Cryptic genetic variation in a heat shock protein shapes the expressivity of a mutation affecting stem cell behaviour in *C. elegans*

Sneha L. Koneru, Mark Hintze[§], Dimitris Katsanos[§] and Michalis Barkoulas[★]

Department of Life Sciences, Imperial College, London SW7 2AZ, United Kingdom.

§ These authors contributed equally to this work.

★Corresponding Author: m.barkoulas@imperial.ac.uk

Running title: Cryptic genetic variation shapes mutation expressivity

Keywords: seam cells, epidermis, stem cells, *C. elegans,* cryptic genetic variation, heat shock protein, *hsp-110,* Wnt signalling

Abstract

A fundamental question in medical genetics is how the genetic background modifies the phenotypic outcome of key mutations. We address this question by focusing on the epidermal seam cells, which display stem cell properties in *Caenorhabditis elegans*. We demonstrate that a null mutation in the GATA transcription factor *egl-18*, which is involved in seam cell fate maintenance, is more tolerated and thus has lower expressivity in the divergent CB4856 isolate from Hawaii than the lab reference strain N2 from Bristol. We identify multiple quantitative trait loci (QTLs) underlying the difference in mutation expressivity between the two isolates. These QTLs reveal cryptic genetic variation, which acts to reinforce seam cell fate through potentiating Wnt signalling. Within one QTL region, a single amino acid deletion in the heat shock protein HSP-110 in CB4856 lowers *egl-18* mutation expressivity. Our work underscores that natural variation in conserved heat shock proteins can shape mutation expressivity.

Introduction

Since the completion of the human genome project, there has been a long-standing interest in using genomic data to predict key phenotypic traits. Understanding the genotype-to-phenotype relationship represents one of the goals of personalised medicine, which aims at exploiting the predictability in this relationship to determine an individual's predisposition to disease or response to therapeutic treatment. Nevertheless, a fundamental challenge in this endeavour stems from the fact that genes exert their effects on a particular phenotype via complex gene-by-gene interactions (Chandler et al., 2013), therefore genetic variation present in the background can modify the genotype-to-phenotype mapping. Even when considering monogenic disorders, genetic modifiers can modulate the penetrance or expressivity of disease-causing loci, thereby influencing the proportion of the population developing the disease (incomplete penetrance) or the severity of the disease for each individual in the population (variable expressivity (Fournier and Schacherer, 2017)).

With regard to model organisms, mutant alleles are commonly studied in a single reference strain. Lab reference strains are highly inbred and have been selected to minimise genetic variation, which can act as confounder in the interpretation of genetic results. However, laboratory-driven evolution can fix alleles that are not widely present in the wild (Gasch et al., 2016; Sittig et al., 2016). In *Caenorhabditis elegans,* the laboratory reference strain N2 (isolated from Bristol, UK), has several adaptations suited to the lab environment (Sterken et al., 2015). For example, N2 carries laboratory-derived alleles for genes such as *npr-1*, which influences a large number of phenotypes or *nath-10*, which affects life-history traits (Duveau and Felix, 2012; Sterken et al., 2015). As a consequence, it is possible that well-studied mutations in N2 can have different phenotypic outcomes in different genetic backgrounds. Such

differences can vary from subtle phenotypic effects to more substantial changes, with a key example here being genes considered essential in one *S. cerevisiae* isolate, while being dispensable for survival in another (Dowell et al., 2010). Investigating how the effects of genetic mutations may vary across genetic backgrounds is important to uncover new gene functions and understand how genetic modifiers can shape phenotypic traits.

We study here this question using seam cell development in *C. elegans* as a simplified model of stem cell patterning (Joshi et al., 2010). The seam cells are lateral epidermal stem cells that contribute to the production of the cuticle, as well as produce neuronal lineages. A number of transcription factors have been shown to play a role in seam cell development (Joshi et al., 2010). Among them, GATA-type transcription factors are evolutionary conserved regulators of proliferation and mutations in these factors have been linked to disease, such as aggressive breast, colorectal and lung cancers (Tremblay et al., 2018). In *C. elegans*, GATA transcription factors play crucial roles in the development of the gut, epidermis and vulva. One particular pair, EGL-18 and its closely related paralogue ELT-6, are thought to specify seam cell fate and are direct targets of the Wnt signalling effector POP-1/TCF (Gorrepati et al., 2013). Mutations in *egl-18* result in precocious seam cell differentiation, which in embryos has been shown to correlate with misexpression of hypodermal markers in seam cells and fusion with the hypodermis (Koh and Rothman, 2001). Furthermore, EGL-18 is required for the seam cell hyperplasia observed upon Wnt signalling hyperactivation, for example following *pop-1* RNAi which results in symmetrisation of normally asymmetric cell divisions expanding the seam cell number (Gleason and Eisenmann, 2010; Gorrepati et al., 2013).

Seam cell development is invariant in the lab reference N2 with stereotypic

symmetric and asymmetric divisions during post-embryonic development giving rise to 16 cells per lateral side (Katsanos et al., 2017; Sulston and Horvitz, 1977). Over recent years, *C. elegans* has been sampled from around the globe (Cook et al., 2017), which offers exciting new opportunities to study how the genetic background affects developmental traits. We have recently shown that seam cell development is robust to standing genetic variation, with genetically divergent *C. elegans* isolates displaying a comparable seam cell number to the lab reference strain (Hintze et al., 2020). Developmental robustness can nevertheless lead to accumulation of genetic variation, which may not manifest phenotypically in wild-type, but can be revealed upon perturbation (Felix and Barkoulas, 2015; Gibson and Dworkin, 2004). This is called cryptic genetic variation and is a hidden source of variation that can influence the penetrance and expressivity of disease-causing loci, so it is linked to the emergence of complex disease (Gibson and Dworkin, 2004; Felix and Wagner, 2008; Paaby and Rockman, 2014).

Mutations in genes affecting seam cell development have been recovered and studied so far only in the N2 background. We sought therefore to investigate how the genetic background influences the phenotypic outcome of mutations disrupting seam cell development. We demonstrate here that a null mutation in the transcription factor *egl-18* results in lower expressivity, and thus higher average seam cell number, in the CB4856 strain from Hawaii compared to N2. Using quantitative genetics, we map the genetic basis of the differential expressivity of *egl-18(ga97)* mutation between N2 and CB4856. We identify a complex genetic architecture with multiple quantitative trait loci (QTLs) affecting mutation expressivity, while being cryptic for seam cell development in wild-type. We show that these QTLs act as potentiators of the effect of the Wnt signalling pathway in seam cell fate maintenance. We finally

126 show that in CB4856 a single amino acid deletion in HSP-110, which is a heat shock
127 70 family member, partially explains the difference in *egl-18(ga97)* mutation
128 expressivity between the two isolates. Our results highlight a paradigm where natural
129 variation in a conserved heat-shock protein is a key determinant of the expressivity of
130 a mutation influencing stem cell behaviour.
131

132 **Results**

133 **The expressivity of a null mutation in the GATA transcription factor *egl-18* is**
134 **lower in CB4856 than N2.**

135 To investigate whether the genetic background could influence the phenotypic
136 outcome of mutations affecting seam cell development, we introduced genetic
137 perturbations targeting known seam cell regulators in N2 into two wild isolates of *C.*
138 *elegans*: the commonly used polymorphic isolate CB4856 from Hawaii, and isolate
139 JU2007, which we had previously sampled from the Isle of Wight in the UK. Over the
140 course of these experiments, we discovered that a null mutation in the GATA
141 transcription factor *egl-18*, which acts downstream of the Wnt signalling pathway to
142 maintain the seam cell fate (Gorrepati et al., 2013; Koh and Rothman, 2001), was more
143 tolerated in the CB4856 background compared to N2 or JU2007 (Figure 1B and Figure
144 S1A). The *egl-18(ga97)* mutation expressivity was therefore lower in CB4856, leading
145 to a higher average seam cell count in the CB4856 population (Figure 1B). Importantly,
146 there was no significant difference in *egl-18(ga97)* mutation expressivity between N2
147 and JU2007, which indicated that the underlying cause was unlikely to relate to the lab
148 domestication of N2 (Figure S1A).

149 The difference in mutation expressivity was more apparent upon strong loss of
150 *egl-18* function, as it was less pronounced upon introgression of the milder *egl-*

151 *18(ok290)* allele (Figure S1A). Perturbing other components of the seam cell network
152 was not successful in revealing reproducible differences in seam cell number between
153 isolates. For example, null mutations in *bro-1*, the CBFβ homologue and binding
154 partner of RNT-1/Runx, or in the fusogen *eff-1,* led to comparable changes in seam
155 cell number between isolates (Figure S1B-C). Therefore, we decided to investigate in
156 more detail the genetic basis underlying the difference in *egl-18(ga97)* mutation
157 expressivity between N2 and CB4856, two isolates for which genomic differences are
158 well-characterised (Thompson et al., 2015; Wicks et al., 2001).

159 Since EGL-18 acts to maintain seam cell fate throughout development
160 (Gorrepati et al., 2013; Koh and Rothman, 2001), the difference in the *egl-18(ga97)*
161 mutation expressivity between N2 and CB4856 could arise from changes in embryonic
162 or post-embryonic development. To distinguish between these two possibilities, we
163 compared seam cell number between *egl-18(ga97)* mutants in the two isolates at the
164 end of the L1 stage. We found that both isolates showed a comparable loss of seam
165 cells at this early stage (Figure 1C). We then compared the spatial distribution of seam
166 cells at the L4 stage between the two isolates carrying the *egl-18(ga97)* mutation. Here
167 we found that *egl-18(ga97)* mutants showed a reduction of seam cells in both
168 backgrounds, however, significantly higher seam cell counts were retained in the
169 CB4856 background throughout the body axis (Figure 1D). There are multiple ways
170 via which an increase in terminal seam cell number can be achieved developmentally
171 in the *egl-18* mutant in CB4856. The first possibility involves the occurrence of
172 additional symmetric divisions, either due to symmetrisation of normally asymmetric
173 cell divisions, where anterior cell daughters maintain the seam cell fate instead of
174 differentiating into hypodermis, or because of ectopic symmetric cell divisions that
175 expand the seam cell pool. Alternatively, an increased maintenance of seam cell fate

176 would reduce precocious cell differentiation, and thus increase terminal seam cell
177 number independently of cell division. While phenotyping large numbers of *egl-*
178 *18(ga97)* animals, we never observed tightly clustered seam cell nuclei, which are
179 indicative of seam cell symmetrisation events (Hintze et al., 2020). These results
180 suggest that *egl-18(ga97)* mutants show higher terminal seam cell number in the
181 CB4856 background compared to N2 due to higher cell fate retention across all cell
182 lineages.

183 One mechanism to explain mutation buffering is through redundancy within the
184 gene network, for example if loss of function in one gene can be compensated by an
185 increase in expression of its closely-related paralogue (Burga et al., 2011; Rossi et al.,
186 2015; Serobyan et al., 2020). In this case, *elt-6,* could compensate for the loss of it
187 paralague *egl-18*, and changes in the degree of compensation between CB4856 and
188 N2 could underlie the observed difference in *egl-18(ga97)* mutation expressivity. It is
189 of note that such genetic compensation here could only be in the form of gene
190 expression regulation *in trans* to the *elt-6* locus, as opposed to regulation *in cis*
191 including changes in ELT-6 protein activity. This is because *egl-18* and *elt-6* are found
192 next to each other so the full *elt-6* locus was transferred from N2 to CB4856 together
193 with the *egl-18(ga97)* mutation. We used single molecule fluorescent *in situ*
194 hybridisation (smFISH) to compare *elt-6* expression between *egl-18(ga97)* mutants in
195 N2 and CB4856 and found no statistically significant difference (Figure 1E and F). This
196 finding suggests that the differential expressivity of the *egl-18(ga97)* mutation between
197 N2 and CB4856 is independent of changes in *elt-6* expression.

198 We then hypothesised that the difference in *egl-18(ga97)* mutation expressivity
199 between the two isolates might reflect changes in Wnt pathway activity. To test this
200 hypothesis, we targeted genes involved in this signalling pathway by RNAi and

201 compared seam cell counts between N2 and CB4856. We found that RNAi against lit-
202 1/NLK decreased seam cell number in both isolates, but N2 was more sensitive to lose
203 seam cells than CB4856 (Figure 1G). CB4856 is known to harbour variation that makes
204 this strain insensitive to germline RNAi (Paaby et al., 2015; Pollard and Rockman,
205 2013; Tijsterman et al., 2002), and seam cell RNAi is also less effective in CB4856
206 than N2 (Figure S1D). Therefore, we cannot formally rule out that the observed
207 difference upon *lit-1* RNAi is independent to inherent changes in RNAi sensitivity
208 between the two isolates. Interestingly, RNAi knockdown of *apr-1/APC* and *pop-1/TCF*
209 increased seam cell number in both isolates compared to control RNAi as expected,
210 but this time CB4856 was more sensitive to gain seam cells than N2 (Figure S1D).
211 This finding suggested an inherent difference between the two isolates in the outcome
212 of Wnt-related perturbation. We compared Wnt pathway activity between wild-type N2
213 and CB4856 animals using an established marker reflecting POP-1 binding in cells
214 where Wnt signalling is activated (Bhambhani et al., 2014), but found no reproducible
215 difference (Figure S1E). Finally, the difference in mutation expressivity between the
216 two isolates was masked upon RNAi of *elt-6* or *pop-1* (Figure S1F), highlighting a
217 requirement for Wnt-signalling for the difference to manifest. Taken together, we
218 conclude that the difference in *egl-18(ga97)* mutation expressivity between N2 and
219 CB4856 is likely to reflect changes that potentiate seam cell fate maintenance by Wnt-
220 signalled cells.

221 Animals carrying *egl-18(ga97)* mutations in both isolates display aberrant vulval
222 phenotypes, such as a protruding vulva and egg-laying defects. This is not surprising
223 because EGL-18 is required for maintenance of vulval precursor cell fate by inhibiting
224 their fusion to hypodermis (Eisenmann and Kim, 2000; Koh et al., 2002). We tested
225 whether the *egl-18(ga97)* mutation shows different expressivity between N2 and

226 CB4856 using vulval cell fate induction as a read-out. We found no statistically
227 significant difference between the two isolates (Figure S1G). We finally tested whether
228 a previously reported polymorphism in the essential RNA cytidine N-acetyltransferase
229 *nath-10* between N2 and other isolates including CB4856*,* which has been shown to
230 influence the differential expressivity of EGF/Ras mutations in vulva development
231 (Duveau and Felix, 2012)*,* is able to modify seam cell number. Here, we found that the
232 CB4856 *nath-10* polymorphism is not sufficient to modify the *egl-18(ga97)* mutation
233 expressivity in the N2 background (Figure S1H). Taken together, these results
234 suggested that the differential expressivity of *egl-18(ga97)* mutations between N2 and
235 CB4856 is specific to the seam and does not rely on previously identified genetic
236 variation that is known to affect epidermal development.
237

238 **Multiple QTLs influence *egl-18(ga97)* mutation expressivity between the two**
239 **isolates.**

240 To discover the genetic basis underlying the differential expressivity of *egl-*
241 *18(ga97)* mutation between N2 and CB4856, we used a quantitative genetic approach.
242 We produced 116 homozygous recombinant inbred lines (RILs), which consisted of
243 shuffled parental genomes and carried the *egl-18(ga97)* mutation in the background,
244 as well as the *scm::GFP* marker to be able to visualise the seam cells (Figure 2A).
245 Upon phenotypic characterisation of the resulting RILs, we found that the seam cell
246 number distribution was continuous, with most RILs displaying an average seam cell
247 number that was intermediate between the two parents (Figure 2B). We also observed
248 transgressive segregation, with some RILs showing a seam cell number that is higher
249 than CB4856 or lower than N2, indicating that multiple genetic loci are likely to affect
250 this trait.

251 We used bulked segregant analysis combined with whole genome sequencing
252 to map the genetic loci involved (Frezal et al., 2018; Michelmore et al., 1991). In this
253 approach, the extremes on each side of the phenotypic distribution that resemble the
254 parental phenotypes (i.e. they show low average seam cell number like N2 or high
255 average seam cell number like CB4856) are pooled and sequenced together as a bulk
256 (Figure 2B). According to the null hypothesis, there should be no statistically significant
257 differences in the SNP frequencies between the low-bulk and high-bulk samples for
258 genomic regions that do not influence the expression of this phenotypic trait. However,
259 deviations are expected for genomic regions that influence positively or negatively the
260 phenotype. Interestingly, we observed significant deviation in SNP frequencies on
261 multiple chromosomes, suggesting the presence of multiple QTLs (Figure 3). In
262 particular, we identified at least four significant QTLs on chromosomes II, III, V and X
263 that are likely to modulate the *egl-18(ga97)* seam cell phenotype (Figure 3A-F). In most
264 cases, the high bulk contained CB4856 alleles in these QTL regions, except for the
265 right arm of chromosome V where the high-bulk contained N2 alleles. Due to the
266 complexity of the experimental design, impaired resolution was anticipated in certain
267 chromosomal areas that were derived from N2 in all RILs. For example, RILs in both
268 groups had a short region on the left arm on chromosome IV from N2, which
269 corresponds to the region where the introgressed *egl-18(ga97)* mutation resides
270 (Figure 3D). Also RILs in both groups had a large portion of chromosome V from N2,
271 which is where the *scm::GFP* transgene is located (Figure 3E). Finally, a region on the
272 left arm of chromosome I was mostly derived from N2 (Figure 3A), because this region
273 harbours a previously known genetic incompatibility between the two parental isolates
274 (Seidel et al., 2011; Seidel et al., 2008).
275 To break-down the genetic composition of individual RILs in the two bulks and

276 study their association with the seam cell number phenotype, we took advantage of
277 known indels between the two parental strains to design markers for genotyping
278 around the QTL regions (Thompson et al., 2015). Upon genotyping of the high and
279 low-bulk RILS, we made the following observations. We found that 87.5% (21 out of
280 24) of high-bulk RILs compared to 10% (1 out of 10) in the low-bulk, carried a shared
281 CB4856 fragment on chromosome II (4.02-7.18 Mb) (Figure S2). We also observed
282 that 92% (22 out of 24) of high-bulk RILs compared to 40% (4 out of 10) in the low
283 bulk, carried an N2 fragment on the right end of chromosome V (18.66 Mb – 20.99 Mb)
284 (Figure S2). In the high bulk RILs, when chromosome II was derived from N2 rather
285 than CB4856 (as in RILs 16 and 73, Figure S2), then CB4856 fragments were both
286 present on chromosomes III (5.46 – 9.59 Mb) and X (3.00 - 5.08 Mb), indicating that
287 these QTLs may act in an additive manner. RILs in the low bulk carrying large CB4856
288 fragments on chromosomes III and X, also contained a CB4856 fragment on
289 chromosome V, which suggested that this chromosome V fragment may antagonise a
290 positive interaction between the QTLs on chromosomes III and X (Figure S2). One RIL
291 from the low-bulk (RIL-65) appeared to be the only exception in these genotype-to-
292 phenotype correlations. This RIL contained a large CB4856 fragment on chromosome
293 II and an N2 fragment on chromosome V, yet it displayed low average seam cell
294 number, which highlights that additional genetic loci, beyond the identified QTLs, are
295 likely to contribute to this complex phenotype.
296
297 **The identified QTLs harbour cryptic genetic variation affecting seam cell**
298 **development through the Wnt pathway.**
299 To investigate the putative effect of these QTLs on seam cell development, we
300 used RIL-28 to create near isogenic lines in wild type and *egl-18(ga97)* mutant N2

301 background. This RIL was chosen because it showed one of the highest average seam
302 cell number in the high bulk and yet contained shorter fragments derived from CB4856
303 on both chromosomes II and III compared to other RILs (Figure S2). By isolating the
304 independent fragments from CB4856 chromosomes II, III and X in N2, we found that
305 fragments on chromosomes II and III individually increased seam cell number
306 compared to the *egl-18(ga97)* N2 mutant phenotype (Figure 4A). The fragment on
307 chromosome X was not sufficient on its own to increase the average seam cell number
308 of the *egl-18(ga97)* mutant in N2, but it was able to act together with QTLs on
309 chromosome II and III to increase the average seam cell number further (Figure 4A).
310 Interestingly, seam cell number in lines containing a combination of two or three QTLs
311 together rescued the mutant phenotype to CB4856 levels, suggesting that the
312 identified QTLs capture a large proportion of the genetic variation influencing this
313 phenotypic trait (Figure 4A). Remarkably, the same QTLs did not alter wild type seam
314 cell development (Figure 4B), so they only acted in the presence of the *egl-18(ga97)*
315 mutation. These results thus highlight the presence of complex genetic variation in *C.*
316 *elegans*, which remains cryptic in wild-type condition but influences seam cell
317 behaviour upon genetic perturbation.

318 Given that the parental strains responded differently to knock-down of Wnt
319 signalling components, we decided to address whether the QTLs affecting *egl-*
320 *18(ga97)* mutation expressivity were related to this difference. To this end, we used
321 RNAi to knock down the expression of *pop-1* in strains containing combinations of the
322 QTL fragments (II, III and X) from CB4856 in a wild-type N2 background. Here we
323 found that *pop-1* RNAi led to significantly higher seam cell number in animals carrying
324 at least two QTL regions compared to N2. This finding highlights that the identified
325 QTLs may promote seam cell fate by potentiating the Wnt signalling pathway.

326

327 **Natural variation in *hsp-110* contributes to the difference in mutation**
328 **expressivity between the two *C. elegans* isolates.**

329 To narrow down the genomic interval of the two major QTLs on chromosomes
330 II and III, we screened for further recombinants after crossing our NILs to N2 and
331 screening for lines in which the presence of a defined CB4856 fragment was sufficient
332 to increase seam cell number of *egl-18(ga97)* mutants in the N2 background. Briefly,
333 with regard to chromosome II, we derived two NILs (strain MBA944 and MBA846),
334 which largely rescued the expressivity of the *egl-18(ga97)* mutant in N2, while they
335 contained mostly distinct genomic regions with a small overlapping region in the middle
336 of approximately 0.14 Mb (Figure S3). Non-overlapping regions in MBA944 and
337 MBA846 contained 106 and 116 genes harbouring natural variation, with 7 genes in
338 the overlap (*dgk-5, del-10, T28D9.1, abch-1, utp-20, wrn-1, C56C10.9*). To narrow
339 down this list further, we reasoned that genes responsible for the difference in *egl-*
340 *18(ga97)* mutation expressivity between N2 and CB4856 were likely to alter seam cell
341 number when targeted by RNAi in an *egl-18(ga97)* mutant background. However, we
342 found no difference in seam cell number in the *egl-18(ga97)* background upon
343 downregulating the 7 genes in the overlap (Figure S4A). These results are therefore
344 more compatible with the existence of two independent QTLs on chromosome II, as
345 opposed to a single QTL in the overlapping region. Consistent with this hypothesis, we
346 found that RNAi knockdown of *egl-27* and *dsh-2,* which are located towards opposite
347 sides of the overlap and harbour natural genetic variation known to play a role in seam
348 cell patterning (Baldwin et al., 2016; Herman et al., 1999; Solari et al., 1999), both
349 decreased seam cell number in the *egl-18(ga97)* mutant background (Figure S4A).
350 Therefore, natural variation in these genes is likely to contribute to the differential

351 mutation expressivity between the two isolates.

352 With regard to chromosome III, we found several loci in the broad QTL region
353 that reduced seam cell number in the *egl-18(ga97)* mutant when targeted by RNAi
354 (Figure S4B). We were then able to narrow down the QTL to a smaller genomic interval
355 of approximately 1.14 Mb, which contained 59 genes in total with known genetic
356 variation (Figure 5 and Table S3) and two strong candidates based on our RNAi
357 experiments. First, *sor-1*, a species-specific polycomb group (PcG) protein that is
358 involved in epigenetic silencing of Hox genes (Zhang et al., 2006). Genetic variation in
359 this gene involved one non-synonymous substitution and one adjacent synonymous
360 mutation, both in exon 7. The second candidate is *hsp-110*, a homolog of HSPA4 and
361 member of the HSP70 gene superfamily (Easton et al., 2000; Nikolaidis and Nei, 2004).
362 Genetic variation in *hsp-110* involved a 3 bp in-frame deletion in exon 5 in CB4856,
363 which deletes a conserved amino acid within the substrate-binding domain of this
364 chaperone (Figure 6A).

365 To validate the impact of these variants, we used CRISPR-mediated genome
366 editing to engineer them into N2 and assess their role on seam cell development both
367 in a wild-type and in an *egl-18* mutant background. Interestingly, we found that CB4856
368 genetic variation in *hsp-110* significantly increased seam cell number when introduced
369 into an *egl-18(ga97)* mutant background in N2 (Figure 6B), while no effect was seen
370 in a wild-type background (Figure 6C). We used smFISH to study the expression of
371 *hsp-110* and found evidence for expression around seam cell nuclei, with no apparent
372 difference between N2 and CB4856 (Figure 6D-E). This indicates that HSP-110 is
373 likely to act cell-autonomously to modify seam cell behaviour. Instead, we observed no
374 effect on seam cell number when we introduced the CB4856 *sor-1* polymorphisms in
375 a wild-type or *egl-18* mutant in the N2 background (Figure 6B-C). Taken together,

these results suggest that natural genetic variation in a conserved heat-shock protein contributes to the difference in *egl-18(ga97)* mutation expressivity between CB4856 and N2.

Discussion

Previous work has shown that the phenotypic effect of mutations can vary depending on the genetic background in various models (Chow et al., 2016; Dworkin et al., 2009; Sittig et al., 2016). For example, the severity of various null mutations in mice depends on the genetic background, with often opposing effects observed among strains (Sittig et al., 2016), indicating that the genetic background can influence the genotype-to-phenotype relationship in a complex way (Chandler et al., 2017). *C. elegans* has low genetic diversity due to its self-fertilization, however the genetic variation present in the background has been shown to influence the outcome of genetic perturbations, such as RNAi-driven phenotypes (Paaby et al., 2015; Vu et al., 2015), although the exact molecular changes in the genetic background that influence developmental outcomes remain poorly understood.. With regard to mutation expressivity, a lab evolved non-synonymous polymorphism in *nath-10* between N2 and the AB1 isolate from Australia has been shown to contribute to the expressivity in a mutation in the epidermal growth factor receptor gene involved in vulval development (Duveau and Felix, 2012).

We find here that the expressivity of the *egl-18(ga97)* mutation is lower in CB4856 compared to the lab reference strain N2, with approximately two more seam cells produced on average in the CB4856 background. GATA factors are important players in animal development and mutations in these factors are known to lead to blood and heart disease, with often variable phenotypic outcomes in human patients

401 (Brambila-Tapia et al., 2017; Tremblay et al., 2018). The difference in *egl-18(ga97)*
402 mutation expressivity is specific to the seam and manifests during post-embryonic
403 development through changes that reinforce seam cell maintenance in a Wnt-
404 dependent manner. Using a quantitative genetic approach, we discover at least four
405 QTLs that modify seam cell number in an *egl-18(ga97)* mutant background, acting both
406 independently, as well as in combination. We pin down the molecular basis of one of
407 these QTLs and demonstrate that a deletion of a single amino acid (D474) in the heat
408 shock protein HSP-110, which has occurred during the evolution of the CB4856 strain,
409 is sufficient to lower mutation expressivity and thus increase the average seam cell
410 number when introduced into N2. HSP110 proteins form a distinct branch of the Hsp70
411 gene superfamily (Dragovic et al., 2006; Easton et al., 2000; Raviol et al., 2006), which
412 includes conserved proteins that assist in protein folding and protect cells from stress
413 (Mayer and Bukau, 2005). HSP110 members are thought to assist in solubilization of
414 protein aggregates through nucleotide exchange activity rather than acting as
415 canonical ATPase chaperones (Rampelt et al., 2012; Scior et al., 2018). This protein
416 has not been studied before in the context of *C. elegans* development. Previous reports
417 have demonstrated a role for HSP-110 in lifespan upon heat shock (Rampelt et al.,
418 2012), as well as recovery from thermal stress in plant pathogenic nematodes, where
419 evolution in *hsp-110* has occurred in *Globodera* species through gene duplication
420 (Jones et al., 2018).

421 Molecular chaperones, like Hsp90, can act as phenotypic capacitors that
422 modulate the genotype-to-phenotype relationship (Zabinsky et al., 2019). Hsp90-
423 mediated buffering of protein folding has been thought to influence the variable
424 expressivity of genetic disease (Karras et al., 2017). In the seam, RNAi knockdown of
425 *hsp-90* or DnaJ chaperones leads to increased variability in seam cell number (Hughes

426 et al., 2019; Katsanos et al., 2017). Differences in heat shock protein expression have
427 also been shown to affect the incomplete penetrance of mutations (Burga et al., 2011;
428 Casanueva et al., 2012). Our work highlights that natural variation in conserved heat
429 shock proteins can shape mutation expressivity. It is of note that mutations in HSP-
430 110 in human patients correlate with variable responses to drug chemotherapy (Collura
431 et al., 2014), so it is conceivable that such mutations may influence the phenotypic
432 outcome of disease-causing loci.

433 EGL-18 and ELT-6 transcription factors act redundantly to promote seam cell
434 fate (Gorrepati et al., 2013). In both CB4856 and N2, EGL-18 appears to be the major
435 seam cell promoting factor based on the severity of the loss of function phenotype,
436 while ELT-6 can compensate to some extent for the absence of EGL-18 activity.
437 Genetic compensation has been shown to occur in various organisms including *C.*
438 *elegans* as a mechanism to buffer against null mutations to protect from developmental
439 failure (Burga et al., 2011; Rossi et al., 2015; Serobyan et al., 2020). However, we did
440 not find changes in *elt-6* expression in an *egl-18* mutant background between the two
441 isolates. We therefore propose a model wherein changes in the seam cell regulatory
442 network through accumulation of natural variation, including the amino acid deletion
443 found within HSP-110 in CB4856, may potentiate the Wnt signalling pathway to
444 promote seam cell fate through ELT-6 or another factor (Figure 6F). This model of
445 reinforced seam cell fate adoption in CB4856 through Wnt signalling is consistent with
446 the increase in seam cell counts observed in a *pop-1* RNAi treatment in CB4856
447 compared to N2. Interestingly, a recent study revealed cryptic genetic variation in the
448 requirement of the POP-1 input in gut differentiation (Torres Cleuren et al., 2019),
449 suggesting that natural variation may more broadly shape Wnt-dependent gene
450 regulatory networks in *C. elegans*. It is of note that a crosstalk between Hsp70 proteins

451 and Wnt signaling has been proposed in other stem cell contexts (Zhang et al., 2016).
452 Based on the position of the deleted amino-acid in CB4856 within the substrate binding
453 domain of HSP-110, we speculate that some altered interaction with a protein partner
454 may increase the strength of the Wnt signaling pathway in promoting the seam cell
455 fate.

456 The amino acid deletion in *hsp-110* is specific to the CB4856 isolate of *C.*
457 *elegans* and is not present in N2, JU2007 and other divergent isolates (CeNDR release
458 20180527). CB4856 belongs to a group of highly polymorphic strains found around
459 geographically isolated Hawaiian islands. These isolates are thought to represent
460 ancestral genetic diversity because they contain approximately three times more
461 diversity than the non-Hawaiian populations (Andersen et al., 2012; Cook et al., 2017;
462 Crombie et al., 2019). Interestingly, we previously reported a genotype-by-environment
463 interaction affecting seam cell number (Hintze et al., 2020). An increase in
464 temperature, from 20 to 25 °C, was reported to increase seam cell number through a
465 symmetrisation of normally asymmetric divisions at the L4 stage in many *C. elegans*
466 isolates, except for CB4856 (Hintze et al., 2020). It is therefore intriguing that CB4856
467 appears to harbour loci that both reinforce seam cell fate maintenance and suppress
468 ectopic seam cell fate acquisition, which highlights that genetic variation in *C. elegans*
469 can influence developmental phenotypes in a complex manner. It is of note that natural
470 variation exists in *hsp-110* and other heat-shock proteins including *hsp-90* (CeNDR),
471 so it remains to be seen whether such variants influence developmental outcomes
472 more broadly.

473 Although natural variation in *hsp-110* contributes to the difference in mutation
474 expressivity, we show that a complex genetic architecture underlies this phenotypic
475 trait so other loci are also involved. Our results suggested two strong candidates for

476 the QTL on chromosome II, that is the Dishevelled protein *dsh-2* and *egl-27*. There
477 are three Dishevelled proteins (MIG-5, DSH-1 and DSH-2) in *C. elegans*. DSH-2
478 and MIG-5 have been shown to act redundantly to regulate seam cell fate by
479 positive regulation of nuclear SYS-1 β-catenin levels (Baldwin et al., 2016). EGL-27
480 is an ortholog of human RERE (arginine-glutamic acid dipeptide repeats protein) and
481 mutations in *egl-27* have been associated with seam cell defects leading to an increase
482 in seam cell number (Herman et al., 1999; Solari et al., 1999). Chromatin-modifying
483 regulators, such as EGL-27, have been proposed to represent highly connected hubs
484 in *C. elegans* genetic interaction networks (Lehner et al., 2006), thereby influencing
485 the phenotypic outcomes of multiple mutations in factors belonging to several
486 unrelated pathways. Therefore, it is conceivable that natural variation in the core Wnt
487 signalling component *dsh-2* and *egl-27* may be involved in the modification of the
488 phenotypic outcome of the *egl-18(ga97)* mutation in the seam. Future work is required
489 to understand the phenotypic consequences of these natural variants and dissect how
490 they may modify the Wnt signalling pathway.
491

492 **Materials and methods**

493 **Strains and genetics**

494 *C. elegans* was maintained on a lawn of *Escherichia coli* strain OP50 seeded on
495 Nematode Growth Medium (NGM) according to standard procedures (Stiernagle,
496 2006). *C. elegans* larvae were synchronised by bleaching gravid hermaphrodites and
497 washing the eggs twice with M9 buffer or by puncturing gravid animals with an injection
498 needle to release eggs in the case of *egl-18* mutants. All strains used in this study are
499 listed in Table S1. Mutations in *egl-18* were introgressed into wild-type strains MBA256
500 and MBA19 (CB4856 and JU2007 respectively carrying the *wIs51* transgene) (Hintze

501 et al., 2020) in a two-step cross. In the first step, MBA256 and MBA19 males were
502 crossed to *egl-18(-)* hermaphrodites carrying the *wIs51* transgene. In the second step,
503 F1 males from the previous cross were crossed to wild isolate hermaphrodites. F2
504 animals from the second cross that were egg-laying defective were isolated and the
505 two-step cross was repeated five times to produce 10 × backcrossed strains. The
506 process was repeated for the introgression of *eff-1(icb4)* into the MBA19 background.
507 The *eff-1(icb4)* mutation will be described elsewhere (Koneru et al. in preparation).
508

509 **RNA interference (RNAi)**

510 RNAi was performed by feeding animals with bacteria expressing double-stranded
511 RNA. Bacterial clones originate from the Ahringer and Vidal libraries, or were custom
512 made (*wrn-1*, *utp-20*). Bacterial clones were grown overnight in liquid LB medium with
513 50 µg/ml ampicillin and 12.5 µg/ml tetracycline. The bacterial cultures were seeded
514 onto filter-sterilised isopropyl β-D-1-thiogalactopyranoside (IPTG)-containing RNAi
515 plates and allowed to dry for 2 – 4 days at room temperature before use. Custom RNAi
516 clones were made by amplifying gene fragments for a gene of interest using primers
517 that contained the following sequences (Fw 5'-AGA CCG GCA GAT CTG ATA TCA
518 TCG ATG- 3', Rev 5'- TCG ACG GTA TCG ATA AGC TTG ATA TCG - 3') to allow
519 Gibson cloning into EcoRI-digested L4440.
520

521 **Phenotypic analysis and microscopy**

522 Seam cell number was quantified in day-1 adults or late L4s. Animals were
523 anaesthetised using 100 µM sodium azide and mounted on a 2% agarose pad. Seam
524 cells were visualized using the *scm::GFP* marker (*wIs51*) (Koh and Rothman, 2001)
525 and the lateral side closer to the lens was counted on a Zeiss compound microscope

526 (AxioScope A1) with 400x magnification.

527 RILs were phenotyped over a period of one week as they were growing at
528 different rates. 116 RILs were scored twice and 38 RILs were scored thrice to make
529 sure that RILs included in the pools for sequencing were the ones with the most
530 reproducible seam cell number (SCN) between replicates. We pooled DNA of the two
531 extreme groups (22 RILs in the low-SCN bulk and 24 RILs in the high-SCN bulk) for
532 QTL mapping.

533 Single molecule mRNA fluorescent *in situ* hybridisation was performed using a
534 Cy5-labelled *elt-6* and *hsp-110* probes (Biomers). Z-stacks with 17-30 slices, each of
535 0.7 µm, were acquired with a 100× oil immersion objective using an Andor iKon M 934
536 CCD camera system on a Nikon Ti Eclipse epifluorescence microscope. Region of
537 interests for spot quantifications were drawn manually around seam cells visualized
538 using the *scm::GFP* marker and a custom script in MATLAB was used to quantify
539 mRNA molecules (Hintze et al., 2020).

540 Confocal images were obtained on a Leica SP5 using an Argon 488nm laser
541 and analysed with Fiji. A region of interest was drawn from the vulva to the tail
542 containing intestinal, vulval and seam cells expressing the POPHHOP marker. The
543 corrected total cell fluorescence (CTCF) was calculated as follows: CTCF = Integrated
544 Density − (Area of ROI × Mean fluorescence of background readings)

545

546 **Quantitative genetics**

547 Recombinant inbred lines (RILs) were generated by crossing hermaphrodites
548 from strain MBA256 (CB4856) and males from strain MBA231 (CB4856). F1 males
549 from the cross were crossed to hermaphrodites from strain MBA290 (N2) containing
550 the *egl-18(ga97)* mutation. Multiple new F1 hermaphrodites, which were egg-laying

551 defective and carried the transgene *vtIs1[dat-1::gfp] wIs51[scm::GFP]* V were picked
552 and allowed to self as cross progeny. 117 F2s which were egg-laying and carried only
553 *wIs51[scm::GFP]* V transgene were picked onto single NGM plates and allowed to self.
554 One hermaphrodite was transferred to a new plate for 10 – 14 generations to establish
555 RILs. RIL-17 did not propagate during the selfing process so 116 RILs were produced
556 in total.

557 We used a bulk segregant analysis approach to discover quantitative trait loci
558 as previously done (Frezal et al., 2018). DNA was extracted from the low and high
559 bulks using a Gentra Puregene Kit (Qiagen®). Whole genome sequencing was
560 performed in an Illumina® Hiseq platform (read length = 2 × 125), aiming to obtain at
561 least 10 million read pairs per sample (25 × coverage). The whole genome sequencing
562 data were analysed through the CloudMap Hawaiian variant mapping pipeline on a
563 local galaxy server maintained in the lab (Minevich et al., 2012). The genotype for low-
564 SCN and high-SCN was extracted into a single vcf file from the whole genome data
565 using vcf combine on galaxy. The combined vcf was converted to tabular format using
566 GATK tools on a bash terminal. Frequency of CB4856-like variation at each SNP
567 position was calculated as the ratio of read counts with CB4856 SNP divided by the
568 total number of reads. For the low-SCN and high-SCN bulk, the SNP positions where
569 the genome quality (as determined by GQ scores < 40) for the sequencing was low
570 were discarded using the QTLseqr package in R (Mansfeld and Grumet, 2018). SNPs
571 where the total read depth was lower than 22 in the low-SCN bulk and lower than 24
572 in the high-SCN bulk were discarded as in (Frezal et al., 2018). We used log-odds
573 ratio, which is the logarithm (base 10) likelihood ratio as a test statistic to evaluate if
574 the deviations observed in SNP frequencies between the low-SCN and high-SCN bulk
575 were statistically significant than expected from a null distribution. We calculated

576 observed log-odds ratio using the following formula: log-odds ratio = $log_{10}\left(\frac{h}{n_h-h}/\frac{l}{n_l-l}\right)$,
577 where l and h are mean SNP frequencies of CB4856 in a 300 kb window with no
578 overlap multiplied by n_l and n_h, n_l and n_h are number of RIL lines pooled for low-SCN
579 and high-SCN bulks, respectively. The log-odds ratios under the null hypothesis were
580 calculated where SNP frequencies for both the bulks were generated by 1 million
581 simulated Bernoulli trails with p = q = 0.5. From these log-odds ratios, we found the
582 two-tailed genome-wide threshold at a significance level (α = 0.05). The observed log-
583 odds ratio was plotted against the genome location, and was compared to the genome-
584 wide thresholds. The presence of a QTL was inferred when the log-odds ratio
585 exceeded the threshold.
586
587 **CRISPR-Cas9 mediated genome editing**
588 *hsp-110* and *sor-1* genetic variants in the N2 background were generated via injection
589 of CAS9 ribonucleoprotein complexes (Paix et al., 2015) at the following final
590 concentrations in 10 µl: tracRNA (0.75 nmol – IDT), custom crRNAs (*hsp-110* target
591 sequence–AAA GTC AAT GGA GCA CGA CC and AAC CTT CTG ATT ACC ACC
592 AT), (*sor-1* – TAG TCA ACG CCC ACC AAG CA, 20 µM, IDT) and Cas9 (1 µg/µl -
593 IDT). The co-injection markers *myo-2::RFP* at 5 ng/µl and *rol-6*(*su1006)* at 40 ng/µl
594 were used to select transgenic animals, and repair templates (6 µM) introducing SmaI
595 sites in the case of *hsp-110* and ClaI in the case of *sor-1* together with the desired
596 CB4856 SNP changes. F1 animals were screened for the co-injection markers,
597 allowed to lay progeny and screened using restriction digests for the introduced sites.
598 Positive lines were Sanger-sequenced to identify lines homozygous for SNP insertions.
599 The co-CRISPR strategy (Arribere et al., 2014) was used to edit *bro-1*. An sgRNA
600 targeting the following sequence (5'- AATCAATATACCTGTCAAGT) was cloned into

601 pU6::unc-119 sgRNA vector by replacing the *unc-119* sgRNA as previously described
602 (Friedland et al., 2013). The recovered icb45 allele is an in-frame deletion of 9 bp
603 (GGAATCAATATA- - - - - - - - -TTGGAATGGT) within the second exon of *bro-1.*
604

605 **Figure legends**

606 **Figure 1. The expressivity of the *egl-18(ga97*) mutation is higher in N2 compared**
607 **to CB4856 resulting in lower seam cell number in the N2 background.**
608 (A) Cartoon depicting the stereotypic seam cell distribution in wild-type animals at the
609 end of larval development, where 8 cells are positioned anterior to the vulva (H0 – V2),
610 2 cells are exactly at the region of the vulva (V3), and 6 cells are found posterior to the
611 vulva (V4 – T). (B) Average terminal seam cell number (SCN) is significantly lower in
612 N2 than CB4856 carrying the *egl-18(ga97)* mutation (one-way ANOVA F (1, 310) =
613 115.02, $p < 2.2 \times 10^{-16}$, n ≥ 150), but is not different in wild-type (one-way ANOVA F
614 (1, 142) = 0.1274, $p = 0.13$, n = 72). (C) No statistically significant difference was found
615 between N2 and CB4856 after the L1 division (one-way ANOVA F (1, 122) = 0.95, p =
616 0.33, n ≥ 60). (D) Comparison of seam cell scorings of *egl-18(ga97)* mutants between
617 N2 and CB4856 when seam cells are classified based on their position (anterior, mid,
618 posterior). Differences in SCN in all three categories are significant between the two
619 isolates (F (1, 310) = 66.41, 4.87, and 61.78, * $p < 0.03$ and *** $p < 0.0001$, n ≥ 150).
620 (E) Representative smFISH image showing *elt-6* expression in posterior seam cells in
621 an *egl-18* mutant at the late L2 stage. Seam cell nuclei are labelled in green due to
622 *scm::GFP* expression and black spots correspond to *elt-6* mRNAs. Scale bar is 20 µm.
623 (F) Quantification of *elt-6* expression by smFISH in N2 and CB4856 strains carrying
624 the *egl-18(ga97)* mutation. No statistically significant difference in the mRNA was
625 found with one-way ANOVA (F (1, 348) = 0.13, $p = 0.72$). Values within bars indicate

626 number of cells analysed. (G) Seam cell number in N2 and CB4856 strains upon
627 knockdown of *lit-1*, *pop-1* and *apr-1*. There is a significant effect of strain on SCN upon
628 knock-down of *lit-1, pop-1* and *apr-1* RNAi, but not in empty-vector control, using one-
629 way ANOVA (F (1, 78) = 32.71, 30.27, 7.3 and 1.37, *** $p < 0.0001$ or ** $p < 0.001$, n
630 = 40). Error bars indicate 95 % confidence intervals.
631
632 **Figure 2: Generation and phenotypic analysis of recombinant inbred lines**
633 (A) Generation of RILs between CB4856 and N2 carrying the *egl-18(ga97)* mutation.
634 Two chromosomes (III and IV) are shown for simplification. (B) Seam cell number
635 quantification in 116 RILs, ranked from lowest to highest average seam cell number.
636 Parental strains are shown in green (N2) and orange (CB4856). The two extremes
637 (low-bulk and high-bulk) of the phenotypic distribution were pooled for whole genome
638 sequencing. Error bars indicate average SCN ± 95 % confidence intervals, n ≥ 77
639 animals per RIL.
640
641 **Figure 3: Bulk segregant analysis of recombinant inbred lines.**
642 (A-F) Proportion of CB4856 SNPs in the sequencing reads of the low-bulk and high-
643 bulk groups is shown along the six chromosomes, from I (A) to X (F). SNP frequencies
644 in low-bulk and high-bulk are shown in green and orange fitted curves, respectively.
645 The curves represent locally weighted scatterplot smoothing (LOESS) regression lines
646 from the allele frequencies at known SNP positions along the chromosomes with a
647 span parameter of 0.1. Log-odds ratio of average SNP frequencies in non-overlapping
648 300 kb windows along the six chromosomes is also shown. The blue dashed lines
649 indicate the thresholds for statistical significance for log-odd ratios at $\alpha = 0.05$. The
650 positions of the *peel-1/zeel-1* element, *egl-18(ga97)* mutation and *scm::GFP*

651 transgene integrations are shown in grey.

652

653 **Figure 4: Phenotypic analysis of near isogenic lines identifies major QTLs.**

654 (A) Seam cell number in near isogenic lines containing individual and combinations of
655 the identified QTLs from CB4856 into the *egl-18(ga97)* N2 background. One-way
656 ANOVA showed that SCN was significantly affected by the strain (F (8, 905) = 16.55,
657 $p < 2.2 \times 10^{-16}$, $n \geq 100$). Note that QTLs on chromosomes II and III, but not on the X,
658 were sufficient to convert the *egl-18(ga97)* mutant expressivity in the N2 background
659 to CB4856 levels. Chromosomes below the graph depict the genotype of the strain.
660 (B) SCN counts in near isogenic lines containing individual QTLs in a wild-type
661 background. One-way ANOVA shows no statistically significant differences in SCN of
662 NILs compared to wild-type with no QTLs (F (3, 396) = 1.34, p = 0.26, n = 100). (C)
663 *pop-1* RNAi in near isogenic lines in a wild-type background containing individual and
664 combinations of the identified QTLs. The same colour code is used as in panel A to
665 define different combinations of QTLs. A significant effect of strain was found on SCN
666 upon *pop-1* RNAi (F (8, 351) = 17.85, p = < 2.2×10^{-16}, n = 40) using one-way ANOVA.
667 Error bars indicate 95% confidence intervals and *** $p < 0.0001$ or * $p < 0.05$,
668 correspond to significant differences by post hoc Dunnett's multiple comparison test
669 when compared N2 (green stars) and CB4856 (orange stars) respectively upon *pop-1*
670 RNAi.

671

672 **Figure 5: Genotype-to-phenotype analysis of NILs carrying genomic fragments**
673 **of chromosome III from CB4856.**

674 (A) Seam cell number in near isogenic lines containing various fragments of
675 chromosome III from CB4856 in the *egl-18(ga97)* N2 background. Vertical black line

676 in the graph distinguishes between two independent experiments. In the both
677 experiments, one-way ANOVA showed that SCN was significantly affected by the
678 strain (F (5, 556) = 26.58, $p < 2.2 \times 10^{-16}$; F (2, 313) = 43.31, $p < 2.2 \times 10^{-16}$
679 respectively, $n \geq 90$). Error bars indicate 95 % confidence intervals and *** $p < 0.0001$
680 corresponds to significant differences by post hoc Tukey HSD compared to N2 (green)
681 and CB4856 (orange) respectively. (B) Genotyping of NILs carrying genomic
682 fragments of chromosome III from CB4856 using genetic markers. Green and orange
683 tiles represent N2 and CB4856 genomes, respectively. Highlighted black-box in the
684 middle indicates the genomic region that is common to the NILs that convert N2-mutant
685 expressivity to CB4856-mutant levels.
686
687 **Figure 6: Natural variation in *hsp-110* contributes to the difference in mutation**
688 **expressivity between N2 and CB4856.**
689 (A) Alignment of the HSP-110 amino acid sequence from *C. elegans*, *C. remanei*, *C.*
690 *brenneri*, *P. pacificus* and *H. sapiens* around the amino acid deletion in CB4856
691 (highlighted in red). (B-C) Comparisons of seam cell counts between animals carrying
692 the CB4856 variants in *hsp-110* and *sor-1* in an N2 background in *egl-18(ga97)*
693 mutants (B) and wild-type (C). Comparison is shown to parental isolates with (B) or
694 without (C) the *egl-18(ga97)* mutation, as well as the introgression of the QTL on
695 chromosome III as a reference for the rescue. Error bars indicate 95% confidence
696 intervals and *** $p < 0.0001$ corresponds to significant differences by one-way ANOVA
697 followed by post hoc Tukey HSD compared to N2 (green stars) and CB4856 (orange
698 stars) respectively. (D-E) Representative smFISH images showing *hsp-110* expression
699 in divided seam cells in N2 or CB4856 at the late L2 stage. Seam cell nuclei are labelled
700 in green due to *scm::GFP* expression and black spots correspond to *hsp-110* mRNAs.

701 (F) Developmental model explaining the difference in *egl-18(ga97)* mutation
702 expressivity between N2 and CB4856. In Wnt-signalled cells, POP-1 acts as a
703 transcriptional activator leading to adoption of seam cell fate through the expression of
704 transcription factors, such as EGL-18 and ELT-6. EGL-18 is a major seam cell
705 promoting factor, so loss of seam cell fate occurs in the *egl-18(ga97)* mutant in both
706 isolates due to differentiation of seam cells. However, in CB4856, seam cell fate is
707 better retained in the absence of EGL-18 because of genetic variation that acts to
708 reinforce seam cell fate acquisition via potentiating the effect of Wnt signalling through
709 ELT-6 and/or other yet unknown factors.
710
711 **Supplemental Figure Legends**
712
713 **Figure S1: Differential mutation expressivity in the seam is specific to *egl-18* and**
714 **requires strong loss-of-*egl-18* function and Wnt signaling.**
715 (A) The differential expressivity of *egl-18(ga97)* between N2 and CB4856 is less
716 pronounced in the weaker allele *egl-18(ok290)*. There was no statistically significant
717 difference in SCN between CB4856 and N2 (p = 0.06), but there was a significant
718 difference in SCN between CB4856 and JU2007 carrying the *ok290* mutation (p =
719 0.0012), n ≥ 28 per strain. (B-C) No difference was found in the expressivity of *bro-1*
720 or *eff-1* mutations between wild isolates. There is no statistically significant difference
721 in SCN between the strains MBA79 (N2) and MBA368 (CB4856) carrying the same
722 putative null allele *icb4* in *eff-1, n* ≥ 37 per strain. Similarly, there is no statistically
723 significant difference in SCN between the strains MBA274 (CB4856) and MBA243
724 (JU2007) carrying the same deletion allele *icb45* of *bro-1*, n = 60 per strain. (D) GFP
725 RNAi is used to assess the efficacy of seam cell RNAi in CB4856 and N2 based on

726 *scm::GFP* expression. A significant difference is revealed in 10% dilutions of the GFP
727 dsRNA bacteria with control bacteria (HT115), but not in 1% dilution of GFP bacteria
728 or when applied non-diluted (100% GFP), n = 40 per strain. (E) Quantification of
729 corrected total fluorescence (CTCF) in N2 and CB4856 wild-type animals carrying
730 the *POPHHOP* reporter. No consistent difference was found among different
731 experimental trials, n ≥ 24 per strain (F) Knockdown of *elt-6* or *pop-1* by RNAi in *egl-*
732 *18(ga97)* animals masks the differential expressivity observed between wild isolates.
733 One-way ANOVA showed that SCN was not affected by strain upon *elt-6* RNAi (F (2,
734 87) = 1.52, p = 0.22), n = 30 per strain or *pop-1* RNAi (F (2, 77) = 0.86, p = 0.43), n ≥
735 20 per strain (G) No significant difference was found in the vulval induction index
736 between N2 and CB4856 carrying the *egl-18(ga97)* mutation, (n > 31 per strain, t-test).
737 (H) The differential expressivity of the *egl-18(ga97)* mutation is not dependent on a
738 previously known polymorphism in *nath-10*. There is a significant effect of strain on
739 SCN (one-way ANOVA F (4, 145) = 77.19, $p < 2.2 \times 10^{-16}$). Post hoc Tukey HSD shows
740 that there is a significant difference between *nath-10(haw6805); egl-18(ga97)* between
741 N2 and CB4856 (p = 0). There was no difference between *nath-10(haw6805); egl-*
742 *18(ga97)* and *nath-10(N2); egl-18(ga97)* (p = 0.98). n = 30 per strain. In all panels,
743 error bars indicate 95 % confidence intervals and *** represent p value < 0.0001, ** p
744 < 0.001 or * p < 0.05 by post hoc Tukey HSD test.
745
746 **Figure S2: Genotyping of the individual RILs included in the low- and high-bulks**
747 **using genetic markers.**
748 Each row represents an individual RIL line and its column a genetic marker on four
749 different chromosomes to infer genetic composition per marker and per RIL. The
750 location of the genetic markers on the chromosomes is shown above the graph.

751 Markers on chromosome II, III, V and X are shown in blue, pink, yellow and green,
752 respectively. Green and orange tiles represent N2 and CB4856 genomes respectively.
753 RILs are ranked in descending order based on number of CB4856 genomic fragments
754 they carry. RIL numbers are colour-coded based on whether they were included in the
755 low-bulk (N2-like, green) or high-bulk (CB4856-like, orange). RIL-28, which was used
756 for producing near isogenic lines containing QTLs on chromosomes II, III and X is
757 highlighted in the middle.
758
759 **Figure S3: Genotype-to-phenotype analysis of NILs carrying genomic fragments**
760 **of chromosome II from CB4856.**
761 (A) Seam cell number in near isogenic lines containing various fragments of
762 chromosome II from CB4856 in the *egl-18(ga97)* N2 background. Vertical black line in
763 the graph distinguishes between two independent experiments. In both experiments,
764 one-way ANOVA showed that SCN was significantly affected by the strain (F (6, 635)
765 = 12.25, $p < 4.5 \times 10^{-13}$; F (5, 622) = 51.18, $p < 2.2 \times 10^{-16}$, $n \geq 90$ animals per strain.
766 Error bars indicate average 95 % confidence intervals and *** $p < 0.0001$ or ** $p <$
767 0.001 correspond to significant differences by post hoc Tukey HSD compared to N2
768 (green stars) and CB4856 (orange stars) respectively. (B) Genotyping of NILs carrying
769 various genomic fragments of chromosome II from CB4856. Each row represents a
770 RIL line and its column its genetic composition at specific genetic markers, the location
771 of which is shown above the graph. Green and orange tiles represent N2 and CB4856
772 genomes, respectively. Black-box in the middle highlights the genomic region that is
773 common to the NILs that convert N2-like to CB4856-like *egl-18(ga97)* phenotype and
774 absent in the NILs that do not. (C) Close-up on chromosome II showing the region that
775 is common between the NILs that convert N2 *egl-18(ga97)* mutant expressivity to

776 CB4856 levels (left). A close up of the region on Chr III, which includes *sor-1* and *hsp-*
777 *110* is shown for comparison (right). In both cases, stars indicate inferred genotype for
778 that marker.
779
780 **Figure S4: RNAi screen to identify candidate genes in the extended QTL regions**
781 **of chromosomes II and III.**
782 (A-B) Seam cell number in mutant N2 and MBA848 (NIL carrying the QTLs on
783 chromosome II, III and X) upon RNAi knockdown of candidate genes. (A) Knockdown
784 of most genes on chromosome II did not have a significant effect on the SCN in mutant
785 in N2 and MBA848, except for two genes (*egl-27* and *dsh-2*), which both had a
786 significant effect. (B) Knockdown of two genes (*sor-1* and *hsp-110*) had a significant
787 effect on the SCN in N2 and MBA848 ($F(5, 234) = 10.48$, $p = 8.37 \times 10^{-9}$; $F(5, 234) =$
788 19.65, $p = 2.47 \times 10^{-16}$, respectively). SCN decreased upon knockdown of *sor-1* and
789 *hsp-110* in both N2 and MBA848 ($p < 0.001$). Knockdown of *lin-9* and *lin-36* showed a
790 significant decrease in SCN only in MBA848. There was a decrease in SCN upon
791 knockdown of *fbn-1* and *kle-2* in both N2 and MBA848. However, these genes were
792 outside the final QTL interval. Error bars indicate standard deviation, n = 40 per strain
793 and *** $p < 0.0001$, ** $p < 0.001$, * $p < 0.05$ correspond to significant differences
794 compared to control RNAi by ANOVA and post hoc Dunnett's multiple comparison test.
795 In both A and B, dashed lines mark independent experiments therefore control
796 treatments are repeated.
797
798 **Supplementary Tables**
799 **Table S1. Strains used in this study.**
800 **Table S2. Primers used in this study (general oligos, genotyping and smFISH).**

801 **Table S3. Genetic variants within QTL regions on chromosome II and III**

802

803 **Acknowledgments**

804 We thank Mingke Pan for technical help with genotyping. Some *C. elegans*
805 strains were provided by the CGC, which is funded by NIH Office of Research
806 Infrastructure Programs (P40 OD010440). This work was funded by the by the
807 European Research Council (ROBUSTNET).

808

809 **References:**

810 Andersen, E.C., Gerke, J.P., Shapiro, J.A., Crissman, J.R., Ghosh, R., Bloom, J.S.,
811 Felix, M.A., and Kruglyak, L. (2012). Chromosome-scale selective sweeps shape
812 *Caenorhabditis elegans* genomic diversity. Nat Genet *44*, 285-290.
813
814 Arribere, J.A., Bell, R.T., Fu, B.X., Artiles, K.L., Hartman, P.S., and Fire, A.Z. (2014).
815 Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in
816 *Caenorhabditis elegans*. Genetics *198*, 837-846.
817
818 Baldwin, A.T., Clemons, A.M., and Phillips, B.T. (2016). Unique and redundant beta-
819 catenin regulatory roles of two Dishevelled paralogs during *C. elegans* asymmetric cell
820 division. J Cell Sci *129*, 983-993.
821
822 Bhambhani, C., Ravindranath, A.J., Mentink, R.A., Chang, M.V., Betist, M.C., Yang,
823 Y.X., Koushika, S.P., Korswagen, H.C., and Cadigan, K.M. (2014). Distinct DNA
824 binding sites contribute to the TCF transcriptional switch in *C. elegans* and *Drosophila*.
825 PLoS Genet *10*, e1004133.
826
827 Brambila-Tapia, A.J.L., Garcia-Ortiz, J.E., Brouillard, P., Nguyen, H.L., Vikkula, M.,
828 Rios-Gonzalez, B.E., Sandoval-Muniz, R.J., Sandoval-Talamantes, A.K., Bobadilla-
829 Morales, L., Corona-Rivera, J.R.*, et al.* (2017). GATA2 null mutation associated with
830 incomplete penetrance in a family with Emberger syndrome. Hematology *22*, 467-471.
831
832 Burga, A., Casanueva, M.O., and Lehner, B. (2011). Predicting mutation outcome from
833 early stochastic variation in genetic interaction partners. Nature *480*, 250-253.
834
835 Casanueva, M.O., Burga, A., and Lehner, B. (2012). Fitness trade-offs and
836 environmentally induced mutation buffering in isogenic *C. elegans*. Science *335*, 82-
837 85.
838
839 Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background
840 check? How genetic background impacts the analysis of mutations, genes, and
841 evolution. Trends Genet *29*, 358-366.

842
843 Chandler, C.H., Chari, S., Kowalski, A., Choi, L., Tack, D., DeNieu, M., Pitchers, W.,
844 Sonnenschein, A., Marvin, L., Hummel, K.*, et al.* (2017). How well do you know your
845 mutation? Complex effects of genetic background on expressivity, complementation,
846 and ordering of allelic effects. PLoS Genet *13*, e1007075.
847
848 Chow, C.Y., Kelsey, K.J., Wolfner, M.F., and Clark, A.G. (2016). Candidate genetic
849 modifiers of retinitis pigmentosa identified by exploiting natural variation in Drosophila.
850 Hum Mol Genet *25*, 651-659.
851
852 Collura, A., Lagrange, A., Svrcek, M., Marisa, L., Buhard, O., Guilloux, A.,
853 Wanherdrick, K., Dorard, C., Taieb, A., Saget, A.*, et al.* (2014). Patients with colorectal
854 tumors with microsatellite instability and large deletions in HSP110 T17 have improved
855 response to 5-fluorouracil-based chemotherapy. Gastroenterology *146*, 401-411 e401.
856
857 Cook, D.E., Zdraljevic, S., Roberts, J.P., and Andersen, E.C. (2017). CeNDR, the
858 *Caenorhabditis elegans* natural diversity resource. Nucleic Acids Res *45*, D650-D657.
859
860 Crombie, T.A., Zdraljevic, S., Cook, D.E., Tanny, R.E., Brady, S.C., Wang, Y., Evans,
861 K.S., Hahnel, S., Lee, D., Rodriguez, B.C.*, et al.* (2019). Deep sampling of Hawaiian
862 *Caenorhabditis elegans* reveals high genetic diversity and admixture with global
863 populations. Elife *8*, eLife 2019;8:e50465
864
865 Dowell, R.D., Ryan, O., Jansen, A., Cheung, D., Agarwala, S., Danford, T., Bernstein,
866 D.A., Rolfe, P.A., Heisler, L.E., Chin, B.*, et al.* (2010). Genotype to phenotype: a
867 complex problem. Science *328*, 469.
868
869 Dragovic, Z., Broadley, S.A., Shomura, Y., Bracher, A., and Hartl, F.U. (2006).
870 Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of
871 Hsp70s. EMBO J *25*, 2519-2528.
872
873 Duveau, F., and Felix, M.A. (2012). Role of pleiotropy in the evolution of a cryptic
874 developmental variation in *Caenorhabditis elegans*. PLoS Biol *10*, e1001230.
875
876 Dworkin, I., Kennerly, E., Tack, D., Hutchinson, J., Brown, J., Mahaffey, J., and Gibson,
877 G. (2009). Genomic consequences of background effects on scalloped mutant
878 expressivity in the wing of *Drosophila melanogaster*. Genetics *181*, 1065-1076.
879
880 Easton, D.P., Kaneko, Y., and Subjeck, J.R. (2000). The hsp110 and Grp1 70 stress
881 proteins: newly recognized relatives of the Hsp70s. Cell Stress Chaperones *5*, 276-
882 290.
883
884 Eisenmann, D.M., and Kim, S.K. (2000). Protruding vulva mutants identify novel loci
885 and Wnt signaling factors that function during *Caenorhabditis elegans* vulva
886 development. Genetics *156*, 1097-1116.
887
888 Felix, M.A., and Barkoulas, M. (2015). Pervasive robustness in biological systems. Nat
889 Rev Genet *16*, 483-496.
890
891 Felix, M.A., and Wagner, A. (2008). Robustness and evolution: concepts, insights and

892 challenges from a developmental model system. Heredity (Edinb) *100*, 132-140.
893
894 Fournier, T., and Schacherer, J. (2017). Genetic backgrounds and hidden trait
895 complexity in natural populations. Curr Opin Genet Dev *47*, 48-53.
896
897 Frezal, L., Demoinet, E., Braendle, C., Miska, E., and Felix, M.A. (2018). Natural
898 Genetic Variation in a Multigenerational Phenotype in *C. elegans*. Curr Biol *28*, 2588-
899 2596 e2588.
900
901 Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco,
902 J.A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat
903 Methods *10*, 741-743.
904
905 Gasch, A.P., Payseur, B.A., and Pool, J.E. (2016). The Power of Natural Variation for
906 Model Organism Biology. Trends Genet *32*, 147-154.
907
908 Gibson, G., and Dworkin, I. (2004). Uncovering cryptic genetic variation. Nat Rev
909 Genet *5*, 681-690.
910
911 Gleason, J.E., and Eisenmann, D.M. (2010). Wnt signaling controls the stem cell-like
912 asymmetric division of the epithelial seam cells during *C. elegans* larval development.
913 Dev Biol *348*, 58-66.
914
915 Gorrepati, L., Thompson, K.W., and Eisenmann, D.M. (2013). C. elegans GATA factors
916 EGL-18 and ELT-6 function downstream of Wnt signaling to maintain the progenitor
917 fate during larval asymmetric divisions of the seam cells. Development *140*, 2093-
918 2102.
919
920 Herman, M.A., Ch'ng, Q., Hettenbach, S.M., Ratliff, T.M., Kenyon, C., and Herman,
921 R.K. (1999). EGL-27 is similar to a metastasis-associated factor and controls cell
922 polarity and cell migration in *C. elegans*. Development *126*, 1055-1064.
923
924 Hintze, M., Koneru, S.L., Gilbert, S.P.R., Katsanos, D., Lambert, J., and Barkoulas, M.
925 (2020). A Cell Fate Switch in the *Caenorhabditis elegans* Seam Cell Lineage Occurs
926 Through Modulation of the Wnt Asymmetry Pathway in Response to Temperature
927 Increase. Genetics *214*, 927-939.
928
929 Hughes, S., Vrinds, I., de Roo, J., Francke, C., Shimeld, S.M., Woollard, A., and Sato,
930 A. (2019). DnaJ chaperones contribute to canalization. J Exp Zool A Ecol Integr Physiol
931 *331*, 201-212.
932
933 Jones, L.M., Eves-van den Akker, S., van-Oosten Hawle, P., Atkinson, H.J., and Urwin,
934 P.E. (2018). Duplication of *hsp-110* Is Implicated in Differential Success of *Globodera*
935 Species under Climate Change. Mol Biol Evol *35*, 2401-2413.
936
937 Joshi, P.M., Riddle, M.R., Djabrayan, N.J., and Rothman, J.H. (2010). *Caenorhabditis*
938 *elegans* as a model for stem cell biology. Dev Dyn *239*, 1539-1554.
939
940 Karras, G.I., Yi, S., Sahni, N., Fischer, M., Xie, J., Vidal, M., D'Andrea, A.D., Whitesell,
941 L., and Lindquist, S. (2017). HSP90 Shapes the Consequences of Human Genetic

942 Variation. Cell *168*, 856-866 e812.
943
944 Katsanos, D., Koneru, S.L., Mestek Boukhibar, L., Gritti, N., Ghose, R., Appleford, P.J.,
945 Doitsidou, M., Woollard, A., van Zon, J.S., Poole, R.J.*, et al.* (2017). Stochastic loss
946 and gain of symmetric divisions in the *C. elegans* epidermis perturbs robustness of
947 stem cell number. PLoS Biol *15*, e2002429.
948
949 Koh, K., Peyrot, S.M., Wood, C.G., Wagmaister, J.A., Maduro, M.F., Eisenmann, D.M.,
950 and Rothman, J.H. (2002). Cell fates and fusion in the *C. elegans* vulval primordium
951 are regulated by the EGL-18 and ELT-6 GATA factors – apparent direct targets of the
952 LIN-39 Hox protein. Development *129*, 5171-5180.
953
954 Koh, K., and Rothman, J.H. (2001). ELT-5 and ELT-6 are required continuously to
955 regulate epidermal seam cell differentiation and cell fusion in *C. elegans*. Development
956 *128*, 2867-2880.
957
958 Lehner, B., Crombie, C., Tischler, J., Fortunato, A., and Fraser, A.G. (2006).
959 Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies
960 common modifiers of diverse signaling pathways. Nat Genet *38*, 896-903.
961
962 Mansfeld, B.N., and Grumet, R. (2018). QTLseqr: An R Package for Bulk Segregant
963 Analysis with Next-Generation Sequencing. Plant Genome *11*
964 *(2):10.3835/plantgenome2018.01.0006*
965
966 Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and
967 molecular mechanism. Cell Mol Life Sci *62*, 670-684.
968
969 Michelmore, R.W., Paran, I., and Kesseli, R.V. (1991). Identification of markers linked
970 to disease-resistance genes by bulked segregant analysis: a rapid method to detect
971 markers in specific genomic regions by using segregating populations. Proc Natl Acad
972 Sci U S A *88*, 9828-9832.
973
974 Minevich, G., Park, D.S., Blankenberg, D., Poole, R.J., and Hobert, O. (2012).
975 CloudMap: a cloud-based pipeline for analysis of mutant genome sequences. Genetics
976 *192*, 1249-1269.
977
978 Nikolaidis, N., and Nei, M. (2004). Concerted and nonconcerted evolution of the Hsp70
979 gene superfamily in two sibling species of nematodes. Mol Biol Evol *21*, 498-505.
980
981 Paaby, A.B., and Rockman, M.V. (2014). Cryptic genetic variation: evolution's hidden
982 substrate. Nat Rev Genet *15*, 247-258.
983
984 Paaby, A.B., White, A.G., Riccardi, D.D., Gunsalus, K.C., Piano, F., and Rockman,
985 M.V. (2015). Wild worm embryogenesis harbors ubiquitous polygenic modifier
986 variation. Elife *4*, eLife 2015;4:e09178
987
988 Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High Efficiency,
989 Homology-Directed Genome Editing in *Caenorhabditis elegans* Using CRISPR-Cas9
990 Ribonucleoprotein Complexes. Genetics *201*, 47-54.
991

Pollard, D.A., and Rockman, M.V. (2013). Resistance to germline RNA interference in a *Caenorhabditis elegans* wild isolate exhibits complexity and nonadditivity. G3 (Bethesda) *3*, 941-947.

Rampelt, H., Kirstein-Miles, J., Nillegoda, N.B., Chi, K., Scholz, S.R., Morimoto, R.I., and Bukau, B. (2012). Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. EMBO J *31*, 4221-4235.

Raviol, H., Sadlish, H., Rodriguez, F., Mayer, M.P., and Bukau, B. (2006). Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. EMBO J *25*, 2510-2518.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M., and Stainier, D.Y. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature *524*, 230-233.

Scior, A., Buntru, A., Arnsburg, K., Ast, A., Iburg, M., Juenemann, K., Pigazzini, M.L., Mlody, B., Puchkov, D., Priller, J.*, et al.* (2018). Complete suppression of Htt fibrilization and disaggregation of Htt fibrils by a trimeric chaperone complex. EMBO J *37*, 282-299.

Seidel, H.S., Ailion, M., Li, J., van Oudenaarden, A., Rockman, M.V., and Kruglyak, L. (2011). A novel sperm-delivered toxin causes late-stage embryo lethality and transmission ratio distortion in *C. elegans*. PLoS Biol *9*, e1001115.

Seidel, H.S., Rockman, M.V., and Kruglyak, L. (2008). Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. Science *319*, 589-594.

Serobyan, V., Kontarakis, Z., El-Brolosy, M.A., Welker, J.M., Tolstenkov, O., Saadeldein, A.M., Retzer, N., Gottschalk, A., Wehman, A.M., and Stainier, D.Y. (2020). Transcriptional adaptation in *Caenorhabditis elegans*. Elife *9*, eLife 2020;9:e50014

Sittig, L.J., Carbonetto, P., Engel, K.A., Krauss, K.S., Barrios-Camacho, C.M., and Palmer, A.A. (2016). Genetic Background Limits Generalizability of Genotype-Phenotype Relationships. Neuron *91*, 1253-1259.

Solari, F., Bateman, A., and Ahringer, J. (1999). The Caenorhabditis elegans genes *egl-27* and *egr-1* are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. Development *126*, 2483-2494.

Sterken, M.G., Snoek, L.B., Kammenga, J.E., and Andersen, E.C. (2015). The laboratory domestication of *Caenorhabditis elegans*. Trends Genet *31*, 224-231.

Stiernagle, T. (2006). Maintenance of *C. elegans*. WormBook, 1-11.

Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Dev Biol *56*, 110-156.

Thompson, O.A., Snoek, L.B., Nijveen, H., Sterken, M.G., Volkers, R.J., Brenchley, R., Van't Hof, A., Bevers, R.P., Cossins, A.R., Yanai, I.*, et al.* (2015). Remarkably

1042 Divergent Regions Punctuate the Genome Assembly of the *Caenorhabditis elegans*
1043 Hawaiian Strain CB4856. Genetics *200*, 975-989.
1044
1045 Tijsterman, M., Okihara, K.L., Thijssen, K., and Plasterk, R.H. (2002). PPW-1, a
1046 PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate
1047 of *C. elegans*. Curr Biol *12*, 1535-1540.
1048
1049 Torres Cleuren, Y.N., Ewe, C.K., Chipman, K.C., Mears, E.R., Wood, C.G., Al-Alami,
1050 C.E.A., Alcorn, M.R., Turner, T.L., Joshi, P.M., Snell, R.G.*, et al.* (2019). Extensive
1051 intraspecies cryptic variation in an ancient embryonic gene regulatory network. Elife *8*,
1052 eLife 2019;8:e48220
1053
1054 Tremblay, M., Sanchez-Ferras, O., and Bouchard, M. (2018). GATA transcription
1055 factors in development and disease. Development *145 : dev164384*
1056
1057 Vu, V., Verster, A.J., Schertzberg, M., Chuluunbaatar, T., Spensley, M., Pajkic, D.,
1058 Hart, G.T., Moffat, J., and Fraser, A.G. (2015). Natural Variation in Gene Expression
1059 Modulates the Severity of Mutant Phenotypes. Cell *162*, 391-402.
1060
1061 Wicks, S.R., Yeh, R.T., Gish, W.R., Waterston, R.H., and Plasterk, R.H. (2001). Rapid
1062 gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. Nat
1063 Genet *28*, 160-164.
1064
1065 Zabinsky, R.A., Mason, G.A., Queitsch, C., and Jarosz, D.F. (2019). It's not magic -
1066 Hsp90 and its effects on genetic and epigenetic variation. Semin Cell Dev Biol *88*, 21-
1067 35.
1068
1069 Zhang, T., Sun, Y., Tian, E., Deng, H., Zhang, Y., Luo, X., Cai, Q., Wang, H., Chai, J.,
1070 and Zhang, H. (2006). RNA-binding proteins SOP-2 and SOR-1 form a novel PcG-like
1071 complex in *C. elegans*. Development *133*, 1023-1033.
1072
1073 Zhang, W., Xue, D., Yin, H., Wang, S., Li, C., Chen, E., Hu, D., Tao, Y., Yu, J., Zheng,
1074 Q.*, et al.* (2016). Overexpression of HSPA1A enhances the osteogenic differentiation
1075 of bone marrow mesenchymal stem cells via activation of the Wnt/beta-catenin
1076 signaling pathway. Sci Rep *6*, 27622.
1077

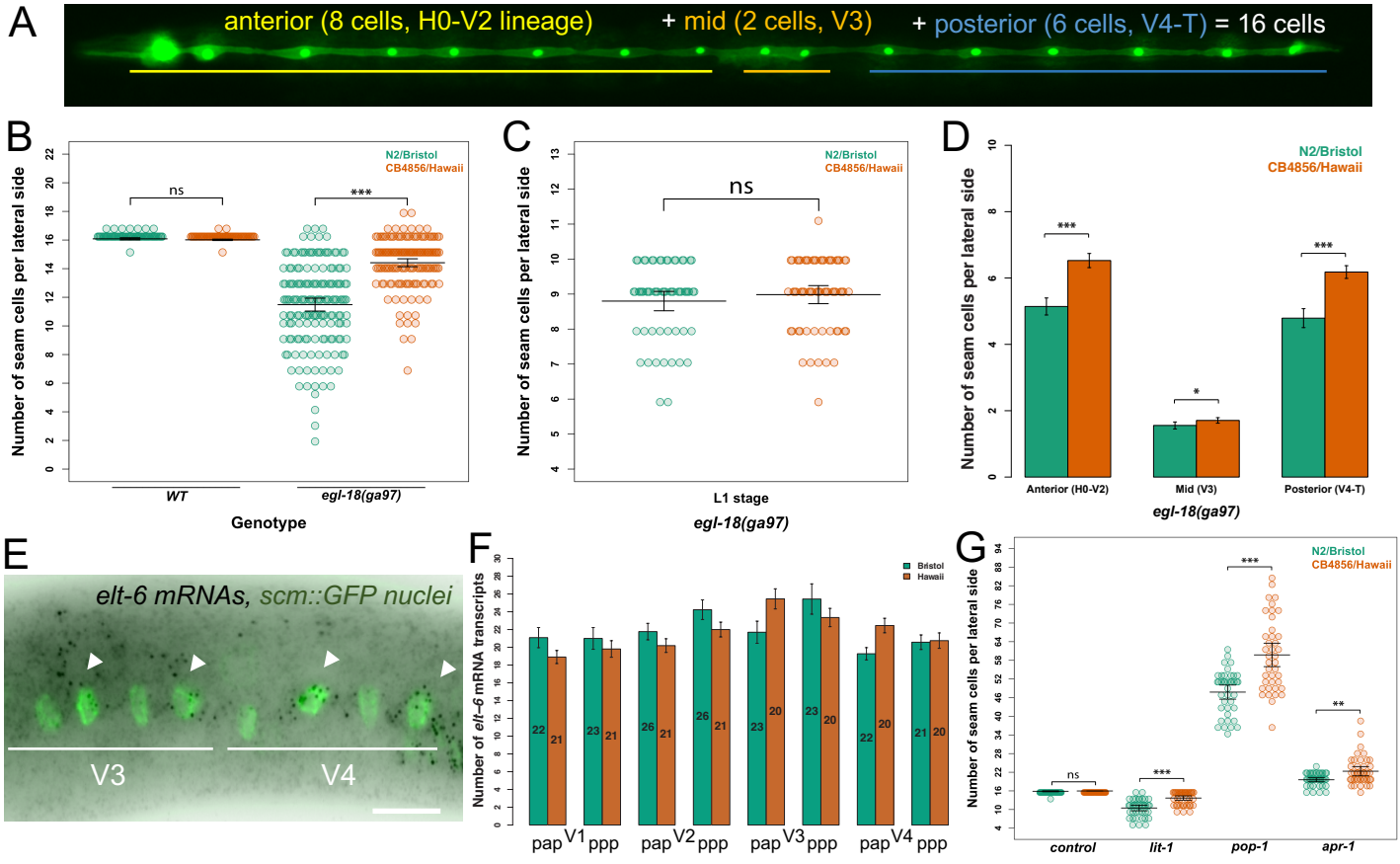

A
anterior (8 cells, H0-V2 lineage)
+ mid (2 cells, V3)
+ posterior (6 cells, V4-T) = 16 cells
B
N2/Bristol
CB4856/Hawaii
ns
***
Number of seam cells per lateral side
WT
egl-18(ga97)
Genotype
C
N2/Bristol
CB4856/Hawaii
ns
Number of seam cells per lateral side
L1 stage
egl-18(ga97)
D
N2/Bristol
CB4856/Hawaii
***
*
***
Number of seam cells per lateral side
Anterior (H0-V2)
Mid (V3)
Posterior (V4-T)
egl-18(ga97)
E
elt-6 mRNAs, scm::GFP nuclei
V3
V4
F
Bristol
Hawaii
Number of elt-6 mRNA transcripts
pap V1 ppp
pap V2 ppp
pap V3 ppp
pap V4 ppp
G
N2/Bristol
CB4856/Hawaii
ns
***
***
**
Number of seam cells per lateral side
control
lit-1
pop-1
apr-1

Figure 1

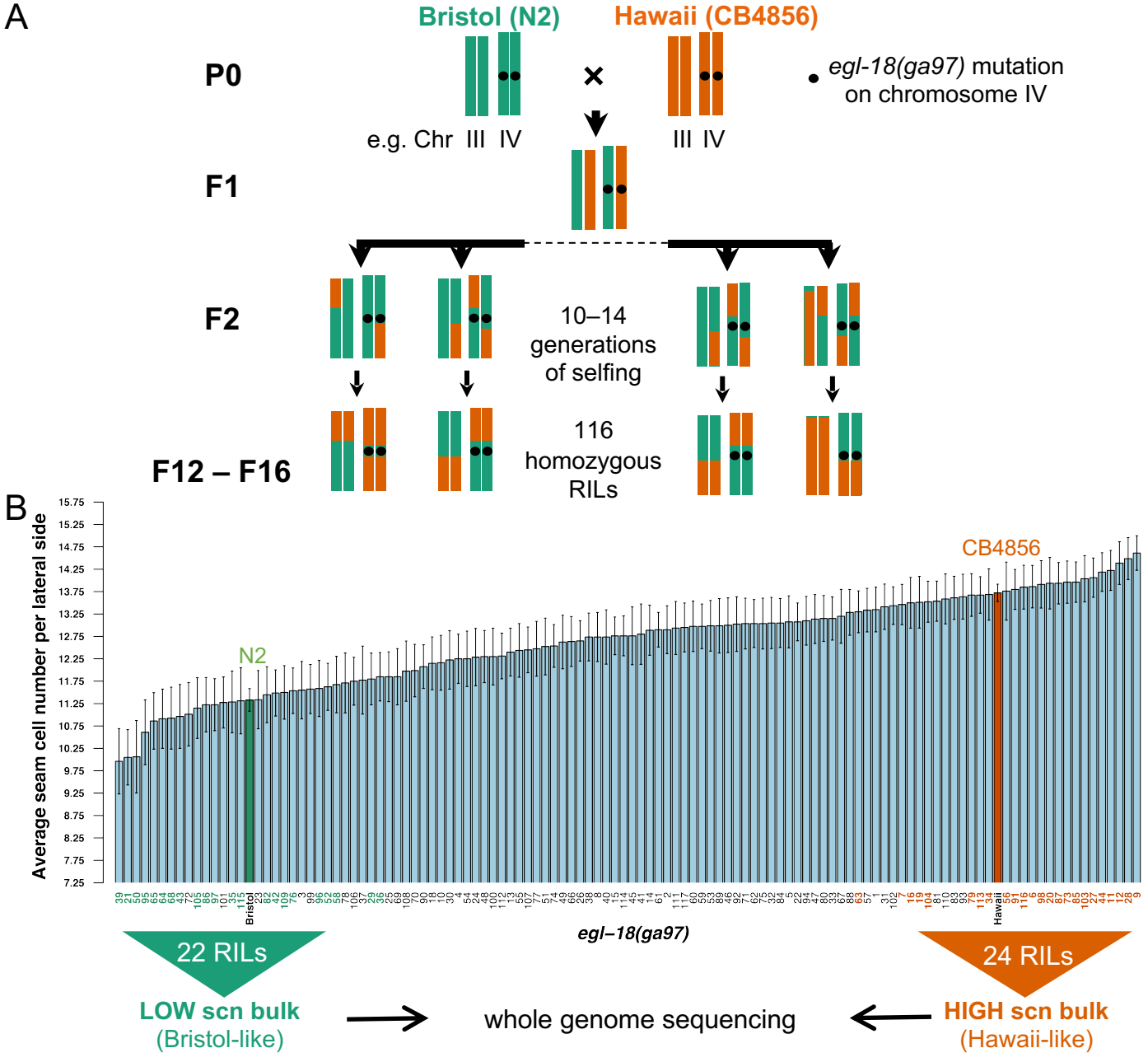

A
Bristol (N2)
Hawaii (CB4856)
P0
egl-18(ga97) mutation on chromosome IV
e.g. Chr III IV
III IV
F1
F2
10–14 generations of selfing
F12 – F16
116 homozygous RILs
B
Average seam cell number per lateral side
N2
CB4856
Bristol
Hawaii
egl–18(ga97)
22 RILs
24 RILs
LOW scn bulk (Bristol-like)
whole genome sequencing
HIGH scn bulk (Hawaii-like)

Figure 2

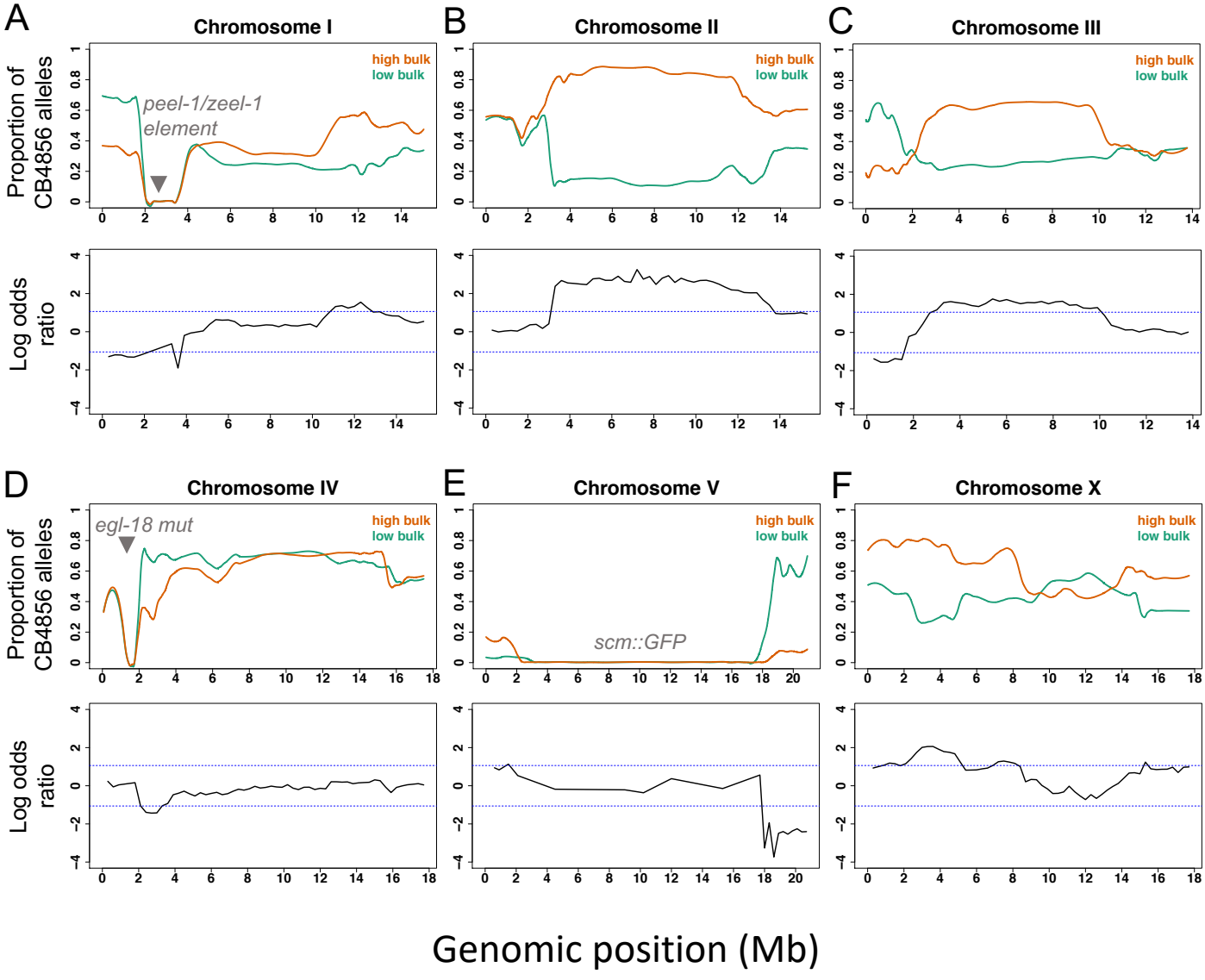

A
Chromosome I
B
Chromosome II
C
Chromosome III
D
Chromosome IV
E
Chromosome V
F
Chromosome X
high bulk
low bulk
peel-1/zeel-1
element
egl-18 mut
scm::GFP
Proportion of
CB4856 alleles
Log odds
ratio
Genomic position (Mb)

Figure 3

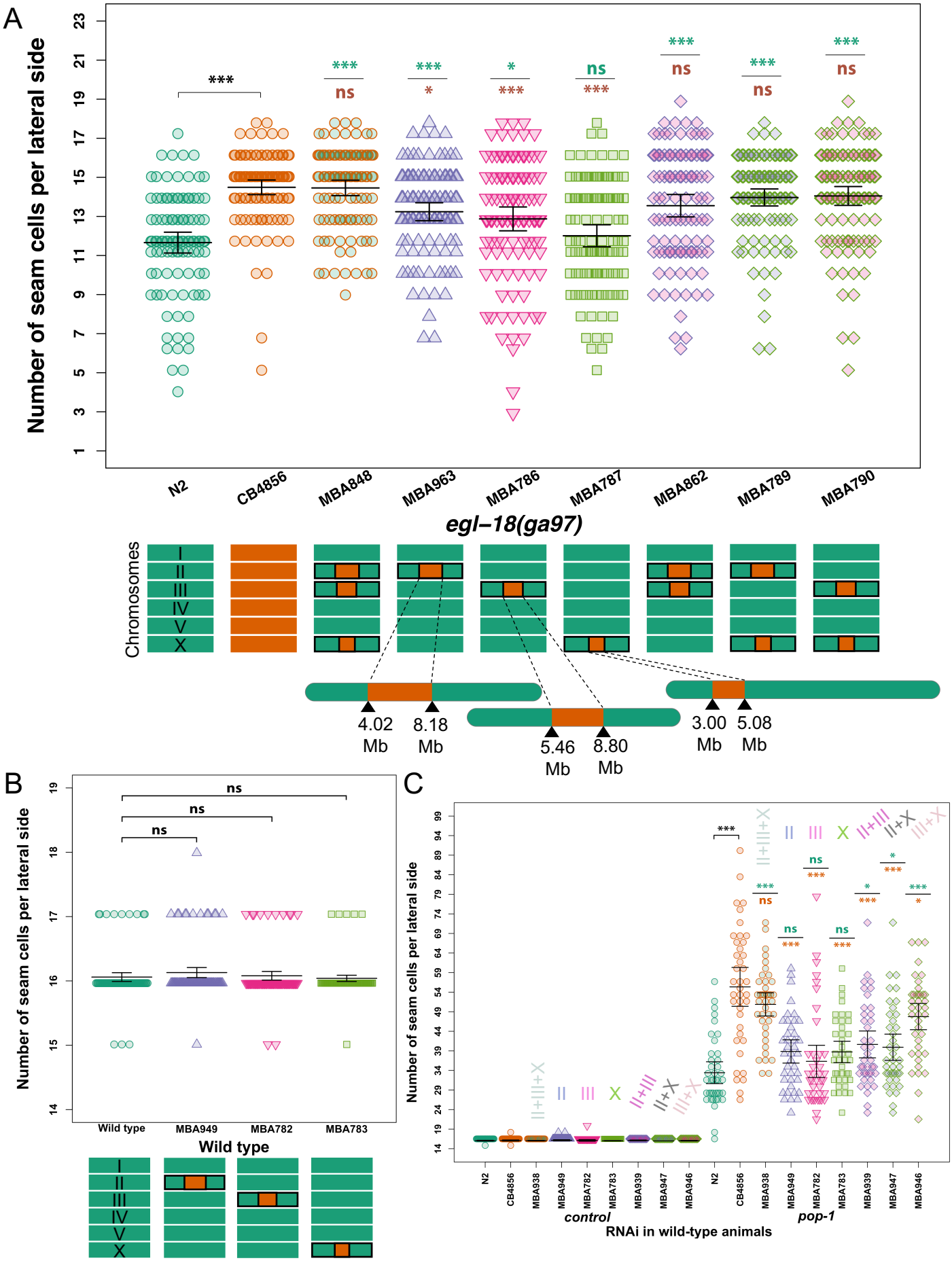

A
Number of seam cells per lateral side
N2
CB4856
MBA848
MBA963
MBA786
MBA787
MBA862
MBA789
MBA790
egl–18(ga97)
Chromosomes
I
II
III
IV
V
X
4.02 Mb
8.18 Mb
5.46 Mb
8.80 Mb
3.00 Mb
5.08 Mb
B
Number of seam cells per lateral side
Wild type
MBA949
MBA782
MBA783
C
Number of seam cells per lateral side
control
pop-1
RNAi in wild-type animals

Figure 4

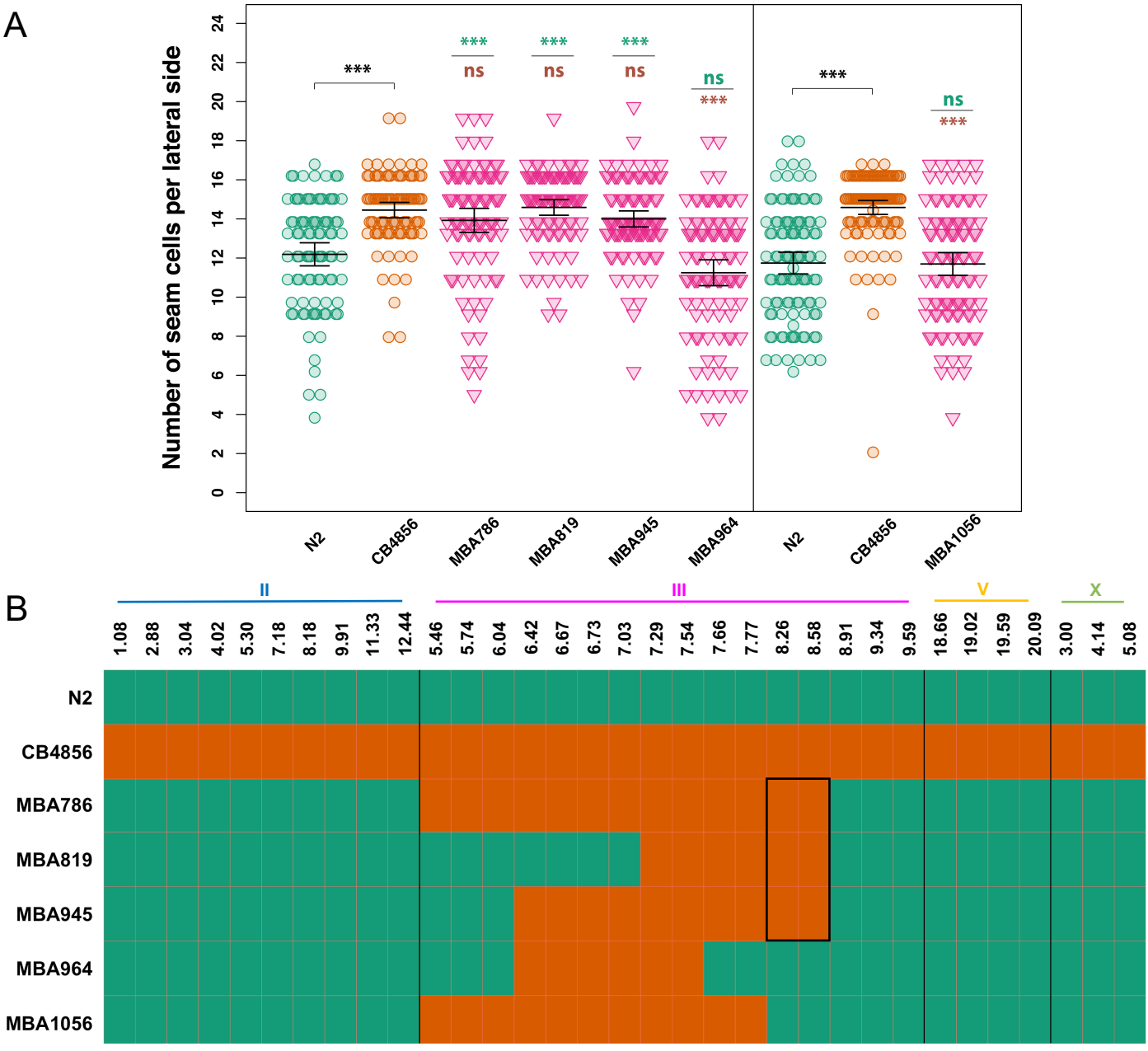

A
Number of seam cells per lateral side
0
2
4
6
8
10
12
14
16
18
20
22
24
***
***
ns
***
ns
***
ns
***
ns
***
***
ns
***
N2
CB4856
MBA786
MBA819
MBA945
MBA964
N2
CB4856
MBA1056
B
II
III
V
X
1.08
2.88
3.04
4.02
5.30
7.18
8.18
9.91
11.33
12.44
5.46
5.74
6.04
6.42
6.67
6.73
7.03
7.29
7.54
7.66
7.77
8.26
8.58
8.91
9.34
9.59
18.66
19.02
19.59
20.09
3.00
4.14
5.08
N2
CB4856
MBA786
MBA819
MBA945
MBA964
MBA1056

Figure 5

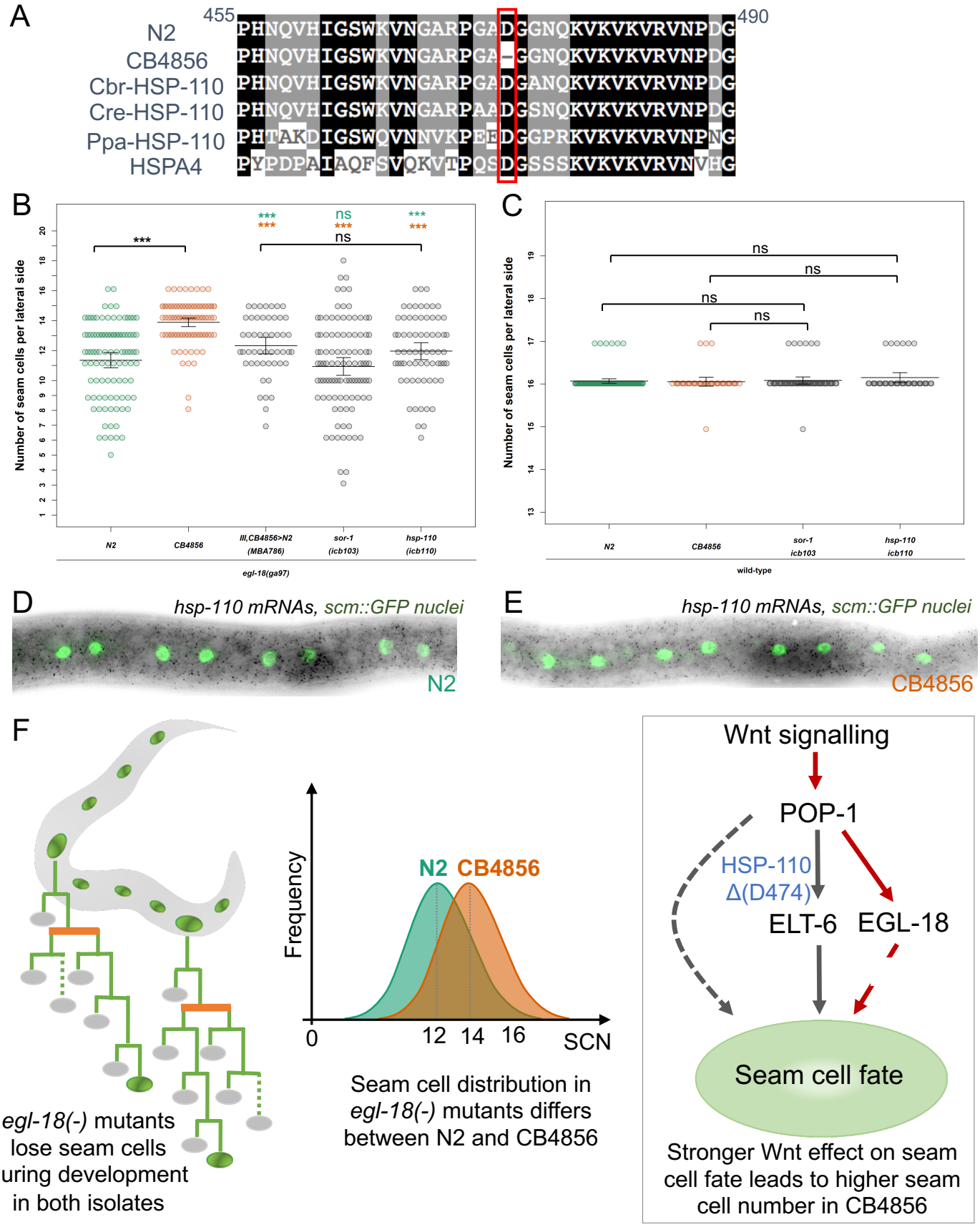

A
455
490
N2
CB4856
Cbr-HSP-110
Cre-HSP-110
Ppa-HSP-110
HSPA4
B
Number of seam cells per lateral side
***
ns
N2
CB4856
III,CB4856>N2 (MBA786)
sor-1 (icb103)
hsp-110 (icb110)
egl-18(ga97)
C
Number of seam cells per lateral side
ns
N2
CB4856
sor-1 icb103
hsp-110 icb110
wild-type
D
hsp-110 mRNAs, scm::GFP nuclei
N2
E
hsp-110 mRNAs, scm::GFP nuclei
CB4856
F
egl-18(-) mutants lose seam cells during development in both isolates
Frequency
N2 CB4856
0
12 14 16
SCN
Seam cell distribution in egl-18(-) mutants differs between N2 and CB4856
Wnt signalling
POP-1
HSP-110 Δ(D474)
ELT-6
EGL-18
Seam cell fate
Stronger Wnt effect on seam cell fate leads to higher seam cell number in CB4856

Figure 6

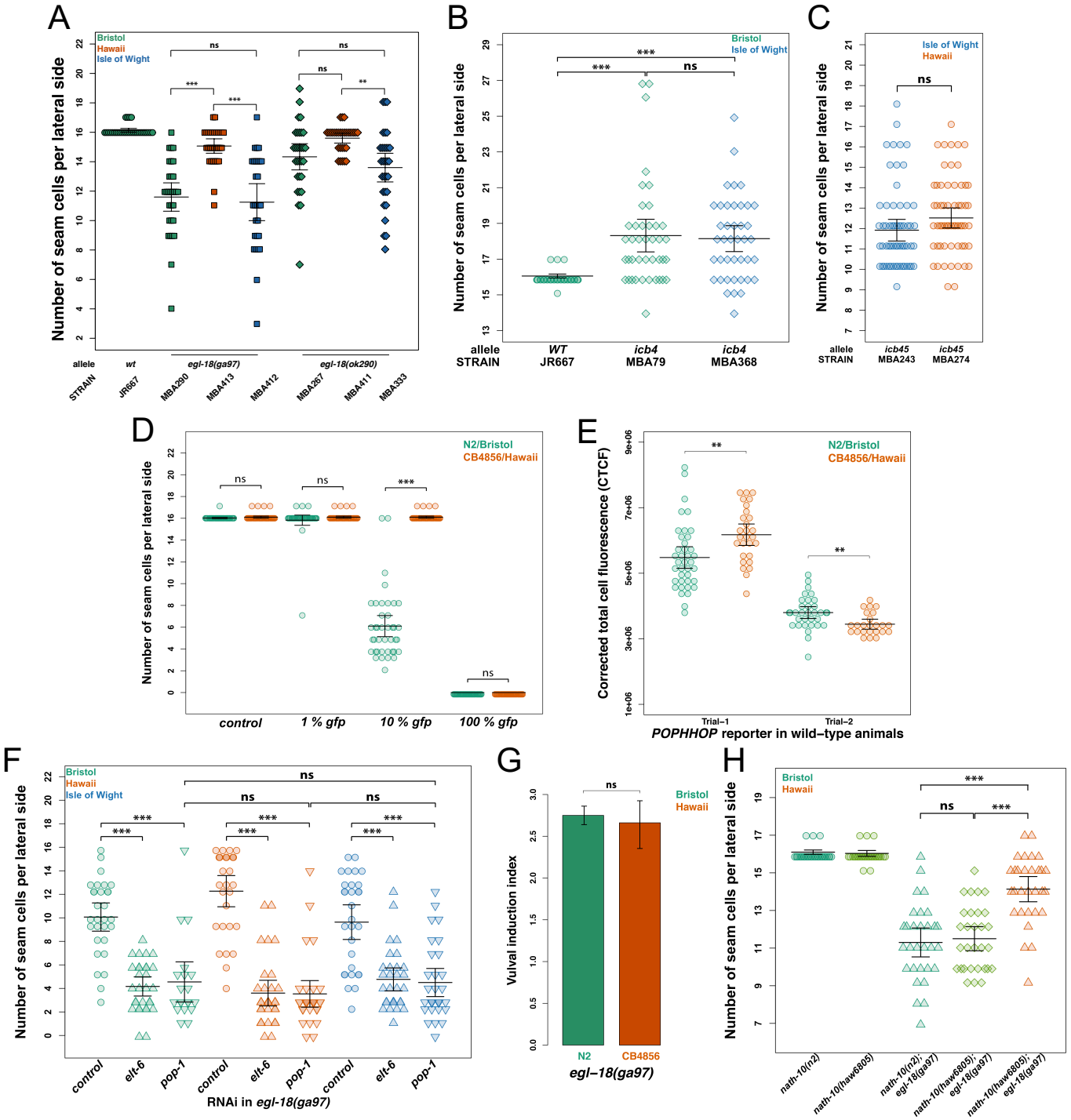

A
Bristol
Hawaii
Isle of Wight
Number of seam cells per lateral side
ns
***
**
allele
STRAIN
wt
JR667
egl-18(ga97)
MBA290
MBA413
MBA412
egl-18(ok290)
MBA267
MBA411
MBA333
B
Bristol
Isle of Wight
Number of seam cells per lateral side
***
ns
allele
STRAIN
WT
JR667
icb4
MBA79
icb4
MBA368
C
Isle of Wight
Hawaii
Number of seam cells per lateral side
ns
allele
STRAIN
icb45
MBA243
icb45
MBA274
D
N2/Bristol
CB4856/Hawaii
Number of seam cells per lateral side
ns
***
control
1 % gfp
10 % gfp
100 % gfp
E
N2/Bristol
CB4856/Hawaii
Corrected total cell fluorescence (CTCF)
**
Trial−1
Trial−2
POPHHOP reporter in wild−type animals
F
Bristol
Hawaii
Isle of Wight
Number of seam cells per lateral side
ns
***
control
elt-6
pop-1
RNAi in egl-18(ga97)
G
Bristol
Hawaii
Vulval induction index
ns
N2
CB4856
egl−18(ga97)
H
Bristol
Hawaii
Number of seam cells per lateral side
***
ns
nath-10(n2)
nath-10(haw6805)
nath-10(n2); egl-18(ga97)
nath-10(haw6805); egl-18(ga97)

Figure S1

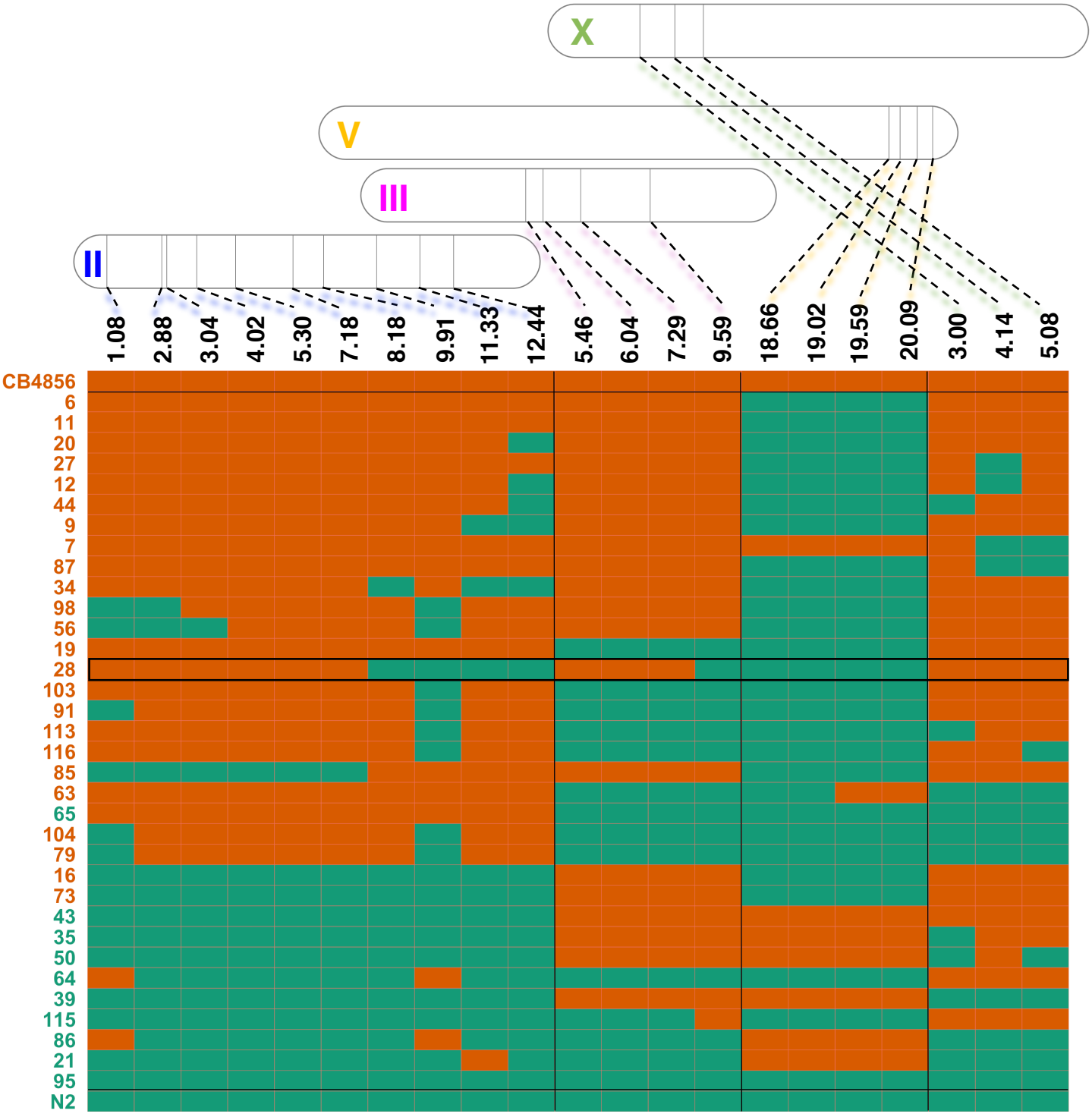

X
V
III
II
1.08
2.88
3.04
4.02
5.30
7.18
8.18
9.91
11.33
12.44
5.46
6.04
7.29
9.59
18.66
19.02
19.59
20.09
3.00
4.14
5.08
CB4856
6
11
20
27
12
44
9
7
87
34
98
56
19
28
103
91
113
116
85
63
65
104
79
16
73
43
35
50
64
39
115
86
21
95
N2

Figure S2

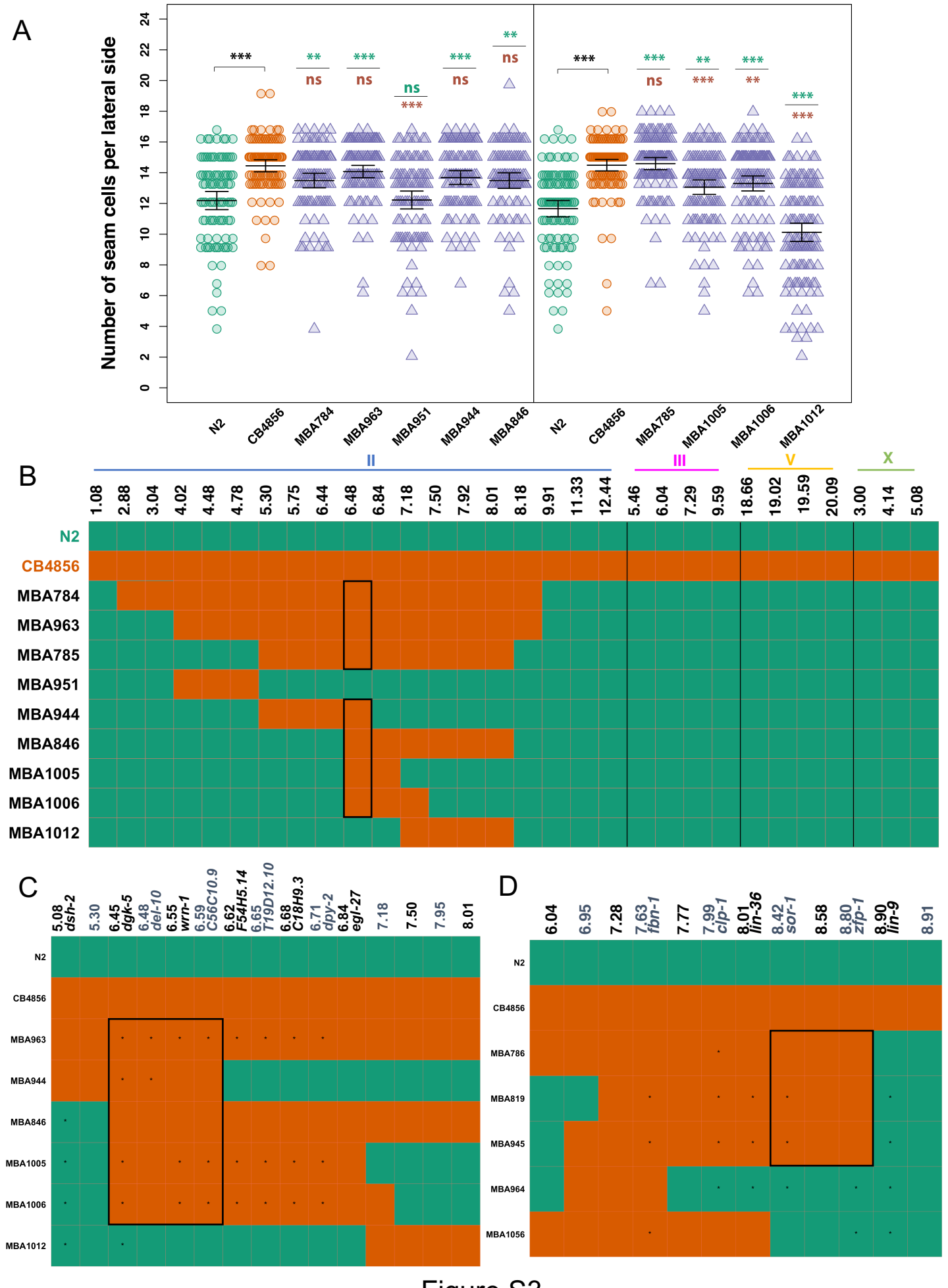
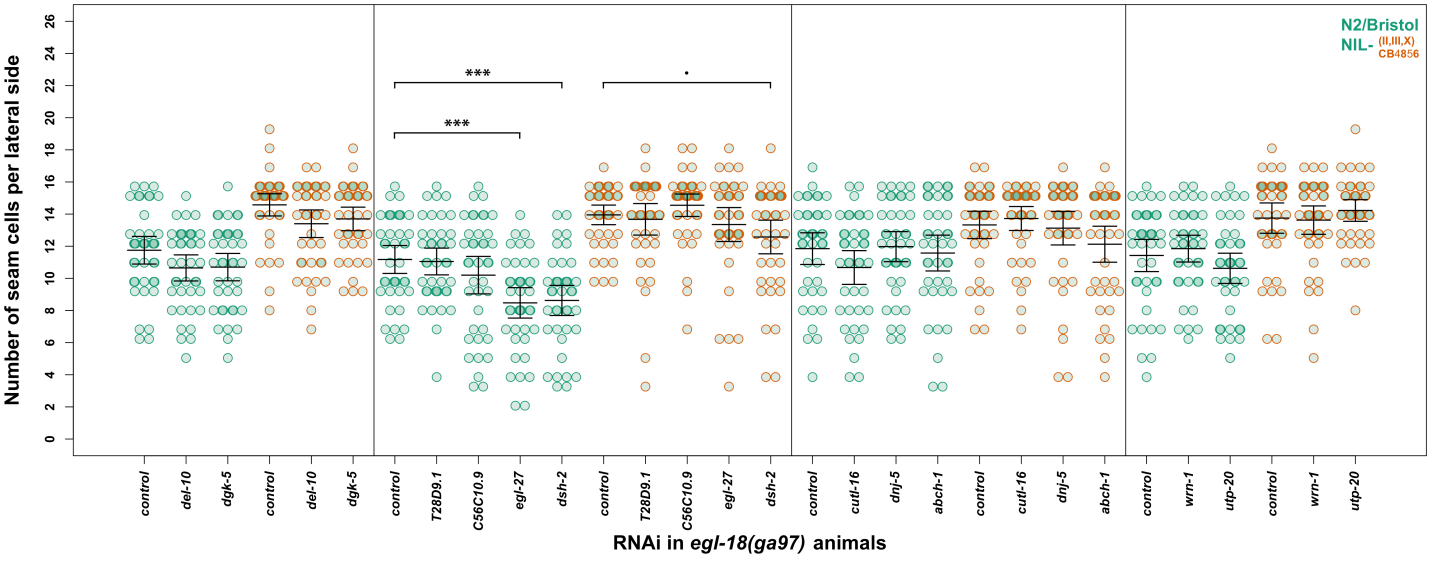

A
Number of seam cells per lateral side
N2
CB4856
MBA784
MBA963
MBA951
MBA944
MBA846
MBA785
MBA1005
MBA1006
MBA1012
B
II
III
V
X
1.08
2.88
3.04
4.02
4.48
4.78
5.30
5.75
6.44
6.48
6.84
7.18
7.50
7.92
8.01
8.18
9.91
11.33
12.44
5.46
6.04
7.29
9.59
18.66
19.02
19.59
20.09
3.00
4.14
5.08
C
5.08 dsh-2
5.30
6.45 dgk-5
6.48 del-10
6.55 wrn-1
6.59 C56C10.9
6.62 F54H5.14
6.65 T19D12.10
6.68 C18H9.3
6.71 dpy-2
6.84 egl-27
7.18
7.50
7.95
8.01
D
6.04
6.95
7.28
7.63 fbn-1
7.77
7.99 clp-1
8.01 lin-36
8.42 sor-1
8.58
8.80 zip-1
8.90 lin-9
8.91
N2
CB4856
MBA786
MBA819
MBA945
MBA964
MBA1056

Figure S3

A


N2/Bristol
NIL- (II,III,X) CB4856
Number of seam cells per lateral side
0
2
4
6
8
10
12
14
16
18
20
22
24
26
***
***
.
control
del-10
dgk-5
control
del-10
dgk-5
control
T28D9.1
C56C10.9
egl-27
dsh-2
control
T28D9.1
C56C10.9
egl-27
dsh-2
control
cutl-16
dnj-5
abch-1
control
cutl-16
dnj-5
abch-1
control
wrn-1
utp-20
control
wrn-1
utp-20
RNAi in egl-18(ga97) animals

B

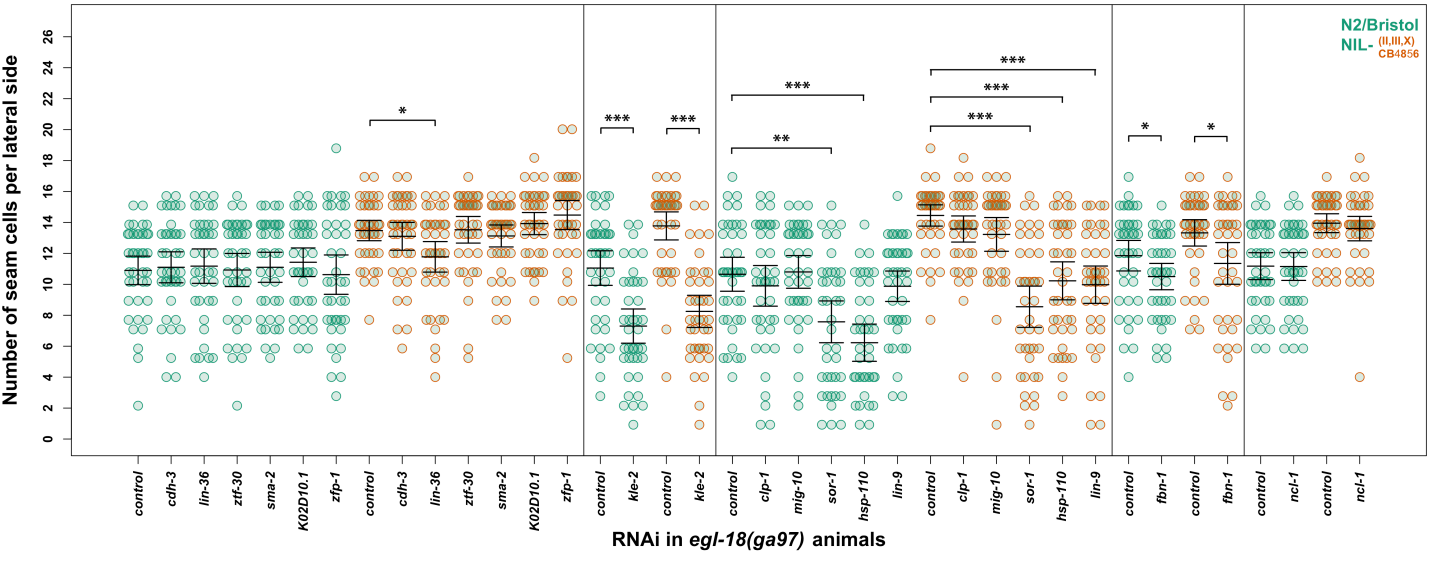

N2/Bristol
NIL- (II,III,X) CB4856
Number of seam cells per lateral side
0
2
4
6
8
10
12
14
16
18
20
22
24
26
*
***
***
***
**
***
***
***
*
*
control
cdh-3
lin-36
ztf-30
sma-2
K02D10.1
zfp-1
control
cdh-3
lin-36
ztf-30
sma-2
K02D10.1
zfp-1
control
kle-2
control
kle-2
control
clp-1
mig-10
sor-1
hsp-110
lin-9
control
clp-1
mig-10
sor-1
hsp-110
lin-9
control
fbn-1
control
fbn-1
control
ncl-1
control
ncl-1
RNAi in egl-18(ga97) animals

Figure S4