Associated DNA Mapping and Genomic Interval Pull-Down Sequencing

Sean M. O'Rourke, John Yochem, Amy A. Connolly, Meredith H. Price, Luke Carter, Joshua B. Lowry,
Douglas W. Turnbull, Nick Kamps-Hughes, Nicholas Stiffler, Michael R. Miller, Eric A. Johnson,
and Bruce Bowerman

File S1

Supporting Data

File S1 is available for download as an Excel file at
<http://www.genetics.org/content/suppl/2011/09/06/genetics.111.134031.DC1>.