Extreme allelic heterogeneity at a *Caenorhabditis elegans* beta-tubulin locus explains natural resistance to benzimidazoles

Steffen R. Hahnel[#,1], Stefan Zdraljevic[#,1,2], Briana C. Rodriguez[1], Yuehui Zhao[3], Patrick T. McGrath[3], and Erik C. Andersen[1,2,4] *

1. Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA
2. Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, IL 60208, USA
3. School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332
4. Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago, IL 60611, USA
# Equal contribution
* Corresponding author

Erik C. Andersen
Assistant Professor of Molecular Biosciences
Northwestern University
Evanston, IL 60208, USA
Tel: (847) 467-4382
Fax: (847) 491-4461
Email: Erik.Andersen@Northwestern.edu

Steffen R. Hahnel, steffen.hahnel@northwestern.edu, ORCID 0000-0001-8848-0691
Stefan Zdraljevic, stefanzdraljevic2018@u.northwestern.edu, ORCID 0000-0003-2883-4616
Briana C. Rodriguez, briana.rodriguez@northwestern.edu
Yuehui Zhao, yzhao349@gatech.edu
Patrick T. McGrath, patrick.mcgrath@biology.gatech.edu, ORCID 0000-0002-1598-3746
Erik C. Andersen, erik.andersen@northwestern.edu, ORCID 0000-0003-0229-9651

Abstract

Benzimidazoles (BZ) are essential components of the limited chemotherapeutic arsenal available to control the global burden of parasitic nematodes. The emerging threat of BZ resistance among nearly all nematode species necessitates the development of novel strategies to identify genetic and molecular mechanisms underlying this resistance. All detection of parasitic helminth resistance to BZ is focused on the genotyping of three variant sites in the orthologs of the β-tubulin gene found to confer resistance in the free-living nematode *Caenorhabditis elegans*. Because of the limitations of laboratory and field experiments in parasitic nematodes, it is difficult to look beyond these three sites, and additional BZ resistance is observed in the field. Here, we took an unbiased genome-wide mapping approach in the free-living nematode species *C. elegans* to identify the genetic underpinnings of natural resistance to the commonly used BZ, albendazole (ABZ). We found a wide range of natural variation in ABZ resistance in natural *C. elegans* populations. In agreement with known mechanisms of BZ resistance in parasites, we find that a majority of the variation in ABZ resistance among wild *C. elegans* strains is caused by variation in the β-tubulin gene *ben-1*. This result shows empirically that resistance to ABZ naturally exists and segregates within the *C. elegans* population, suggesting that selection in natural niches could enrich for resistant alleles. We identified 25 distinct *ben-1* alleles that are segregating at low frequencies within the *C. elegans* population, including many novel molecular variants. Population genetic analyses indicate that *ben-1* variation arose multiple times during the evolutionary history of *C. elegans* and provide evidence that these alleles likely occurred recently because of local selective pressures. Additionally, we find purifying selection at all five β-tubulin genes, despite predicted loss-of-function resistants variants in *ben-1*, indicating that BZ resistance in natural niches is a stronger selective pressure than loss of one β-tubulin gene. Furthermore, we use genome-editing to show that the most common parasitic nematode β-tubulin allele that confers BZ resistance, F200Y, confers resistance in *C. elegans*. Importantly, we identified a novel genomic region that is correlated with ABZ resistance in the *C. elegans* population but independent of *ben-1* and the other β-tubulin loci, suggesting that there are multiple mechanisms underlying BZ resistance. Taken together, our results establish a population-level resource of nematode natural diversity as an important model for the study of mechanisms that give rise to BZ resistance.

Author summary

Nematode parasites have a tremendous impact on human health with almost two billion people infected worldwide. The control of nematode infections relies mainly on the efficacy of a limited repertoire of anthelmintic compounds, including the benzimidazoles (BZ). Already a significant problem in veterinary medicine, increasing evidence exists for the development of BZ resistance in nematodes that infect humans. Laboratory screens and field surveys identified β-tubulin genes as major determinants of BZ resistance in nematodes but detailed population-wide genetic analyses of resistance mechanisms are only just beginning. Therefore, we took advantage of the free-living model organism *Caenorhabditis elegans* to study the genetic basis of resistance to the commonly used BZ, albendazole (ABZ) in a natural nematode population. Performing genome-wide association mappings, we were able to identify extreme heterogeneity in the β-tubulin gene *ben-1* as a major determinant of ABZ resistance. Moreover, our study provided new insights into the effects of missense and loss-of-function alleles at this locus, and how anthelmintic resistance could have developed within a natural nematode population.

Introduction

72

73 Parasitic nematodes have a tremendous impact on global health and socio-economic development,

74 especially in the developing world [1]. They are among the most widespread human pathogens, and

75 almost two billion people are estimated to suffer from infection of one or multiple nematode species [1,2].

76 The main endemic areas of nematode infections are highly correlated with tropical and subtropical regions

77 worldwide. Because of their detrimental impact on human health, several nematode infections belong to a

78 class of diseases designated by the World Health Organization (WHO) as Neglected Tropical Diseases

79 (NTDs). Altogether, the loss of disability-adjusted life years (DALY) caused by parasitic nematodes is

80 conservatively estimated to be 10 million DALYs per year, which ranks them among the top of all NTDs

81 [1]. Apart from this drastic impact on human health, several nematode species infect a variety of key crops

82 and livestock causing substantial economic losses throughout the world [3].

83

84 Global control of nematode infections relies on the efficacy of a limited repertoire of anthelmintic

85 drugs, including benzimidazoles (BZ). BZs are frequently used in mass drug administration (MDA)

86 programs to treat parasitic nematode infections in endemic regions. However, the long-term success of

87 these MDA programs is limited by the persistence and re- emergence of nematode infections in these

88 regions. The most significant of these factors is the high reinfection rate caused by long-term parasitic

89 nematode reservoirs, which necessitates frequent deworming of affected communities [2]. As a

90 consequence of continuous drug-pressure and relaxation cycles, parasite populations have the potential to

91 develop anthelmintic resistance. This resistance is already a significant problem in veterinary medicine [4],

92 and resistance in the parasitic nematodes that infect humans is becoming increasingly evident. For

93 example, reduced cure rates have been observed for soil-transmitted helminthiases, which might be

94 attributable to BZ resistance [5–7]. To face this growing threat, detailed knowledge of anthelmintic

95 resistance mechanisms is required to improve diagnostic tools for field surveys and educate drug-

96 treatment strategies in MDA programs.

97

98 *In vitro* studies have shown that BZs inhibit the polymerization of microtubules [8–10]. Mutagenesis

99 screens that selected for BZ resistance in *Saccharomyces cerevisiae* [11] and *Caenorhabditis elegans*

100 [12] have independently identified numerous β-tubulin mutant alleles. These findings were then translated

101 to veterinary-relevant nematodes, where three major single-nucleotide variants (SNVs) in parasitic

102 nematode β-tubulin genes were found to be highly correlated with BZ resistance. The most common

103 variant causes a tyrosine to phenylalanine amino acid change at position 200 (F200Y) [13]. Other SNVs

104 linked to BZ resistance have been described at positions 167 (F167Y) and 198 (E198A) in several

105 nematode species [14,15]. Although all three of these missense variants have been shown to reduce BZ

106 binding affinity to purified β-tubulins *in vitro* [16,17], the BZ β-tubulin binding site remains experimentally

107 uncharacterized. Furthermore, these missense variants do not explain all of the BZ resistance observed in

108 the field [4,18]. This discrepancy might indicate that additional, non-target related components such as

109 drug efflux pumps and detoxification enzymes might contribute to BZ resistance [4,19,20].

110

111   Additional genetic variants associated with natural BZ resistance can be identified using

112 quantitative genetic approaches that consider genotypic and phenotypic variation present within a wild

113 population of parasitic nematodes. However, parasitic nematodes are not easily amenable to quantitative

114 genetic approaches because of their complicated life cycles, their poorly annotated reference genomes,

115 and limited molecular and genetic tools [21–23]. Recent successes in the hookworm parasites found

116 genomic intervals that underlie ivermectin resistance [21,24], but similar approaches have not been

117 applied to BZ resistance yet. By contrast, the free-living nematode species *C. elegans* has a short life

118 cycle, a well annotated reference genome [25–27], and an abundance of molecular and genetic tools for

119 the characterization of BZ responses [28–30]. A recent example that applied quantitative genetic

120 approaches to investigated BZ responses in *C. elegans* used a mapping population generated between

121 two genetically divergent strains to identify a quantitative trait loci (QTL) linked to BZ resistance [30].

122 Importantly, the QTL identified in this study did not overlap with β-tubulin genes, suggesting that

123 quantitative genetic approaches in *C. elegans* can be used to discover novel mechanisms associated with

124 BZ resistance.

125

126   In the present study, we leveraged the power of *C. elegans* natural diversity to perform genome-

127 wide association (GWA) mappings for BZ resistance. We used a set of 249 wild *C. elegans* isolates

128 available through the *C. elegans* Natural Diversity Resource (CeNDR) [27] to identify genomic regions that

129 contribute to albendazole (ABZ) resistance. ABZ is a broadly administered BZ used to treat parasitic

130 nematode infections in humans and livestock [4,30]. We show that a major source of ABZ resistance in the

131 *C. elegans* population is driven by putative loss-of-function (LoF) variants in the *ben-1* locus. Notably, we

132 found 25 distinct *ben-1* alleles with low minor allele frequencies (MAF) that contribute to ABZ resistance.

133 We show that these putative *ben-1* LoF variants arose independently during the evolutionary history of the

134 *C. elegans* species, which suggests that local BZ selective pressures might have contributed to the

135 extreme allelic heterogeneity at this locus. We next made use of the extensive molecular toolkit available

136 in *C. elegans* to verify that the introduction of the *ben-1* LoF alleles do not result in any detectable fitness

137 consequences in standard laboratory conditions, but does confer ABZ resistance. Taken together, these

138 results suggest that *C. elegans* and likely free-living stages of parasitic species encounter natural

139 compounds that promote BZ resistance through selection for standing or *de novo* variation in conserved

140 nematode-specific β-tubulin genes.

Results

Genetically distinct *C. elegans* natural isolates respond differently to ABZ

We investigated the differences in responses across a panel of *C. elegans* wild strains to one of the most commonly used BZ compounds (albendazole, ABZ) in human and veterinary medicine [4,27,37]. To test the effects of ABZ treatment on *C. elegans*, we exposed four genetically divergent *C. elegans* isolates to various ABZ concentrations and measured the number and length of progeny the animals produced. We used these two traits because they represent endpoint measures for ABZ efficacy against human parasitic nematodes. After four days of exposure to ABZ, we detected a decrease in brood size and progeny length for all four assayed strains (Supplemental Figure 1; Supplemental data 1 and 2). Additionally, this assay revealed differential ABZ sensitivity among these wild strains, as measured by brood size and progeny length. For subsequent genome-wide association (GWA) mapping experiments, we used an ABZ concentration of 12.5 µM, which is a concentration that induced robust differences in ABZ responses among the four assayed wild strains.

Natural variation in *C. elegans* ABZ responses maps to multiple genomic regions, including the *ben-1* locus

To identify genomic loci that underlie strain-specific ABZ responses, we performed genome-wide association (GWA) mappings using HTA fitness data obtained from exposing 209 *C. elegans* to either DMSO (control) or ABZ and DMSO conditions (Supplemental data 3 and 4). We employed two statistical methods to perform GWA mappings using the regressed animal length (q90.TOF) and normalized brood size (norm.n) traits, a single-marker and a gene-burden approach [39,40]. Using the single-marker approach [32,35–37], we identified three distinct quantitative trait loci (QTL) that explained variation in animal length among wild isolates exposed to ABZ, whereas the brood-size trait did not map to any significant genomic loci (Supplemental data 5 and 6). Two of the animal-length QTL we identified are located on chromosome II, and the third is located on chromosome V (Figure 1A). The QTL on the left arm of chromosome II spans from 25 kb to 3.9 Mb and has a peak-marker position at 458 kb. The second QTL

168 on chromosome II has a peak-marker position at 11 Mb and spans from 10.17 Mb to 11.6 Mb. The QTL on

169 the right arm of chromosome V spans from from 18.04 Mb to 18.68 Mb, with the peak marker located at

170 18.35 Mb (Supplemental Figure 2; Supplemental data 5). Remarkably, the β-tubulin gene *ben-1*, shown to

171 be the major determinant for ABZ resistance in the *C. elegans* laboratory strain N2 [12,13], did not map

172 using this single marker-based approach (Figure 1A).

173

174   To complement the single-marker mapping described above, we performed a gene-burden

175 approach to map BZ resistance [39,40]. Compared to the single-marker approach that includes SNVs with

176 a minimum 5% minor allele frequency among all strains, the gene-burden approach incorporates all rare

177 strain-specific variation within a gene for association testing [39,56]. With this alternative mapping

178 strategy, we identified the same two QTL on chromosome II that we identified with the single-marker

179 approach. In addition to the chromosome II QTL, we found a significant association between variation at

180 the *ben-1* locus on chromosome III and animal-length variation in response to ABZ (Figure 1B). These

181 results suggest that no variants are shared above 5% minor allele frequency among the wild isolates at

182 the *ben-1* locus (chromosome III, 3,537,688-3,541,628 bp). An additional significant gene was detected on

183 the right arm of chromosome I (Supplemental data 7 and 8).

184

185 ***C. elegans* ABZ resistance correlates with extreme allelic heterogeneity at the *ben-1* locus**

186 To explain the differences between the single-marker and gene-burden based GWA mapping results with

187 respect to *ben-1*, we investigated the natural variation found at this genomic locus in more detail (Figure 2;

188 Supplemental data 9). We first used the *snpeff* function of the *cegwas* package [27,57] to look for SNVs in

189 *ben-1* with predicted moderate-to-high effects on gene function [27,57]. These moderate-to-high impact

190 variants include missense variants, splice donor and acceptor variants, and alternative start and stop

191 codons that might disrupt the open reading frame of *ben-1*.

192

193   In the set of 209 strains used for the GWA mapping, we identified 19 strains with moderate-to-high

194 impact variants in the *ben-1* locus, as predicted by *snpEff* [57], including 13 strains with amino-acid

195 substitutions. Amino-acid substitutions in β-tubulin genes are important markers for BZ resistance in

196 parasitic nematodes and are hypothesized to reduce binding affinity of BZs to β-tubulin [16,17]. In the

197 *C. elegans* panel of wild strains, we identified six strains that have the F200Y mutation in BEN-1, which is

198 known to be the most common BZ resistance marker in livestock parasites [13]. Additionally, we detected

199 novel amino-acid substitutions in BEN-1, which have not previously been associated with ABZ resistance

200 in nematodes. These missense variants include an A185P substitution variant, found in two strains, as

201 well as the E69G, Q131L, S145F, M257I, and D404N substitution variants, all of which were unique to

202 single strains. In contrast to the amino-acid variants F200Y, E198A, and F167Y found in parasitic

203 nematodes, which cluster at the putative ABZ binding site [58–60], these novel variants are more

204 distributed throughout the protein structure (Supplemental Figure 3). We classified a strain as resistant to

205 ABZ if its animal-length phenotype after ABZ exposure was greater than 50% of the difference between

206 the most and least ABZ-resistant strains (Figure 2A). Eleven of the twelve strains with amino-acid

207 substitutions are more resistant to ABZ than strains with no variation at the *ben-1* locus. Similarly, strains

208 with predicted high-impact variants in *ben-1* are resistant to ABZ treatment. The strains with high-impact

209 variants include five strains with unique premature stop codons and one strain with a predicted splice-

210 donor variant at the end of exon 1. However, many ABZ-resistant strains in the *C. elegans* population do

211 not contain any of these rare and common variants. Therefore, these strains either have other types of

212 deleterious variants at the *ben-1* locus or are resistant to ABZ through a distinct mechanism. To

213 differentiate these two possibilities, we manually curated all of the strains raw sequence read alignment

214 files (BAM files), publicly available through the CeNDR website [27]. This in-depth investigation revealed

215 extreme allelic heterogeneity at the *ben-1* locus in the natural *C. elegans* population. Twenty-seven strains

216 had either rare deletions or insertions in *ben-1* (22 deletions, 5 insertions; Supplemental data 9).

217 Interestingly, a subset of individuals with missense variants described above were found to be in perfect

218 linkage disequilibrium (LD) with nearby deletion alleles. All six strains with the F200Y allele share the

219 same deletion of approximately 160 bp that partially removes exons 3 and 4. The close proximity of the

220 annotated F200Y to this 160 bp deletion suggested that the F200Y variant is actually an error in read

221 alignment, which we confirmed by manual review of the BAM files of corresponding strains (JU751,

222 JU830, JU2581, JU2587, JU2593, JU2829) (genome browser on CeNDR) [27]. The majority of paired-end

223 reads at this genomic position that lead to the call of the F200Y allele in *ben-1*, map with their mate reads

224   to the locus of the β-tubulin gene *tbb-2*, which is in relative close genomic distance (chromosome III,

225   4,015,769 - 4,017,643 bp). Additionally, the amino acid substitution (E69G), which is only observed in a

226   single isotype (KR314), co-occurs with a deletion in exon 2 that is predicted to cause a frameshift in the

227   *ben-1* open reading frame.

228

229   Of all the strains with *ben-1* SNVs and indels, only three individuals were not classified as resistant

230   to ABZ treatment. The least resistant of these three strains was the CB4856 strain, which contains a nine

231   base pair deletion near the end of the *ben-1* coding sequence. Because this in-frame deletion variant is at

232   the end of the coding sequence, we hypothesized that CB4856 still contains a functional copy of the *ben-1*

233   gene and this variant likely does not confer resistance to ABZ. The next least resistant strain is QG2075,

234   which shares a four base pair deletion in exon 1 of *ben-1* with the ABZ-resistant isotype WN2033. During

235   the growth phase of our HTA assay, we noted that the QG2075 strain has a slow-growth phenotype in

236   normal growth conditions, which is likely confounding our phenotypic measurements in ABZ. Finally, the

237   JU2862 strain has a unique four base pair deletion in exon 4 of *ben-1*, that is predicted to cause a

238   frameshift in the *ben-1* open reading frame. We note that JU2862 is right at our arbitrary ABZ-resistance

239   threshold and upon re-phenotyping with higher replication might be classified as ABZ resistant.

240

241   Among the remaining ABZ-resistant strains, we identified a putative transposon insertion in exon 5

242   of *ben-1* in the JU3125 strain. The genomic origin of this putative transposon insertion is from position

243   17.07 Mb on chromosome X and corresponds to a cut and paste DNA transposon Tc5B, which is part of

244   the TcMar-Tc4 transposon superfamily [61,62]. Finally, the MY518 strain is resistant to ABZ but did not

245   contain any of the above classes of variation. However, we did identify a 1 kb inversion that spans exon 1

246   and the promoter region of *ben-1*. Remarkably, all structural variants present in *ben-1* are only present in

247   one or few wild strains and are mostly predicted to cause loss of *ben-1* function (Figure 2). Altogether, the

248   putative loss-of-function variants described above explain 73.8% of the phenotypic variation present in the

249   natural *C. elegans* population.

250

251

252

Within-species selective pressures at the *ben-1* locus

254 The presence of multiple low-frequency *ben-1* alleles that confer resistance to ABZ treatment suggests

255 that selection might have acted on this locus within the *C. elegans* population. To determine if this

256 hypothesis is possible, we calculated the Ka/Ks ratio between *ben-1* genes from *C. elegans* and from two

257 diverged nematode species *C. remanei* and *C. briggsae* [63]. The Ka/Ks ratio between the *ben-1* coding

258 sequences of *C. elegans* and these two diverged species is ~0.008, which indicates that the evolution of

259 this locus is constrained across species. However, among the 249 wild isolates within the *C. elegans*

260 species [27], we identified 10 synonymous and 22 nonsynonymous, stop-gained, or splice-site variants in

261 the *ben-1* locus when compared to the N2 reference genome [25]. We note that the synonymous variants

262 identified in *ben-1* are primarily found in wild strains that contain putative loss-of-function alleles. These

263 results indicate that, despite the evolutionary constraint we observe at the *ben-1* locus across nematode

264 species, an excess of potentially adaptive mutations (~2.2X nonsynonymous variants) have arisen within

265 the *C. elegans* species. These results are consistent with our estimates of Tajima's *D* [64] at the *ben-1*

266 locus. When we considered all variant types across coding and non-coding regions of *ben-1*, we found

267 Tajima's *D* to be -2.02. When we only consider putative loss-of-function variants in the *ben-1* locus, the

268 Tajima's *D* estimate is -2.59. This large negative value of Tajima's *D* likely reflects the high number of rare

269 *ben-1* alleles present in the *C. elegans* population. We next calculated Tajima's *D* for the genomic region

270 surrounding *ben-1* (Figure 3A; Supplemental data 10). These results show a strong negative dip in

271 Tajima's *D* (< -1.5) between 3.50 and 3.55 Mb, which is the region directly surrounding *ben-1* (3.537 -

272 3.541 Mb). This observation indicates that the *C. elegans* population has undergone a population

273 expansion or that the region surrounding *ben-1* may have been subject to selective pressures. To

274 differentiate between these two possibilities, we calculated Fay and Wu's *H* [65] and Zeng's *E* [66] for *ben-*

275 *1* (Figure 3A; Supplemental data 11). We noticed that these two statistics showed peaks around the *ben-1*

276 locus, which do not correspond to our observations of multiple low minor allele frequency alleles.

277 However, when we consider *ben-1* coding variation and non-coding variation separately, we found the

278 H_{coding} statistic to be 0.34 and the $H_{noncoding}$ statistic to be -5.5, which indicate that when considering high

279 frequency derived alleles, the coding sequence of this locus is evolving neutrally and that there are many

280   high-frequency derived alleles in the intronic and UTR regions. By contrast, we found the E_{coding} statistic,

281   which considers low and high frequency alleles, to be -1.7 and the $E_{noncoding}$ to be 3.4. Taken together,

282   these results indicate that recent selective pressures have acted on the *ben-1* locus.

283

284   If localized selective pressures, possibly caused by increased levels of environmental BZ, drove

285   the loss of *ben-1* function in a subset of wild strains, then we could expect to see geographic clustering of

286   resistant strains. When we looked at the distribution of strains with putative *ben-1* loss-of-function variants,

287   we observed no trend in the sampling locations of these individuals (Figure 3B). However, we do note that

288   highly diverse strains sampled from various locations on the Pacific Rim [27,37,38] do not harbor any

289   putative *ben-1* loss-of-function alleles (Figure 3B). This diverse set of strains is hypothesized to represent

290   the ancestral state of *C. elegans*, because they do not show signs of chromosome-scale selective sweeps

291   throughout their genomes [38]. The lack of *ben-1* loss-of-function variants in these ancestral strains

292   suggests that variation in *ben-1* arose after individuals in the species spread throughout the world. We

293   next considered that strains with putative *ben-1* loss-of-function alleles might have been isolated from

294   similar local environments and substrates. Strains that are predicted to have a functional *ben-1* gene have

295   been isolated from a greater diversity of environmental sampling locations (Supplemental Figure 4A) and

296   sampling substrates (Supplemental Figure 4B) than strains with predicted loss-of-function variants in *ben-*

297   *1*. However, a hypergeometric test for enrichment within specific locations and substrates showed no

298   signs of significant enrichment. This lack of enrichment might be caused by biases in global coverage of

299   *C. elegans* sampling and low sampling density. Nevertheless, the observation that the diversity of *ben-1*

300   alleles arose independently on various branches of the *C. elegans* phylogeny lends support to the

301   hypothesis that local selective pressures have acted on the *ben-1* locus (Figure 3C).

302

303   **ben-1 natural variants confer BZ resistance to sensitive *C. elegans* strains**

304   The majority of the variants we observe at the *ben-1* locus are predicted to result in the loss of gene

305   function. Our findings in *C. elegans* are in contrast to findings in parasitic nematode populations that

306   describe the F200Y and other missense variants as the major variants contributing to BZ resistance [13].

307   To test whether the putative loss-of-function variants in *ben-1* that we observe in the *C. elegans* population

308   confer the same level of BZ resistance as the known F200Y allele, we introduced a *ben-1* deletion in the

309   N2 strain using CRISPR/Cas9 (referred to as Del). In addition to the Del strain, we introduced the F200Y

310   *ben-1* allele into the N2 strain (referred to as F200Y) to test whether this allele confers BZ resistance in

311   *C. elegans* and to compare the levels of BZ resistance between the loss-of-function and missense alleles.

312   We exposed the N2 parental, Del, and F200Y strains to 12.5 μM ABZ using the HTA described above.

313   Both the Del and F200Y variants conferred ABZ resistance to the otherwise sensitive N2 parental strain

314   (Figure 4A; Supplemental Figure 5; Supplemental Data 12 and 13). Interestingly, we found no significant

315   difference in BZ resistance between the Del and F200Y strains (*p*-value = 0.99, TukeyHSD). These results

316   show that putative loss-of-function variants of *ben-1* and known parasitic BZ resistance alleles confer BZ

317   resistance in otherwise isogenic *C. elegans* genetic backgrounds.

318

319   To complement the results from the liquid-based HTA experiments, we performed a plate-based

320   competition assay. In the competition experiments, we individually competed the F200Y, Del, and parental

321   N2 strains against an N2 strain that contains a barcode sequence (PTM229). We quantified the relative

322   allele frequencies of the barcoded strain for the first, third, fifth, and seventh generations of the competition

323   assay (Figure 4B; Supplemental Data 14 and 15). Throughout the competition assay, and for all strains

324   tested, the relative frequencies of the barcoded strain did not significantly deviate from the initial frequency

325   when grown on DMSO plates. These results suggest that in standard laboratory conditions, the BEN-1

326   F200Y and *ben-1* deletion alleles do not have fitness consequences. We observed the same trend when

327   we competed the N2 and the barcoded N2 strains on ABZ plates. However, the allele frequencies of the

328   barcoded strain dropped to ~20% when competed against strains that contain either of the two *ben-1*

329   alleles on ABZ plates (relative fitness w = ~ 1.3, both). These two independent assays show that the two

330   *ben-1* alleles confer BZ resistance with no negative fitness consequence under standard laboratory growth

331   conditions.

332

333   **Additional genomic intervals contribute to ABZ resistance in the *C. elegans* population**

334   Above, we showed that extreme allelic heterogeneity at the *ben-1* locus explains 73.8% of the phenotypic

335   variation in response to ABZ treatment. To identify potential genomic loci that contribute to the remaining

336   26.2% of the observed BZ response variation, we statistically corrected the animal length phenotype in

337   response to ABZ treatment by using the presence of a putative loss-of-function variant at the *ben-1* locus

338   in a strain as a covariate for linear regression. After correcting for variation at the *ben-1* locus, we

339   performed GWA mapping using both the single-marker (Figure 5A; Supplemental Data 16, 17, and 18)

340   and gene-burden (Supplemental Figure 6; Supplemental Data 19, 20, and 21) approaches. Both of these

341   approaches resulted in the disappearance of the two QTL on chromosome II and the QTL on chromosome

342   V (Figure 5A, Supplemental Figure 7). This result suggests that the cumulative variation at the *ben-1* locus

343   is in complex interchromosomal linkage disequilibrium (LD) with these three loci (Supplemental Figure 7),

344   despite the loci on chromosomes II and V not being in strong LD with each other (Supplemental Figure

345   2B). Interestingly, using the single-marker GWA-mapping approach, we found an additional genomic locus

346   on chromosome X that was significantly associated with ABZ resistance (Figure 5A; Supplemental Figure

347   8). However, the gene-burden based approach did not result in significant associations between ABZ

348   resistance and rare variation within any genes. These results suggest that ABZ resistance in the

349   *C. elegans* population is caused by at least two loci because additional genetic variation present in the

350   *C. elegans* population contributes to phenotypic variation in response to ABZ treatment.

Discussion

351 ABZ is a broadly administered BZ used to treat parasitic nematode infections in humans and livestock

353 [4,67]. Already a significant problem in veterinary medicine, the heavy reliance of ABZ and other BZ

354 compounds in MDA programs, the small repertoire of other anthelmintic compounds, and high rates of re-

355 infection of parasitic nematodes in endemic regions around the world have raised the fear of the

356 emergence of BZ resistance among parasitic nematode populations that infect humans [4,68]. Though the

357 BZ-resistance alleles found in model organisms through classical genetic screens have had little direct

358 translatability to parasitic nematodes, these classic alleles have been instrumental toward the elucidation

359 of the mechanism of action of BZs [4,69]. Recent advances in sequencing technologies have enabled

360 researchers to take a quantitative genetics approach to search for novel mechanisms of anthelmintic

361 resistance in natural parasitic and non-parasitic nematodes [27,29,30,70–72]. In the present study, we

362 leveraged genetic diversity within the *C. elegans* population to study the genetic basis of ABZ resistance

363 within this species.

364

365 To identify the genetic basis of ABZ resistance in *C. elegans*, we used two unbiased genome-wide

366 association approaches, single-marker [35,36] and gene-burden [40] based mappings. We found that the

367 *C. elegans* population harbors extreme allelic heterogeneity at the *ben-1* locus, and that this variation

368 contributes to differential ABZ resistance among wild isolates (Figure 1B, Figure 2A). The variants in the

369 *ben-1* locus that contribute to ABZ resistance are present at low allele frequencies (<0.05 MAF) within the

370 global *C. elegans* population and arose independently during the evolutionary history of the species

371 (Figure 3). As a result of this complex demographic history, common variants (>0.05 MAF) used as

372 genetic markers in the mappings near the *ben-1* locus are not in LD with the *ben-1* alleles that confer ABZ

373 resistance. Therefore, we were unable to detect an association between the *ben-1* locus and ABZ

374 response with the single-marker based mapping approach. We note that burden-based mapping

375 approaches are only possible in species with well annotated genomes and genome-wide genetic variation,

376 such as *C. elegans* and other model organisms. To extend this type of analysis to parasitic species, more

377 effort will need to be made toward improving parasitic nematode reference genomes and the accumulation

378 of genome-wide variation resources [73,74]. In addition to the strong association between ABZ responses

379   and variants in the *ben-1* locus, we identified a novel ABZ-response QTL on chromosome X after

380   accounting for variants in the *ben-1* locus (Figure 5). The identification of this QTL suggests that

381   mechanisms independent of β-tubulin could account for BZ resistance in the *C. elegans* and in parasitic

382   nematode populations.

383

384   The majority (22/25) of *ben-1* alleles that correlate with ABZ resistance are predicted to result in

385   the loss of *ben-1* function (Figure 2). These results suggest that the function of the *ben-1* gene is not

386   essential to the survival of *C. elegans* in natural habitats. This observation is in close agreement with

387   previous work using the laboratory-derived N2 strain, which showed that putative *ben-1* loss-of-function

388   alleles do not confer observable defects in standard laboratory growth conditions [12]. However, our

389   analysis of Ka/Ks between the *C. elegans ben-1* and the distantly related *C. briggsae* and *C. remanei* β-

390   tubulin coding sequences suggest that evolution of this gene is highly constrained. A possible explanation

391   for this discrepancy is that the *ben-1* gene has a highly specialized function in natural settings and that

392   survival in the presence of a strong selective pressure, such as the presence of increased concentrations

393   of environmental BZs, outweighs the necessity of this specialization. A possible source of environmental

394   BZs are microbes that *C. elegans* might contact in their natural habitat. For example, 5,6-

395   dimethylbenzimidazole is a BZ derivative that is produced naturally by prokaryotes [75]. However, whether

396   these BZ derivatives bind to and inhibit β-tubulins remains unclear. Alternatively, natural selection on

397   *C. elegans ben-1* might occur by contamination of nematode niches by synthetic BZs produced by

398   humans. Initially developed as fungicides in the 1960s, several BZs, including thiabendazole, benomyl,

399   and carbendazim are extensively used to treat crops, fruits, and grains [76]. A third source of

400   environmental BZs might be runoff from livestock farms because they are used extensively to prevent

401   parasite infections in these animals [77–80]. The observation that putative *ben-1* loss-of-function alleles

402   are only found at low frequencies in the *C. elegans* population might suggest that individuals with these

403   alleles are eventually removed from the population, though we see no evidence for a decrease in fitness in

404   laboratory competitions experiments, which might not recapitulate natural settings (Figure 4). Taken

405   together, our observations indicate that the independent putative *ben-1* loss-of-function alleles might have

406   arose recently within the *C. elegans* population. Considering these findings in retrospect, it is extremely

407 fortunate that N2 was used for initial studies of BZ sensitivity. Rapid progress on the mechanism of action

408 for BZ compounds from the free-living *C. elegans* model to parasites might have not occurred otherwise,

409 strongly illustrating the pitfalls associated with the study of a single genetic background to elucidate

410 anthelmintic resistance.

411

412 The remaining (3/25) *ben-1* alleles we found to be correlated with ABZ resistance in the *C. elegans*

413 population result in missense variants. These missense variants are S145F, A185P, and M257I and are

414 found in one, two, and one *C. elegans* strain, respectively. Remarkably, substitutions of the amino acid

415 residues A185 (A185S) and M257 (M257L) were previously described to confer BZ resistances in *Tapesia*

416 *yallundae* and *Aspergillus nidulans*, respectively [81,82]. Of these two previously identified residues, M257

417 is postulated to directly interact with BZs, because it is in close three-dimensional proximity to the known

418 BZ interacting F200 residue [60]. The S145F allele is in close proximity to the highly conserved GGGTGS

419 motif of the GTP binding and hydrolysis site [83]. A fourth missense mutation Q131L, which is present in a

420 strain that was not phenotyped in our mapping study because it grows slowly under normal laboratory

421 conditions, is located near the β-tubulin/α-tubulin interaction interface [84,85]. Upon re-phenotyping this

422 isotype with the Q131L variant, we found that it is resistant to ABZ treatment (Supplemental Figure 9). A

423 final missense variant (D404N) present in the *C. elegans* population was not correlated with ABZ

424 resistance (Figure 2). Though theoretical structure-based evidence suggests that most of the missense

425 variants in the *C. elegans* population associated with ABZ resistance cause a non-functional β-tubulin,

426 further experiments are necessary to confirm these results.

427

428 The high prevalence of putative *ben-1* loss-of-function variants in the natural *C. elegans* population

429 stands in stark contrast to the relative paucity of allelic diversity found in orthologous β-tubulin genes in

430 parasitic nematodes. Perhaps parasitic nematode species do not have similarly high levels of functional

431 redundancy of β-tubulins as has been observed in *C. elegans*. The main *C. elegans* β-tubulins are *tbb-1*

432 and *tbb-2*, which are both ubiquitously highly expressed, functionally redundant in laboratory conditions,

433 and contain a tyrosine at position 200 [86–88]. The co-expression of *tbb-1*, *tbb-2*, and *ben-1* in the nervous

434 system [89] suggests that only one β-tubulin with F200 is required to confer sensitivity to BZs. Two other

435   genes that encode for β-tubulins in *C. elegans* are *mec-7* and *tbb-4*. Both genes are only expressed in a

436   subset of neuronal cells involved in chemo- and mechanosensation, and both have a phenylalanine at

437   amino acid position 200 [88,90–92]. *tbb-6* encodes for a highly diverged *C. elegans* β-tubulin

438   (Supplemental Figure 10) that could be a stress-resistant tubulin because it is highly expressed during the

439   unfolded protein response [93]. Despite having phenylalanine at amino acid position 200, these tubulins

440   (TBB-4, MEC-7, and TBB-6) are likely not involved in BZ response because we see no difference in ABZ-

441   resistance between control and ABZ conditions for BZ-resistant strains (Supplemental Data 12). Of the six

442   β-tubulin genes described above, we only observe variation in *ben-1* and *tbb-6*, both of which have low-

443   frequency variants with high predicted functional effects. These observations are also shown by estimates

444   of Tajima's *D* at these loci (Supplemental Figure 11).

445

446   The β-tubulin repertoire among parasitic nematode species is highly diverse likely because of

447   multiple gene duplication events [94]. For example, *H. contortus* has four known β-tubulin genes, of which

448   *Hco-tbb-iso-1* and *Hco-tbb-iso-2* are thought to be the targets of BZ, and *Hco-tbb-iso-3* and *Hco-tbb-iso-4*

449   are closely related in sequence and expression pattern to *C. elegans tbb-4* and *mec-7* [94]. Despite the

450   presence of a phenylalanine at amino acid position 200 in all four of the β-tubulins, the focus of the

451   parasitology community is on missense variants present in *Hco-tbb-iso-1*. However, our data argue that it

452   is important to consider the high level of sequence similarity among the β-tubulin genes. We see that

453   nearby structural variants, like deletions, can alter sequence read alignments and misannotate

454   orthologous β-tubulin gene sequences to artifactually create these variant sites. Careful read alignments

455   and high read depth will be necessary to interpret variation across highly divergent parasites with

456   orthologous β-tubulin genes. Additionally, similar to our observations of *ben-1* in *C. elegans*, deletion

457   alleles of *H. contortus Hco-tbb-iso-2* have been observed in few field studies, but their importance for BZ

458   resistance remains unclear [95,96]. This observation led Kwa and colleagues [95,97] to hypothesize that

459   BZ resistance requires two steps. First, mutation of F200 to tyrosine in one β-tubulin isoform, and second,

460   deletion of the second β-tubulin that contains F200. This hypothesis aligns well with our observation that

461   the single F200Y change in *C. elegans* BEN-1 confers BZ resistance in the presence of TBB-1 (Y200) and

462   TBB-2 (Y200). An additional layer of complexity comes from recent evidence in *Trichuris trichiura* that non-

463 BEN-1 β-tubulins might also play a role in BZ resistance in diverse parasitic species [98]. Our results

464 showing that BZ resistance is a complex trait within the *C. elegans* population (Figure 5) and others'

465 similar observations within parasitic nematode populations necessitates further investigation into the

466 genetic and molecular mechanisms that underlie this trait.

467

468

469

470 ## Material and Methods

471

472 ### *C. elegans* strains

473 Animals were cultured at 20°C on modified nematode growth medium (NGM), containing 1% agar and

474 0.7% agarose [31] and seeded with OP50 *E. coli*. Prior to each assay, strains were passaged for at least

475 four generations without entering starvation or encountering dauer-inducing conditions [31]. For the

476 genome-wide association (GWA) studies, 249 wild isolates from CeNDR (version 20170531) were used as

477 described previously [27,32]. Construction of *ben-1* allele-replacement and *ben-1* deletion strains are

478 described in the corresponding section. All strain information can be found in Supplemental table 1.

479

480 ### High-throughput fitness assay

481 The high-throughput fitness assays (HTA) were performed as described previously [32] with the exception

482 that each strain was assayed in four technical replicates that consisted of four independent bleach

483 synchronization steps across two days with two independent drug and control preparations. In short,

484 strains were propagated for four generations on agar plates, followed by bleach synchronization. The

485 embryos were titered to 96-well microtiter plates at a final concentration of approximately one embryo per

486 microliter of K medium [33] with modified salt concentrations (10.2 mM NaCl, 32 mM KCl, 3 mM $CaCl_2$, 3

487 mM $MgSO_4$). After overnight incubation, hatched L1 larvae were fed with 5 mg/mL HB101 bacterial lysate

488 (Pennsylvania State University Shared Fermentation Facility, State College, PA) and cultured for two days

489 until they reached the L4 larval stage. Using the large particle flow cytometer COPAS BIOSORT (Union

490 Biometrica, Holliston MA), three L4 larvae per well were sorted into new microtiter plates containing

491 modified K medium, 10 mg/mL HB101 lysate, 50 µM kanamycin, and either 12.5 µM albendazole (ABZ)

492 dissolved in 1% DMSO or 1% DMSO alone as a control. During the following four-day incubation, animals

493 were allowed to mature and to produce offspring. Fitness parameters, including traits for animal length and

494 brood size, were measured for each population under drug and control conditions using the COPAS

495 BIOSORT platform after 96 hours. To facilitate body straightening for more accurate length

496 determinations, animals were treated with sodium azide (50 mM) immediately before measurement.

497

498

Processing of fitness traits for genetic mapping

500 Fitness parameters measured by the COPAS BIOSORT platform were processed using the R package

501 *easysorter* [34] with some modifications to incorporate technical replicates into the analysis. The

502 *easysorter* package is specifically developed for this type of data and includes reading and modifying of

503 the raw data, pruning of anomalous data points, and regression of control and experimental phenotypes.

504 Measured parameters include time-of-flight (animal length), extinction (optical density), and total object

505 count (brood size) for each well. In short, the function *read_data* reads in raw phenotype data and runs a

506 support vector machine to identify and eliminate air bubbles, which can be confused with nematodes. In

507 the next step, the function *remove_contamination* removes data obtained from microtiter wells that were

508 manually identified to contain bacterial or fungal contamination. To generate summarized statistics for

509 each well, the function *sum_plate* calculates the 10th, 25th, 50th, 75th, and 90th quantiles for all fitness

510 parameters obtained. In this process, brood size is normalized by the number of animals sorted originally

511 in each well. Next, biologically impossible data points passing certain cut-offs (n > 1000, n < 5, norm.n >

512 350) are eliminated (function *bio_prune*). We removed outlier replicates if they were outside 1.8 times the

513 standard deviation of that strain's median phenotype of a particular trait. To adjust differences among

514 replicates of each assay, a linear model was applied using the formula (phenotype ~ experiment + assay),

515 which replaced the *easysorter* function *regress* (*assay = TRUE*). The experiment component of the linear

516 model corresponds to two independent drug preparations of the same strains and the assay component

517 corresponds to blocks of different strains performed across multiple weeks. In addition, the *bamf_prune*

518 function of the *easysorter* package was skipped, because previous outlier removal based on technical

519 replicates made this step unnecessary. Finally, drug-specific phenotype data is calculated using the

520 *regress* (*assay = FALSE*) function of the *easysorter* package. This function fits a linear model with the

521 formula (phenotype ~ control phenotype) to account for any differences in population parameters present

522 in the control DMSO-only conditions.

523

524

525

Dose-response experiments

Albendazole (Sigma Aldrich Product #A4673) (ABZ) dose-response experiments were performed on a set of four genetically divergent *C. elegans* strains, including the laboratory strain N2 and three different wild isolates (CB4856, JU775, DL238), to determine suitable drug concentrations for subsequent GWA experiments. To this end, strain-specific drug responses were measured in four technical replicates using the HTA as described above at concentrations of 3.125 µM, 6.25 µM, 12.5 µM, and 25 µM ABZ. A suitable drug concentration for subsequent GWA experiments was selected based on the lowest concentration in which we observed a significant difference in ABZ response among strains for brood size and animal length (Supplemental data 1 and 2**)**.

Genome-wide association (GWA) mappings

The GWA mappings were performed on the processed ABZ HTA phenotype data of 209 *C. elegans* wild isolates. In short, the *easysorter* processed phenotype data was analysed using the *cegwas* R package for association mapping [27]. This package uses the EMMA algorithm for performing association mapping and correcting for population structure [35], which is implemented by the *GWAS* function in the rrBLUP package [36]. In detail, the GWAS function was used with the following command: rrBLUP::GWAS (min.MAF = 0.05, P3D = FALSE). The kinship matrix used for association mapping was generated using whole-genome high-quality single-nucleotide variants (SNVs) [37] and the *A.mat* function from the rrBLUP package. All SNVs included in the marker set for GWA mapping had a minimum 5% minor allele frequency in the 240 isotype set [38]. Quantitative trait loci (QTL) were defined by at least one SNV that passed the Bonferroni-corrected threshold and were processed further using fine mapping, as described previously [32]. Computational fine mapping of the genomic regions of interest was performed as described previously [32]. We used the following linear model to correct for the presence of a putative *ben-1* LoF variant: $lm(animal\ length \sim (ben-1\ LoF))$. The list of strains that were considered to have putative LoF variants and the *ben-1*-corrected phenotype data are presented (Supplemental data 16). We did not include CB4856 as a strain with a putative LoF variant because it has a 9 bp in-frame deletion near the end of the gene. We performed single-marker mappings as described above and gene-burden mappings as described below using these regressed phenotypes (Supplemental data 17-21).

554

Burden Testing

556 Burden test analyses were performed using RVtests [39] and the variable-threshold method [40]. We

557 called SNV using bcftools [41] with settings previously described [27,37,39]. We next performed

558 imputation using BEAGLE v4.1 [42] with *window* set to 8000, *overlap* set to 3000, and *ne* set to 17500.

559 Within RVtests, we set the minor allele frequency range from 0.003 to 0.05 for burden testing.

560

Generation of *ben-1* allele replacement and deletion strains

562 All *ben-1* edited strains assayed in this study were generated in the N2 background by CRISPR/Cas9-

563 mediated genome editing, using a co-CRISPR approach [43] and Cas9 ribonucleoprotein (RNP) delivery

564 [44]. For the *ben-1* F200Y allele replacement strains, sgRNAs for *ben-1* and *dpy-10* were ordered from

565 Synthego (Redwood City, CA) and injected at final concentrations of 5 µM and 1 µM, respectively. Single-

566 stranded oligodeoxynucleotides (ssODN) templates for homology-directed repair (HDR) of *ben-1* and *dpy-*

567 *10* (IDT, Skokie IL) were used at final concentrations of 6 µM and 0.5 µM, respectively. Cas9 protein (IDT,

568 Product #1074182) was added to the injection mixture at a concentration of 5 µM and incubated with all

569 other components for ten minutes at room temperature prior to injection. All concentrations used for the

570 sgRNA-mediated allele replacement were adapted from the work of Prior and colleagues [45]. N2 *ben-1*

571 deletion strains were generated using two *ben-1* specific crRNAs synthesized by IDT (Skokie, IL), which

572 targeted exon 2 and exon 4. For the injection mixture, *ben-1* crRNAs were used at final concentration of

573 8.3 µM each, mixed with *dpy-10* crRNA and tracrRNA (IDT, Product #1072532) at final concentrations of

574 1.2 µM and 17.6 µM, respectively, and incubated at 95°C for five minutes. After cooling to room

575 temperature, Cas9 protein was added at a final concentration of 15.25 µM (IDT Product #1074181) and

576 incubated for five minutes before *dpy-10* ssODN was added to a concentration of 5 µM.

577

578 RNP injection mixtures were microinjected into the germline of young adult hermaphrodites (P0)

579 and injected animals were singled to fresh 6 cm NGM plates 18 hours after injection. Two days later, F1

580 progeny were screened, and animals expressing a Rol phenotype were transferred to new plates and

581 allowed to generate progeny (F2). Afterwards, F1 animals were genotyped by PCR. For the *ben-1*(F200Y)

582 allele replacement, ssODN repair templates contained the desired edit and a conservative change of the

583 PAM site to prevent sgRNA:Cas9 cleavage. In addition, the PAM change introduced a new restriction site.

584 PCRs were performed using the primers oECA1297 and oECA1298, and PCR products were incubated

585 with BTsCI restriction enzyme (R0647S, New England Biolabs, Ipswich, MA) to identify successfully edited

586 animals by differential band patterns. For genotyping of *ben-1* deletion strains, the primers oECA1301 and

587 oECA1302 were used and successful deletions were identified by shorter PCR products. Non-Rol progeny

588 (F2) of F1 animals positive for the desired edits were propagated on separate plates to generate

589 homozygous progeny. F2 animals were genotyped afterwards, and PCR products were Sanger

590 sequenced for verification. Generated *ben-1*(F200Y) replacement strains and *ben-1* deletion strains were

591 phenotyped for ABZ response using the HTA as described above. All oligonucleotide sequences are listed

592 in the supplement (Supplemental table 2).

593

594 **Competition assays**

595 Pairwise multi-generation competition assays were performed between a *ben-1* wild-type strain PTM229

596 and the *ben-1* edited strains, containing either the F200Y allele replacement or a *ben-1* deletion. All strains

597 were generated in the N2 background. For the assay, strains were bleach-synchronized and embryos

598 were transferred to 10 cm NGM plates. 48 hours later, seven L4 larvae per strain were transferred to 50

599 fresh 6 cm NGM plates containing either 1.25 µM ABZ for drug selection or DMSO as control. The ABZ

600 concentration was determined by dose-response assays beforehand as the lowest concentration that

601 caused a developmental delay in PTM229 and N2 when added to NGM. For each strain combination, 50

602 plates were grown, representing 10 technical replicates of five independent populations. Plates were

603 grown for one week until starvation. Animals were transferred to fresh plates by cutting out a 0.5 cm³ agar

604 chunk. After every culture transfer, starved animals were washed off the plates with M9, and DNA was

605 collected using the Qiagen DNeasy Kit (Catalog #69506). Allele frequencies of PTM229 compared to wild-

606 type N2 or the *ben-1* edit strains in each replicate populations were measured using Taqman analysis in a

607 Bio-Rad QX200 digital droplet PCR system. Digital PCR was performed following the standard protocol

608 provided by Bio-Rad with the absolute quantification method. To calculate the relative fitness w of the

609 competitive strains, we used linear regression to fit the relative allele frequencies at the first, third, fifth,

610   and seventh weeks into a one-locus genic selection model [46]. All oligonucleotide sequences are listed in

611   the supplement (Supplemental table 2).

612

Computational modelling of *ben-1* variants

614   We obtained the predicted BEN-1 peptide sequence from WormBase [25]. We used the online utility

615   PHYRE2 to generate a homology model of the predicted BEN-1 peptide [47]. We specified the intensive

616   homology search option for running PHYRE2. Visualization of the BEN-1 homology model and highlighting

617   of the variants of interest was performed using PyMol [48].

618

Statistical analyses

620   Phenotype data are shown as Tukey box plots and p-values were used to assess significant differences of

621   strain phenotypes in allele-replacement experiments. All calculations were performed in R using the

622   *TukeyHSD* function on an ANOVA model with the formula (*phenotype ~ strain*). P-values less than 0.05

623   after Bonferroni correction for multiple testing were considered to be significant.

624
625   ### Population Genetics

626   Sliding window analysis of population genetic statistics was performed using the PopGenome package in

627   R [49,50]. All sliding window analyses were performed using the imputed SNV VCF available on the

628   CeNDR website with the most diverged isotype XZ1516 set to the outgroup [27,42,51]. Window size was

629   set to 100 SNVs with a slide distance of one SNV. The Tajima's *D* calculation, using all of the manually

630   curated variants in the *ben-1* locus, was performed using modified code from the developers of the LDhat

631   package [52]. Estimates of Ka/Ks were performed using an online service

632   (http://services.cbu.uib.no/tools/kaks). Multiple sequence alignments were performed using MUSCLE [53]

633   through the online service (https://www.ebi.ac.uk/Tools/msa/muscle/). The β-tubulin neighbor-joining

634   phylogeny construction was performed using Jalview 2 [54]. Linkage disequilibrium (LD) of QTL markers

635   was calculated using the *genetics* [55] package in R. LD calculations are reported as $r = -D / sqrt(p(A) *$

636   $p(a) * p(B) * p(b))$ where $D = p(AB) - p(A) * p(B)$. $r = -D / sqrt(p(A) * p(a) * p(B) * p(b))$

637

638

Acknowledgements

640   We thank Shannon C. Brady, Kyle Siegel, and members of the Andersen lab for editing the manuscript for

641   flow and content. We thank members of the Andersen lab for making reagents used in the experiments

642   presented in the manuscript.

References

648

649 1. Hotez PJ, Alvarado M, Basáñez M-G, Bolliger I, Bourne R, Boussinesq M, et al. The global burden of
650   disease study 2010: interpretation and implications for the neglected tropical diseases. PLoS Negl
651   Trop Dis. 2014;8: e2865.

652 2. Lustigman S, Prichard RK, Gazzinelli A, Grant WN, Boatin BA, McCarthy JS, et al. A research agenda
653   for helminth diseases of humans: the problem of helminthiases. PLoS Negl Trop Dis. 2012;6: e1582.

654 3. Charlier J, van der Voort M, Kenyon F, Skuce P, Vercruysse J. Chasing helminths and their economic
655   impact on farmed ruminants. Trends Parasitol. 2014;30: 361–367.

656 4. Kotze AC, Hunt PW, Skuce P, von Samson-Himmelstjerna G, Martin RJ, Sager H, et al. Recent
657   advances in candidate-gene and whole-genome approaches to the discovery of anthelmintic
658   resistance markers and the description of drug/receptor interactions. Int J Parasitol Drugs Drug
659   Resist. 2014;4: 164–184.

660 5. De Clercq D, Sacko M, Behnke J, Gilbert F, Dorny P, Vercruysse J. Failure of mebendazole in
661   treatment of human hookworm infections in the southern region of Mali. Am J Trop Med Hyg.
662   1997;57: 25–30.

663 6. Albonico M, Bickle Q, Ramsan M, Montresor A, Savioli L, Taylor M. Efficacy of mebendazole and
664   levamisole alone or in combination against intestinal nematode infections after repeated targeted
665   mebendazole treatment in Zanzibar. Bull World Health Organ. 2003;81: 343–352.

666 7. Humphries D, Simms BT, Davey D, Otchere J, Quagraine J, Terryah S, et al. Hookworm infection
667   among school age children in Kintampo north municipality, Ghana: nutritional risk factors and
668   response to albendazole treatment. Am J Trop Med Hyg. 2013;89: 540–548.

669 8. Neff NF, Thomas JH, Grisafi P, Botstein D. Isolation of the beta-tubulin gene from yeast and
670   demonstration of its essential function in vivo. Cell. 1983;33: 211–219.

671 9. Laclette JP, Guerra G, Zetina C. Inhibition of tubulin polymerization by mebendazole. Biochem
672   Biophys Res Commun. 1980;92: 417–423.

673 10. Ireland CM, Gull K, Gutteridge WE, Pogson CI. The interaction of benzimidazole carbamates with
674   mammalian microtobule protein. Biochem Pharmacol. 1979;28: 2680–2682.

675 11. Thomas JH, Neff NF, Botstein D. Isolation and characterization of mutations in the beta-tubulin gene
676   of Saccharomyces cerevisiae. Genetics. 1985;111: 715–734.

677 12. Driscoll M, Dean E, Reilly E, Bergholz E, Chalfie M. Genetic and molecular analysis of a
678   Caenorhabditis elegans beta-tubulin that conveys benzimidazole sensitivity. J Cell Biol. 1989;109:
679   2993–3003.

680 13. Kwa MS, Veenstra JG, Roos MH. Benzimidazole resistance in Haemonchus contortus is correlated
681   with a conserved mutation at amino acid 200 in beta-tubulin isotype 1. Mol Biochem Parasitol.
682   1994;63: 299–303.

683 14. Silvestre A, Cabaret J. Mutation in position 167 of isotype 1 beta-tubulin gene of Trichostrongylid
684   nematodes: role in benzimidazole resistance? Mol Biochem Parasitol. 2002;120: 297–300.

685 15. Ghisi M, Kaminsky R, Mäser P. Phenotyping and genotyping of Haemonchus contortus isolates
686   reveals a new putative candidate mutation for benzimidazole resistance in nematodes. Vet Parasitol.
687   2007;144: 313–320.

688 16. Lacey E, Gill JH. Biochemistry of benzimidazole resistance. Acta Trop. 1994;56: 245–262.

689   17.   Lubega GW, Prichard RK. Specific interaction of benzimidazole anthelmintics with tubulin: high-affinity
690   binding and benzimidazole resistance in Haemonchus contortus. Mol Biochem Parasitol. 1990;38:
691   221–232.

692   18.   Krücken J, Fraundorfer K, Mugisha JC, Ramünke S, Sifft KC, Geus D, et al. Reduced efficacy of
693   albendazole against Ascaris lumbricoides in Rwandan schoolchildren. Int J Parasitol Drugs Drug
694   Resist. 2017;7: 262–271.

695   19.   Yilmaz E, Ramünke S, Demeler J, Krücken J. Comparison of constitutive and thiabendazole-induced
696   expression of five cytochrome P450 genes in fourth-stage larvae of Haemonchus contortus isolates
697   with different drug susceptibility identifies one gene with high constitutive expression in a multi-
698   resistant isolate. Int J Parasitol Drugs Drug Resist. 2017;7: 362–369.

699   20.   Jones LM, Flemming AJ, Urwin PE. NHR-176 regulates cyp-35d1 to control hydroxylation-dependent
700   metabolism of thiabendazole in Caenorhabditis elegans. Biochem J. 2015;466: 37–44.

701   21.   Doyle SR, Bourguinat C, Nana-Djeunga HC, Kengne-Ouafo JA, Pion SDS, Bopda J, et al. Genome-
702   wide analysis of ivermectin response by Onchocerca volvulus reveals that genetic drift and soft
703   selective sweeps contribute to loss of drug sensitivity. PLoS Negl Trop Dis. 2017;11: e0005816.

704   22.   Holden-Dye L, Walker RJ. Anthelmintic drugs and nematicides: studies in Caenorhabditis elegans.
705   WormBook. 2014; 1–29.

706   23.   Hunt PW, Kotze AC, Knox MR, Anderson LJ, McNally J, LE Jambre LF. The use of DNA markers to
707   map anthelmintic resistance loci in an intraspecific cross of Haemonchus contortus. Parasitology.
708   2010;137: 705–717.

709   24.   Stear MJ, Boag B, Cattadori I, Murphy L. Genetic variation in resistance to mixed, predominantly
710   Teladorsagia circumcincta nematode infections of sheep: from heritabilities to gene identification.
711   Parasite Immunol. 2009;31: 274–282.

712   25.   Stein L, Sternberg P, Durbin R, Thierry-Mieg J, Spieth J. WormBase: network access to the genome
713   and biology of Caenorhabditis elegans. Nucleic Acids Res. 2001;29: 82–86.

714   26.   C. elegans Sequencing Consortium. Genome sequence of the nematode C. elegans: a platform for
715   investigating biology. Science. 1998;282: 2012–2018.

716   27.   Cook DE, Zdraljevic S, Roberts JP, Andersen EC. CeNDR, the Caenorhabditis elegans natural
717   diversity resource. Nucleic Acids Res. 2016;45: D650–D657.

718   28.   Strange K. Drug Discovery in Fish, Flies, and Worms. ILAR J. 2016;57: 133–143.

719   29.   Ghosh R, Andersen EC, Shapiro JA, Gerke JP, Kruglyak L. Natural variation in a chloride channel
720   subunit confers avermectin resistance in C. elegans. Science. 2012;335: 574–578.

721   30.   Zamanian M, Cook DE, Zdraljevic S, Brady SC, Lee D, Lee J, et al. Discovery of unique loci that
722   underlie nematode responses to benzimidazoles [Internet]. 2017. doi:10.1101/116970

723   31.   Andersen EC, Bloom JS, Gerke JP, Kruglyak L. A variant in the neuropeptide receptor npr-1 is a
724   major determinant of Caenorhabditis elegans growth and physiology. PLoS Genet. 2014;10:
725   e1004156.

726   32.   Zdraljevic S, Strand C, Seidel HS, Cook DE, Doench JG, Andersen EC. Natural variation in a single
727   amino acid substitution underlies physiological responses to topoisomerase II poisons. PLoS Genet.
728   2017;13: e1006891.

729   33.   Boyd WA, Smith MV, Freedman JH. Caenorhabditis elegans as a model in developmental toxicology.

Methods Mol Biol. 2012;889: 15–24.

34. Shimko TC, Andersen EC. COPASutils: An R Package for Reading, Processing, and Visualizing Data from COPAS Large-Particle Flow Cytometers. PLoS One. 2014;9: e111090.

35. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. Genetics. 2008;178: 1709–1723.

36. Endelman JB. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. The Plant Genome Journal. 2011;4: 250.

37. Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, Noble LM, et al. The Genetic Basis of Natural Variation in Caenorhabditis elegans Telomere Length. Genetics. 2016;204: 371–383.

38. Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, et al. Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nat Genet. 2012;44: 285–290.

39. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics. 2016;32: 1423–1426.

40. Price AL, Kryukov GV, de Bakker PIW, Purcell SM, Staples J, Wei L-J, et al. Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet. 2010;86: 832–838.

41. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011;27: 2987–2993.

42. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. Am J Hum Genet. 2016;98: 116–126.

43. Kim H, Ishidate T, Ghanta KS, Seth M, Conte D Jr, Shirayama M, et al. A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. Genetics. 2014;197: 1069–1080.

44. Paix A, Folkmann A, Rasoloson D, Seydoux G. High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. Genetics. 2015;201: 47–54.

45. Prior H, Jawad AK, MacConnachie L, Beg AA. Highly Efficient, Rapid and Co-CRISPR Independent Genome Editing in Caenorhabditis elegans. G3: Genes|Genomes|Genetics. 2017; g3.300216.2017.

46. Zhao Y, Long L, Xu W, Campbell RF, Large EL, Greene JS, et al. Laboratory evolution from social to solitary behavior in the N2 reference strain is unnecessary for its fitness advantages [Internet]. bioRxiv. 2018. p. 309997. doi:10.1101/309997

47. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10: 845–858.

48. De Lano WL, California U. PyMOL: An Open-Source Molecular Graphics Tool. Available: http://www.ccp4.ac.uk/newsletters/newsletter40/11_pymol.pdf

49. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. Mol Biol Evol. Oxford University Press; 2014;31: 1929–1936.

50. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014. 2016.

51. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27: 2156–2158.

770   52. McVean G, Awadalla P, Fearnhead P. A Coalescent-Based Method for Detecting and Estimating
771   Recombination From Gene Sequences. Genetics. Genetics; 2002;160: 1231–1241.

772   53. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic
773   Acids Res. 2004;32: 1792–1797.

774   54. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple
775   sequence alignment editor and analysis workbench. Bioinformatics. 2009;25: 1189–1191.

776   55. Warnes G, Gorjanc G, Leisch F, Man M. Genetics: Population genetics. R package version 1.3. 6.
777   2012.

778   56. Kishore SP, Blank E, Heller DJ, Patel A, Peters A, Price M, et al. Modernizing the World Health
779   Organization List of Essential Medicines for Preventing and Controlling Cardiovascular Diseases. J
780   Am Coll Cardiol. 2018;71: 564–574.

781   57. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and
782   predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila
783   melanogaster strain w1118; iso-2; iso-3. Fly . 2012;6: 80–92.

784   58. Aguayo-Ortiz R, Méndez-Lucio O, Medina-Franco JL, Castillo R, Yépez-Mulia L, Hernández-Luis F, et
785   al. Towards the identification of the binding site of benzimidazoles to β-tubulin of Trichinella spiralis:
786   insights from computational and experimental data. J Mol Graph Model. 2013;41: 12–19.

787   59. Aguayo-Ortiz R, Méndez-Lucio O, Romo-Mancillas A, Castillo R, Yépez-Mulia L, Medina-Franco JL,
788   et al. Molecular basis for benzimidazole resistance from a novel β-tubulin binding site model. J Mol
789   Graph Model. 2013;45: 26–37.

790   60. Robinson MW, McFerran N, Trudgett A, Hoey L, Fairweather I. A possible model of benzimidazole
791   binding to β-tubulin disclosed by invoking an inter-domain movement. J Mol Graph Model. 2004;23:
792   275–284.

793   61. Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, et al. The Dfam database of repetitive
794   DNA families. Nucleic Acids Res. 2016;44: D81–9.

795   62. Oosumi T, Garlick B, Belknap WR. Identification of putative nonautonomous transposable elements
796   associated with several transposon families in Caenorhabditis elegans. J Mol Evol. 1996;43: 11–18.

797   63. Li WH, Wu CI, Luo CC. A new method for estimating synonymous and nonsynonymous rates of
798   nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol
799   Evol. 1985;2: 150–174.

800   64. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.
801   Genetics. 1989;123: 585–595.

802   65. Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. Genetics. 2000;155: 1405–1413.

803   66. Zeng K, Fu Y-X, Shi S, Wu C-I. Statistical tests for detecting positive selection by utilizing high-
804   frequency variants. Genetics. 2006;174: 1431–1439.

805   67. Furtado LFV, de Paiva Bello ACP, Rabelo ÉML. Benzimidazole resistance in helminths: From
806   problem to diagnosis. Acta Trop. 2016;162: 95–102.

807   68. Clarke NE, Doi SAR, Wangdi K, Chen Y, Clements ACA, Nery SV. Efficacy of anthelminthic drugs
808   and drug combinations against soil-transmitted helminths: a systematic review and network meta-
809   analysis. Clin Infect Dis. 2018; doi:10.1093/cid/ciy423

810 69. Gilleard JS. Understanding anthelmintic resistance: the need for genomics and genetics. Int J
811   Parasitol. 2006;36: 1227–1239.

812 70. Large EE, Xu W, Zhao Y, Brady SC, Long L, Butcher RA, et al. Selection on a Subunit of the NURF
813   Chromatin Remodeler Modifies Life History Traits in a Domesticated Strain of Caenorhabditis
814   elegans. PLoS Genet. Public Library of Science; 2016;12: e1006219.

815 71. Redman E, Sargison N, Whitelaw F, Jackson F, Morrison A, Bartley DJ, et al. Introgression of
816   ivermectin resistance genes into a susceptible Haemonchus contortus strain by multiple
817   backcrossing. PLoS Pathog. 2012;8: e1002534.

818 72. Valentim CLL, Cioli D, Chevalier FD, Cao X, Taylor AB, Holloway SP, et al. Genetic and Molecular
819   Basis of Drug Resistance and Species-Specific Drug Action in Schistosome Parasites. Science.
820   American Association for the Advancement of Science; 2013;342: 1385–1389.

821 73. Howe KL, Bolt BJ, Cain S, Chan J, Chen WJ, Davis P, et al. WormBase 2016: expanding to enable
822   helminth genomic research. Nucleic Acids Res. 2016;44: D774–80.

823 74. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite - a comprehensive resource
824   for helminth genomics. Mol Biochem Parasitol. 2017;215: 2–10.

825 75. Crofts TS, Men Y, Alvarez-Cohen L, Taga ME. A bioassay for the detection of benzimidazoles reveals
826   their presence in a range of environmental samples. Front Microbiol. 2014;5.
827   doi:10.3389/fmicb.2014.00592

828 76. Oliver RP, Geoff Hewitt H. Fungicides in Crop Protection, 2nd Edition. CABI; 2014.

829 77. Prchal L, Podlipná R, Lamka J, Dědková T, Skálová L, Vokřál I, et al. Albendazole in environment:
830   faecal concentrations in lambs and impact on lower development stages of helminths and seed
831   germination. Environ Sci Pollut Res Int. 2016;23: 13015–13022.

832 78. Perruchon C, Pantoleon A, Veroutis D, Gallego-Blanco S, Martin-Laurent F, Liadaki K, et al.
833   Characterization of the biodegradation, bioremediation and detoxification capacity of a bacterial
834   consortium able to degrade the fungicide thiabendazole. Biodegradation. 2017;28: 383–394.

835 79. Cycoń M, Mrozik A, Piotrowska-Seget Z. Bioaugmentation as a strategy for the remediation of
836   pesticide-polluted soil: A review. Chemosphere. 2017;172: 52–71.

837 80. de Oliveira Neto OF, Arenas AY, Fostier AH. Sorption of thiabendazole in sub-tropical Brazilian soils.
838   Environ Sci Pollut Res Int. 2017;24: 16503–16512.

839 81. Hawkins NJ, Fraaije BA. Predicting Resistance by Mutagenesis: Lessons from 45 Years of MBC
840   Resistance. Front Microbiol. 2016;7: 1814.

841 82. Oakley BR. Tubulins in Aspergillus nidulans. Fungal Genet Biol. 2004;41: 420–427.

842 83. Hesse J, Thierauf M, Ponstingl H. Tubulin sequence region beta 155-174 is involved in binding
843   exchangeable guanosine triphosphate. J Biol Chem. 1987;262: 15472–15475.

844 84. McKean PG, Vaughan S, Gull K. The extended tubulin superfamily. J Cell Sci. 2001;114: 2723–2733.

845 85. Löwe J, Li H, Downing KH, Nogales E. Refined structure of alpha beta-tubulin at 3.5 A resolution. J
846   Mol Biol. 2001;313: 1045–1057.

847 86. Lu C, Srayko M, Mains PE. The Caenorhabditis elegans microtubule-severing complex MEI-1/MEI-2
848   katanin interacts differently with two superficially redundant beta-tubulin isotypes. Mol Biol Cell.
849   2004;15: 142–150.

850   87. Ellis GC, Phillips JB, O'Rourke S, Lyczak R, Bowerman B. Maternally expressed and partially
851   redundant beta-tubulins in Caenorhabditis elegans are autoregulated. J Cell Sci. 2004;117: 457–464.

852   88. Hutter H, Suh J. GExplore 1.4: An expanded web interface for queries on Caenorhabditis elegans
853   protein and gene function. Worm. 2016;5: e1234659.

854   89. Hurd DD. Tubulins in C. elegans. WormBook. 2018; 1–34.

855   90. Hao L, Thein M, Brust-Mascher I, Civelekoglu-Scholey G, Lu Y, Acar S, et al. Intraflagellar transport
856   delivers tubulin isotypes to sensory cilium middle and distal segments. Nat Cell Biol. 2011;13: 790–
857   798.

858   91. Bounoutas A, O'Hagan R, Chalfie M. The multipurpose 15-protofilament microtubules in C. elegans
859   have specific roles in mechanosensation. Curr Biol. 2009;19: 1362–1367.

860   92. Savage C, Hamelin M, Culotti JG, Coulson A, Albertson DG, Chalfie M. mec-7 is a beta-tubulin gene
861   required for the production of 15-protofilament microtubules in Caenorhabditis elegans. Genes Dev.
862   1989;3: 870–881.

863   93. Munkácsy E, Khan MH, Lane RK, Borror MB, Park JH, Bokov AF, et al. DLK-1, SEK-3 and PMK-3
864   Are Required for the Life Extension Induced by Mitochondrial Bioenergetic Disruption in C. elegans.
865   PLoS Genet. 2016;12: e1006133.

866   94. Saunders GI, Wasmuth JD, Beech R, Laing R, Hunt M, Naghra H, et al. Characterization and
867   comparative analysis of the complete Haemonchus contortus β-tubulin gene family and implications
868   for benzimidazole resistance in strongylid nematodes. Int J Parasitol. 2013;43: 465–475.

869   95. Kwa MS, Kooyman FN, Boersema JH, Roos MH. Effect of selection for benzimidazole resistance in
870   Haemonchus contortus on beta-tubulin isotype 1 and isotype 2 genes. Biochem Biophys Res
871   Commun. 1993;191: 413–419.

872   96. Beech RN, Prichard RK, Scott ME. Genetic variability of the beta-tubulin genes in benzimidazole-
873   susceptible and -resistant strains of Haemonchus contortus. Genetics. 1994;138: 103–110.

874   97. Prichard R. Genetic variability following selection of Haemonchus contortus with anthelmintics. Trends
875   Parasitol. 2001;17: 445–453.

876   98. Demeler J, Krüger N, Krücken J, von der Heyden VC, Ramünke S, Küttler U, et al. Phylogenetic
877   characterization of β-tubulins and development of pyrosequencing assays for benzimidazole
878   resistance in cattle nematodes. PLoS One. 2013;8: e70212.

Figure list:

Supplemental Figure 1: Dose response of four genetically divergent *C. elegans* strains to ABZ exposure

Figure 1: GWA mappings of 209 *C elegans* wild isolates in response to ABZ

Supplemental Figure 2: A) Wild isolate phenotypes split by the the presence of the ALT or REF allele at the QTL identified using the single-marker mapping approach. B) LD of peak markers identified with the single-marker mapping approach.

Figure 2: A) ABZ resistance phenotypes colored by the presence of variation in the *ben-1* locus. B) Detailed figure showing natural variation found in *ben-1* among *C. elegans* wild isolates.

Supplemental Figure 3: 3D structure of *C. elegans ben-1* homology model, highlighting the observed missense variants present in the *C. elegans* population.

Figure 3: A) Neutrality statistics of genomic regions surrounding the *ben-1* locus. B) Global sampling locations of *C. elegans* wild isolates. C) *C. elegans* phylogeny.

Supplemental Figure 4: Distribution of sampling A) environment and B) substrate where *C. elegans* isolates with putative LoF (ALT) or REF alleles at the *ben-1* locus.

Figure 4: Fitness of F200Y and deletion alleles of the *ben-1* gene using the A) high-throughput and B) competitive growth assays.

Supplemental Figure 5: HTA results for all *ben-1* edit strains in the presence of ABZ

Figure 5: A) Genome-wide single-marker mapping of ABZ length phenotype after correcting isotype phenotypes based on the presence of putative *ben-1* LoF variation. B) Fine-mapping of chromosome X QTL.

Supplemental Figure 6: Genome-wide gene-burden mapping of ABZ length phenotype after correcting isotype phenotypes based on the presence of putative *ben-1* LoF variation.

Supplemental Figure 7: Wild isolate phenotypes split by the presence of the ALT or REF allele at the QTL identified using the single-marker mapping approach. We highlight the presence of putative *ben-1* LoF variants.

Supplemental Figure 8: Wild isolate phenotypes after *ben-1* LoF regression split by the isotype genotype at the chromosome X QTL peak marker.

906 **Supplemental Figure 9:** HTA re-phenotyping results for selected *C. elegans* wild strains in the presence

907 of ABZ

908 **Supplemental Figure 10**: Neighbor joining phylogenetic tree of the six *C. elegans* β-tubulins, the *C.*

909 *remanei* and *C. briggsae ben-1* orthologs, the *Haemonchus contortus ben-1* ortholog, and the *Necator*

910 *americanus ben-1* ortholog.

911 **Supplemental Figure 11**: Tajima's *D* for the genomic regions surrounding *ben-1* orthologs in the

912 *C. elegans* genome.

Supplemental Data

913

914

915 **TS1**: HTA ABZ dose response data_raw(.Rda)

916 **TS2**: HTA ABZ dose response data_processed(.tsv)

917 **TS3**: HTA phenotype data of *C. elegans* wild strains_raw(.Rda)

918 **TS4**: HTA phenotype data of *C. elegans* wild strains_processed(.tsv)

919 **TS5**: GWA single marker mapping data_processed(.tsv)

920 **TS6**: GWA single marker fine mapping data(.tsv)

921 **TS7:** GWA gene-burden mapping phenotype data(.ped)

922 **TS8**: GWA gene-burden mapping data(.tsv)

923 **TS9**: Positions and effects for all *ben-1* variants(.tsv)

924 **TS10**: Summary of population genetics data(.Rda)

925 **TS11**: Tajima's D, Fay's Wu and Zeng's E for *ben-1* (.tsv)

926 **TS12:** HTA phenotype data of *C. elegans ben-1* mutants_summarized(.tsv)

927 **TS 13**: HTA phenotype data of *C. elegans ben-1* mutants_processed(.tsv)

928 **TS 14**: competition assay data_raw(.tsv)

929 **TS 15**: relative fitness calculations(.xml)

930 **TS 16**: *ben-1* regressed with *ben-1* covariate(.tsv)

931 **TS 17**: GWA single-marker mapping data after *ben-1* regression(.tsv)

932 **TS 18**: GWA single-marker fine mapping data after *ben-1* regression(.tsv)

933 **TS 19:** GWA gene-burden mapping phenotype data after ben-1 regression(.ped)

934 **TS 20:** GWA gene-burden mapping data after ben-1 regression(.assoc)

935 **TS 21:** GWA gene-burden mapping data after ben-1 regression(.tsv)

936 **TS 22**: Tajima's D for other *C. elegans* beta-tubulins (.tsv)

937 **TS 23**: HTA re-phenotyping of *C. elegans* wild strains(.Rda)

938 **Supplemental Table 1**: strains(.csv)

939 **Supplemental Table 2**: Oligonucleotide sequences(.csv)

940

941

942
943
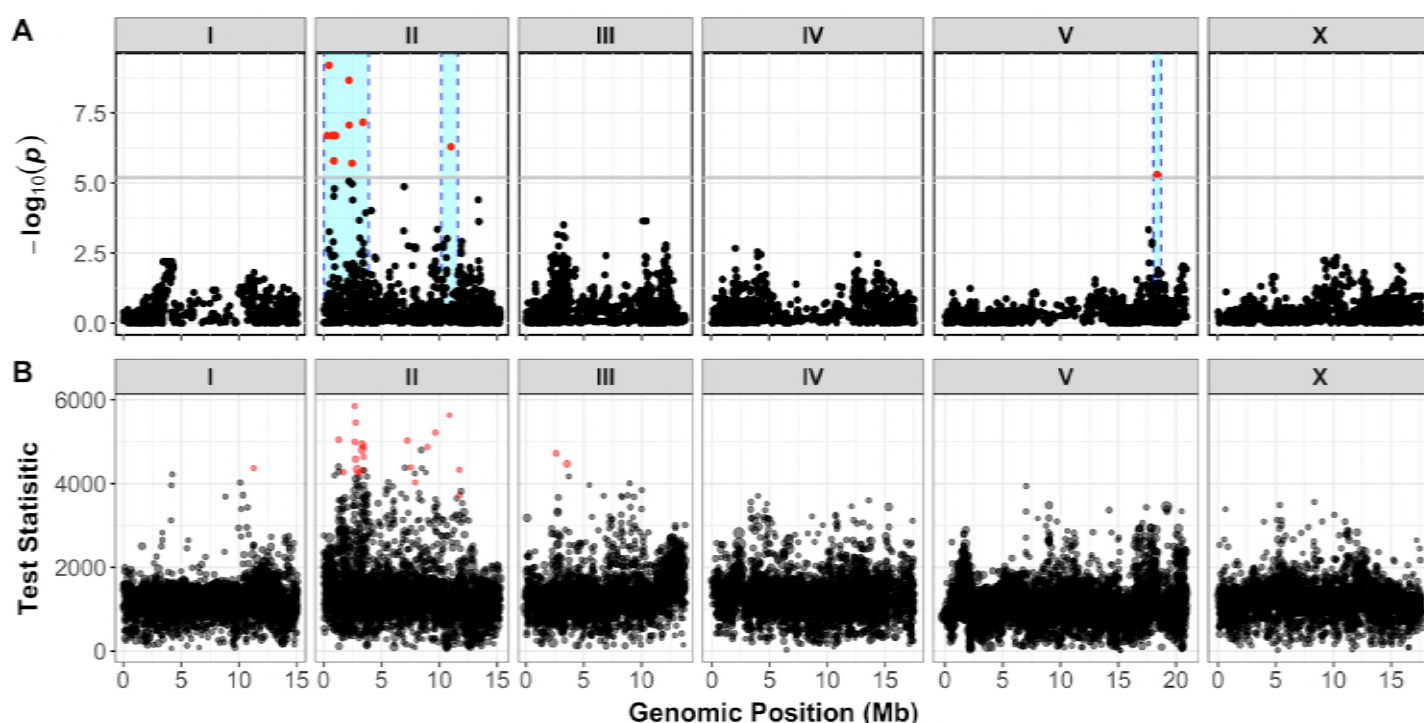
Figure 1: GWA mappings of 209 C elegans wild isolates in response to ABZ

A) The single-marker based Manhattan plot for animal length (q90.TOF) in the presence of ABZ is shown. Each dot represents an SNV that is present in at least 5% of the assayed population. The genomic location of each SNV is plotted on the x-axis, and the statistical significance of the correlation between genotype and phenotype is plotted on the y-axis. SNVs are colored red if they pass the genome-wide Bonferroni-corrected significance threshold, which is shown by a gray horizontal line. Genomic regions of interest are represented by cyan rectangles surrounding each QTL. In total, three QTL were identified, located on chr II and chr V. B) A genome-wide manhattan plot of animal length (q90.TOF) after ABZ exposure based on the gene-burden mapping approach is shown. Each dot represents a single *C. elegans* gene with its genomic location plotted on the x-axis, and the statistical significance of the correlation between genotype and phenotype plotted on the y-axis. Genes passing the burden test statistical significance threshold are colored in red. The burden test analysis identified genes significantly correlated with ABZ resistance on chr I, chr II, and chr III.

957
958
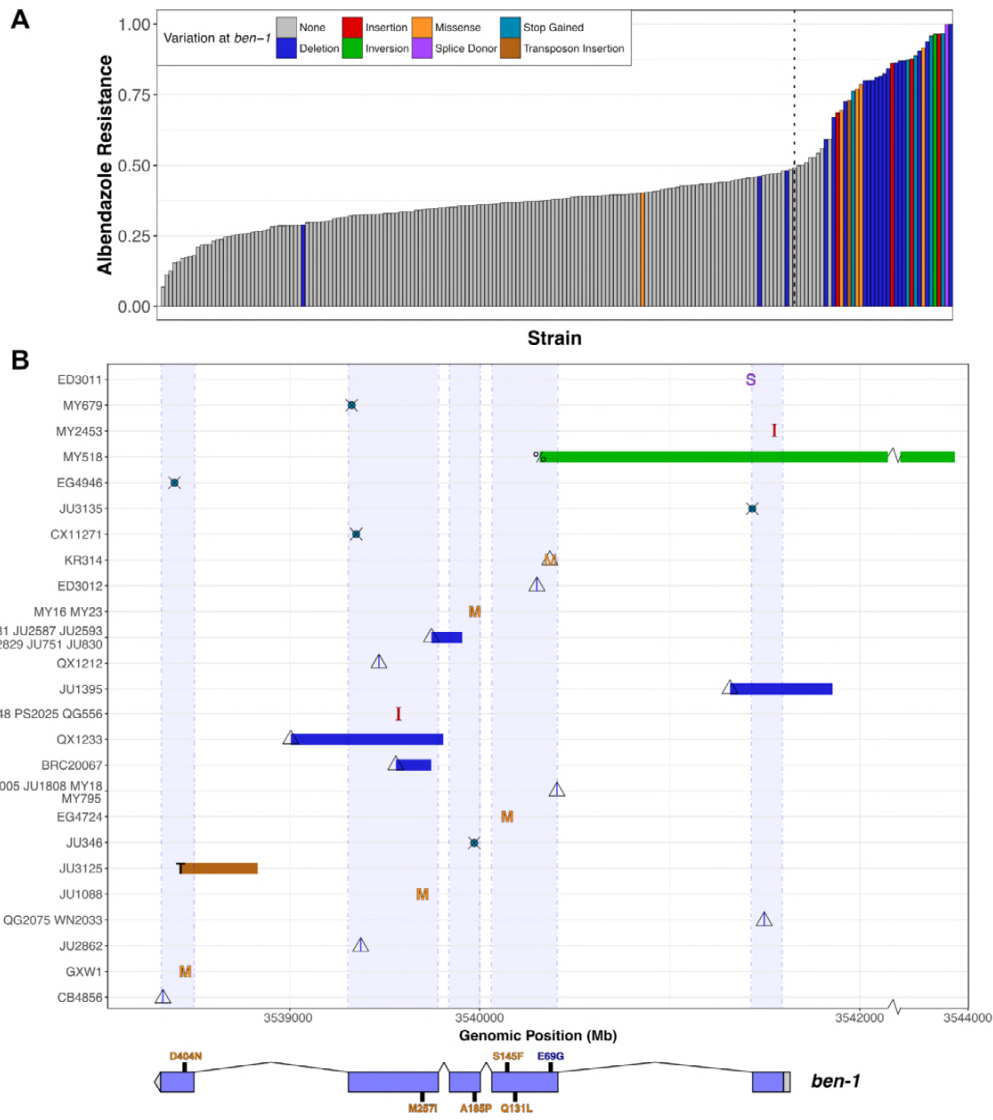
Figure 2: Variants in *ben-1* are highly correlated with ABZ resistance

A) A bar plot for the phenotypic response of *C. elegans* strains under ABZ exposure is shown. Each bar

represents a single wild strain included in the HTA. Strains are sorted by their relative resistance to ABZ

based on the mean animal length (q90.TOF) from four replicate measures. All strains that were found to

have variants with predicted moderate-to-high impact and/or structural variation are colored by their

specific type of variant. Strains similar to the N2 reference genome with respect to the *ben-1* locus are

shown in grey. The black dotted line marks the point halfway point between the most and least resistant

ABZ strains. B) An overview of variants found in the *ben-1* locus among *C. elegans* wild strains is shown.

The genomic position of each variant is shown on the x-axis (S: splice donor variant; M: missense variant

leading to amino acid substitution; crossed circle (⊗): alternative stop codon; triangle: deletion; I: insertion;

969   percent sign: inversion; T: transposon insertion. Colors correspond to colors in (A). All strains with *ben-1*

970   variants are listed on the y-axis in increasing order by their length in ABZ. For variants shared by two or

971   more strains, the mean phenotype value of all corresponding strains determines their placement on the y-

972   axis. A representation of the *ben-1* gene model is shown below the variant summary panel, including 5'

973   and 3' UTRs (grey rectangle and bar), five exons (blue bars), and four introns (thin lines). Additionally,

974   missense variants leading to amino acid substitutions are highlighted with their exact position and the

975   corresponding amino acid exchange in the single letter code. Missense variants are colored orange if the

976   corresponding strains do not contain any other *ben-1* variation. Missense variants that are always

977   associated with deletion variants are colored blue.

978

979
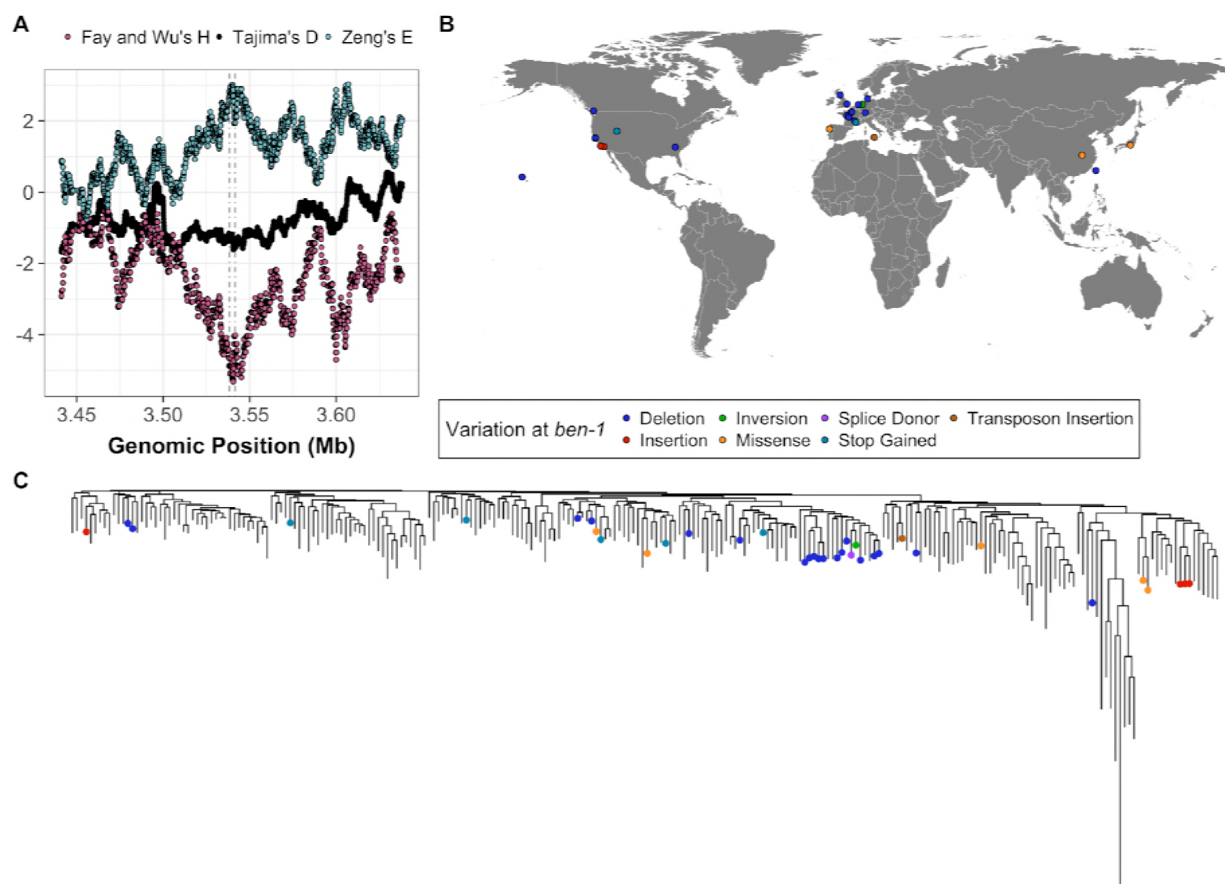
980
981 **Figure 3: Population-level summary of ben-1 variants**

982 A) Fay and Wu's *H* (pink), Tajima's *D* (black), and Zeng's *E* (blue) for the genomic region surrounding *ben-*

983 *1*. Genomic position is shown on the x-axis, and the value for each displayed statistic is shown on the y-

984 axis. The neutrality statistics were calculated using a sliding window approach (100 SNVs window size

985 and 1 SNV slide distance) using only SNV data. B) The global distribution of strains that contain moderate-

986 to-high predicted variation in *ben-1*. Each dot corresponds to the sampling location of an individual strain

987 and is colored by the type of variant discovered in the *ben-1* locus. C) The genome-wide phylogeny of 249

988 *C. elegans* strains showing that variation in the *ben-1* locus occurred independently multiple times during

989 the evolutionary history of the species. The dots on individual branch nodes correspond to strains with

990 variation in *ben-1* and have the same color code as in panel B.
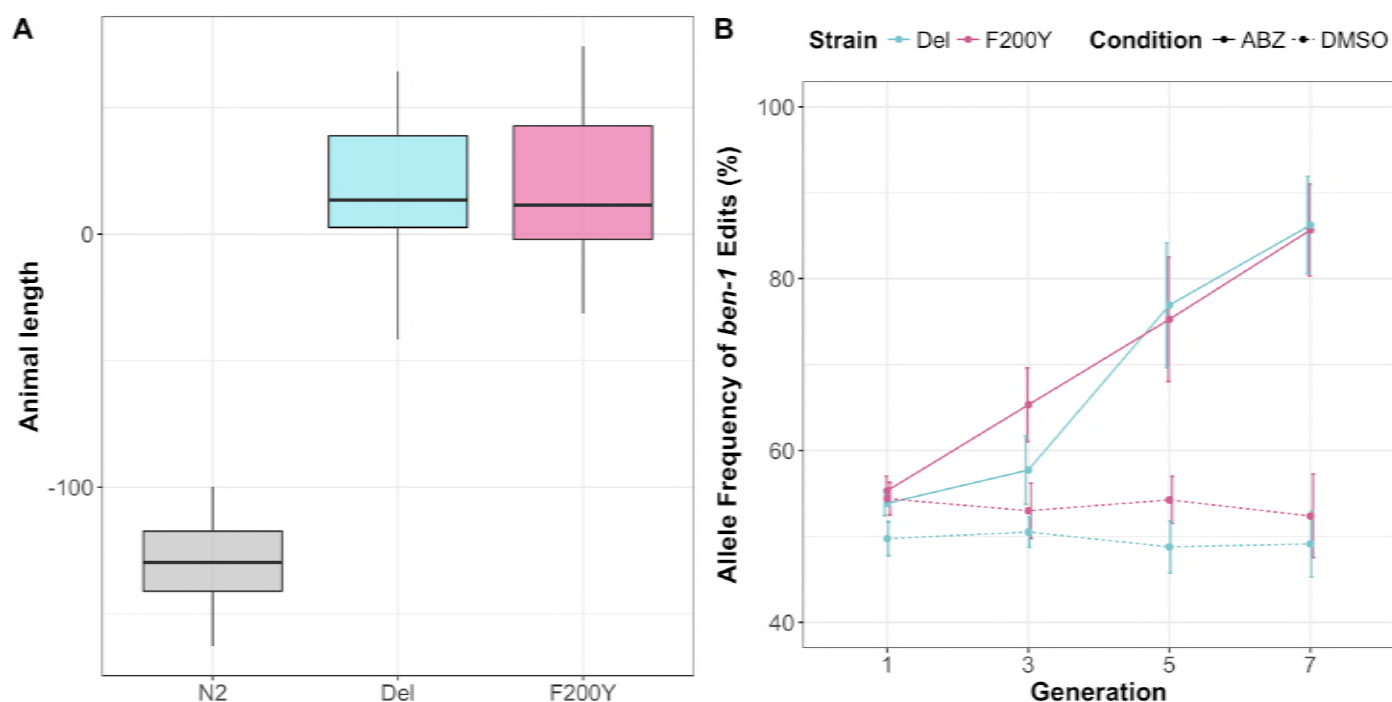
Figure 4: F200Y allele replacement and *ben-1* deletion confer resistance to the ABZ sensitive N2 strain

A) Tukey box plots of the animal length phenotypes for at least 100 replicates of the generated *ben-1* allele-replacement strains after ABZ treatment is shown. The y-axis represents the animal length phenotype after correcting for growth in DMSO conditions. Both *ben-1* allele strains are significantly more resistant to ABZ treatment than the N2 strain (*p*-value < 1E-10, TukeyHSD). **B)** The results from a multi-generation competition experiment between a barcoded N2 strain and the *ben-1* allele-replacement strains are shown. The generation number is shown on the x-axis, and the allele frequency of the *ben-1* allele-replacement strains is shown on the y-axis.

1003

1004
1005
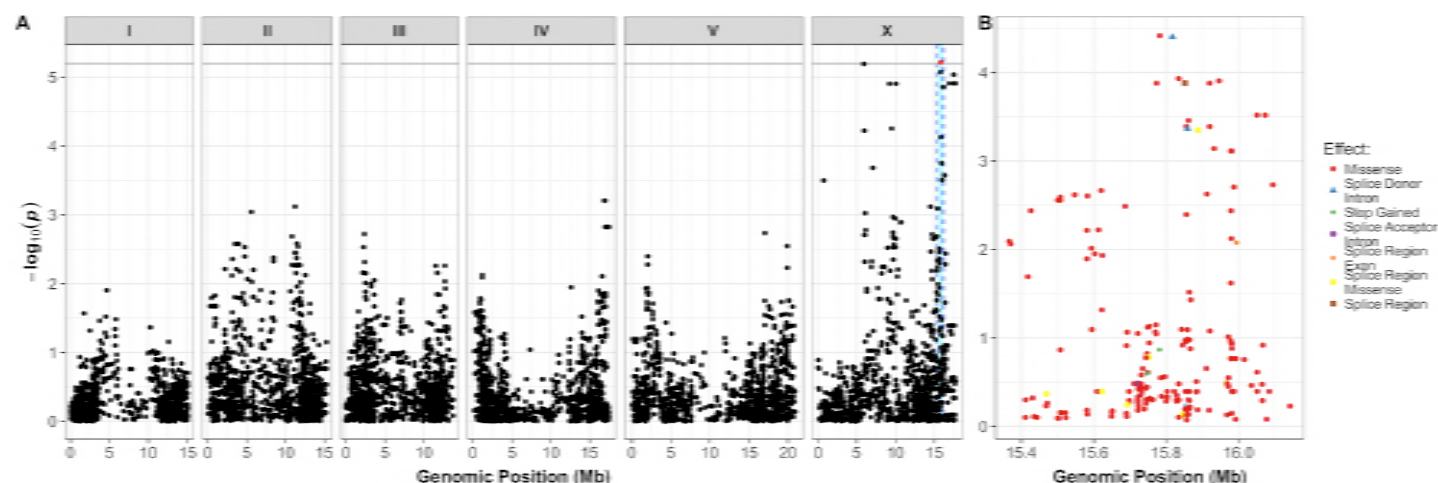
1006 **Figure 5: Regression of the putative *ben-1* LoF variants identifies a novel QTL on chromosome X**

1007 A) Manhattan plot of the single-marker based GWA mapping showing that a nominally significant QTL on

1008 chromosome X is associated with residual phenotypic variation in response to ABZ treatment. Each dot

1009 represents an SNV that is present in at least 5% of the assayed population. The genomic location of each

1010 SNV is plotted on the x-axis, and the statistical significance is plotted on the y-axis. SNVs are colored red

1011 if they pass the Bonferroni-corrected significance threshold, which is shown by a gray horizontal line. The

1012 genomic region of interest is represented by cyan rectangle surrounding the chromosome X QTL. B) Fine-

1013 mapping of the chromosome X QTL showing all variants with moderate-to-high predicted effects within the

1014 chromosomes X QTL region of interest. The genomic position in Mb is shown on the x-axis, and the

1015 $-log10(p)$ value, which represents the strength of the association between residual ABZ resistance and

1016 the plotted variant, is shown on the y-axis. Dots are colored by the type of variant.

1017
1018
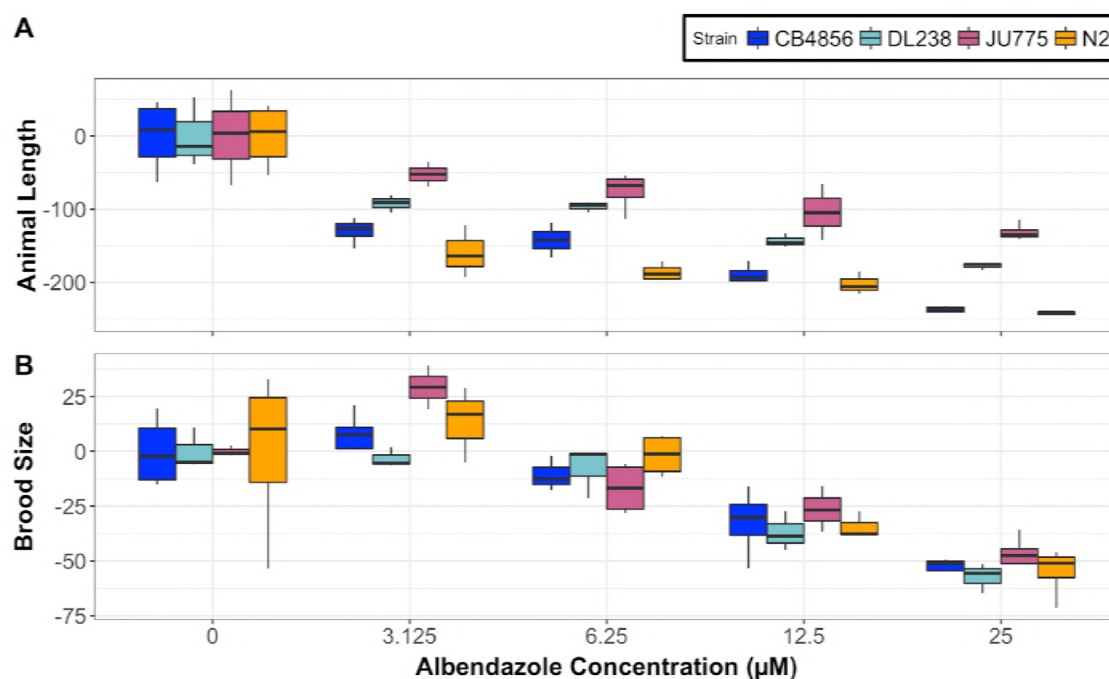
Supplemental Figure 1:

**Dose response of four genetically divergent *C. elegans* strains after ABZ exposure**

The box plots show a representative dose-response experiment on four *C. elegans* strains for A) animal length (q90.TOF) and B) brood size (norm.n). The ABZ concentration is plotted on the x-axis, and individual replicate trait values subtracted from the mean in DMSO control conditions is plotted on the y-axis. Each box represents four technical replicates.

1027

1028
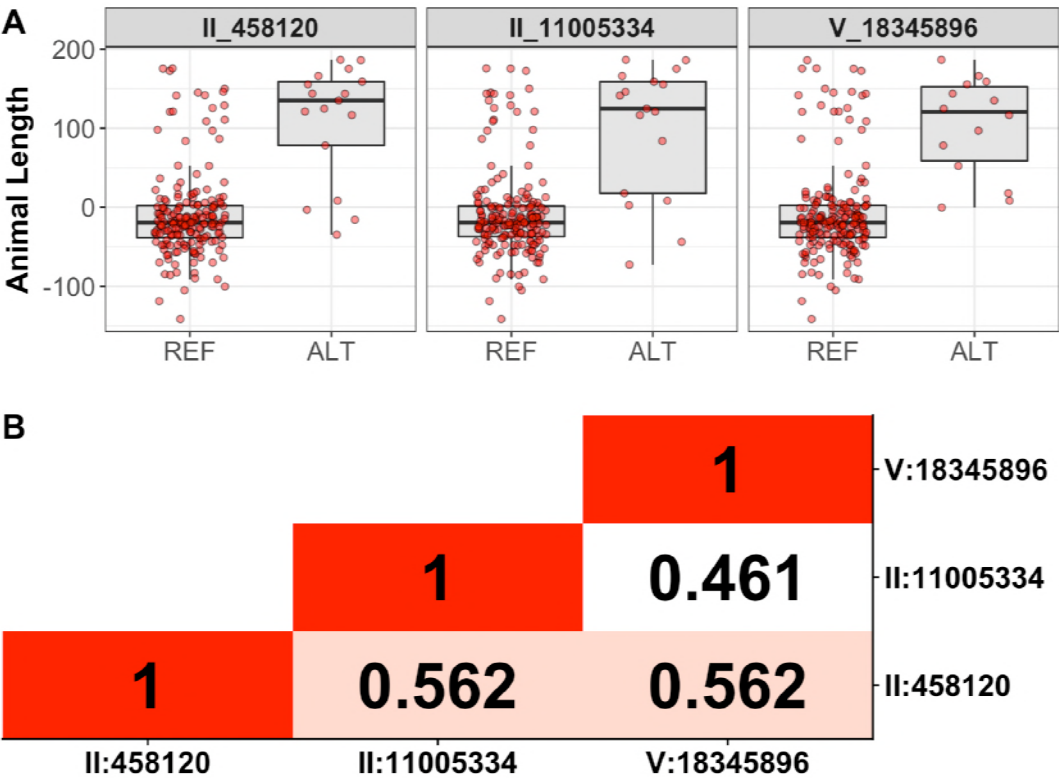
1029   **Supplemental Figure 2:**

1030   **Single-marker mapping summary**

1031   A) Regressed animal length (q90.TOF) phenotypes in the presence of ABZ. Each dot represents the mean

1032   regressed animal length of four replicates. Strains are grouped by the presence of the REF or ALT

1033   genotype at the peak QTL marker identified in the single-marker GWA mapping approach. B) Linkage

1034   disequilibrium (LD) as measured by the correlation coefficient between peak QTL markers in A. The

1035   formula for the correlation coefficient $r = -D / sqrt(p(A) * p(a) * p(B) * p(b))$.

1036

1037
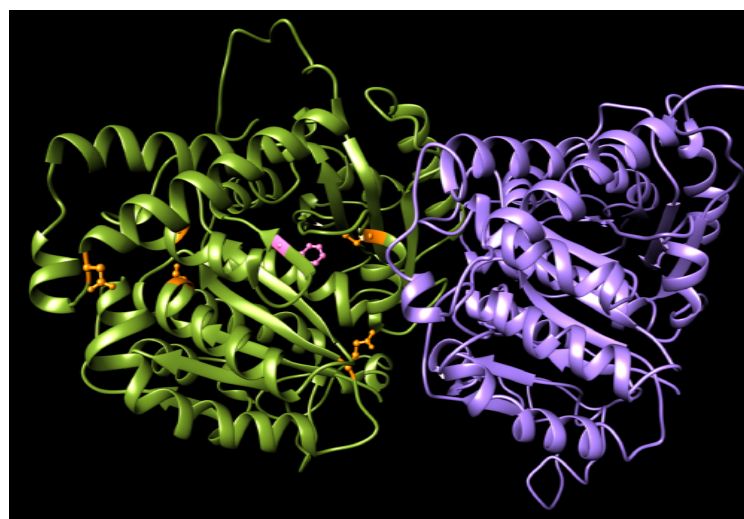
1038 **Supplemental Figure 3:**

1039 **Amino acid substitutions found in *C. elegans* wild strains are distributed throughout the BEN-1**

1040 **structure**

1041 An *in silico* model of the BEN-1 structure (green), binding to an alpha-tubulin (purple) is shown. Novel

1042 identified amino acid substitutions among *C. elegans* wild strains are highlighted in color. Most alleles

1043 were correlated with ABZ resistance (orange). The F200Y mutation, known as a major BZ resistance

1044 marker in parasitic nematodes is shown in pink.

1045

1046
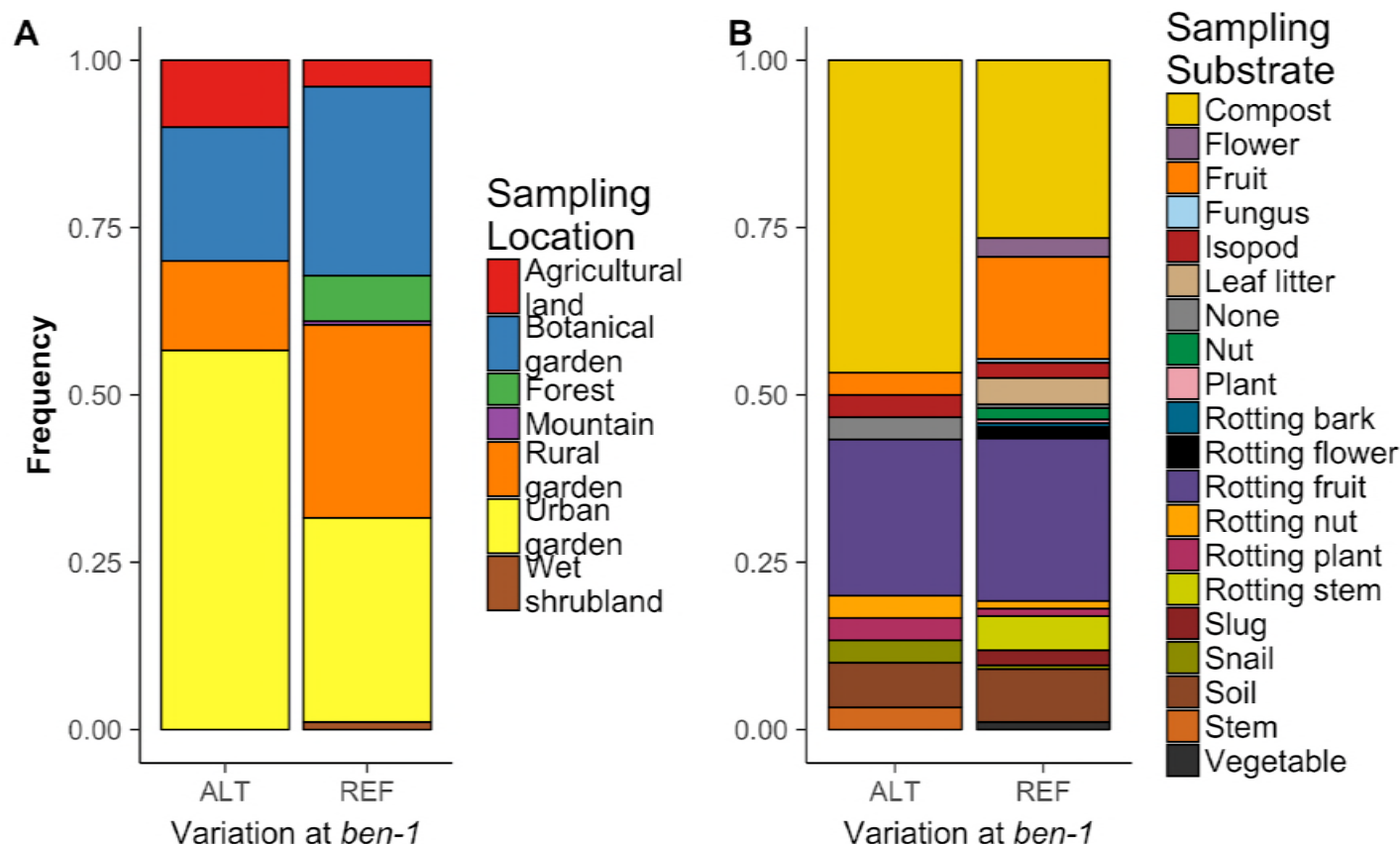
Supplemental Figure 4:

Sampling environment and substrate of wild *C. elegans* isolates

The fractions of wild *C. elegans* isolates sampled in a given A) location and on a given B) substrate are

shown. Colors for the stacked bar plots correspond to different A) sampling locations and B) substrates.

1051

1052

1053
1054
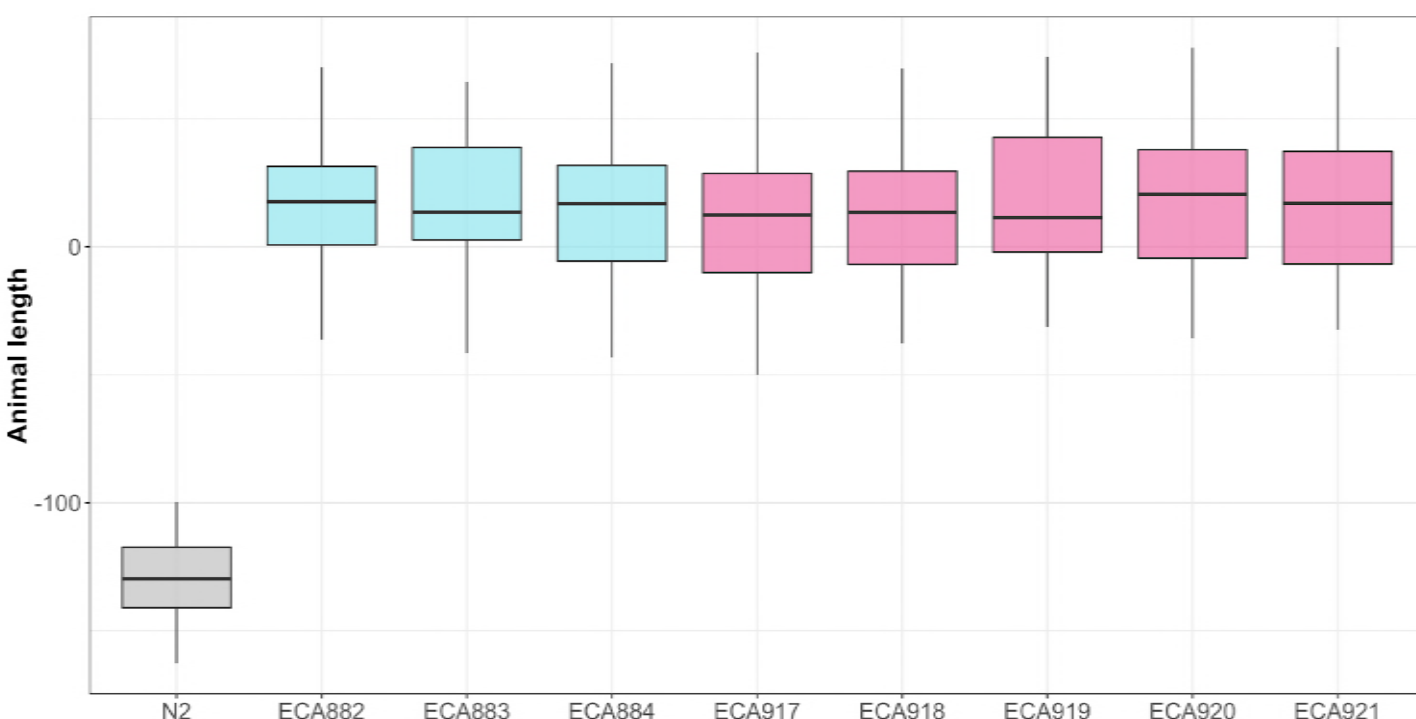
1055 **Supplemental Figure 5:**

1056 **Complete *ben-1* CRISPR allele HTA assay**

1057 Tukey box plots of the animal length phenotypes of the generated *ben-1* allele-replacement strains after

1058 ABZ treatment is shown. Blue boxes correspond to independent F200Y allele strains and pink boxes

1059 correspond to independent Del strains. The y-axis represents the animal-length phenotype after correcting

1060 for growth in DMSO conditions.

1061

1062

1063

1064
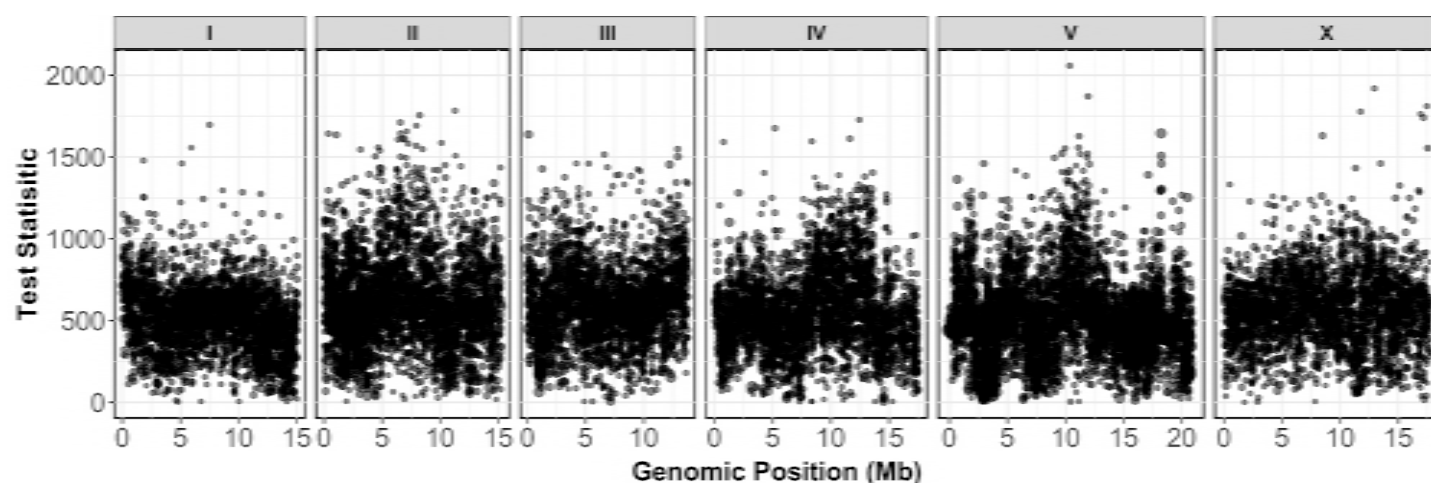
1065 **Supplemental Figure 6:**

1066 **Gene-burden mapping of putative *ben-1* LoF regressed phenotype**

1067 Gene-burden manhattan plot of the *ben-1* corrected animal length (q90.TOF) after ABZ exposure. Animal-

1068 length phenotypes in the presence of ABZ were adjusted based on the presence of a putative loss-of-

1069 function variant in *ben-1*. Each dot represents a single gene of the *C. elegans* genome with its genomic

1070 location plotted on the x-axis, and the test statistic plotted on the y-axis. Genes passing the burden test

1071 statistical significance threshold are colored in red.

1072

1073

1074
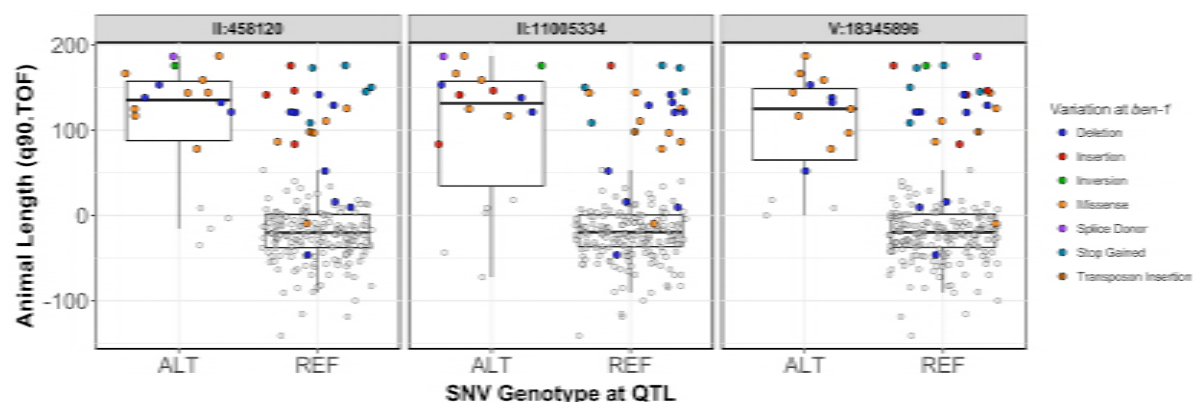
1075 **Supplemental Figure 7:**

1076 **Phenotypes of isolates split by genotypes at peak QTL markers identified using the single-marker**

1077 **mapping approach**

1078 Regressed animal length (q90.TOF) phenotypes in the presence of ABZ. Each dot represents the mean

1079 regressed animal length of four replicates per strain. Strains are grouped by the presence of the REF or

1080 ALT genotype at the peak QTL marker identified in the single-marker GWA mapping approach. Dot colors

1081 correspond to the identified variant class in the *ben-1* locus.

1082

1083

1084

1085
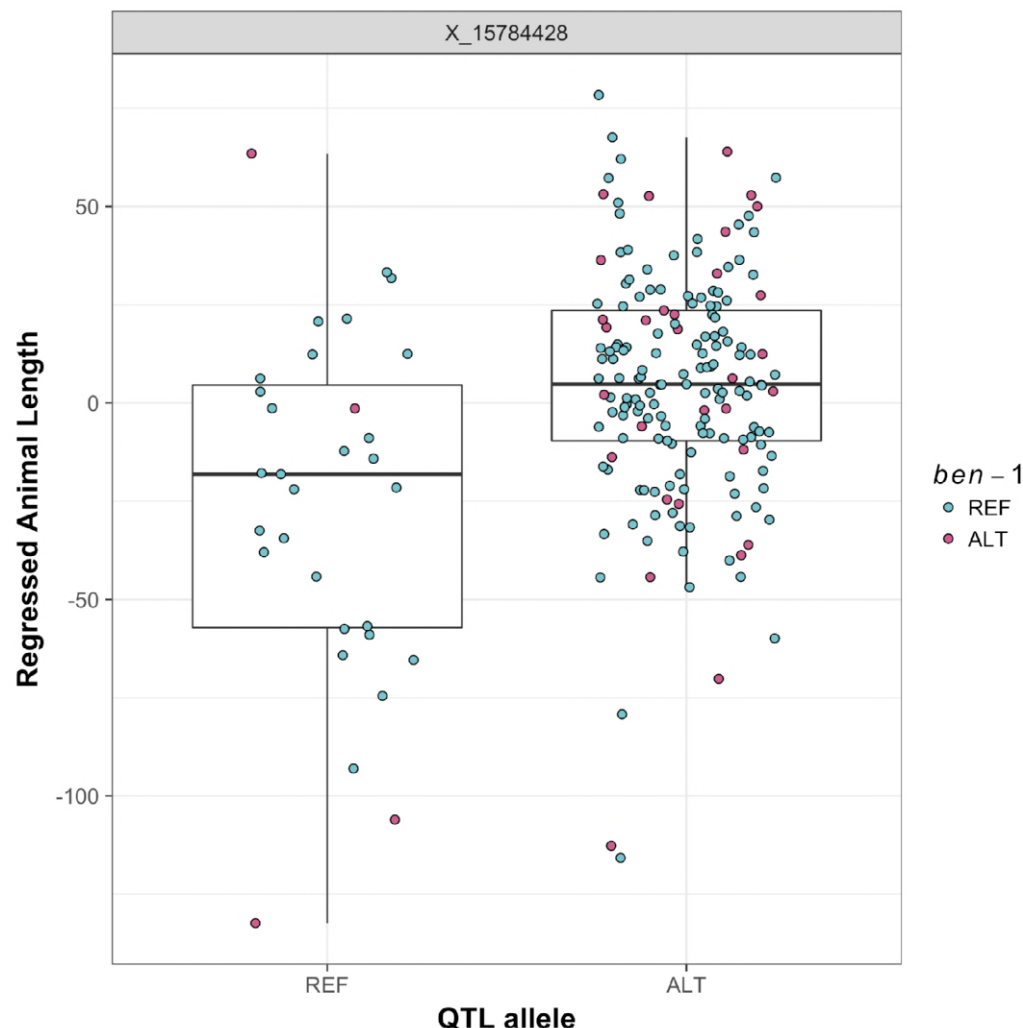
1086 **Supplemental Figure 8:**

1087 **Animal length phenotypes after correcting for the presence of putative *ben-1* loss-of-function**

1088 **variants**

1089 Animal length phenotypes in the presence of ABZ after correction for the presence of a putative loss-of-

1090 function variant in *ben-1*. Each dot represents the mean regressed animal length of four replicates per

1091 strain. Strains are split into groups based on the presence of the REF or ALT allele at the chromosome X

1092 QTL peak marker. Dots are colored by the presence of a putative loss-of-function variant in *ben-1*, where

1093 pink corresponds to strains with a putative loss-of-function variant, and blue corresponds to strains with no

1094 putative loss-of-function variants.

1095
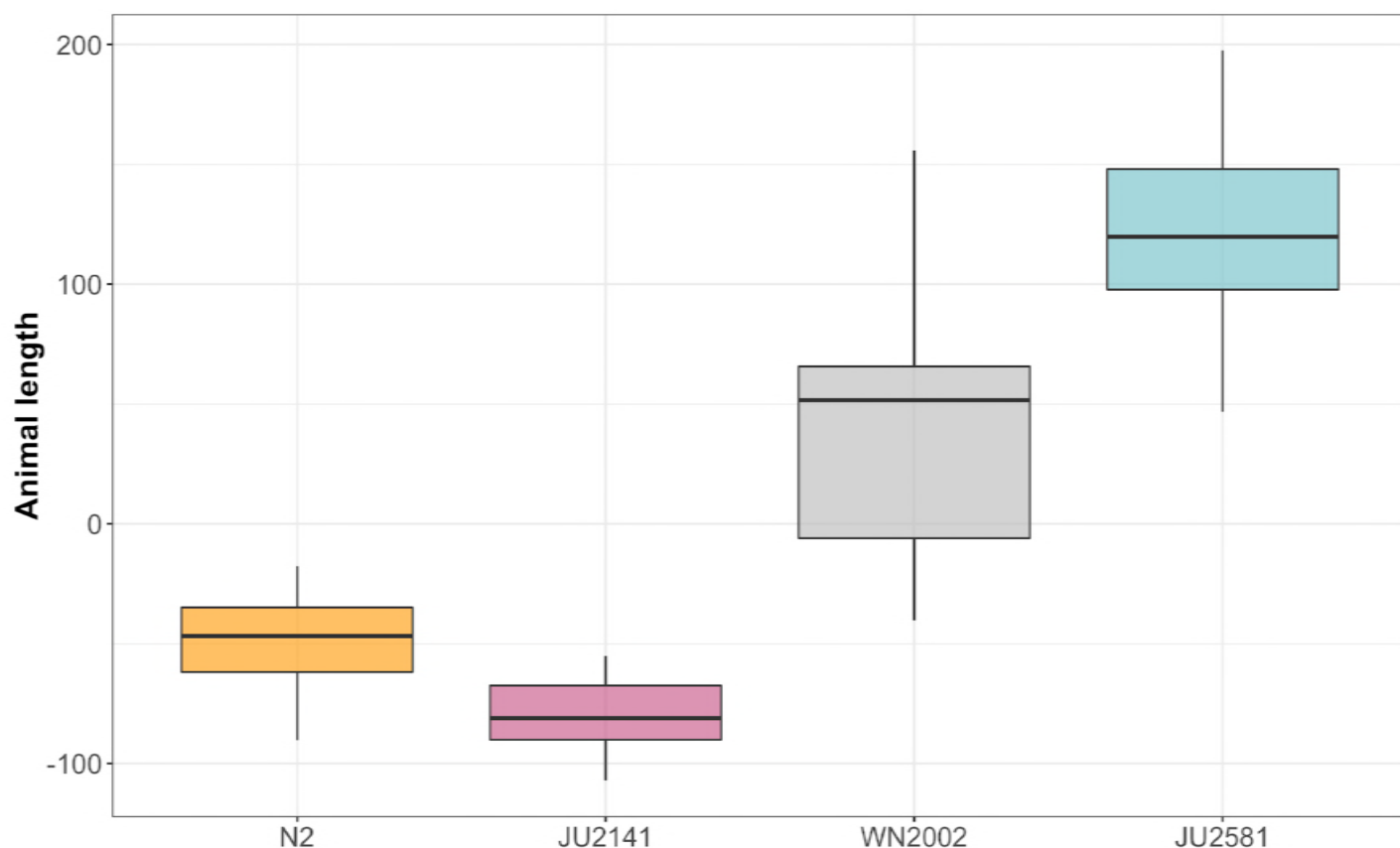
Supplemental Figure 9:

**Re-phenotyping of a slow-growing *ben-1* variant strain**

Tukey box plots of the animal length phenotypes of N2 and three wild isolates in the presence of ABZ. The y-axis represents the animal length phenotype after correcting for growth in DMSO conditions.

1103
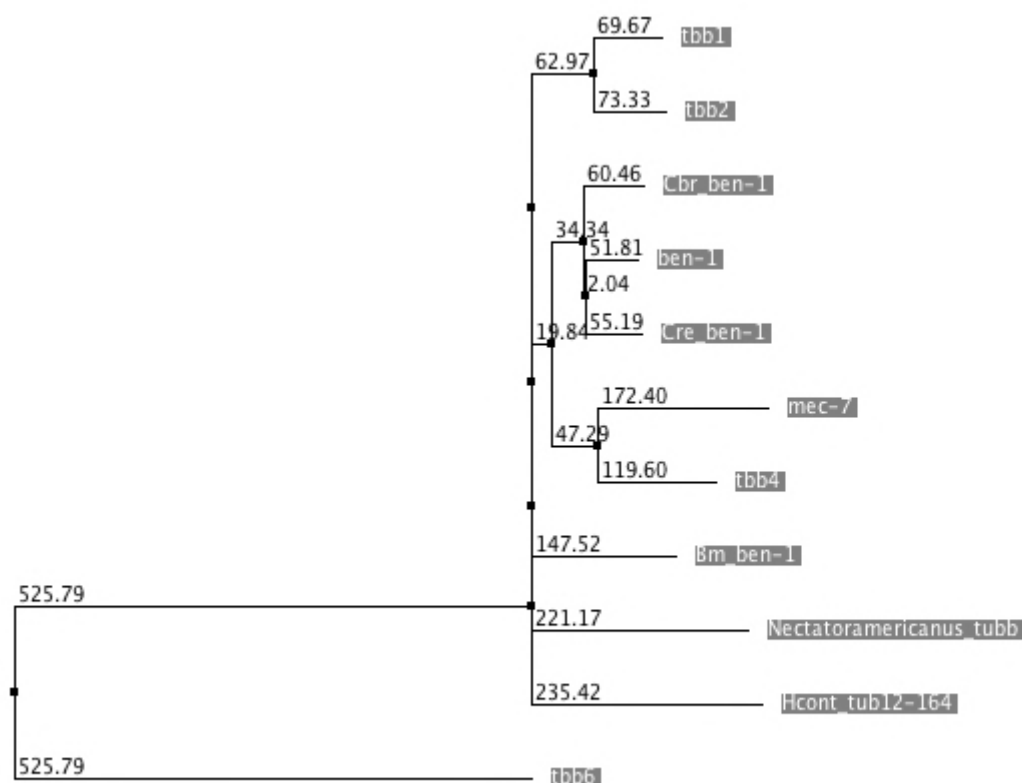
Supplemental Figure 10: Phylogenetic analysis of β-tubulins

Neighbor-joining tree of the amino acid sequences of the six known *C. elegans* β-tubulins and paralogs in

C. remanei (Cre), *C. briggsae* (Cbr), *H. contortus* (Hcont), and *Necator americanus*.

1107

1108
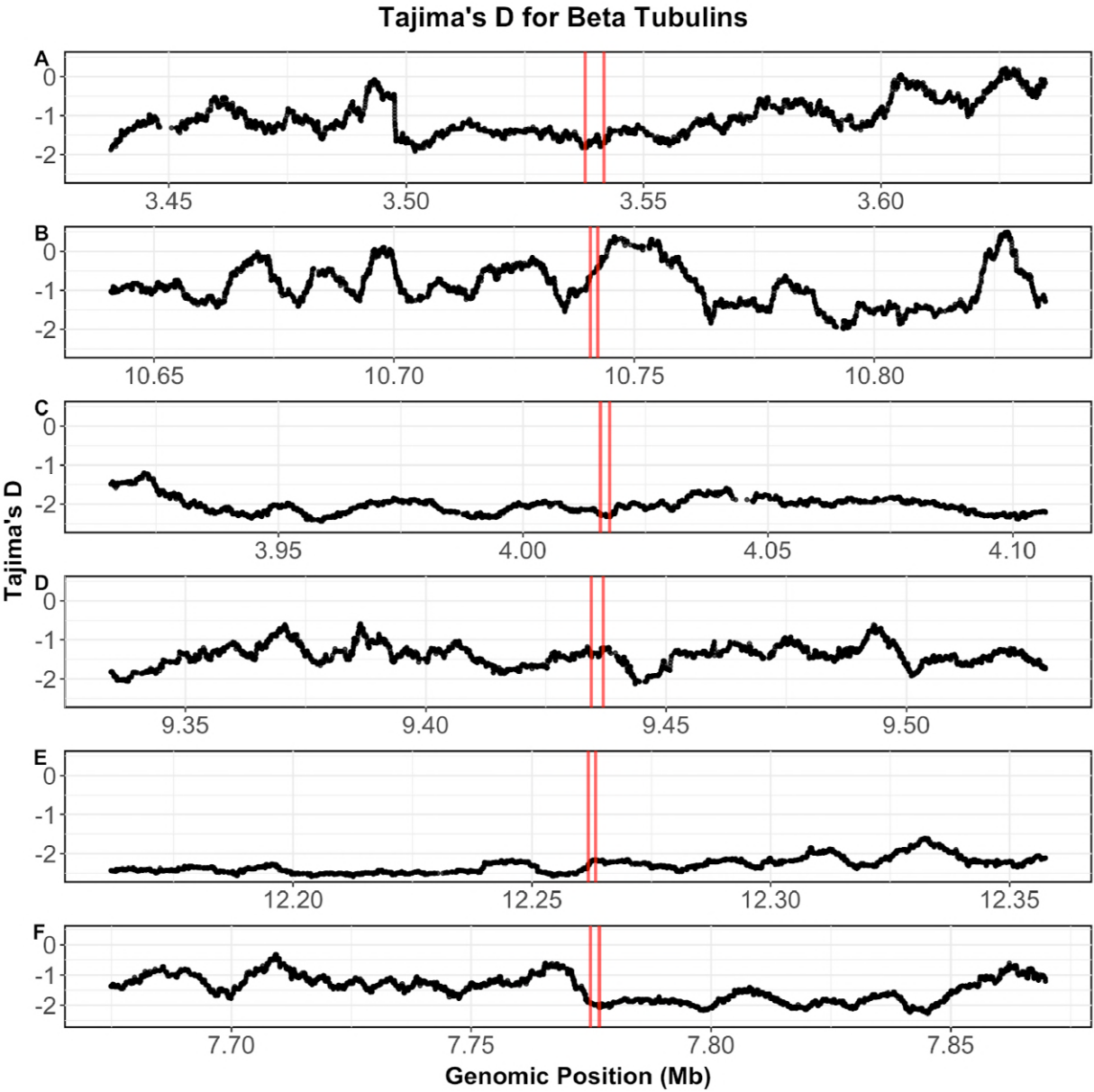
1109
1110

1111 **Supplemental Figure 11:**

1112 **Tajima's *D* at the *ben-1* locus and other β-tubulin genes**

1113 Tajima's *D* calculated from SNV data of genomic regions surrounding A) *ben-1*, B) *tbb-1*, C) *tbb-2*, D) *tbb-*

1114 *4*, E) *tbb-6,* and F) *mec-7*. Genomic position in Mb is plotted on the x-axis, and Tajima's *D* is plotted on the

1115 y-axis. Tajima's *D* was calculated using a sliding window with a 100 SNV window and a one-SNV step

1116 size. Two red bars for each panel correspond to the start and end positions for the corresponding gene.