Supporting Information

Konrad et al. *"Mutational and Transcriptional Landscape of Spontaneous Gene Duplications and Deletions in Caenorhabditis elegans."*

SI Experimental Procedures

Study System. *Caenorhabditis elegans* is a self-fertilizing nematode where the majority of individuals are hermaphrodites. The species has a short generation time of 3.5 days from egg to egg-laying adult at 20°C in the laboratory. The species is particularly amenable to MA studies as selfing rapidly drives populations to homozygosity (1–3). Another immense advantage of this system is the ability of nematode cultures to survive long-term cryogenic storage at −86°C, enabling direct comparisons between experimentally evolved lines and ancestral genotypes (4).

Spontaneous MA experiment in *C. elegans* at Varying Population Sizes. The size of these bottlenecks was ensured through careful counting of randomly chosen L4 larvae selected to establish a new generation every four days. All populations were grown on NGM plates (Nematode Growth Medium) at 20˚C. Twenty MA lines were propagated through single individual descent, while ten lines were bottlenecked to ten individuals, and five lines were bottlenecked to 100 individuals every generation (*Fig. S1A*). The lines were propagated through 409 generations or until extinction under standard laboratory conditions. Populations of size N = 1 and 10 were maintained on 60×15 mm petri dishes seeded with 250μl of *Escherichia coli* (OP50) in YT medium. N = 100 populations were housed on 90 ×15 mm petri dishes seeded with 750μl OP50 in YT medium.

Theoretical Underpinnings of the MA experiment. The fitness effect of a mutation can range continuously from the lethal to deleterious to neutral to beneficial. In small populations, beneficial mutations can be lost and detrimental mutations fixed by chance, or genetic drift. As the population size increases, the importance of genetic drift for the loss or fixation of mutations is diminished. The magnitude of the fitness effect of an individual mutation, s, also has bearing on the influence of drift on the fate of mutations. In small populations, mutations with a small reduction in fitness can reach fixation by genetic drift whereas mutations with a large reduction in fitness are more likely to get lost due to natural selection. Thus, the loss or fixation of mutations depend upon both their selection coefficients (s) and the effective population size, N_e. It has been shown that for sexually reproducing diploids, the dynamics of mutations with $|s| << 1/2N_e$ are dominated by random genetic drift (5, 6). Conversely, the dynamics of mutations with $|s| >> 1/2N_e$ are governed by natural selection. Furthermore, given the obligate self-fertilization mode of reproduction in hermaphroditic species such as *C. elegans*, complete inbreeding results in a 50% reduction in N_e relative to the census population size N (or N_C) (7, 8). Hence, from the perspective of population-genetic theory, the genetic effective population sizes of our three treatments of $N = 1$, 10 and 100 hermaphrodites actually correspond to $N_e = 1$, 5 and 50 individuals, respectively (*Fig. S1B*). For lines bottlenecked each generation at $N = 1$, 10 and 100 individuals, the fate of mutations with selection coefficients less than 0.5, 0.1, and 0.01 is expected to be dominated by genetic drift and they will accumulate at the neutral rate although they may not be neutral with respect to absolute fitness (9, 10). Therefore, small populations subjected to attenuated selection and an increased magnitude of genetic drift can potentially accumulate mutations with extremely large effects in addition to ones with moderate to very slight effects. The demarcation between the behavior of mutations with $2N_es < 1$, which should

be dominated by genetic drift and $2N_e s > 1$, which are under greater influence of natural selection, is not sharply defined and some mutations with a larger fitness cost than $1/2N_e$ could be fixed by genetic drift. Nonetheless, the differences in populations size in these MA experiments alters the relative importance of genetic drift versus natural selection in the fixation or loss of mutations, with genetic drift having the greatest influence in $N = 1$ lines and diminishing in strength with increasing population size.

gDNA Extraction and Whole-Genome Sequencing. One individual of the pre-MA ancestral line was selected to serve as the control. The 86 individual worms were allowed to go through several cycles of self-fertilization to generate enough offspring necessary for extracting sufficient gDNA. Sequencing preparation followed the methodologies previously described (11–13). The PureGene Genomic DNA Tissue Kit (QIAGEN no. 158622) was used for genomic DNA isolation with a supplemental nematode protocol. gDNA quality was checked on 1% agarose gels via electrophoresis and the concentration was determined by a BR Qubit assay (Invitrogen) and a Nanodrop spectrophotometer (Thermo Fisher). Sonication of $2\mu g$ of each DNA sample in $85\mu l$ TE buffer yielded a target DNA fragment length of 200-400 bp, which were subsequently end-repaired by the NEBNext end repair module (New England BioLabs) and purified using Agencourt AMPure XP beads (Beckman Coulter Genomics). 3′ adenine overhangs were added (AmpliTaq DNA Polymerase Kit, Life Technologies), and custom pre-annealed Illumina adapters were ligated to the fragments. PCR amplification was performed via Kapa Hifi DNA Polymerase (Kapa Biosystems) using Illumina's paired-end genomic DNA primers with 8 bp barcodes. Size-fractionation of the PCR products was done on 6% PAGE gels and 300-400 bp fractions were selected for excision. Gel extraction of the fragments was performed via diffusion

at 65˚C and gel filtration (NanoSep, Pall Life Sciences), followed by a final purification step with Agencourt AMPure beads. The Agilent HS Bioanalyzer and HS Qbit assays were used to assess DNA quality and quantity. The multiplexed DNA libraries were sequenced on Illumina HiSeq sequencers at the Northwest Genomics Center (University of Washington), using default quality filters for the sequence reads. The sequence reads from the 86 genomes were demultiplexed and stored as raw sequence reads in individual FASTQ files for further analysis.

oaCGH-based Copy-number Variant Detection. Custom designed microarrays (071114_CE2_WG_CGH_T and 120618_ Cele_WS230_JK_CGH) were manufactured by Roche NimbleGen Inc. Every subarray of the custom 3-plex microarrays consists of 720,000 50-mer oligonucleotide probes. The selection of probes followed Maydan et al. (14) and included both coding and non-coding DNA as well as preexisting duplications (15, 16). For $N = 1$ lines, single individuals were expanded to generated worms for gDNA extraction. Because larger populations are expected to possess greater genetic variation, we surveyed the genomes of two and four individuals from each of the $N = 10$ and $N = 100$ MA lines, respectively. Segmentation analysis followed the techniques of Maydan et al. (14). A bottom-up algorithm was applied for the segmentation, in which case adjacent segments were merged until the neighboring segments failed to reach a preset similarity (*t*-test calculated) threshold. Following the segmentation, the putative segments were further filtered based on their average $\log_2$ratios and *p*-values (*t*-test), and sorted into duplications and deletions, as previously described (14). An initial estimation of minimum length and the breakpoint locations of deletion and duplication events identified above was based on the sequence length contained between the first and last probe within the given

CNV event. The exact breakpoint locations were based on DNA sequence alignment data as described below.

Identification of Copy-Number and Structural Variants from WGS Data. The raw reads in FASTQ format were aligned to the reference N2 genome (version WS247; www.wormbase.org) (17) via the Burrows-Wheeler Aligner (BWA Version 0.5.9) (18) with default settings. Additional alignments were generated using Phaster using default settings, but allowing for insertions and deletions of 300 bp. Potential PCR and optical duplicates were removed using the sort and rmdup utilities in the SAMtools package (Version 0.1.18) (19). Additionally, the Phaster and BWA alignments were analyzed separately. Due to potential contamination in two lines of size $N = 1$, only 18 $N = 1$ lines were included in the final analysis (1A-1H, 1J-1K, and 1M-1T). Five copy-number variation (CNV) and structural variation (SV) detection tools were employed, including paired-end and split read analyses using DELLY (20), Pindel (21), and RAPTR-SV (22) and the read depth approach CNVnator (23) and CNV-seq (24). While all five programs were used to detect putative CNVs/SVs in the BWA and Phaster alignments independently, calls made by Delly, Pindel, and RAPTR-SV based on Phaster alignments were given more weight for breakpoint determination due to improved split read delineation.

CNVnator was run with two different bin sizes. Bins of 50 bp allowed for the detection and verification of small duplication and deletion events, while bins of size 1,000 bp helped in the verification and detection of larger-scale events. DELLY, Pindel and RAPTR-SV were run using mainly default parameters, although the minimum CNV size for Pindel was set to 50 bp in order to attain a manageable list of putative CNVs. Low-quality calls were removed from the list

of putative calls, while high quality calls were initially retained, even if the breakpoints listed were not exact. Any calls pre-existing in the ancestral strain were also removed from the list of putative variants, save for those which were ancestral multi-copy regions which either gained an extra copy, or lost one throughout the MA phase. CNV-seq was run with a window size of 500 bp, a minimum extent of four windows of 250 bp overlap, a $\log_2$-ratio threshold of 0.6, and a *p*-value threshold of 10^{-5}.

The initial list of putative CNVs was comprised of all putative CNVs identified via oaCGH. These events were verified and filtered by comparison to calls made by CNVnator, CNV-seq, DELLY, Pindel, and RAPTR-SV, as well as manual visualization of the putative variant in the Integrated Genome Viewer (IGV) (25). A second list of putative CNV events (not detected via oaCGH) was comprised of all high-quality additional CNV calls made by any of the four sequence-based analysis programs (CNVnator, Pindel, DELLY, and RAPTR-SV). These variant calls were considered for further analysis only if at least two of the programs above supported the call, and if CNV-seq indicated that a copy-number increase or decrease with respect to the ancestor had, in fact, taken place.

Breakpoints for all copy-number changes were first determined using DELLY, Pindel, and RAPTR-SV to identify multiple split reads verifying the same breakpoints. If such high-quality breakpoints were not available because the breakpoints fell into repetitive regions thereby rendering split reads unavailable or unreliable, putative CNV boundaries were examined with respect to the mapping of the read segment that was split off, and the orientation and location of the read mates. MUSCLE (26) was used to generate alignments of the regions flanking the

duplication or deletion of interest and visual refinement of the alignments was sometimes necessary to identify the precise location of the breakpoints. If neither of the above applied, read depth cut-offs were identified visually in IGV, using calls made by CNV-seq, CNVnator, and the oaCGH as guides. In order to finalize the data set of duplications and deletions, SVELTER (27) was used as a final tool in order to determine if any duplications were likely candidates for inverted duplications or more complex events.

RNA Library Preparation, Sequencing, and Analysis of Transcript Abundance. One individual was isolated from every population of size $N = 1$ and the ancestral control, while two and three individuals were isolated from populations of size $N = 10$ and $N = 100$, respectively. The isolated individuals were sequestered on to NGM plates containing OP50 lawns and kept at 20ºC. For each of these 54 individuals, three offspring at the L4 larval stage were isolated from the F_1 generation to serve as biological replicates in the expressional analysis (*Fig. S5*), yielding a total of 162 samples for tissue collection and RNA sequencing at the L1 larval stage. The isolated individuals were sequestered on to NGM plates containing OP50 lawns and maintained at 20ºC. All 162 sample populations were kept for another three generations to yield enough tissue for RNA extraction. Synchronized populations of L1 larvae were generated through the collection of gravid eggs from adults using a standard bleaching protocol. Hatched L1 larvae were collected for RNA extraction and total RNA was isolated via the Qiagen RNeasy Mini Kit. The RNA quality of the samples was evaluated using the Nanodrop 2000, Qubit 3.0 Fluorometer, and an Agilent RNA Analyzer. RNA sequencing libraries for each sample were prepared with the Illumina TruSeq RNA library Prep Kit v2 using standard procedures at the Texas A&M University Genomics and Bioinformatics Services Center. The RNA was fragmented and

Illumina adapters were annealed for amplification. A Qiagen Gel Extraction Kit was used to isolate size selected cDNA fragments. Sequencing was performed using an Illumina HiSeq 4000 platform using default quality filters for the sequencing reads. Sequencing reads were demultiplexed and prefiltered based on default Illumina QC protocols, reads with abnormally short insert lengths were filtered out, and adapters were removed from the reads. The final raw reads were stored in FASTQ files.

TopHat (28, 29) was used to align sequencing reads to the protein-coding transcriptome of *C. elegans* (Wormbase reference N2 genome version WS247). The "very sensitive" bowtie2 algorithm within TopHat was used for alignment with a maximum of one mismatch in the anchor region for each spliced alignment and a minimum and maximum intron length of 20 and 3,000 bp, respectively. Relative transcript abundance for each protein-coding gene was estimated via Cufflinks (29, 30) and gene annotations from the N2 genome version WS247 using default settings. All following analyses were focused on FPKM values calculated on the per gene level. For each individual isolated from one of the original 33 MA populations or the ancestral line, the relative transcript abundance (FPKM) from the three replicates were averaged to get a mean relative transcript abundance for each gene.

Annotation and Characterization of Duplications and Deletions. Gene duplicates and deletions were annotated based on the GFF file available for the N2 reference genome of *C. elegans* (version WS247; www.wormbase.org) (17) using a custom script. Genes were annotated based on their assigned gene types: protein-coding, rRNA, tRNA, pseudogene, and other non-protein coding genes. Initially, protein-coding genes that were duplicated or deleted in their

entirety were designated as *complete* gene duplicates or deletions, while those which were only partially duplicated or deleted were classified as *partial*. In cases where a duplication was immediately followed by a partial loss of copies within the same MA line (which was evident from split read and read pair analysis), only the genes remaining in multiple copies at the end of the MA experiment were considered for inclusion into mutation rates.

Mutation rates (μ) were estimated as duplications or deletions per genome per generation ($M_{CNV_i} = \frac{E_{CNV}}{g}$), and as duplications or deletions per protein-coding gene per generation ($\mu_{CNV_i} = \frac{G_{CNV}}{G_{total^*}\, g}$) for each individual line, *i*, sequenced, where E_{CNV} refers to the number of duplications, deletions, or complex events within the line, *g* refers to the number of generations through which a specific line was propagated, G_{CNV} refers to the number of genes affected by a duplication or deletion within the line, and G_{total} refers to the total number of protein-coding genes in the genome (20,724 protein-coding genes in the N2 reference genome: version WS247). For populations of size $N > 1$, an average mutation rate was attained by averaging across all individuals sequenced for a given MA line (four and five individuals for MA lines of size $N = 10$ and 100, respectively). Further, to attain average mutation rates across all replicate MA lines within a given population size treatment, the line-specific mutation rates were averaged within each population size treatment: $\mu_{CNV} = \frac{\sum_{i=1}^{N} \mu_{CNV_i}}{N}$, where CNV_i refers to the line-specific mutation rate, and *N* refers to the total number of replicates within a given population size treatment (18, 10, and 5 lines for populations of size $N = 1$, 10, and 100, respectively). The number of generations through which each MA line was propagated differed between the $N = 1$ lines (*Table S7*), as some MA lines required frequent backups due to significant fitness decline or

went extinct. Mutations rates were estimated separately for duplications and deletions in previously unduplicated genomic regions (referred to as unique duplications or unique deletions), and for those in pre-existing duplicated regions (referred to CNV gains or CNV losses). Differences in the distribution of mutation rates and length distributions were assessed via correlation tests (Kendall's correlation on log mutation rates and Pearson's correlation on log lengths) in R (31). Permutation tests were performed by resampling the mutation rates and lengths into the population sizes without replacement and recalculating the corresponding correlation coefficients for each analysis every iteration. 100,000 iterations were performed for each permutation test to identify the 95% confidence intervals for the correlation coefficients.

Protein-coding genes which were only partially duplicated were further subdivided into true *partial* duplications and *chimeric* duplicates. True *partial* duplicates were those whose coding region was only partially duplicated, and not fused with another potential ORF (Fig. 4*A*). In contrast, *chimeric* duplicates were those where a partially duplicated open reading frame was fused to another ORF, which most likely arose due to the generation of two *partial* duplicates by the same duplication event, which are fused at the insertion site of the duplicate (Fig. 4*A*). *Chimerics* were inferred in two cases: from duplications giving rise to two *partial* duplicates (one per breakpoint) which remained in tandem order without interspersing sequence between the duplication products, and from *partial* duplicates generated by inverted duplications which remained adjacent but inverted order.

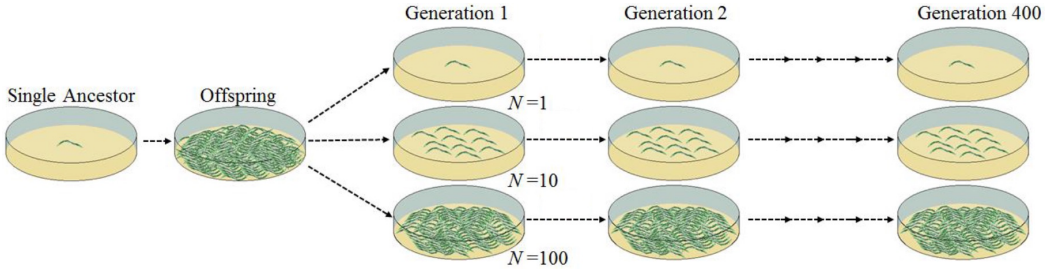
Intrachromosomal location was assigned based on five chromosomal regions as previously described (32) namely, left tip, left arm, core, right arm, and right tip. Expected

proportions of CNV breakpoints within each region was estimated based on the proportion of the genome falling within each category. An alternative expected number considering differences in recombination between these regions was also incorporated by normalizing the above length-normalized expected values by recombination rates of each of the five regions for each chromosome. For statistical analyses, the right and left tips and arms were combined into one tip and one arm category, yielding three chromosomal regions in total. Each breakpoint from a duplication or deletion event (two in each case) was mapped to one of those chromosomal regions in order to account for cases where a CNV spans the boundary between two chromosomal regions.

Differential transcription analyses were performed for *complete* gene duplicates only. Ancestral transcript abundance for affected genes was calculated by averaging the FPKM estimates from the three replicates sequenced via RNA-Seq for the ancestral line. These ancestral FPKM values were used to compare ancestral transcript abundance of duplicated genes between MA lines of different sizes. Furthermore, FPKM ratios were calculated between each gene duplicated within a given line and the average transcript abundance of the same gene in all lines bearing the gene in its ancestral state (neither duplicated nor deleted). FPKM ratios were only calculated if the average FPKM for the gene of interest in its ancestral state was equal to or above 1.

All statistical tests were performed in R (31).

A



B

Spectrum of mutations accumulating in lines	$N = 1$, $N_e = 1$	$N = 10$, $N_e = 5$	$N = 100$, $N_e = 50$
$0.1 < s < 0.5$ (10 - 50%)	✔		
$0.01 < s < 0.1$ (1 – 10%)	✔	✔	
$s < 0.01$ (< 1%)	✔	✔	✔

Fig. S1. Schematic and theoretical underpinnings of the *C. elegans* spontaneous MA experiment with three population size treatments. (A) All 35 lines are descended from a single *N2* hermaphrodite ancestor whose additional descendants were expanded for two generations and cryopreserved as ancestral, pre-MA controls. The maintenance of lines at varying N enables manipulation of the strength of selection. After t generations, the N =1 lines are expected to have independently accumulated mutations, leading to a mean decline in fitness relative to the ancestral control and increased among-line variance. The larger population size treatments (N =10 and 100 worms) will accumulate mutations whose fates will be determined by the fitness effects of the mutation and the strength of natural selection operating in these genetic backgrounds. (B) The maintenance of experimental lines at varying N (and effective population size, N_e) permits the manipulation of the strength of selection across different treatments. This enables the subdivision of the spectrum of mutational effects into a wide range of successively narrower classes across the three N_e treatments. Lines maintained at N_e = 1 are expected to accumulate mutations with selection coefficient, s, < 0.5 (very large effect to neutral). N_e = 5 lines experience minimal levels of selection and are predicted to accumulate mutations with s < 0.1 (moderately large effect to neutral). N_e = 50 lines are subjected to the greatest intensity of selection and are predicted to accumulate mutations with s < 0.01 (slight effect to neutral).

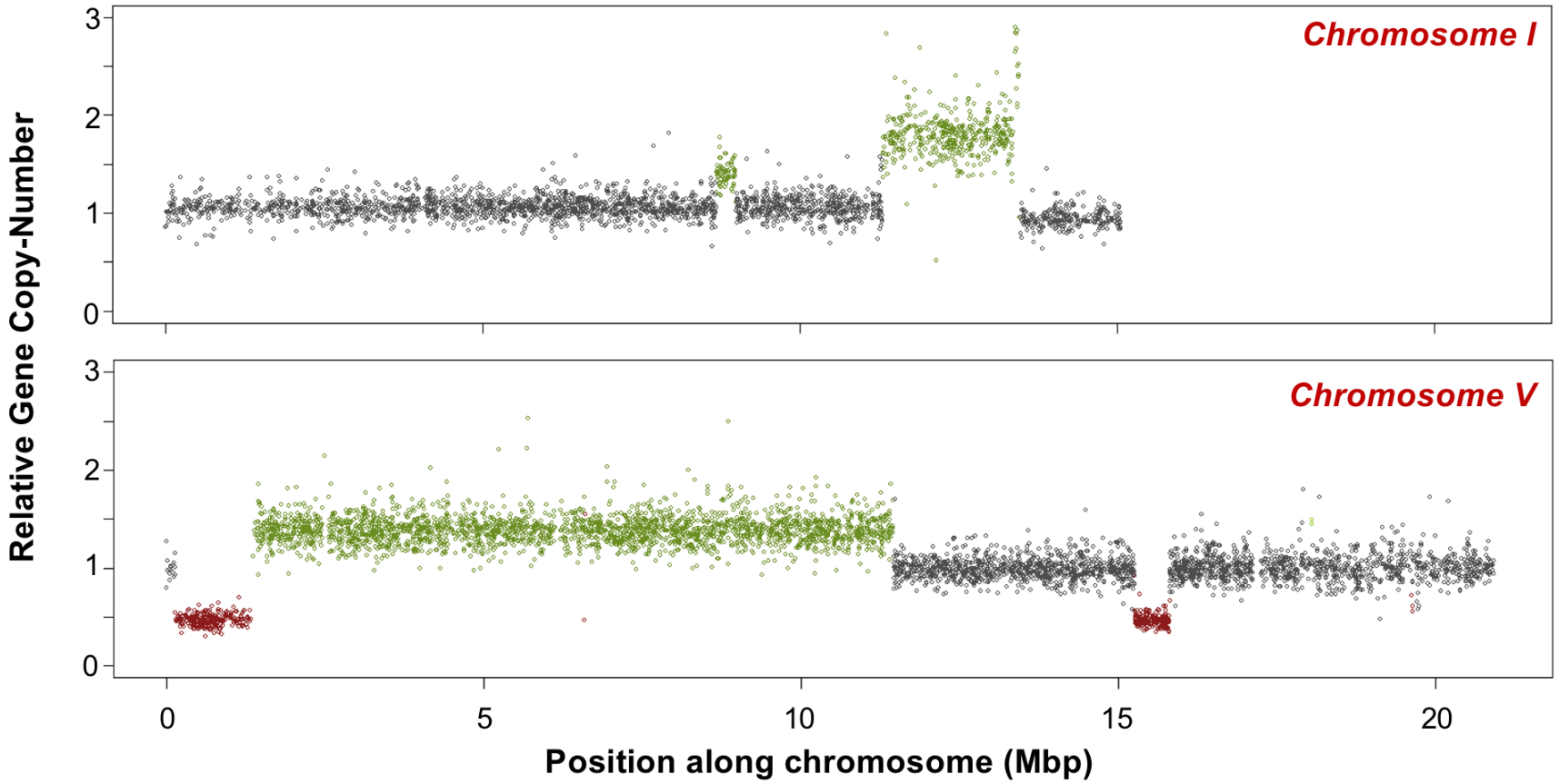


Fig. S2. Relative copy-number ratio between genes of MA line 1T and the ancestor for chromosomes I (top) and V (bottom). Duplicated and deleted genes are shown in green and red, respectively, and all other genes are in gray. Relative copy-number for each gene was calculated independently for 1T and the ancestor by normalizing gene coverage by the genome-wide coverage of the corresponding line. The ratio between the relative copy-number of 1T and the ancestor was calculated for each gene. Each ratio was further normalized by dividing it with the average gene copy-number ratio of unduplicated/undeleted genes. Large scale duplications and deletions detected by both arrays and other sequence-based methods are clearly visible.

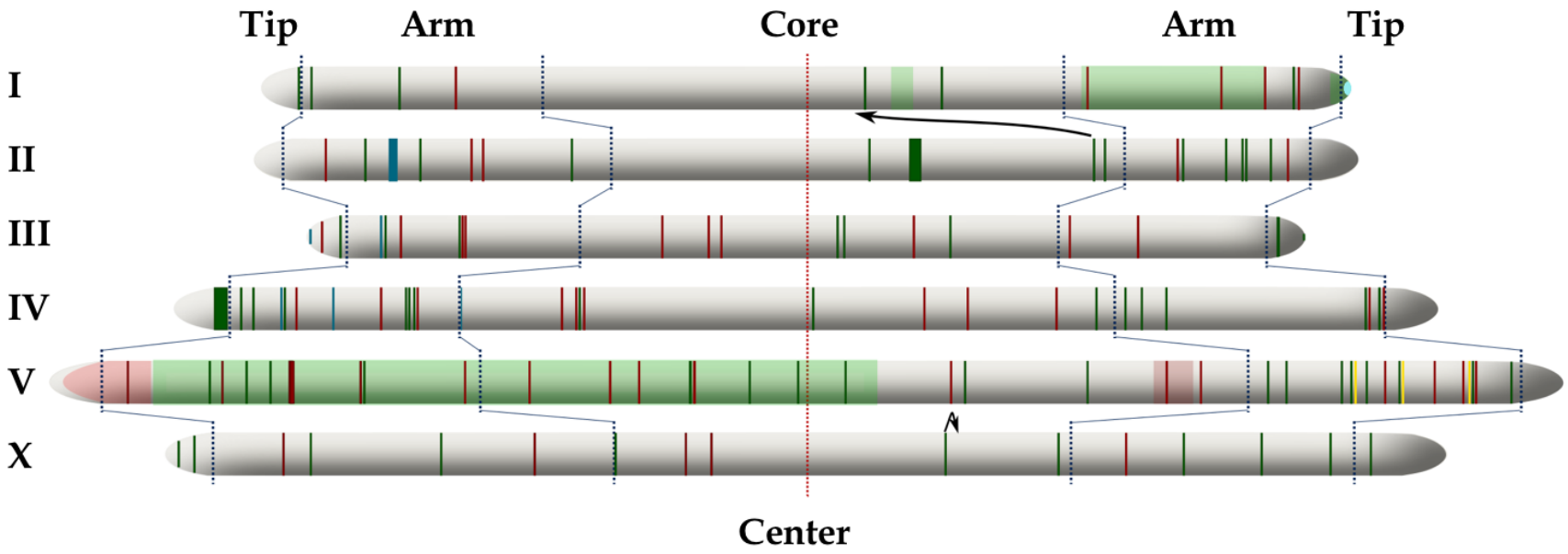


Fig. S3. The chromosomal and regional location of each duplication (green) and deletion (red) across all MA lines. Hypermutable regions and complex events are illustrated in yellow and teal, respectively. Copy-number gains of rDNA genes at one tip of chromosome I are displayed in light blue.

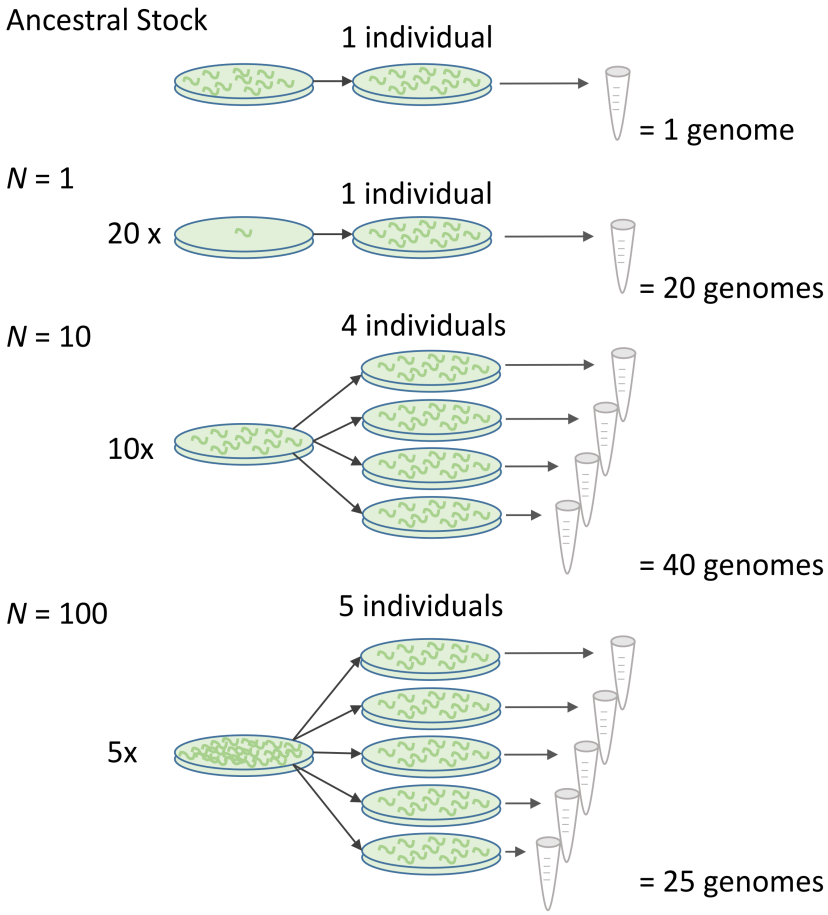


Fig. S4. Whole-genome sequencing to yield 86 *C. elegans* genomes, including that of the ancestral control and 35 MA lines following ~409 MA generations. Multiple individuals were sequenced for MA lines maintained at larger population sizes (N = 10 and 100).

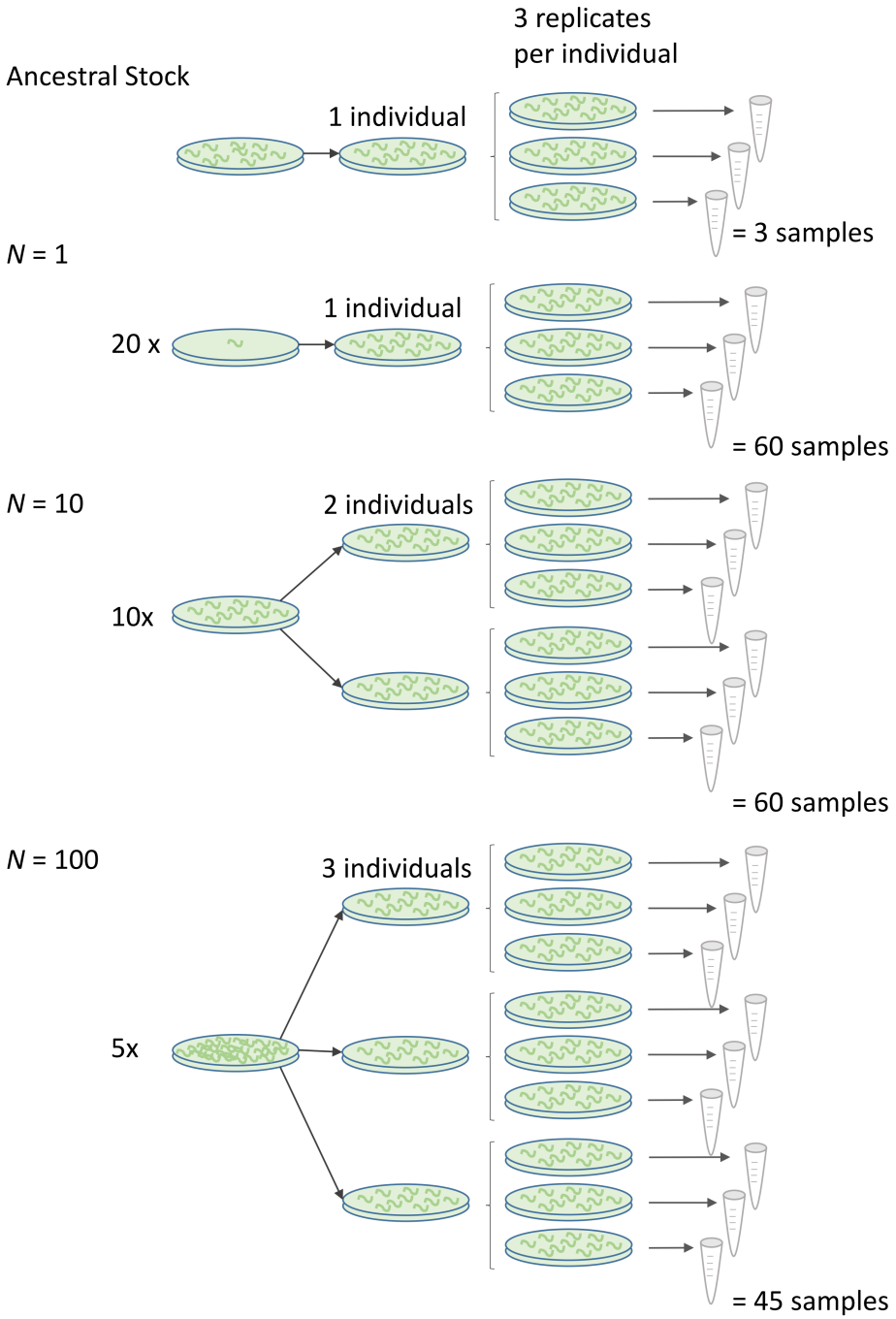


Fig. S5. Experimental design for RNA-Seq at the L1 larval stage. One individual was isolated from every population of size $N = 1$, while two and three individuals were isolated from populations of size $N = 10$ and $N = 100$, respectively. Additionally, one individual from the ancestral pre-MA control was isolated for sequencing. For each of these 56 individuals, three pre-adult L4 offspring were isolated from the F_1 generation to establish three biological replicates yielding a total of 168 lines for L1 tissue collection and RNA-Seq.

Table S1. List of duplications and copy-number gains identified in the *C. elegans* MA lines. Details for all duplications are summarized with the corresponding detection methods used. Start and end positions indicate the boundaries of the duplication. Type refers to the class of duplication; 'DUP' represents the duplication of a gene existing in single-copy form in the ancestral control, 'CNV-Gain' refers to additional copy-number increases of a gene already duplicated in the ancestral control, 'INV' refers to an inversion event. Span indicates the length of the duplication tract. Array Avg. Log_2-Ratio values are indicated and also represent a confirmation of the duplication event via oaCGH. Check marks indicate that a given variant caller (column headers) verified the presence of the CNV (DELLY, CNVnator, RAPTR-SV, Pindel, CNV-seq). The SplitRD column indicates the presence of split reads.

MA Line	Individual	Chromosome	Start	End	Type	Span (bp)	Array Avg. Log_2-Ratio	Array *p*-value	DELLY	CNVnator	RAPTR-SV	Pindel	CNV-seq	SplitRD
1A	1	IV	1,491,426	1,502,092	DUP	10,667	0.77	0.00E+00	√	√	√		√	√
1B	1	I	1,890,742	1,902,823	DUP	12,082	0.91	0.00E+00	√	√			√	√
1B	1	III	13,775,397	13,783,592	CNV-Gain	8,196	1.44	0.00E+00		√	√	√	√	√
1C	1	II	2,241,381	2,252,027	DUP	10,647	1.02	0.00E+00	√	√	√		√	
1C	1	V	19,628,302	19,638,846	CNV-Gain	10,545	0.74	0.00E+00		√			√	√
1C	1	X	6,179,685	6,180,981	DUP	1,297	0.89	2.41E-08	√	√	√		√	
1D	1	II	14,062,836	14,075,924	DUP	13,089	0.97	0.00E+00	√	√	√		√	√
1D	1	V	16,834,786	16,840,667	DUP	5,882	1.00	0.00E+00	√	√	√		√	√
1E	1	V	19,628,301	19,638,841	CNV-Gain	10,541	0.55	0.00E+00	√	√			√	
1F	1	V	17,858,099	17,870,256	DUP-INV	12,158	0.75	0.00E+00	√	√	√		√	√
1F	1	V	18,703,462	18,725,500	CNV-Gain	22,039	0.43	0.00E+00	√	√	√		√	√
1G	1	IV	3,170,028	3,175,871	CNV-Gain/INV	5,844	0.64	6.66E-16		√			√	√
1H	1	II	11,604,654	11,612,116	DUP	7,463	0.93	0.00E+00	√	√	√		√	
1H	1	V	2,666,647	2,677,587	DUP	10,941	0.83	2.22E-16	√	√	√		√	
1H	1	V	2,995,626	2,999,375	DUP	3,750	0.69	3.11E-15		√			√	
1H	1	V	20,208,424	20,218,338	DUP	9,915	0.93	2.22E-16	√	√	√		√	√
1J	1	I	14,772,019	15,059,500	DUP	287,482	0.96	0.00E+00		√			√	√
1J	1	V	14,328,511	14,348,442	DUP	19,932	0.93	0.00E+00		√			√	√
1J	1	X	14,065,154	14,067,623	DUP	2,470	0.85	4.37E-09	√	√	√		√	

1K	1	IV	13,132,618	13,138,209	CNV-Gain	5,593	0.53	2.22E-16		√				
1M	1	II	11,758,483	11,759,418	CNV-Gain	936	0.57	1.94E-07	√	√				
1M	1	V	9,642,547	9,656,316	DUP	13,770	0.92	0.00E+00	√	√	√		√	
1M	1	V	18,047,556	18,055,941	CNV-Gain	16,500	0.80	2.22E-16		√			√	
1M	1	V	19,628,302	19,638,846	CNV-Gain	10,545	0.69	0.00E+00	√	√	√			
1N	1	IV	13,699,635	13,708,460	DUP	8,826	---	---	√	√	√		√	
1N	1	V	2,153,728	2,171,422	DUP	17,695	---	---	√	√	√		√	
1O	1	V	18,703,462	18,725,500	CNV-Gain	22,039	0.75	0.00E+00	√	√	√		√	
1P	1	II	1,477,555	1,494,758	DUP	17,204	0.92	0.00E+00	√	√	√		√	
1P	1	II	9,047,662	9,218,406	DUP	170,745	0.93	0.00E+00	√	√	√		√	
1Q	1	III	416,789	424,370	DUP	7,582	0.67	0.00E+00					√	√
1R	1	II	8,483,391	8,485,504	DUP/INV	2,114	0.91	1.88E-08	√	√	√		√	
1R	1	III	2,072,788	2,074,728	DUP	1,941	0.81	9.01E-08		√			√	
1R	1	X	3,755,108	3,757,743	CNV-Gain	2,636	0.66	9.07E-12	√	√			√	
1S	1	IV	507,687	701,904	DUP	194,218	1.00	0.00E+00		√			√	
1S	1	X	16,101,121	16,111,314	DUP	10,194	0.88	4.44E-16	√	√	√		√	
1T	1	I	8,684,377	8,990,908	DUP/INV	306,532	0.43	0.00E+00	√	√	√			√
1T	1	I	11,303,876	13,455,916	DUP	2,152,041	0.74	0.00E+00		√			√	
1T	1	II	13,442,444	13,455,009	DUP	12,566	1.28	0.00E+00	√	√	√		√	
1T	1	IV	16,465,126	16,470,875	DUP	5,750	0.75	0.00E+00		√			√	
1T	1	V	1,358,715	11,455,976	DUP	10,097,262	0.30	0.00E+00	√				√	
1T	1	V	18,047,556	18,055,941	CNV-Gain	9,006	0.56	0.00E+00		√				
10A	1,2,3,4	I	9,371,845	9,383,731	DUP	11,887	0.71	4.44E-16	√	√	√		√	√
10A	1,2,3,4	II	12,845,876	12,851,456	DUP	5,581	0.70	2.86E-14	√	√	√		√	√
10A	1,2,3,4	IV	3,301,776	3,303,533	CNV-Gain	1,049	0.61	5.40E-11		√			√	
10A	1,2,3,4	V	10,314,272	10,322,679	DUP	8,408	0.87	0.00E+00	√	√	√		√	√
10A	1,2,3,4	X	10,759,688	10,777,482	DUP	17,795	0.88	0.00E+00	√	√	√		√	√
10B	1,2,3,4	II	4,346,667	4,375,131	DUP	28,465	0.86	0.00E+00	√	√	√		√	√
10B	1	II	13,668,689	13,671,209	DUP	2,885	0.82	5.62E-06		√				√
10B	1,2,3,4	IV	16,652,907	16,674,262	DUP	21,356	0.65	0.00E+00	√	√	√		√	√
10B	1,2,3,4	V	18,047,556	18,055,941	CNV-Gain	9,385	0.73	0.00E+00		√			√	
10D	1,2,3,4	III	8,852,596	8,861,460	DUP	8,865	0.94	0.00E+00		√			√	

10E	1	III	7,290,616	7,322,981	DUP	32,366	0.81	0.00E+00		√	√		√	√
10F	1,2,3,4	I	671,745	674,753	DUP	3,009	0.83	2.89E-11	√		√		√	√
10F	1,2,3,4	III	7,384,627	7,405,305	DUP	20,679	0.77	0.00E+00	√	√	√		√	√
10F	1,2,3,4	V	10,975,075	10,981,752	DUP	6,678	1.01	0.00E+00	√	√	√		√	√
10F	1,2,3,4	V	19,699,111	19,701,469	DUP/INV	2,359	1.32	2.50E-11		√				√
10F	1,2,3,4	X	16,662,716	16,691,668	DUP/INV	28,953	1.07	4.44E-16	√	√	√		√	√
10G	1,2,3,4	V	17,087,547	17,093,113	DUP	5,567	0.73	1.62E-14	√	√	√		√	√
10G	1,2,3,4	X	331,769	343,185	DUP	11,417	0.67	0.00E+00	√	√	√		√	√
10H	1,2,3,4	I	498,632	515,066	DUP	16,435	0.79	0.00E+00	√	√	√		√	√
10H	1,2,3,4	II	13,723,799	13,732,766	DUP	8,968	0.90	0.00E+00	√	√	√		√	
10H	1,2,3,4	IV	12,731,376	12,734,914	DUP	3,539	0.65	4.16E-06	√	√	√		√	√
10H	1,2,3,4	V	18,047,556	18,055,941	CNV-Gain	8,386	1.08	2.53E-05					√	
10H	1,2,3,4	V	18,703,462	18,725,500	CNV-Gain	22,039	1.22	0.00E+00	√	√	√		√	√
10H	1,2,3,4	V	19,628,302	19,638,846	CNV-Gain	10,545	1.22	0.00E+00	√	√			√	√
10I	1,2,3,4	V	18,200,657	18,220,972	DUP	20,316	0.52	0.00E+00	√	√	√		√	√
10I	1,2,3,4	V	18,655,157	18,659,860	DUP	4,704	0.56	4.44E-16	√	√	√		√	√
10J	1,2,3,4	IV	5,573,447	5,577,327	DUP	3,881	0.89	9.20E-09	√	√	√		√	√
10J	1,2,3,4	V	8,813,209	8,854,407	CNV-Gain	38,250	0.51	0.00E+00		√			√	
10J	1,2,3,4	V	18,703,462	18,725,500	CNV-Gain	22,039	0.51	2.22E-16	√	√	√		√	√
100A	1,2,3,4,5	IV	8,808,951	8,820,643	DUP	11,693	0.56	6.66E-16	√	√	√		√	√
100A	1,2,3,4,5	X	15,148,341	15,156,990	DUP	8,650	0.77	0.00E+00	√	√	√		√	√
100B	1,2,3,4,5	IV	13,361,611	13,370,521	DUP	8,911	0.86	0.00E+00	√	√	√		√	√
100B	1,2,3,4,5	V	4,297,741	4,299,258	DUP	1,518	0.83	2.34E-08		√			√	
100B	1,2,3,4,5	V	17,988,149	17,990,367	DUP	2,219	0.94	4.19E-13	√	√	√		√	
100C	4	I	8,311,992	8,313,600	DUP	1,609	0.76	1.60E-08	√	√	√		√	√
100C	1,2,3,4,5	III	1,038,709	1,047,573	DUP	8,865	0.89	0.00E+00	√	√	√		√	√
100C	1,2,3,4,5	IV	883,969	885,882	DUP	1,914	0.93	4.04E-04	√	√	√	√	√	√
100D	1,2,3,4,5	I	14,228,067	14,232,862	DUP	4,796	0.90	0.00E+00	√	√	√		√	√
100D	1,2,3,4,5	X	113,370	124,038	DUP	10,669	0.73	0.00E+00	√	√	√		√	√
100D	1,2,3,4,5	X	1,946,357	1,954,508	CNV-Gain	8,571	0.70	0.00E+00		√			√	
100E	1,2,3,4,5	IV	1,054,613	1,076,182	DUP	21,570	0.84	4.44E-16	√	√	√		√	√

100E	1	V	12,630,113	12,651,002		DUP	20,890	0.78	0.00E+00		√			√	√
100E	1,2,3,4,5	X	12,327,311	12,344,353		DUP	17,043	1.36	0.00E+00	√	√	√		√	√

Table S2. List of deletions and copy-number losses identified in the *C. elegans* MA lines. Details for all deletions are summarized with the corresponding detection methods used. Start and end positions indicate the boundaries of the deletion. Type refers to the class of deletion; 'DEL' represents the deletion of a gene existing in single-copy form in the ancestral control, 'CNV-Loss' refers to copy-number decreases of a gene already duplicated in the ancestral control, 'INV' refers to an inversion event. Span indicates the length of the deletion tract. Array Avg. Log_2-Ratio values are indicated and also represent a confirmation of the deletion event via oaCGH. Check marks indicate that a given variant caller (column headers) verified the presence of the CNV (DELLY, CNVnator, RAPTR-SV, Pindel, CNV-seq). The SplitRD column indicates the presence of split reads.

MA Line	Individual	Chromosome	Start	End	Type	Span (bp)	Array Avg. Log_2-Ratio	Array *p*-value	DELLY	CNVnator	RAPTR-SV	Pindel	CNV-seq	SplitRD
1A	1	III	10,504,419	10,504,983	DEL	566	-3.72	9.04E-04		√		√		√
1A	1	V	2,328,000	2,343,592	DEL	15,593	-3.41	2.22E-16	√	√			√	√
1A	1	V	18,703,462	18,725,500	CNV-Loss	22,039	-0.52	0.00E+00	√	√	√		√	√
1B	1	V	18,703,462	18,725,500	CNV-Loss	22,039	-0.65	0.00E+00					√	√
1B	1	V	19,142,663	19,143,016	DEL	354	-2.67	4.68E-09				√		√
1B	1	X	7,155,417	7,160,896	DEL	5,480	-4.99	0.00E+00	√	√	√	√	√	√
1B	1	X	7,510,116	7,523,788	CNV-Loss	13,674	-0.96	0.00E+00					√	√
1C	1	I	11,386,307	11,395,050	DEL/INV	8,744	-4.48	0.00E+00		√			√	√
1D	1	III	158,848	162,360	CNV-Loss	3,513	-0.81	2.66E-15					√	
1D	1	V	12,436,894	12,439,536	DEL	2,643	-2.76	2.39E-09		√			√	
1E	1	III	1,250,401	1,267,050	CNV-Loss	16,650	-1.38	0.00E+00		√			√	
1E	1	III	8,344,593	8,352,550	CNV-Loss	7,958	-0.46	0.00E+00		√			√	
1E	1	V	4,244,084	4,244,803	DEL	720	---	---	√	√	√	√	√	
1E	1	V	18,703,462	18,725,500	CNV-Loss	22,039	0.43	0.00E+00	√	√	√		√	
1H	1	II	925,851	996,807	DEL	70,957	-3.96	0.00E+00	√	√	√		√	
1J	1	V	8,855,304	8,897,399	CNV-Loss	42,096	-0.76	0.00E+00		√			√	
1J	1	X	5,057,126	5,059,598	DEL	2,473	-2.65	2.09E-11		√			√	√
1K	1	I	13,833,573	13,852,339	DEL	18,767	-3.83	0.00E+00		√	√	√	√	√
1K	1	V	3,280,876	3,349,125	DEL	68,250	-2.56	0.00E+00		√			√	

1K	1	V	18,047,556	18,055,941	CNV-Loss	16,500	-0.46	0.00E+00		√			√	
1K	1	V	19,628,302	19,638,846	CNV-Loss	10,545	-0.74	0.00E+00	√					√
1M	1	II	2,950,118	2,951,005	DEL	888	-2.27	4.69E-07	√	√	√		√	√
1M	1	V	18,703,462	18,725,500	CNV-Loss	22,039	-0.44	0.00E+00		√				√
1M	1	V	19,541,101	19,541,214	DEL	114	-1.73	1.91E-02		√		√		
1N	1	I	13,231,770	13,232,375	DEL	606	---	---	√	√	√	√	√	
1N	1	III	5,683,068	5,687,599	DEL	4,532	---	---	√	√	√	√	√	
1N	1	V	7,706,610	7,716,867	DEL	10,258	---	---	√	√	√	√	√	
1Q	1	IV	10,350,537	10,351,598	CNV-Loss	1,062	-0.70	4.27E-06					√	
1Q	1	IV	10,949,693	10,951,150	CNV-Loss	1,458	-0.74	2.78E-09					√	
1R	1	I	2,665,172	2,666,157	DEL	986	-3.72	6.61E-06	√	√	√	√		√
1R	1	IV	5,635,106	5,635,368	DEL	263	---	---	√	√	√	√		
1S	1	V	15,902,781	15,907,135	CNV-Loss	4,355	-1.46	2.66E-15		√			√	
1T	1	I	14,300,280	14,303,743	CNV-Loss	3,994	-0.61	0.00E+00		√				
1T	1	IV	16,515,654	16,518,892	CNV-Loss	3,781	-0.64	0.00E+00		√			√	
1T	1	V	134,256	1,358,715	CNV-Loss	1,224,460	-0.93	2.22E-16		√			√	
1T	1	V	5,695,358	5,695,787	CNV-Loss	430	-0.94	0.00E+00			√		√	
1T	1	V	6,587,512	6,594,832	CNV-Loss	7,321	-0.83	0.00E+00			√		√	
1T	1	V	15,258,626	15,809,375	CNV-Loss	550,750	-0.97	0.00E+00		√			√	
1T	1	V	19,628,302	19,638,846	CNV-Loss	10,545	-0.85	0.00E+00					√	√
10A	1,2,3,4	III	2,110,889	2,116,009	CNV-Loss	5,121	-0.73	1.56E-05		√			√	
10A	1,2,3,4	IV	12,178,739	12,186,970	CNV-Loss	8,500	-0.81	0.00E+00		√			√	
10B	1,2,3,4	IV	5,555,357	5,556,903	CNV-Loss	1,547	-0.62	9.53E-12		√			√	
10C	1,2,3,4	IV	2,821,058	2,828,228	CNV-Loss	7,171	-0.66	0.00E+00					√	√
10D	1,2,3,4	II	14,304,215	14,308,448	CNV-Loss	4,234	-0.79	0.00E+00		√			√	
10D	1,2,3,4	V	18,703,462	18,725,500	CNV-Loss	22,039	-0.56	0.00E+00					√	√
10E	1,2,3,4	III	4,868,161	4,868,432	DEL	272	---	---	√		√	√		√
10F	1,2,3,4	II	12,770,142	12,770,446	DEL	305	-3.29	7.29E-05	√	√	√	√		√
10F	1,2,3,4	III	11,448,761	11,449,280	DEL	520	---	---	√	√	√	√		√
10F	1,2,3,4	IV	3,301,776	3,303,533	CNV-Loss	1,758	-0.60	9.97E-07		√			√	
10F	1,2,3,4	V	18,703,462	18,725,500	CNV-Loss	22,039	-1.26	0.00E+00					√	√

10F	1,2,3,4	V	19,628,302	19,638,846	CNV-Loss	10,545	-0.86	0.00E+00	√					√
10G	1,2,3,4	V	18,047,556	18,055,941	CNV-Loss	8,386	-0.86	2.22E-16					√	
10I	1,2,3,4	V	1,016,652	1,016,777	DEL	126	---	---	√			√		√
10J	1,2,3,4	V	18,456,488	18,461,466	DEL	4,979	-2.10	5.89E-12	√	√	√	√	√	√
10J	1,2,3,4	X	1,568,326	1,572,549	DEL	4,224	-3.35	4.44E-16	√	√	√	√	√	√
100A	1,2,3,4,5	II	3,111,346	3,111,550	DEL	205	-1.95	4.62E-04	√		√	√		√
100A	3	IV	5,327,952	5,334,336	DEL	6,385	-0.98	1.92E-12		√		√	√	√
100A	1	V	3,263,626	3,321,375	DEL	57,750	-1.53	0.00E+00		√			√	
100A	1,2,3,4,5	V	8,110,102	8,110,204	DEL	103	---	---	√		√	√		√
100A	1,2,3,4,5	V	19,628,302	19,638,846	CNV-Loss	10,545	-0.81	0.00E+00		√			√	√
100B	1,2,3,4,5	III	2,110,889	2,116,009	CNV-Loss	5,121	-1.00	0.00E+00					√	√
100B	1,2,3,4,5	IV	1,647,932	1,648,348	DEL	417	---	---	√	√	√	√		√
100B	1,2,3,4,5	V	19,711,626	19,712,625	CNV-Loss	1,000	-0.98	1.49E-10		√			√	
100C	1,2,3,4,5	V	15,426,625	15,435,582	CNV-Loss	8,958	-0.46	0.00E+00	√	√	√		√	√
100C	1,2,3,4,5	X	13,264,348	13,265,264	DEL	917	-2.47	1.03E-03	√	√	√	√	√	√
100D	1,2,3,4,5	V	18,703,462	18,725,500	CNV-Loss	22,039	-0.58	0.00E+00		√	√		√	√
100E	1,2,3,4,5	III	5,507,055	5,509,326	CNV-Loss	2,272	-0.87	3.55E-08					√	
100E	1,2,3,4,5	IV	16,712,189	16,713,716	CNV-Loss	1,528	-0.78	7.94E-12					√	

Table S3. List of complex CNVs identified in the *C. elegans* MA lines. Details for all complex CNVs are summarized with the corresponding detection methods used, and their various individual (sub-) CNVs which together constitute the event. Start and end positions indicate the boundaries of the sub-events comprising the larger complex event. Type refers to the class of sub-event; 'DUP' represents the duplication of a gene existing in single-copy form in the ancestral control, 'CNV-Gain' refers to additional copy-number increases of a gene already duplicated in the ancestral control, 'DEL' represents the deletion of a gene existing in single-copy form in the ancestral control, 'CNV-Loss' refers to copy-number decreases of a gene already duplicated in the ancestral control, 'INV' refers to an inversion event. Total span (bp) indicates the length of the entire complex event; sub-span refers to the length of the individual sub-CNV. Array Avg. Log_2-Ratio values are indicated and also represent a confirmation of the duplication/deletion event via oaCGH. Check marks indicate that a given variant caller (column headers) verified the presence of the complex CNV (DELLY, CNVnator, RAPTR-SV, Pindel, CNV-seq, SVELTER). The SplitRD column indicates the presence of split reads for each of the sub-CNVs.

MA Line	Individual	Chromosome	Start	End	Sub-Type	Total Span (bp)	Sub-Span (bp)	Array Avg. Log_2-Ratio	Array *p*-value	DELLY	CNVnator	RAPTR-SV	Pindel	CNV-seq	SplitRD	SVELTER
1B	1	II	1,812,432	1,817,620	DEL	123,283	5,189	-3.49	1.37E-11		√			√	√	
			1,913,641	1,927,858	DUP/INV		18,103	0.89	0.00E+00	√	√			√	√	
			1,931,830	1,935,714	DUP			1.21	2.22E-16		√			√	√	
			1,931,830	1,935,714	CNV-Gain		3,885				√			√	√	
1B	1	III	13,378,306	13,395,314	DUP	17,009	17,009	1.45	0.00E+00		√			√	√	
			13,379,116	13,395,314	CNV-Gain		16,199				√			√	√	
1J	1	IV	3,942,336	3,946,150	CNV-Gain	15,558	4,096	---	---	√	√	√		√	√	DEL-DUP
			3,953,798	3,957,893	CNV-Gain			0.83	4.66E-15	√	√	√		√	√	
			3,946,151	3,949,588	CNV-Loss		3,438	-3.25	0.00E+00		√			√	√	
1O	1	III	979,540	979,778	DUP	1,888	239	---	---	√		√			√	
			977,891	978,943	DEL		1,208	-2.69	1.72E-06		√			√	√	
			979,124	979,278	DEL			---	---		√			√	√	
1Q	1	IV	1,465,372	1,466,929	DUP/INV	21,288	1,558	0.59	2.08E-05	√						
			1,445,642	1,465,097	DEL		19,456	-3.39	0.00E+00		√			√	√	
100B	1	III	131	495	DEL	14,800	680	-1.76	5.57E-03		√	√	√		√	

			599	913	DEL						√	√	√		√		
			1,429	14,930	DUP		14,931	0.96	0.00E+00		√			√	√		
100D	1,2,3,4,5	IV	2,162,953	2,165,969	DUP	3,017	2,841	---	---	√	√	√		√	√		
			2,164,750	2,164,925	CNV-Loss		176	---	---	√		√	√				

Table S4. List of 161 copy-number changes detected in 33 spontaneous MA lines of *C. elegans* maintained at three differing population sizes. Duplication and deletion events are further subdivided into (i) unique, previously unduplicated sequence, and (ii) copy-number gains or losses in ancestral multi-copy sequence (CNV). Copy-number changes are provided as raw counts as well as normalized counts (multiplying each event by the fraction of individuals of the population within which the variant was found). The normalized event counts are used to calculate the per genome mutation rates. Counts of protein-coding and other genes per population size were normalized in a similar fashion. Mutation rates per protein-coding gene were calculated using the normalized gene counts. Lengths of mutations were limited to the net length of copy-number increase or decrease at the end of the MA experiment. Hence, in the case of a complex rearrangement, the true length of a duplication may be underestimated, leading to a conservative estimate for the per gene duplication and deletion rates.

		Count of mutational events, normalized by variant frequency (raw counts)		
		***N* = 1**	***N* =10**	***N* =100**
Simple	Total Events	81	44.5 (46)	23.8 (27)
Duplications	Total	41	28.5 (31)	12.4 (14)
	Unique	29	20.5 (22)	11.4 (13)
	CNV	12	8.0 (9)	1.0 (1)
Deletions	Total	40	15.0 (15)	11.4 (13)
	Unique	19	6.0 (6)	4.4 (6)
	CNV	21	9.0 (9)	7.0 (7)
Complex	Total Events	5	–	1.2 (2)
Duplications	Total	7	–	1.2 (2)
	Unique	4	–	1.2 (2)
	CNV	3	–	–
Deletions	Total	4	–	1.2 (2)
	Unique	3	–	0.2 (1)
	CNV	1	–	1.0 (1)
Event/Genome/Generation ($\mu \times 10^{-3}$)				
Duplications	All duplications	6.50	6.97	6.06
	Unique	4.72	5.02	5.57
	CNV	1.78	1.96	0.49
Deletions	All deletions	6.12	3.67	5.57
	Unique	2.71	1.47	2.15
	CNV	3.40	2.20	3.42
Complex	Complex Events	0.71	–	0.98
All CNV	All CNV	13.32	10.64	12.62
Protein-coding genes affected by mutational events (normalized by frequency)				
Unique genes duplicated	Total	3,277.00	58.25	32.00
	Complete	3,236.00	34.00	20.40
	Partial	41.00	24.25	11.60
Duplication of preexisting	Total	39.0	29.0	2.0
CNV genes	Complete	18.0	12.0	2.0
	Partial	21.0	7.0	–
Unique genes deleted	Total	92.0	6.0	8.4
	Complete	66.0	2.0	4.6

		Count of mutational events, normalized by variant frequency (raw counts)		
		***N* = 1**	***N* =10**	***N* = 100**
	Partial	26.0	4.0	3.8
Deletion of preexisting	Total	496.0	21.0	15.0
CNV genes	Complete	475.0	11.0	6.0
	Partial	21.0	10.0	9.0
Span of Mutational Events				
All duplications	Median	10,647	9,385	8,756
	Minimum	936	1,049	1,518
	Maximum	10,097,262	38,250	21,570
Unique duplications	Median	11,512	8,917	8,856
	Minimum	1,297	2,359	1,518
	Maximum	10,097,262	32,366	21,570
Copy-Number Gains	Median	10,541	10,545	8,571
	Minimum	936	1,049	8,571
	Maximum	22,039	38,250	8,571
All deletions	Median	7,321	4,607	2,272
	Minimum	114	126	103
	Maximum	1,224,260	22,039	57,750
Unique deletions	Median	2,558	413	667
	Minimum	114	126	103
	Maximum	70,957	4,979	57,750
Copy-Number Losses	Median	10,545	7,779	5,121
	Minimum	430	1,547	1,000
	Maximum	1,224,460	22,039	22,039
Mutation Rate ($\mu \times 10^{-5}$) (/ protein-coding gene/generation; based on 20,724 protein-coding genes)				
$\mu_{duplication}$				
	Total	2.883	0.103	0.080
	Unique duplications	2.856	0.069	0.076
	Copy-number gains	0.027	0.034	0.005
$\mu_{deletion}$				
	Total	0.500	0.032	0.055
	Unique deletions	0.073	0.007	0.020
	Copy-number losses	0.427	0.025	0.035

Table S5. Distribution of copy-number changes across chromosomes. Chromosome V harbours a disproportionately greater number of copy-number changes relative to other chromosomes. This is primarily due to copy-number gains and losses in pre-existing duplications. The expected number of copy-number changes for the *G*-tests is the product of the relative size of each chromosome and the number of independent copy-number changes of any given type.

	Chromosome						***G***	***df***	***p-value***
	I	**II**	**III**	**IV**	**V**	**X**			
Expected	**15%**	**15%**	**14%**	**17%**	**21%**	**18%**			
Observed									
All events	14 (8.8%)	17 (10.6%)	20 (12.5%)	30 (18.8%)	64 (40.0%)	15 (9.0%)	37.4	5	5.01×10^{-7}
All CNV	1 (1.6%)	2 (3.3%)	7 (11.5%)	13 (21.3%)	35 (57.4%)	3 (4.9%)	55.2	5	1.17×10^{-10}
All unique	12 (13.0%)	15 (16.3%)	10 (10.9%)	14 (15.2%)	29 (31.5%)	12 (13.0%)	6.7	5	0.24
All duplications	9 (10.5%)	12 (14.0%)	7 (8.1%)	16 (18.6%)	32 (37.2%)	10 (11.6%)	14.7	5	0.01
All deletions	5 (8.1%)	5 (8.1%)	10 (16.1%)	5 (8.1%)	32 (51.6%)	5 (8.1%)	33.0	5	3.76×10^{-6}
CNV-Gain	0 (0.0%)	1 (4.3%)	1 (4.3%)	5 (21.7%)	14 (60.9%)	2 (8.7%)	4.5	5	1.72×10^{-4}
CNV-Loss	1 (2.6%)	1 (2.6%)	6 (15.8%)	8 (21.1%)	21 (55.3%)	1 (2.6%)	34.8	5	1.67×10^{-6}
Unique duplications	9 (14.3%)	11 (17.5%)	6 (9.5%)	11 (17.5%)	18 (28.6%)	8 (12.7%)	3.7	5	0.59
Unique deletions	3 (10.3%)	4 (13.8%)	4 (13.8%)	3 (10.3%)	11 (37.9%)	4 (13.8%)	5.0	5	0.41
Complex	1 (14.3%)	0 (0.0%)	3 (42.9%)	3 (42.9%)	0 (0.0%)	0 (0.0%)	12.1	5	0.30

Table S6. Intrachromosomal distribution of copy number break-points. Chromosomal regions identified as tips, cores and arms are based on Rockman and Kruglyak (26). The number of breakpoints deviates significantly from random expectations based on the cumulative size of different regions, and the combined size and recombination frequency.

	Chromosomal Region			**Expectation by Length**			**Expectation by Length and Recombination**		
	Tip	***Arm***	***Core***	***G***	***df***	***p-value***	***G***	***df***	***p-value***
MA Population Size Treatment									
$N = 1$	14	119	49	28.42	2	6.74×10^{-7}	∞	2	$< 2.2 \times 10^{-16}$
$N = 10$	6	60	22	17.42	2	1.65×10^{-4}	∞	2	$< 2.2 \times 10^{-16}$
$N = 100$	8	32	18	6.49	2	3.89×10^{-2}	∞	2	$< 2.2 \times 10^{-16}$
All MA Lines	28	211	89	48.79	2	2.55×10^{-11}	∞	2	$< 2.2 \times 10^{-16}$
Expected by Length	0.07	0.47	0.46						
Expected by Length and Recombination	0.00	0.20	0.80						

Table S7. Summary of population size (*N*), generations of propagation, sequence coverage per genome, and average sequence coverage across individuals per line.

Line	*N*	MA Generations	Average Line Coverage	Individual Sequenced	Coverage
Pre-MA Ancestral Control	N/A	N/A	30	1	30
MA 1A	1	389	35	1	36
MA 1B	1	397	32	1	32
MA 1C	1	392	35	1	35
MA 1D	1	326	33	1	33
MA 1E	1	368	27	1	27
MA 1F	1	386	29	1	29
MA 1G	1	346	27	1	27
MA 1H	1	281	34	1	34
MA 1J	1	396	42	1	42
MA 1K	1	349	21	1	21
MA 1M	1	395	29	1	29
MA 1N	1	363	29	1	29
MA 1O	1	397	17	1	17
MA 1P	1	393	27	1	27
MA 1Q	1	381	30	1	30
MA 1R	1	386	35	1	35
MA 1S	1	291	23	1	23
MA 1T	1	305	18	1	18
MA 10A	10	409	19	1	23
				2	21
				3	17
				4	17
MA 10B	10	409	13	1	15
				2	11
				3	16
				4	10
MA 10C	10	408	9	1	8
				2	12
				3	8
				4	8
MA 10D	10	408	13	1	8
				2	16
				3	14
				4	16
MA 10E	10	409	16	1	16
				2	18
				3	13
				4	16

MA 10F	10	409	13	1	10
				2	12
				3	15
				4	14
MA 10G	10	409	17	1	18
				2	18
				3	16
				4	16
MA 10H	10	408	18	1	17
				2	18
				3	17
				4	18
MA 10I	10	409	19	1	15
				2	19
				3	10
				4	32
MA 10J	10	408	18	1	16
				2	20
				3	16
				4	21
MA 100A	100	409	17	1	15
				2	19
				3	16
				4	17
				5	18
MA 100B	100	409	17	1	19
				2	17
				3	16
				4	18
				5	15
MA 100C	100	409	17	1	16
				2	16
				3	15
				4	17
				5	19
MA 100D	100	409	17	1	16
				2	18
				3	16
				4	16
				5	17
MA 100E	100	409	17	1	17
				2	18
				3	16
				4	16
				5	18

SI References

1. Vassilieva LL, Hook AM, Lynch M (2000) The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. *Evolution* 54:1234–1246.
2. Baer CF, et al. (2005) Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. *Proc Natl Acad Sci USA* 102:5785–5790.
3. Katju V, LaBeau EM, Lipinski KJ, Bergthorsson U (2008) Sex change by gene conversion in a *Caenorhabditis elegans fog-2* mutant. *Genetics* 180:669–672.
4. Lewis JA, Fleming JT. 1995. Basic culture methods. *Methods Cell Biol* 48:3–29.
5. Kimura, M (1962) On the probability of fixation of mutant genes in a population. *Genetics* 47:713-719.
6. Kimura M (1983) *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge).
7. Crow J, Kimura M (1970). *An Introduction to Population Genetics Theory*. (Harper & Row, New York).
8. Pollak E (1987) On the theory of partially inbreeding finite population. I. Partial selfing. *Genetics* 117:353–360.
9. Ohta T (1972) Fixation probability of a mutant influenced by random fluctuation of selection intensity. *Genet. Res.* 19:33–38.
10. Ohta T (1992) The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* 23:263–286.
11. Flibotte S, et al. (2010) Whole-genome profiling of mutagenesis in *Caenorhabditis elegans*. *Genetics* 185:431–441.
12. Thompson O, et al. (2013) The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res* 23:1749–1762.

13. Konrad A, et al. (2017) Mitochondrial mutation rate, spectrum and heteroplasmy in *Caenorhabditis elegans* spontaneous mutation accumulation lines of differing population size. *Mol Biol Evol* 34:1319–1334.

14. Maydan JS, Lorch A, Edgley ML, Flibotte S, Moerman DG (2010) Copy number variation in the genomes of twelve natural isolates of *Caenorhabditis elegans*. *BMC Genomics* 11:62.

15. Maydan JS, et al. (2007) Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res* 17:337–347.

16. Farslow JC, et al. (2015) Rapid increase in frequency of gene copy-number variants during experimental evolution in *Caenorhabditis elegans. BMC Genomics* 16:1044.

17. Harris TW, et al. (2010) WormBase: a comprehensive resource for nematode research. *Nucleic Acids Res* 38 (Database issue):D463–D467.

18. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

19. Li H, et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25:2078–2079.

20. Rausch T, et al. (2012) DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* 28:i333–i339.

21. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z (2009) Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25:2865–2871.

22. Bickhart DM, et al. (2015) RAPTR-SV: a hybrid method for the detection of structural variants. *Bioinformatics* 31:2084–2090.

23. Abyzov A, Urban AE, Snyder M, Gerstein M (2011) CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 21:974–984.

24. Xie C, Tammi MT (2009) CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* 10:80.

25. Thorvaldsdóttir H, Robinson JT, Mesirov JP (2011) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14:178–192.

26. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.

27. Zhao X, Emery SB, Myers B, Kidd JM, Mills RE (2016) Resolving complex structural genomic rearrangements using a randomized approach. *Genome Biol* 17:126.

28. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111.

29. Trapnell C, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protocols* 7:562–578.

30. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol* 28:511–515.

31. R Core Development Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

32. Rockman MV, Kruglyak L. (2009) Recombinational landscape and population genomics of *Caenorhabditis elegans. PLoS Genet* 5:e1000419.