Genome structural variation in natural *Caenorhabditis elegans* populations

Kyle J. Lesack, James D. Wasmuth*

Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada

* Corresponding author: jwasmuth@ucalgary.ca

Abstract

Structural variation, involving large alterations in chromosome structure, drives genetic diversification and the emergence of new phenotypes. These changes are widespread in natural populations and play an important role in adaptation and speciation. For many species, research has been limited to laboratory adapted strains and experimental evolution, which may not reflect the diversity of structural variants in the wild. Furthermore, technological limitations have proved to be a major barrier to accurate and comprehensive variant calling. In this study, we used PacBio sequencing data from 14 wild *Caenorhabditis elegans* strains to characterize structural variants in a natural population. With long-reads, we overcame limitations associated with short-read approaches and leveraged population-level data to further refine the accuracy of variant calls. We found that large, rapidly evolving gene families, such as GPCR, F-Box, and C-Type lectin, were prominent among the variants predicted to have phenotypic consequences. These results shed light on the significant role of structural variants in the evolution of *Caenorhabditis elegans* in its wild habitats and the limitations of treating N2 as the reference wild-type.

Keywords: structural variation; *Caenorhabditis elegans*; environmental adaptation

Introduction

Structural variants (SVs) are large-scale alterations in a genome—such as deletions, duplications, and inversions—that affect gene content, order, or dosage. These variants can have profound phenotypic effects by disrupting genes, altering gene dosage, or rearranging regulatory elements. While often deleterious, SVs can also drive adaptive evolution through mechanisms such as gene duplication followed by neofunctionalization or chromosomal inversions that suppress recombination and maintain beneficial allele combinations (Faria et al., 2019; Katju and Bergthorsson, 2013; Kirkpatrick, 2010).

Despite their importance, SVs remain poorly characterized in most species, especially at the population level. This is partly due to the technical limitations of short-read DNA sequencing, which struggles to resolve large or complex variants, particularly in repetitive regions (Koboldt, 2020; Lesack et al., 2022). Long-read sequencing has substantially improved the resolution of SVs and providing more reliable alignments (Smolka et al., 2024).

In the nematode *Caenorhabditis elegans*, much of our knowledge about SVs comes from laboratory strains and mutation accumulation experiments. These studies have revealed a high rate of spontaneous duplications and the formation of complex alleles (Farslow et al., 2015; Konrad et al., 2018). For example, a recombinant inbred line showed increased fitness due to a partial duplication and inversion of five different regions of *rcan-1*, a gene which regulates thermotactic behavior (Zhao et al., 2020). However, laboratory strains, such as N2, carry numerous derived alleles and pseudogenised genes, limiting their relevance for understanding variation in natural populations (Andersen and Rockman, 2022; Sterken et al., 2015).

43 Comparative studies between N2 and wild isolates, particularly the Hawaiian strain CB4856,
44 have revealed extensive structural divergence (Lee et al., 2021; Vergara et al., 2014). In 12
45 natural strains, more than 5% of the N2 genome is affected by SVs, with CB4856 showing over
46 79,000 insertion/deletion events, more than 800 unaligned segments, and 61 hyper-divergent
47 regions (Lee et al., 2021; Thompson et al., 2015).

48 The *Caenorhabditis elegans* Natural Diversity Resource (CeNDR) has expanded access
49 to genomic data from wild strains (Cook et al., 2017). This collection has facilitated the
50 discovery of genetic loci underlying xenobiotic metabolism and anthelmintic drug resistance
51 (Hahnel et al., 2018; Zdraljevic et al., 2019).

52 In this study, we used PacBio long-read sequencing data from 14 wild *C. elegans* strains
53 to generate a high-confidence, population-scale map of SVs. We identified thousands of
54 deletions, duplications, and inversions—many affecting genes involved in environmental
55 response, stress, and immune function.
56

57 # Results and Discussion

58 We identified 14,373 deletions, 3,149 duplications, and 1,748 inversions in 14 wild *C.*
59 *elegans* strains (Table S1). SVs were unevenly distributed across the genome (Figure S1).
60 Deletions and duplications were enriched at the autosomal arms, whereas inversions were more
61 evenly distributed. These patterns are consistent with previous findings; SVs preferentially occur
62 in low gene density, high recombination regions (Cutter et al., 2009). Chromosome arms may
63 promote SV formation through recombination-based mechanisms while protecting essential
64 genes located in denser chromosome centres (Carvalho and Lupski, 2016). Among all SVs, 64%
65 were singletons—variants unique to a single strain—suggesting high inter-strain diversity (Table
66 S2). Hawaiian strains had the highest number and proportion of singleton variants while Western
67 European strains had the fewest (Table 1). The high ancestral diversity in Hawaiian strains is
68 known, with reduced diversity in non-Hawaiian lineages due to selective sweeps (Crombie et al.,
69 2019). Conversely, 25 variants were found in all 14 strains: 15 deletions, two duplications, and
70 eight inversions (Table S2; Figures S2, S3, S4).

71 To prioritise variants with likely functional consequences, we focused on high-impact
72 SVs. This subset included 1,485 deletions, 214 duplications, and three inversions, collectively
73 affecting 1,258 genes (Figure 1; Tables S3, S4, S5). Deletions were enriched for splicing
74 disruptions, transcript ablations, and frameshift variants (Table 2). Duplications primarily
75 resulted in transcript amplifications, and all three high impact inversions were classified as
76 stop/loss mutations. Functional enrichment revealed strong overrepresentation in categories
77 related to environmental stress, including GPCRs, F-box proteins, C-type lectins (CLECs), and
78 detoxification enzymes (Figure 2; Table S6). Many of these genes were in hyper-variable

79 regions, which have been shown to maintain ancestral variation through balancing selection (Lee

80 et al., 2021).

81

Strain	Region	Deletions			Duplications			Inversions		
		Total	Singleton	Singleton Proportion	Total	Singleton	Singleton Proportion	Total	Singleton	Singleton Proportion
DL238	Hawaii	2343	804	0.34	485	209	0.43	333	95	0.29
ECA396	Hawaii	3412	1643	0.48	519	300	0.58	399	140	0.35
QX1794	Hawaii	2363	756	0.32	459	186	0.41	311	74	0.24
XZ1516	Hawaii	4007	2440	0.61	443	266	0.60	519	301	0.58
ECA36	New Zealand	3522	1671	0.47	586	308	0.53	398	147	0.37
EG4725	Western Europe	1913	290	0.15	452	94	0.21	280	42	0.15
JU1400	Western Europe	1692	225	0.13	417	120	0.29	240	34	0.14
JU2526	Western Europe	2257	452	0.20	442	96	0.22	309	55	0.18
JU2600	Western Europe	1519	290	0.19	385	87	0.23	238	39	0.16
JU310	Western Europe	1366	170	0.12	385	72	0.19	217	16	0.07
MY2147	Western Europe	1279	217	0.17	364	85	0.23	199	29	0.15
MY2693	Western Europe	1251	97	0.08	385	84	0.22	258	35	0.14
NIC2	Western Europe	457	80	0.18	248	77	0.31	150	12	0.08
NIC526	Western Europe	1544	179	0.12	402	78	0.19	257	22	0.09

82 **Table 1:** Singleton variants among individual *C. elegans* strains.

83

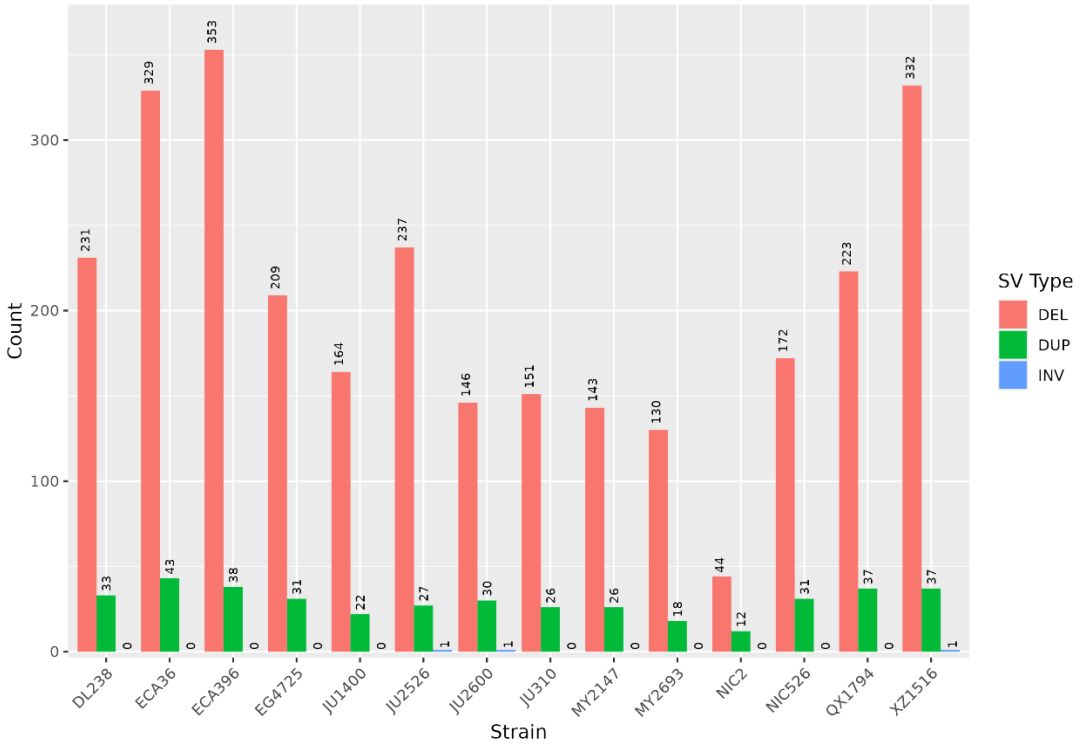

Figure 1: Number of structural variants determined to have high functional impact.

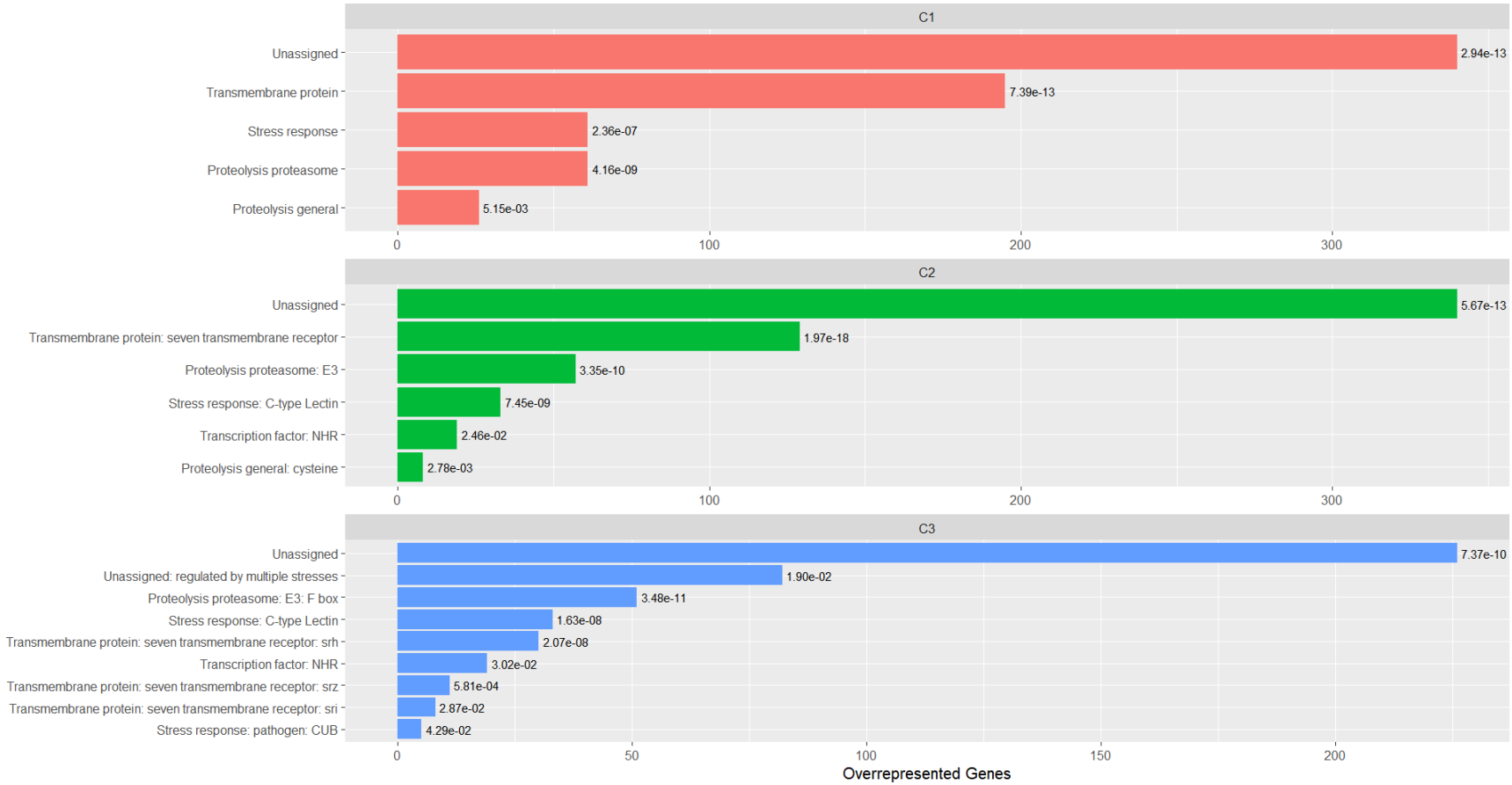

Figure 2: Statistically significant overrepresented WormCat terms for putative high impact SVs. The red, green, and blue colours represent the categories for the C1, C2, and C3 levels, respectively.

91

SV Type	Term	Description‡	Count
Deletion (n = 1,485)	splice_donor_variant	A splice variant that changes the 2-base region at the 5' end of an intron	510
	splice_acceptor_variant	A splice variant that changes the 2-base region at the 3' end of an intron	503
	transcript_ablation	A feature ablation whereby the deleted region includes a transcript feature	240
	frameshift_variant	A sequence variant which causes a disruption of the translational reading frame, because the number of nucleotides inserted or deleted is not a multiple of three	168
	stop_lost	A sequence variant where at least one base of the terminator codon (stop) is changed, resulting in an elongated transcript	89
	start_lost	A codon variant that changes at least one base of the canonical start codon	16
	stop_gained	A sequence variant whereby at least one base of a codon is changed, resulting in a premature stop codon, leading to a shortened transcript	2
Duplication (n = 241)	transcript_amplification	A feature amplification of a region containing a transcript	240
	start_lost	A codon variant that changes at least one base of the canonical start codon	1
Inversion (n=3)	start_lost	A codon variant that changes at least one base of the canonical start codon	3

92 **Table 2:** Predicted Consequences of High Severity Structural Variants.

93 ‡ Consequence descriptions downloaded on May 22, 2023 from:
94 https://useast.ensembl.org/info/genome/variation/prediction/predicted_data.html

95

96 **GPCRs and sensory behaviour**

97 Seven-transmembrane G-protein-coupled receptors (GPCRs), involved in
98 chemosensation and behaviour, were among the most affected gene families. Of the 195 genes in
99 the "Transmembrane protein" category, 120 were GPCRs. An additional GPCR, *srh-208*, was
100 categorized under the "Proteolysis proteasome" category. The SRH family was enriched, with 31
101 genes, which is consistent with previous reports of rapid expansion and diversification in *C.*

102 *elegans* (Thomas and Robertson, 2008; Vidal et al., 2018). A duplication of *srh-234* was present
103 in two strains. The expression of *srh-234* is regulated by the chemosensation of food (Gruner et
104 al., 2014). The deletion of *str-208* was noteworthy given that this gene is mplicated in erratic
105 movement (Daigle et al., 2022). While the loss of this gene may affect locomotion, it is not yet
106 known whether the associated structural variants confer an adaptive benefit, impose a cost, or are
107 selectively neutral.

108 **F-box genes and immune adaptation**

109 SVs affected 87 proteolysis-related genes, including 44 F-box genes. The F-box gene
110 family is considerably expanded in *C. elegans* relative to other *Caenorhabditis* species with
111 several members under positive selection in the non-Hawaiian strains (Ma et al., 2021).
112 Additional immune-related SVs may reflect adaptations to local microbial environments. For
113 example, *hecw-1*, deleted in ECA-396 is an E3 ubiquitin ligase that modulate pathogen
114 avoidance (Chang et al., 2011).

115 **C-type lectins and pathogen response**

116 Thirty-one C-type lectin (CLEC) genes were affected by SVs. CLECs are a rapidly
117 evolving gene family involved in innate immunity and pathogen recognition (Pees et al., 2016).
118 Five CLECs were affected by deletions—*clec-2, clec-10, clec-122, clec-143*, and *clec-174*—all
119 known to be modulated by bacterial pathogens (Pan et al., 2021; Sahu et al., 2012; White and
120 Herman, 2018). An additional CLEC, *clec-189,* was predicted as being deleted in EG4725 and
121 JU2526, partially deleted in QX1794 and XZ1516, and duplicated in EG4725 and JU2526.
122 Although it awaits functional annotation it may represent a novel immune effector (Pan et al.,
123 2021). The prediction of complete deletions and duplications spanning a single gene within the
124 same strains indicates that even long-read SV callers require further improvements.

Detoxification and xenobiotic response

Several families of detoxification genes were enriched in SVs, including cytochrome P450s (CYPs), glutathione S-transferases (GSTs), UDP-glucuronosyltransferases (UGTs), and nuclear hormone receptors (NHRs). Two caffeine inducible CYPs, *cyp-33e2* and *cyp-35a4* harbored deletions; the latter gene was also duplicated in QX1794 and is associated with the regulation of fat storage (Aarnio et al., 2011; Min et al., 2015). Other deletion-affected CYP genes included: *cyp-33c5* and *cyp-13a5*, which are involved in cadmium detoxification; and *cyp-33d3*, which is inducible by ethidium bromide and low pH (Cong et al., 2020; Cui et al., 2007). A deletion was also identified in *ugt-14*, which is upregulated in response to benzimidazole-class anthelmintics (Stasiuk et al., 2019).

Evolutionary and Functional Implications

Our results highlight the utility of long-read sequencing for characterizing SVs at a population scale. The high diversity of SVs in wild *C. elegans* populations—particularly those in gene families related to sensory perception, immunity, and detoxification—emphasizes the limitations of relying solely on N2 as a wild-type reference (Andersen and Rockman, 2022; Sterken et al., 2015). Many of the identified SVs may confer adaptive advantages, particularly in geographically and microbially diverse environments.

Importantly, many SVs overlap with hyper-divergent regions—genomic hotspots for adaptive diversity—shown in *C. elegans* and the parasitic *Heligmosomoides polygyrus* to contain genes involved in sensory perception and stress response (Lee et al., 2021; Stevens et al., 2023).

Going forward, hybrid approaches that integrate long- and short-read sequencing and transcriptomic data from wild isolates will help resolve the regulatory and phenotypic consequences of SVs (Zhang et al., 2022). In particular, eQTL mapping that includes SVs could

148 offer insights into complex traits not explained by single-nucleotide variation alone (Boyling et
149 al., 2022). Additionally, functional studies targeting SVs genes currently categorized as
150 “Unassigned” may reveal previously uncharacterized molecular mechanisms.
151
152

Materials and Methods

Strains

We analysed PacBio sequencing data from 14 *C. elegans* strains (DL238, ECA36, ECA396, EG4725, JU1400, JU2526, JU2600, JU310, MY2147, MY2693, NIC2, NIC526, QX1794, XZ1516) sourced from CeNDR and collected from Hawaii, New Zealand, and Western Europe (Figure S5).

Sequence alignment and structural variant calling

Reads were aligned to the N2 reference genome using NGMLR (v0.2.7), and alignments were sorted using Picard (v2.27.5) (Broad Institute, 2025; Sedlazeck et al., 2018). SVs were called using Sniffles2 (v2.0.7), filtered to include only deletions, duplications, and inversions ≥ 100bp (Smolka et al., 2024). SV calls were refined using Iris (v1.0.04) for improved breakpoint resolution and merged across strains using Jasmine (v1.1.5) and Iris (v1.0.4) (Kirsche et al., 2023).

SV annotation and functional profiling

We used Ensembl's Variant Effect Predictor (v.109.3) to predict the functional consequences of each SV (McLaren et al., 2016). SVs classified as "high impact" were selected for functional enrichment using easyGSEA (v1.3.1) and annotated using WormCat (v2) (Cheng et al., 2021; Holdorf et al., 2020). Genes were assigned to statistically significant terms across WormCat's hierarchical levels (FDR ≤ 0.05), though no single level was exclusively used due to trade-offs in resolution and coverage. Gene family membership was assigned based on the annotations in the WormBase database.

174 **Data presentation**

175 SV densities along chromosomes were plotted using karyoploteR (v1.26.0) (Gel and
176 Serra, 2017). SV allele frequencies were visualized UpSetPlot (v0.8.0) (Lex et al., 2014). Only
177 the largest intersections—25 duplications, 20 for deletions and inversions—were displayed. The
178 remaining plots were created using ggplot2 (v.3.4.2) (Wickham, 2016).

179

180 # Data availability

181 The code, and Snakemake pipeline, used in this project is freely available:

182 https://github.com/kyleLesack/sv_analysis_cendr_pacbio

183

184 # Acknowledgments

185 We thank Dr. Erik Andersen and members of the Andersen lab (Johns Hopkins University) for

186 discussions on using CeNDR datasets. We also acknowledge the high-performance computing

187 resources by the Faculty of Veterinary Medicine and Research Computing at the University of

188 Calgary.

189

190 # Funding

191 This work was funded through grants to JDW from the Natural Sciences and Engineering

192 Research Council of Canada (NSERC; #04589-2020).

193

194

References

Aarnio, V., Lehtonen, M., Storvik, M., Callaway, J.C., Lakso, M., Wong, G., 2011. Caenorhabditis Elegans Mutants Predict Regulation of Fatty Acids and Endocannabinoids by the CYP-35A Gene Family. Front. Pharmacol. 2. https://doi.org/10.3389/fphar.2011.00012

Andersen, E.C., Rockman, M.V., 2022. Natural genetic variation as a tool for discovery in *Caenorhabditis* nematodes. Genetics 220, iyab156. https://doi.org/10.1093/genetics/iyab156

Boyling, A., Perez-Siles, G., Kennerson, M.L., 2022. Structural Variation at a Disease Mutation Hotspot: Strategies to Investigate Gene Regulation and the 3D Genome. Front. Genet. 13, 842860. https://doi.org/10.3389/fgene.2022.842860

Broad Institute, 2025. broadinstitute/picard.

Carvalho, C.M.B., Lupski, J.R., 2016. Mechanisms underlying structural variant formation in genomic disorders. Nat Rev Genet 17, 224–238. https://doi.org/10.1038/nrg.2015.25

Chang, H.C., Paek, J., Kim, D.H., 2011. Natural polymorphisms in C. elegans HECW-1 E3 ligase affect pathogen avoidance behaviour. Nature 480, 525–529. https://doi.org/10.1038/nature10643

Cheng, X., Yan, J., Liu, Y., Wang, J., Taubert, S., 2021. eVITTA: a web-based visualization and inference toolbox for transcriptome analysis. Nucleic Acids Research 49, W207–W215. https://doi.org/10.1093/nar/gkab366

Cong, Y., Yang, H., Zhang, P., Xie, Y., Cao, X., Zhang, L., 2020. Transcriptome Analysis of the Nematode Caenorhabditis elegans in Acidic Stress Environments. Front. Physiol. 11, 1107. https://doi.org/10.3389/fphys.2020.01107

Cook, D.E., Zdraljevic, S., Roberts, J.P., Andersen, E.C., 2017. CeNDR, the Caenorhabditis elegans natural diversity resource. Nucleic Acids Res 45, D650–D657. https://doi.org/10.1093/nar/gkw893

Crombie, T.A., Zdraljevic, S., Cook, D.E., Tanny, R.E., Brady, S.C., Wang, Y., Evans, K.S., Hahnel, S., Lee, D., Rodriguez, B.C., Zhang, G., Van Der Zwagg, J., Kiontke, K., Andersen, E.C., 2019. Deep sampling of Hawaiian Caenorhabditis elegans reveals high genetic diversity and admixture with global populations. eLife 8, e50465. https://doi.org/10.7554/eLife.50465

Cui, Y., McBride, S.J., Boyd, W.A., Alper, S., Freedman, J.H., 2007. Toxicogenomic analysis of Caenorhabditis elegans reveals novel genes and pathways involved in the resistance to cadmium toxicity. Genome Biol 8, R122. https://doi.org/10.1186/gb-2007-8-6-r122

Cutter, A.D., Dey, A., Murray, R.L., 2009. Evolution of the Caenorhabditis elegans Genome. Molecular Biology and Evolution 26, 1199–1234. https://doi.org/10.1093/molbev/msp048

Daigle, A.T., Deiss, T.C., Melde, R.H., Bergthorsson, U., Katju, V., 2022. Bergerac strains of *Caenorhabditis elegans* revisited: expansion of Tc *1* elements imposes a significant genomic and fitness cost. G3 Genes|Genomes|Genetics 12, jkac214. https://doi.org/10.1093/g3journal/jkac214

Faria, R., Johannesson, K., Butlin, R.K., Westram, A.M., 2019. Evolving Inversions. Trends in Ecology & Evolution 34, 239–248. https://doi.org/10.1016/j.tree.2018.12.005

238 Farslow, J.C., Lipinski, K.J., Packard, L.B., Edgley, M.L., Taylor, J., Flibotte, S., Moerman,
239 D.G., Katju, V., Bergthorsson, U., 2015. Rapid Increase in frequency of gene copy-
240 number variants during experimental evolution in Caenorhabditis elegans. BMC
241 Genomics 16, 1044. https://doi.org/10.1186/s12864-015-2253-2
242 Gel, B., Serra, E., 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes
243 displaying arbitrary data. Bioinformatics 33, 3088–3090.
244 https://doi.org/10.1093/bioinformatics/btx346
245 Gruner, M., Nelson, D., Winbush, A., Hintz, R., Ryu, L., Chung, S.H., Kim, K., Gabel, C.V., Van
246 Der Linden, A.M., 2014. Feeding State, Insulin and NPR-1 Modulate Chemoreceptor
247 Gene Expression via Integration of Sensory and Circuit Inputs. PLoS Genet 10,
248 e1004707. https://doi.org/10.1371/journal.pgen.1004707
249 Hahnel, S.R., Zdraljevic, S., Rodriguez, B.C., Zhao, Y., McGrath, P.T., Andersen, E.C., 2018.
250 Extreme allelic heterogeneity at a Caenorhabditis elegans beta-tubulin locus explains
251 natural resistance to benzimidazoles. PLoS Pathog 14, e1007226.
252 https://doi.org/10.1371/journal.ppat.1007226
253 Holdorf, A.D., Higgins, D.P., Hart, A.C., Boag, P.R., Pazour, G.J., Walhout, A.J.M., Walker,
254 A.K., 2020. WormCat: An Online Tool for Annotation and Visualization of
255 *Caenorhabditis elegans* Genome-Scale Data. Genetics 214, 279–294.
256 https://doi.org/10.1534/genetics.119.302919
257 Katju, V., Bergthorsson, U., 2013. Copy-number changes in evolution: rates, fitness effects and
258 adaptive significance. Front. Genet. 4. https://doi.org/10.3389/fgene.2013.00273
259 Kirkpatrick, M., 2010. How and Why Chromosome Inversions Evolve. PLoS Biol 8, e1000501.
260 https://doi.org/10.1371/journal.pbio.1000501
261 Kirsche, M., Prabhu, G., Sherman, R., Ni, B., Battle, A., Aganezov, S., Schatz, M.C., 2023.
262 Jasmine and Iris: population-scale structural variant comparison and analysis. Nat
263 Methods 20, 408–417. https://doi.org/10.1038/s41592-022-01753-3
264 Koboldt, D.C., 2020. Best practices for variant calling in clinical sequencing. Genome Med 12,
265 91. https://doi.org/10.1186/s13073-020-00791-w
266 Konrad, A., Flibotte, S., Taylor, J., Waterston, R.H., Moerman, D.G., Bergthorsson, U., Katju, V.,
267 2018. Mutational and transcriptional landscape of spontaneous gene duplications and
268 deletions in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. U.S.A. 115, 7386–7391.
269 https://doi.org/10.1073/pnas.1801930115
270 Lee, D., Zdraljevic, S., Stevens, L., Wang, Y., Tanny, R.E., Crombie, T.A., Cook, D.E., Webster,
271 A.K., Chirakar, R., Baugh, L.R., Sterken, M.G., Braendle, C., Félix, M.-A., Rockman,
272 M.V., Andersen, E.C., 2021. Balancing selection maintains hyper-divergent haplotypes in
273 Caenorhabditis elegans. Nat Ecol Evol 5, 794–807. https://doi.org/10.1038/s41559-021-
274 01435-x
275 Lesack, K., Mariene, G.M., Andersen, E.C., Wasmuth, J.D., 2022. Different structural variant
276 prediction tools yield considerably different results in Caenorhabditis elegans. PLOS
277 ONE 17, e0278424. https://doi.org/10.1371/journal.pone.0278424
278 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H., 2014. UpSet: Visualization of
279 Intersecting Sets. IEEE Trans Vis Comput Graph 20, 1983–1992.
280 https://doi.org/10.1109/TVCG.2014.2346248

281 Ma, F., Lau, C.Y., Zheng, C., 2021. Large Genetic Diversity and Strong Positive Selection in F-
282 Box and GPCR Genes among the Wild Isolates of Caenorhabditis elegans. Genome Biol
283 Evol 13, evab048. https://doi.org/10.1093/gbe/evab048
284 McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P.,
285 Cunningham, F., 2016. The Ensembl Variant Effect Predictor. Genome Biol 17, 122.
286 https://doi.org/10.1186/s13059-016-0974-4
287 Min, H., Kawasaki, I., Gong, J., Shim, Y.-H., 2015. Caffeine Induces High Expression of cyp-
288 35A Family Genes and Inhibits the Early Larval Development in Caenorhabditis elegans.
289 Molecules and Cells 38, 236–242. https://doi.org/10.14348/molcells.2015.2282
290 Pan, W., Huang, X., Guo, Z., Nagarajan, R., Mylonakis, E., 2021. Identification and Functional
291 Analysis of Cytokine-Like Protein CLEC-47 in Caenorhabditis elegans. mBio 12,
292 e02579-21. https://doi.org/10.1128/mBio.02579-21
293 Pees, B., Yang, W., Zárate-Potes, A., Schulenburg, H., Dierking, K., 2016. High Innate Immune
294 Specificity through Diversified C-Type Lectin-Like Domain Proteins in Invertebrates. J
295 Innate Immun 8, 129–142. https://doi.org/10.1159/000441475
296 Sahu, S.N., Lewis, J., Patel, I., Bozdag, S., Lee, J.H., LeClerc, J.E., Cinar, H.N., 2012. Genomic
297 Analysis of Immune Response against Vibrio cholerae Hemolysin in Caenorhabditis
298 elegans. PLoS ONE 7, e38200. https://doi.org/10.1371/journal.pone.0038200
299 Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., Schatz,
300 M.C., 2018. Accurate detection of complex structural variations using single-molecule
301 sequencing. Nat Methods 15, 461–468. https://doi.org/10.1038/s41592-018-0001-7
302 Smolka, M., Paulin, L.F., Grochowski, C.M., Horner, D.W., Mahmoud, M., Behera, S., Kalef-
303 Ezra, E., Gandhi, M., Hong, K., Pehlivan, D., Scholz, S.W., Carvalho, C.M.B., Proukakis,
304 C., Sedlazeck, F.J., 2024. Detection of mosaic and population-level structural variants
305 with Sniffles2. Nat Biotechnol 1–10. https://doi.org/10.1038/s41587-023-02024-y
306 Stasiuk, S.J., MacNevin, G., Workentine, M.L., Gray, D., Redman, E., Bartley, D., Morrison, A.,
307 Sharma, N., Colwell, D., Ro, D.K., Gilleard, J.S., 2019. Similarities and differences in the
308 biotransformation and transcriptomic responses of Caenorhabditis elegans and
309 Haemonchus contortus to five different benzimidazole drugs. Int J Parasitol Drugs Drug
310 Resist 11, 13–29. https://doi.org/10.1016/j.ijpddr.2019.09.001
311 Sterken, M.G., Snoek, L.B., Kammenga, J.E., Andersen, E.C., 2015. The laboratory
312 domestication of Caenorhabditis elegans. Trends in Genetics 31, 224–231.
313 https://doi.org/10.1016/j.tig.2015.02.009
314 Stevens, L., Martínez-Ugalde, I., King, E., Wagah, M., Absolon, D., Bancroft, R., Gonzalez de la
315 Rosa, P., Hall, J.L., Kieninger, M., Kloch, A., Pelan, S., Robertson, E., Pedersen, A.B.,
316 Abreu-Goodger, C., Buck, A.H., Blaxter, M., 2023. Ancient diversity in host-parasite
317 interaction genes in a model parasitic nematode. Nat Commun 14, 7776.
318 https://doi.org/10.1038/s41467-023-43556-w
319 Thomas, J.H., Robertson, H.M., 2008. The Caenorhabditis chemoreceptor gene families. BMC
320 Biol 6, 42. https://doi.org/10.1186/1741-7007-6-42
321 Thompson, O.A., Snoek, L.B., Nijveen, H., Sterken, M.G., Volkers, R.J.M., Brenchley, R., Van'T
322 Hof, A., Bevers, R.P.J., Cossins, A.R., Yanai, I., Hajnal, A., Schmid, T., Perkins, J.D.,
323 Spencer, D., Kruglyak, L., Andersen, E.C., Moerman, D.G., Hillier, L.W., Kammenga,
324 J.E., Waterston, R.H., 2015. Remarkably Divergent Regions Punctuate the Genome

Assembly of the *Caenorhabditis elegans* Hawaiian Strain CB4856. Genetics 200, 975–989. https://doi.org/10.1534/genetics.115.175950

Vergara, I.A., Tarailo-Graovac, M., Frech, C., Wang, J., Qin, Z., Zhang, T., She, R., Chu, J.S., Wang, K., Chen, N., 2014. Genome-wide variations in a natural isolate of the nematode Caenorhabditis elegans. BMC Genomics 15, 255. https://doi.org/10.1186/1471-2164-15-255

Vidal, B., Aghayeva, U., Sun, H., Wang, C., Glenwinkel, L., Bayer, E.A., Hobert, O., 2018. An atlas of Caenorhabditis elegans chemoreceptor expression. PLoS Biol 16, e2004218. https://doi.org/10.1371/journal.pbio.2004218

White, C.V., Herman, M.A., 2018. Transcriptomic, Functional, and Network Analyses Reveal Novel Genes Involved in the Interaction Between Caenorhabditis elegans and Stenotrophomonas maltophilia. Front. Cell. Infect. Microbiol. 8, 266. https://doi.org/10.3389/fcimb.2018.00266

Wickham, H., 2016. ggplot2, Use R! Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-24277-4

Zdraljevic, S., Fox, B.W., Strand, C., Panda, O., Tenjo, F.J., Brady, S.C., Crombie, T.A., Doench, J.G., Schroeder, F.C., Andersen, E.C., 2019. Natural variation in C. elegans arsenic toxicity is explained by differences in branched chain amino acid metabolism. eLife 8, e40260. https://doi.org/10.7554/eLife.40260

Zhang, G., Roberto, N.M., Lee, D., Hahnel, S.R., Andersen, E.C., 2022. The impact of species-wide gene expression variation on Caenorhabditis elegans complex traits. Nat Commun 13, 3462. https://doi.org/10.1038/s41467-022-31208-4

Zhao, Y., Long, L., Wan, J., Biliya, S., Brady, S.C., Lee, D., Ojemakinde, A., Andersen, E.C., Vannberg, F.O., Lu, H., McGrath, P.T., 2020. A spontaneous complex structural variant in rcan-1 increases exploratory behavior and laboratory fitness of Caenorhabditis elegans. PLoS Genet 16, e1008606. https://doi.org/10.1371/journal.pgen.1008606