Extensive intraspecies cryptic variation in an ancient embryonic gene regulatory network

Yamila N. Torres Cleuren[1,2]*, Chee Kiang Ewe[2], Kyle C. Chipman[2], Emily Mears[1], Cricket G. Wood[2], Coco A.E. Al-Alami[1], Melissa R. Alcorn[2], Thomas L. Turner[3], Pradeep M. Joshi[2], Russell G. Snell[1], and Joel H. Rothman[1,2,3]

1. School of Biological Sciences, University of Auckland, Auckland, New Zealand

2. Department of MCD Biology and Neuroscience Research Institute, University of California Santa Barbara, CA, USA

3. Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, CA, USA

*Current address: Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, 5020 Norway

Author for correspondence: joel.rothman@lifesci.ucsb.edu

Keywords: SKN-1, *C. elegans*, endoderm, GWAS, development, gene regulatory networks, Genotype-By-Sequencing, cryptic variation

ABSTRACT

Innovations in metazoan development arise from evolutionary modification of gene regulatory networks (GRNs). We report widespread cryptic variation in the requirement for two key regulatory inputs, SKN-1/Nrf2 and MOM-2/Wnt, into the *C. elegans* endoderm GRN. While some natural variants show a nearly absolute requirement for these two regulators, in others, most embryos differentiate endoderm in their absence. GWAS and analysis of recombinant inbred lines reveal multiple genetic regions underlying this broad phenotypic variation. We observe a reciprocal trend, in which genomic variants, or knockdown of endoderm regulatory genes, that result in a high SKN-1 requirement often show low MOM-2/Wnt requirement and *vice-versa,* suggesting that cryptic variation in the endoderm GRN may be tuned by opposing requirements for these two key regulatory inputs. These findings reveal that while the downstream components in the endoderm GRN are common across metazoan phylogeny, initiating regulatory inputs are remarkably plastic even within a single species.

38 **INTRODUCTION**

39 While the core regulatory machinery that specifies embryonic germ layers and major
40 organ identity in the ancestor of modern animals is largely conserved in all extant animals,
41 GRN architecture must be able to accommodate substantial plasticity to allow for
42 evolutionary innovation in developmental strategies, changes in selective pressures, and
43 genetic drift [1,2]. Genetic variation, often with neutral effects on fitness, provides for
44 plasticity in GRN structure and implementation [2]. Although studies of laboratory strains of
45 model organisms with a defined genetic background have been highly informative in
46 identifying the key regulatory nodes in GRNs that specify developmental processes [3–5],
47 these approaches do not reveal the evolutionary basis for plasticity in these networks. The
48 variation and incipient changes in GRN function and architecture can be discovered by
49 analyzing phenotypic differences resulting from natural genetic variation present in distinct
50 isolates of a single species [6–8].

51 The endoderm has been proposed to be the most ancient of the three embryonic
52 germ layers in metazoans [9,10], having appeared prior to the advent of the Bilateria about
53 600 Mya [11]. It follows, therefore, that the GRN for endoderm in extant animals has
54 undergone substantial modifications over the long evolutionary time span since its
55 emergence. However, the core transcriptional machinery for endoderm specification and
56 differentiation appears to share common mechanisms across metazoan phylogeny. For
57 example, cascades of GATA-type transcription factors function to promote endoderm
58 development not only in triploblastic animals but also in the most basal metazoans that
59 contain a digestive tract [12–16]. Among the many observations supporting a common
60 regulatory mechanism for establishing the endoderm, it has been found that the endoderm-
61 determining GATA factor, END-1, in the nematode *C. elegans*, is sufficient to activate

62 endoderm development in cells that would otherwise become ectoderm in Xenopus [17]. This
63 indicates that the role of GATA factors in endoderm development has been preserved since
64 the nematodes and vertebrates diverged from a common ancestor that lived perhaps 600
65 Mya.

66 To assess the genetic basis for evolutionary plasticity and cryptic variation underlying
67 early embryonic germ layer specification, we have analyzed the well-described GRN for
68 endoderm specification in *C. elegans*. The E cell, which is produced in the very early *C. elegans*
69 embryo, is the progenitor of the entire endoderm, which subsequently gives rise exclusively
70 to the intestine. The EMS blastomere at the four-cell stage divides to produce the E founder
71 cell and its anterior sister, the MS founder cell, which is the progenitor for much of the
72 mesoderm [18]. Both E and MS fates are determined by maternally provided SKN-1, an
73 orthologue of the vertebrate Nrf2 bZIP transcription factor [19–21]. In the laboratory N2
74 strain, elimination of maternal SKN-1 function (through either knockdown or knockout)
75 results in fully penetrant embryonic lethality as a result of misspecification of EMS cell
76 descendants. In these embryos, the fate of MS is transformed to that of its cousin, the
77 mesectodermal progenitor C cell. E cells similarly adopt a C cell-like fate in a majority, but not
78 all, of these embryos [19]. SKN-1 initiates mesendoderm development via the GRN in E and
79 MS cells in part by activating zygotic expression of the MED-1/2 divergent GATA transcription
80 factors [22,23]. This event mobilizes a cascade of GATA factors in the E cell lineage that
81 ultimately direct intestinal differentiation [21,24–26].

82 This differential requirement for SKN-1 in endoderm (E) and mesoderm (MS)
83 development is determined by its combinatorial action with triply redundant Wnt, MAPK, and
84 Src signaling systems, which act together to polarize EMS [27–30]. MOM-2/Wnt acts through

85 the MOM-5/Frizzled receptor, mobilizing WRM-1/β-catenin, resulting in its cytoplasmic
86 accumulation in the posterior side of EMS. WRM-1, together with LIT-1/NLK kinase, alters
87 both the nucleocytoplasmic distribution and activity of the Wnt effector POP-1/Tcf [30–32],
88 converting it from a repressor of endoderm in the MS cell lineage to an activator in the E cell
89 lineage [33–38]. Loss of MOM-2 expression in the laboratory N2 strain results in a partial
90 gutless phenotype, while removal of both MOM-2 and SKN-1, through either knockdown or
91 knockout, leads to a completely penetrant loss of gut [30], revealing their genetically
92 redundant roles.

93 The regulatory relationship between SKN-1 and POP-1, the effector of Wnt signaling,
94 shows substantial variation even in species that diverged 20-40 million years ago, suggesting
95 significant evolutionary plasticity in this key node in the endoderm GRN. *C. elegans* embryos
96 lacking maternal POP-1 always make gut, both in the normal E cell lineage and in the MS cell
97 lineage. However, in embryos lacking both SKN-1 and POP-1, endoderm is virtually never
98 made, implying that these two factors constitute a Boolean "OR" logic gate. In contrast,
99 removal of either SKN-1 or POP-1 alone in *C. briggsae* causes >90% of embryos to lack gut,
100 indicative of an "AND" logic gate (Fig. 1A, B) [39].

101 In this study, we sought to determine whether the plasticity in regulatory logic of the
102 two major inputs into endoderm development is evident within a single species. The
103 availability of many naturally inbred variants (isotypes) of *C. elegans* that show widespread
104 genomic variation [40–42], provides a genetically rich resource for investigating potential
105 quantitative variation in developmental GRNs. We report here that the requirement for
106 activation of the endoderm GRN by SKN-1 or MOM-2, but not POP-1, is highly variable
107 between natural *C. elegans* isolates, and even between closely related isotypes. Genome-

108 wide association studies (GWAS) in isolates from the natural populations and targeted
109 analysis of recombinant inbred lines (RILs), revealed that a multiplicity of loci and their
110 interactions are responsible for the variation in the developmental requirement for SKN-1
111 and MOM-2. We found a complex, but frequently reciprocal requirement for SKN-1 and
112 MOM-2 among variants underlying these phenotypes: loci associated with a high
113 requirement for SKN-1 in endoderm development tend to show a more relaxed requirement
114 for MOM-2 and *vice-versa*. Consistent with this finding, three other endoderm regulatory
115 factors, RICT-1, PLP-1, and MIG-5, show similar inverse relationships between these two GRN
116 inputs. These findings reveal that the activation of the GRN network for a germ layer, one of
117 the most critical early developmental switches in embryos, is subject to remarkable genetic
118 plasticity within a species and that the dynamic and rapid change in network architecture
119 reflects influences distributed across many genetic components that affect both SKN-1 and
120 Wnt pathways.

121

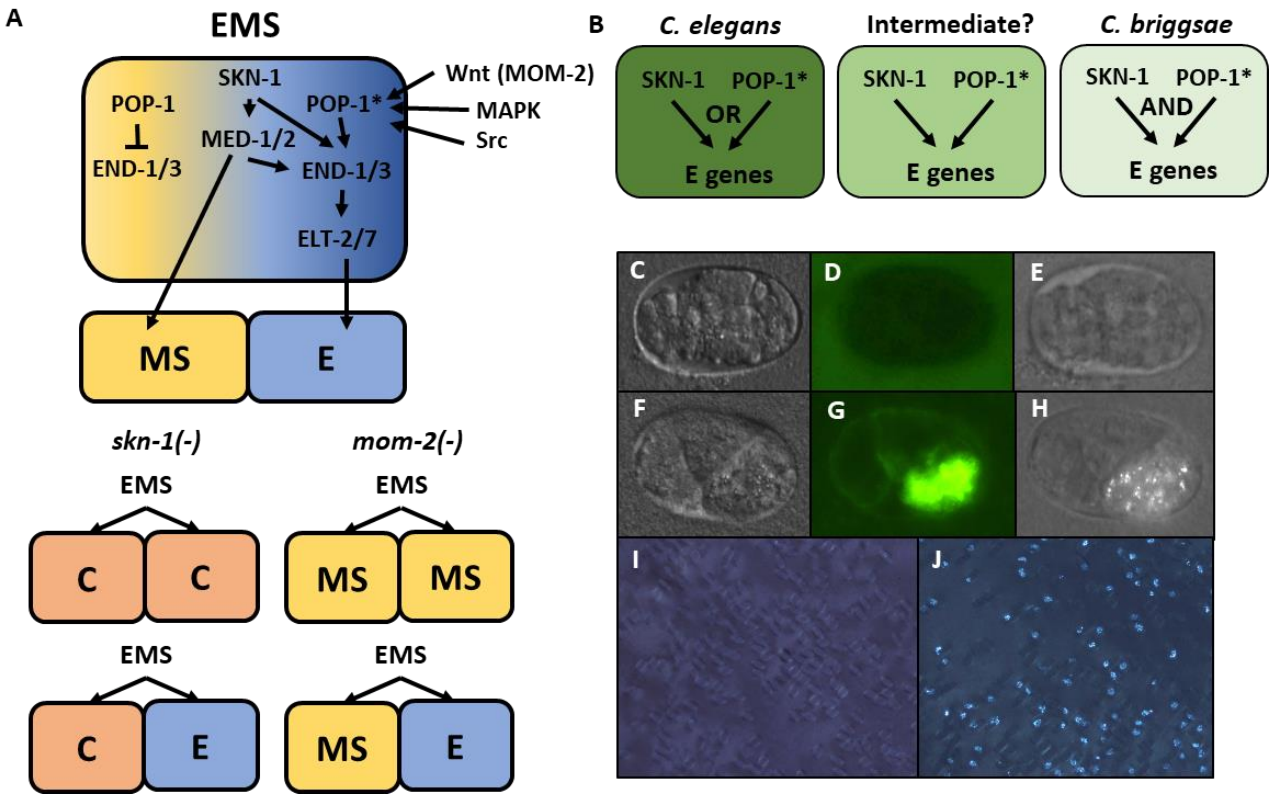


Fig 1. Endoderm regulatory pathway and scoring of gut differentiation.

A) Under normal conditions, signaling from the posterior P_2 cell (Wnt, MAPK and Src) results in asymmetric cortical localization of Wnt signaling pathway components in EMS leading to POP-1 asymmetry in the descendants of EMS, with high levels of nuclear POP-1 in anterior MS and low levels of nuclear POP-1 in the posterior, E, daughter cell. In the anterior MS cell, high nuclear POP-1 represses the END genes, allowing SKN-1 to activate MS fate. In the posterior E cell, which remains in contact with P2, POP-1 is converted to an activator and, along with SKN-1, activates the END genes, resulting in endoderm fate. Loss of *skn-1,* either by RNAi or in loss-of-function mutants, causes 100% of the embryos to arrest; in 70% of the arrested embryos, EMS gives rise to two C-like cells, while in the remaining 30% only MS is converted to a C fate; the posterior daughter retains its E fate. Loss of *mom-2* leads to 100% embryonic arrest with a partially penetrant E→MS cell fate transformation, resulting in two MS-like daughter cells in ~72% of the embryos. (B) Regulatory logic of SKN-1 and POP-1 in E specification in *C. elegans, C. briggsae* and a hypothetical intermediate state. POP-1* denotes the activated state. (C-H) Gut visualization in embryos affected by *skn-1* RNAi. (C-E) arrested embryos without endoderm, (F-H) arrested embryos with endoderm. (C, F) DIC images of arrested embryos ~12 hours after egg laying. (D, G) the same embryos expressing the gut-specific *elt-2::GFP* reporter, and (E,H) birefringent gut granules under polarized light. All embryos showing gut birefringence also show *elt-2::GFP* expression. (I, J) Fields of arrested *skn-1(RNAi)* embryos in wild isolate strains JU1491 (I) and JU440 (J), which reflect the extremes in the spectrum of requirement of SKN-1 in gut development at 0.9% and 60%, respectively.

143 **MATERIALS AND METHODS**

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
strain, strain background (C. elegans)	Wild isolate refer to supplementary file 1	CGC		
strain, strain background (C. elegans)	JJ185	CGC		dpy-13(e184) skn-1(zu67) IV; mDp1 (IV; f).
strain, strain background (C. elegans)	JR3666	This study		(elt-2::GFP) X ; (ifb-2::CFP) IV
strain, strain background (C. elegans)	EU384	CGC		dpy-11(e1180) mom-2(or42) V/nT1 [let-?(m435)] (IV;V).
strain, strain background (C. elegans)	JJ1057	CGC		pop-1(zu189) dpy-5(e61)/hT1 I; him-5(e1490)/hT1 V.
strain, strain background (C. elegans)	KQ1366	CGC		rict-1(ft7) II.
strain, strain background (C. elegans)	SU351	CGC		mig-5(rh94)/mIn1 [dpy-10(e128) mIs14] II.
strain, strain background (C. elegans)	RB1711	CGC		plp-1(ok2155) IV.
strain, strain background (C. elegans)	JR3493-JR3590	This study		N2xMY16 RILs

strain, strain background (C. elegans)	MT3414	CGC		dpy-20(e1282) unc-31(e169) unc-26(e205) IV.
strain, strain background (C. elegans)	DA491	CGC		dpy-20(e1282) unc-30(e191) IV.
strain, strain background (C. elegans)	JR2750	[43]		bli-6(Sc16)unc-22(e66)/unc-24(e138)fus-1(w13) dpy-20(e2017)IV
strain, strain background (C. elegans)	JR3812 (NIL 1)	This study		NIL N2XMY16
strain, strain background (C. elegans)	JR3813 (NIL 2)	This study		NIL N2XMY16
strain, strain background (C. elegans)	JR3814 (NIL 3)	This study		NIL N2XMY16
strain, strain background (C. elegans)	JR3815 (NIL 4)	This study		NIL N2XMY16
strain, strain background (C. elegans)	JR3816 (NIL 5)	This study		NIL N2XMY16
strain, strain background (C. elegans)	JR3817 (NIL 6)	This study		NIL N2XMY16
antibody	MH33	DSHB		
antibody	AHP418	Serotec Bio-Rad		
antibody	ab150116	Abcam		Goat Anti-Mouse IgG H&L (Alexa Fluor® 594)
antibody	ab150077	Abcam		Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488)
software, algorithm	R v 3.2.3	The R Foundation		
software, algorithm	PLINK	http://pngu.mgh.harvard.edu/purcell/plink/		

144

145 ***C. elegans* strains and maintenance**

146 All wild isolates, each with a unique haplotype [42], were obtained from the Caenorhabditis
147 Genetics Center (CGC) (see Supplementary File 1). Worm strains were maintained as
148 described [44] and all experiments were performed at 20°C unless noted otherwise.

149 **RNAi**

150 Feeding-based RNAi experiments were performed as described [45]. RNAi clones were
151 obtained from either the Vidal [46] or Ahringer libraries [47]. RNAi bacterial strains were
152 grown at 37°C in LB containing 50 µg/ml ampicillin. The overnight culture was then diluted
153 1:10. After 4 hours of incubation at 37°C, 1 mM of IPTG was added and 60µl was seeded onto
154 35mm agar plates containing 1 mM IPTG and 25 µg/ml carbenicillin. Seeded plates were
155 allowed to dry and used within five days. Five to 10 L4 animals were placed on RNAi plate. 24
156 hours later, they were transferred to another RNAi plate and allowed to lay eggs for four or
157 12 hours (12 hours for *skn-1* RNAi and four hours for the other RNAi). The adults were then
158 removed, leaving the embryos to develop for an extra 7-9 hours. Embryos were quantified
159 and imaged on an agar pad using a Nikon Ti-E inverted microscope. We chose to perform RNAi
160 on agar plates to maximize sensitivity, robustness, and reproducibility of the assay, as liquid
161 culture RNAi can introduce variability owing to aggregation and settling of bacteria, which
162 affects RNAi efficacy [48]. In addition, performing RNAi on agar plates allowed us to collect
163 large numbers of embryos with which to quantify gut formation (as described below).

164 **Antibody staining**

165 The embryonic gut cells and nuclei of all cells were stained with MH33 (mouse anti-IFB-2,
166 deposited to the DSHB by Waterston, R.H.) and AHP418 (rabbit anti-acetylated histone H4,

167 Serotec Bio-Rad) respectively. Fixation and permeabilization were carried out as described
168 previously [49]. Goat anti-mouse Alexa Fluor® 594 and goat anti-rabbit Alexa Fluor® 488
169 secondary antibodies were used at 1:1000 dilution.

170 **Quantification of endoderm specification**

171 Gut was scored by the presence of birefringent gut granule in arrested embryos [50,51]. For
172 *skn-1(RNAi)*, the laboratory strain N2, which shows invariable ~30% of embryos with
173 endoderm, was used as a control for all experiments.

174 **Introgression of *skn-1(zu67)*, *pop-1(zu189)*, and *mom-2(or42)* alleles into wild isolate**
175 **backgrounds**

176 To introgress *skn-1(zu67)* into wild isolates (WI), males from the wild isolate strains were
177 crossed to JJ186 *dpy-13(e184) skn-1(zu67) IV; mDp1 (IV;f)* hermaphrodites. mDp1 is a free
178 duplication maintained extrachromosomally that rescues the Dpy and lethal phenotypes of
179 *dpy-13(e184)* and *skn-1(zu67)* respectively. mDp1 segregates in a non-Mendelian fashion and
180 animals that have lost the free duplication are Dpy and produce dead offspring. Wild type F1
181 hermaphrodites that have lost the free duplication, as determined by the presence of 1/4 Dpy
182 progeny in the F2 generation, were selected. 10 single non-Dpy F2 hermaphrodite
183 descendants from F1 animals heterozygous for *skn-1(zu67)* (2/3 of which are expected to be
184 of the genotype *WI dpy-13(+) skn-1(+) / dpy-13(e184) skn-1(zu67)* were backcrossed to their
185 respective parental wild strain. 10 F3 hermaphrodites were picked to individual plates. Half
186 of the F3 cross progeny are expected to be heterozygous for *dpy-13(e184) skn-1(zu67)*, as
187 evidenced by presence of F4 Dpy progeny that produced dead embryos. Non-Dpy siblings
188 were used to continue the introgression as described. This strategy was repeated for at least

189 five rounds of introgression. The embryonic gutless phenotype in the progeny of the Dpy
190 animals was quantified.

191 Similarly, to introgress *pop-1(zu189)* or *mom-2(or42)* alleles into wild isolates, JJ1057
192 *pop-1(zu189) dpy-5(e61)/hT1 I; him-5(e1490)/hT1V* or EU384 *dpy-11(e1180) mom-2(or42)*
193 *V/nT1 [let-?(m435)] (IV;V)* were used, respectively. The mutant strain was crossed to the wild
194 isolates. Non-Dpy F2 animals heterozygous for the chromosomal mutation were selected and
195 backcrossed to their respective parental wild strain for at least four rounds of introgression
196 for *pop-1* and seven rounds for *mom-2*. The embryonic gutless phenotype in the progeny of
197 the Dpy animals was quantified, as above.

198 **Statistical Analyses: GWAS**

199 All data were analyzed and plotted using R software v 3.2.3 (https://www.r-
200 project.org/). GWAS for both phenotypes was performed using *C. elegans* wild isolates and
201 a previously published SNP map containing 4,690 SNPs [42] with the EMMA R package. P-
202 values were calculated using mixed model analysis [52] (emma.REML.t() function) and
203 identity-by-state (IBS) kinship matrix to account for population structure. For *skn-1* and *mom-*
204 *2* RNAi phenotypic data, a genome-wide permutation-based FDR was also calculated for the
205 EMMA results from 10,000 permuted values [53,54].

206 **Phylogenetic and geographical analyses**

207 Phylogenetic trees were constructed from 4,690 polymorphisms using R package "ape"
208 [55]. Neighbor-joining algorithm based on pairwise distances was used. Phylogenetic signal
209 (Pagel's λ statistics) was measured using "phylosig()" function in phytools R package [56,57].
210 Statistical significance of λ was obtained by comparing the likelihood a model accounting for

211 the observed λ with the likelihood of a model that assumes complete phylogenetic
212 independence.

213 Geographic information for strains were obtained from Andersen *et al.* [42], available in
214 Supplementary File 1, together with the corresponding *skn-1(RNAi)* and *mom-2(RNAi)*
215 phenotypes.

216 **Correlation Analysis**

217 To test for the relationship between *mom-2 (RNAi)* and *skn-1 (RNAi)* phenotypic data, the
218 differences between median phenotypic values for each SNP were calculated independently
219 on a genome-wide level for the wild isolates. In order to correct for LD, SNPs were pruned
220 with PLINK (http://pngu.mgh.harvard.edu/purcell/plink/) [58] and only a subset of SNPs was
221 used for the correlation analyses. Outliers were removed from the calculations by using z-
222 score with a cutoff of 1.96 (i.e., 95% of values fall within ±1.96 in a normal distribution).

223 **RIL construction and Genotype-By-Sequencing (GBS)**

224 Recombinant inbred lines (RILs) were created by crossing an N2 hermaphrodite and an MY16
225 male. 120 F2 progeny were cloned to individual plates and allowed to self-fertilize for 10
226 generations. A single worm was isolated from each generation to create inbred lines. A total
227 of 95 lines were successfully created and frozen stocks were immediately created and kept at
228 -80°C (Supplementary File 2), prior to DNA sequencing.

229 DNA was extracted using Blood and Tissue QIAGEN kit from worms from each of the RILs
230 grown on four large NGM plates (90x15mm) with OP50 *E. coli* until starved (no more than a
231 day). Samples were submitted in 96-well plate format at 10 ng/µl < n < 30 ng/µl. GBS libraries
232 were constructed using digest products from ApeKI (GWCGC), using a protocol modified from

233 [59]. After digestion, the barcoded adapters were ligated and fragments < 100bp were
234 sequenced as single-end reads using an Illumina HiSeq 2000 lane (100 bp, single-end reads).

235 SNP calling was performed using the GBSversion3 pipeline in Trait Analysis by aSSociation,
236 Evolution and Linkage (TASSEL) [60]. Briefly, fastq files were aligned to reference genome
237 WS252 using BWA v. 0.7.8-r455 and SNPs were filtered using vcftools [61]. Samples with
238 greater than 90% missing data and SNPs with minor allele frequencies (mAF) of <1% were
239 excluded from analysis, identifying 27,396 variants.

240 **QTL mapping using R/qtl**

241 Variants identified by GBS pipeline were filtered to match the SNPs present in the parental
242 MY16 strain (using vcftools –recode command), and variants were converted to a 012 file
243 (vcftools –012 command). Single-QTL analysis was performed in R/QTL [62] using 1770
244 variants and 95 RILs. Significant QTL were determined using Standard Interval Mapping
245 (scanone() "em") and genome-wide significance thresholds were calculated by permuting the
246 phenotype (N =1,000). Change in log-likelihood ratio score of 1.5 was used to calculate 95%
247 confidence intervals and define QTL regions [63]. SNP data for the RILs and their
248 corresponding phenotypes used in analysis are shown in Supplementary Files 2 and 3.

249 **Creation of Near-Isogenic Lines (NILs)**

250 Three N2-derived mutant strains were used to introgress regions from chromosome IV from
251 N2 into the MY16 strain background and vice-versa. For both types of crosses, N2 was always
252 used as the maternal line. The following strains were used:

253 - DA491: *dpy-20(e1282) unc-30(e191) IV.*

254 ● JR2750: *bli-6(sc16) unc-22(e66)/unc-24(e138)fus-1(w13) dpy-20(e2017) IV*. Worms
255 segregating Bli Unc were selected for crosses.
256 ● MT3414: *dpy-20(e1282) unc-31(e169) unc-26(e205) IV*.

257 To introgress the N2 region into the MY16 genetic background, hermaphrodites from the N2-
258 derived strains containing genetic markers flanking the genomic region of interest were
259 crossed with MY16 males (Supplementary Fig. 7). After successful mating, 10 F1
260 heterozygotes were isolated and allowed to self. After 24 hours, the F1 adults were removed
261 from the plate, and the F2 hermaphrodites left to develop to young adults. F2 animals
262 homozygous for the region being introgressed were selected as young adults and crossed
263 with MY16 males. This process was repeated until ten rounds of introgression were
264 completed. These new lines were preserved at -80C.

265 Introgression of MY16 region into an N2 background began with the same initial cross as
266 above. F1 heterozygous males were crossed with N2 hermaphrodites containing phenotypic
267 markers near the region being introgressed (Supplementary Fig. 7). After successful mating,
268 the F1 parents were removed and the F2 generation was left to develop until heterozygous
269 males were visible. F2 heterozygous males were crossed with hermaphrodites from the N2-
270 derived strain. This process was repeated until ten successful introgressions were completed.
271 To homozygose the introgressed MY16 regions, worms were singled and allowed to self until
272 a stable wildtype population was obtained. These new lines preserved at -80C.

273 NILs were genotyped to test for correct introgression of the desired regions by Sanger
274 sequencing of 10 markers spaced along chromosome IV (carried out by the Centre for
275 Genomics and Proteomics, University of Auckland). Upon confirmation of their genetic

276 identity, one NIL was used to further dissect the QTL region by segregating the visual markers

277 (Dpy and Unc).

278 **RESULTS**

279 **Extensive natural cryptic variation in the requirement for SKN-1 in endoderm specification**
280 **within the *C. elegans* species**

281 The relationship between SKN-1 and Wnt signaling through POP-1 in the endoderm
282 GRN has undergone substantial divergence in the *Caenorhabditis* genus [39]. While neither
283 input alone is absolutely required for endoderm specification in *C. elegans*, each is essential
284 in *C. briggsae*, which has been estimated to have diverged from *C. elegans* ~20-40 Mya
285 [64,65]. In contrast to the *C. elegans* N2 laboratory strain, removal of either SKN-1 or POP-1
286 alone results in fully penetrant conversion of the E founder cell fate into that of the
287 mesectodermal C blastomere and of E to MS fate, respectively, in *C. briggsae* [39]. These
288 findings revealed that the earliest inputs into the endoderm GRN are subject to substantial
289 evolutionary differences between these two species (Fig. 1B). We sought to determine
290 whether incipient evolutionary plasticity in this critical node at the earliest stages of
291 endoderm development might be evident even within a single species of the *Caenorhabditis*
292 genus by assessing their requirement in *C. elegans* wild isolates and testing whether the
293 quantitative requirements of each input were correlated.

294 Elimination of detectable maternal SKN-1 from the laboratory N2 strain by either a
295 strong (early nonsense) chromosomal mutation (*skn-1(zu67)*), or by RNAi knockdown, results
296 in a partially penetrant phenotype: while the E cell adopts the fate of the C cell in the majority
297 of embryos, and gut is not made, ~30% of arrested embryos undergo strong gut
298 differentiation, as evidenced by the appearance of birefringent, gut-specific rhabditin
299 granules, or expression of *elt-2::GFP*, a marker of the developing and differentiated intestine
300 (Fig. 1C-H). In our experimental conditions, we found that RNAi of *skn-1* in different N2-

301 derived mutant strains gave highly reproducible results: 100% of the embryos derived from
302 *skn-1(RNAi)*-treated mothers arrest (n>100,000) and 32.0 ± 1.9% of the arrested embryos
303 exhibited birefringent gut granules (Fig. 2A; Supplementary Fig. 1) over many trials by
304 separate investigators. We found that the LSJ1 laboratory strain, which is derived from the
305 same original source as N2, but experienced very different selective pressures in the
306 laboratory owing to its constant propagation in liquid culture over 40 years [66], gave virtually
307 identical results to that of N2 (31.0% ± s.d 1.2%), implying that SKN-1-independent endoderm
308 formation is a quantitatively stable trait. The low variability in this assay, and high number of
309 embryos that can be readily examined (≥500 embryos per experiment), provides a sensitive
310 and highly reliable system with which to analyze genetic variation in the endoderm GRN
311 between independent *C. elegans* isolates.

312 To assess variation in SKN-1 requirement within the *C. elegans* species, we analyzed
313 the outcome of knocking down SKN-1 by RNAi in 96 unique *C. elegans* wild isolates [42].
314 Owing to their propagation by self-fertilization, each of the isolates (isotypes) is a naturally
315 inbred clonal population that is virtually homozygous and defines a unique haplotype. The
316 reported estimated population nucleotide diversity averages 8.3×10^{-4} per bp [42], and we
317 found that a substantial fraction (29/97) of isotypes were quantitatively indistinguishable in
318 phenotype between the N2 and LSJ1 laboratory strains (Fig. 2A, Supplementary File 1). We
319 found that all strains, with the exception of the RNAi-resistant Hawaiian CB4856 strain,
320 invariably gave 100% embryonic lethality *with skn-1(RNAi)*. However, we observed dramatic
321 variation in the fraction of embryos with differentiated gut across the complete set of strains,
322 ranging from 0.9% to 60% (Fig. 2A). Repeated measurements with >500 embryos per replicate
323 per strain revealed very high reproducibility (Supplementary Fig. 1), indicating that even small
324 differences in the fraction of embryos generating endoderm could be reproducibly measured.

325 Further, we found that some wild isolates that were subsequently found to have identical
326 genome sequences also gave identical results.

327 Although birefringent and autofluorescent rhabditin granules have been used as a
328 marker of gut specification and differentiation in many studies [50,51], it is conceivable that
329 the variation in fraction of embryos containing this marker that we observed might reflect
330 variations in gut granule formation rather than in gut differentiation *per se*. We note that
331 embryos from all strains showed a decisive "all-or-none" phenotype: i.e., they were either
332 strongly positive for gut differentiation or completely lacked gut granules, with virtually no
333 intermediate or ambiguous phenotypes. A threshold of gene activity in the GRN has been
334 shown to account for such an all-or-none switch in gut specification [22,67,68]. This
335 observation is inconsistent with possible variation in gut granule production: if SKN-1-
336 depleted embryos were defective in formation of the many granules present in each gut cell,
337 one might expect to observe gradations in numbers or signal intensity of these granules
338 between gut cells or across a set of embryos. Nonetheless, we extended our findings by
339 analyzing the expression of the gut-specific intermediate filament IFB-2, a marker of late gut
340 differentiation, in selected strains representing the spectrum of phenotypes observed (Fig.
341 2B). As with gut granules, we found that embryos showed all-or-none expression of IFB-2. In
342 all cases, we found that the fraction of embryos containing immunoreactive IFB-2 was not
343 significantly different (Fisher's exact test, p-values > 0.05) from the fraction containing gut
344 granules, strongly suggesting that the strains vary in endoderm specification *per se* and
345 consistent with earlier studies of SKN-1 function [19,22].

346 Although we found that *skn-1(RNAi)* was 100% effective at inducing embryonic
347 lethality in all strains (with the exception of the RNAi-defective Hawaiian strain, CB4856), it is

348 conceivable that, at least for the strains that showed a weaker phenotype than for N2 (i.e.,
349 higher number of embryos specifying endoderm), the variation observed between strains was
350 attributable to differences in RNAi efficacy rather than in the endoderm GRN. Indeed, studies
351 with N2 and CB4856 showed that germline RNAi sensitivity is a quantitative trait, involving
352 the Argonaute-encoding *ppw-1* gene and additional interacting loci present in some wild
353 isolates [69,70]. To address this possibility, we introgressed the strong loss-of-function *skn-*
354 *1(zu67)* chromosomal mutation into five wild isolates whose phenotypes spanned the
355 spectrum observed (ranging from 2% of embryos with differentiated gut for MY16 to 50% for
356 MY1) (Fig. 2C). In all cases, we found that introgression of the allele through five rounds of
357 backcrosses resulted in a quantitative phenotype that was similar or indistinguishable from
358 that observed with *skn-1(RNAi)*. The phenotypes of the introgressed allele were significantly
359 different (p-values <0.01) from that of the parental N2 *skn-1(zu67)* strain, except for DL238,
360 whose *skn-1(RNAi)* phenotype was indistinguishable from that of N2. The results obtained by
361 introgression from four of the isotypes (CX11262, DL238, EG4724 and MY1), were not
362 statistically different (Student t-test, p-values >0.05) from the corresponding RNAi
363 knockdown results (Fig. 2C) (i.e., the phenotype was suppressed or enhanced relative to N2
364 in these genetic backgrounds to the same extent as with *skn-1(RNAi))*. However, while the
365 MY16 *skn-1(zu67)* strain shifted in the predicted direction (i.e., became stronger) as
366 compared to the N2 strain, it showed a weaker phenotype than was evident by RNAi
367 knockdown, even following eight rounds of introgression. Regardless, diminished RNAi
368 efficacy in MY16 cannot explain the large difference between the *skn-1(RNAi)* phenotype of
369 N2 and MY16, as the latter phenotype is, in fact, much stronger, not weaker, than the former.
370 As described below, we identified a modifier locus in the MY16 strain that is closely linked to
371 the *skn-1* gene; it therefore seems likely that the N2 chromosomal segment containing this

372 modifier was carried with the *skn-1(zu67)* mutation through the introgression crosses,
373 thereby explaining the somewhat weaker phenotype of the introgressed allele in MY16. We
374 conclude that the extreme variation in *skn-1(RNAi)* phenotype between the wild isolates
375 tested results from *bona fide* cryptic variation in the endoderm GRN, rather than differences
376 in RNAi efficacy.

377 We note that the strength of *skn-1(RNAi)* phenotype does not correlate with
378 phylogenetic relatedness between the strains (Pagel's $\lambda = 0.42$, p-value = 0.14). For example,
379 while some closely related strains (e.g., MY16 and MY23) showed a similar phenotype, other
380 very closely related strains (e.g., JU1491 and JU778) showed phenotypes on the opposite ends
381 of the phenotypic spectrum (Fig. 3A). We also did not observe any clear association between
382 geographical distribution and *skn-1 (RNAi)* phenotype (Fig. 3B). These findings suggest that
383 the initiating inputs into the endoderm GRN is subject to rapid intraspecies evolutionary
384 divergence.

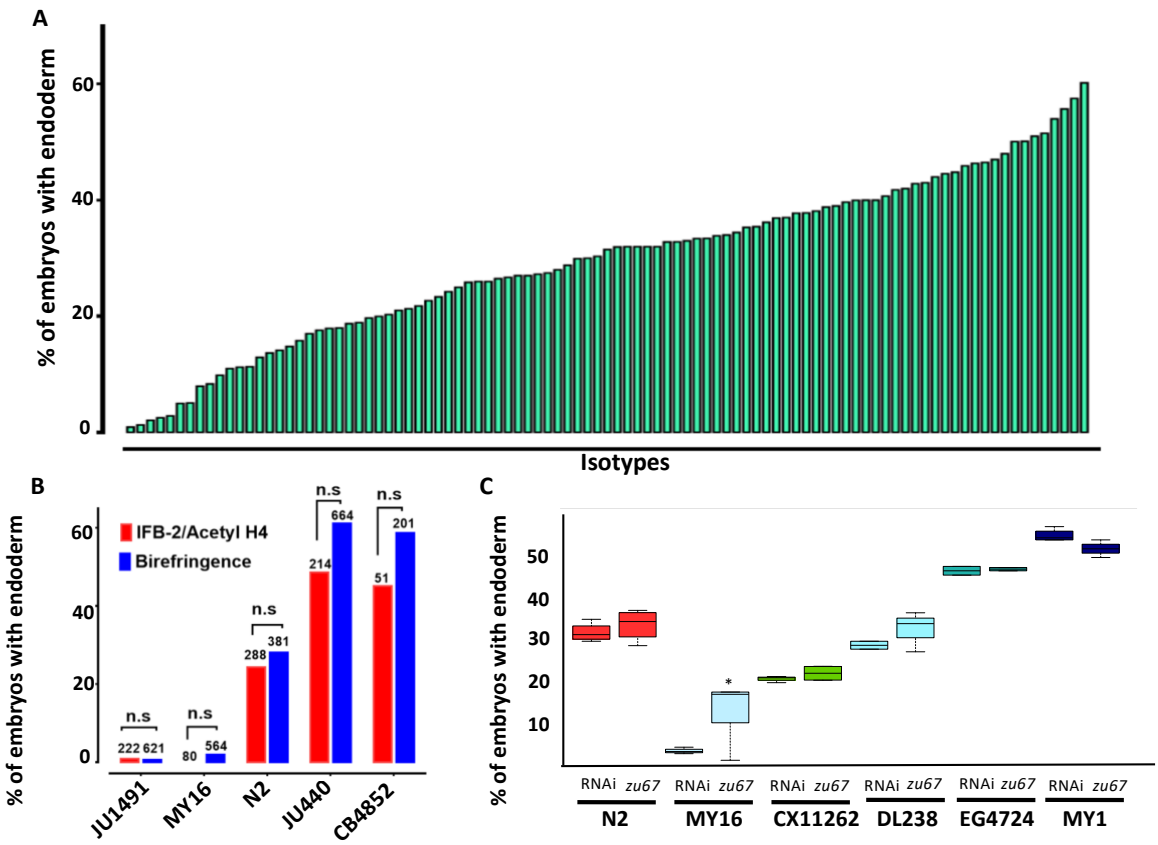


Fig. 2: Quantitative effects of loss of *skn-1* on endoderm formation.

(A) Spectrum of *skn-1(RNAi)* effects across the *C. elegans* isolates. The effects of *skn-1(RNAi)* are quantified as the average percentage of arrested embryos with endoderm (y-axis). All wild isolates treated with *skn-1(RNAi)* resulted in 100% embryonic arrest (n >500 embryos per replicate per isotype and at least two replicates per isotype). (B) Comparison of *skn-1(RNAi)* phenotype using two different gut markers (birefringent gut granules and MH33 staining of IFB-2) in five different genetic backgrounds. In all cases, no significant statistical difference was found between the two quantitative methods. Fisher's exact test (NS p-value>0.05). (C) Comparison of *skn-1(RNAi)* and *skn-1(zu67)* effects on endoderm development in six different genetic backgrounds. For each color-coded strain, the first value is of the *skn-1(RNAi)* results (five replicates), while the second is the result for the *skn-1(zu67)* allele introgression (10 replicates). For all strains (with the exception of MY16), no significant statistical difference was found between the RNAi knockdown and corresponding *skn-1(zu67)* allele effects on endoderm development. Student t-test (NS p-value > 0.05, * p-value < 0.05).

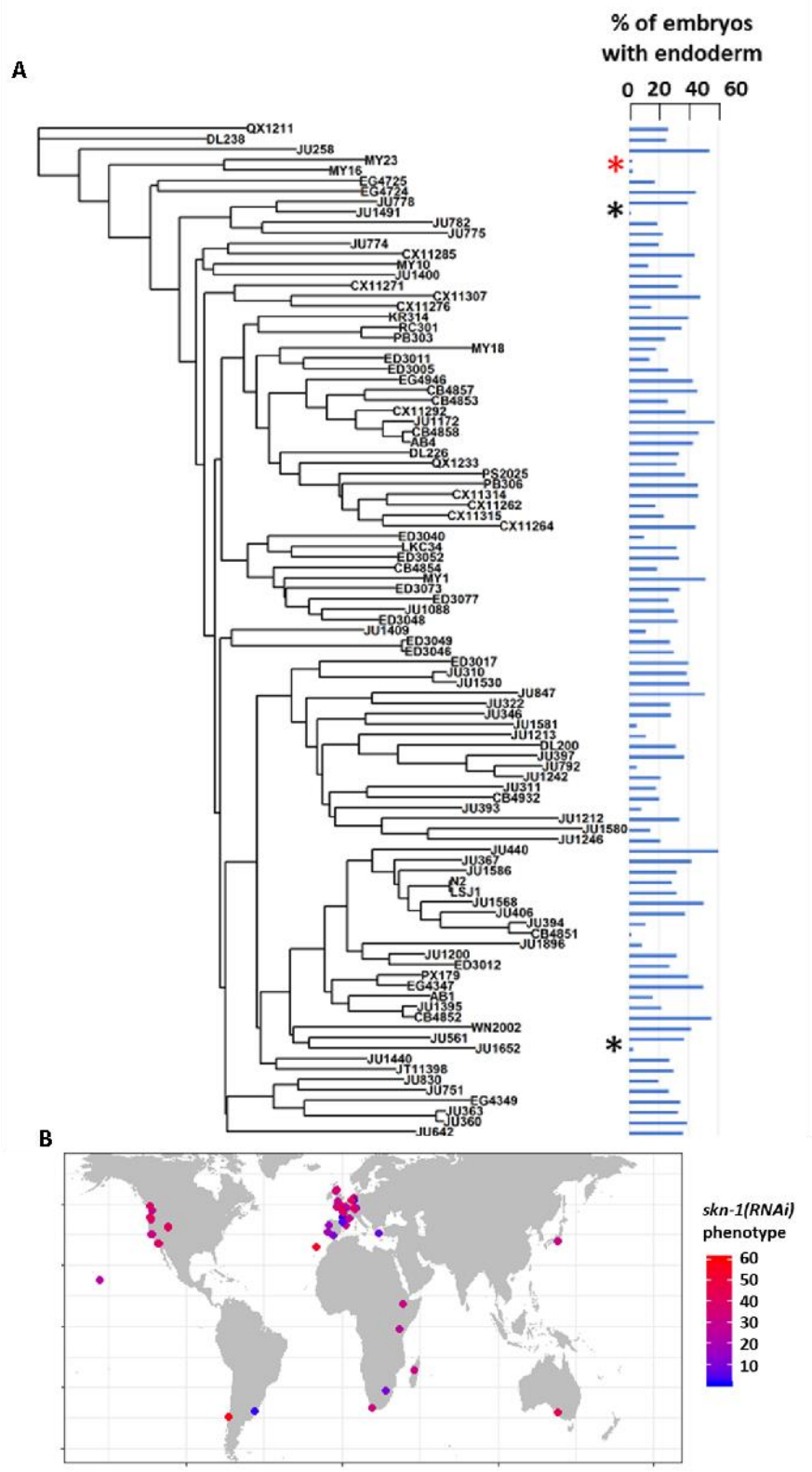


Fig. 3: SKN-1 requirement does not correlate with genotypic relatedness or geographical location. (A) *skn-1(RNAi)* phenotype of 97 isolates arranged with respect to the neighbor-joining tree constructed using 4,690 SNPs and pseudo-rooted to QX1211. Red asterisk indicates an example of closely related strains (MY23 and MY16) with similar phenotype, while black asterisks indicate example sister strains (JU778 and JU1491; JU561 and JU1652) with distinct phenotype. Phylogenetic relatedness and phenotype are not significantly correlated (Pagel's λ = 0.42, p-value = 0.14). (B) Worldwide distribution of *skn-1(RNAi)* phenotype across 97 wild isolates. Each circle represents a single isotype.

407 **Cryptic variation in the quantitative requirement for MOM-2/Wnt, but not POP-1, in**
408 **endoderm development**

409 The switch in the relationship of the SKN-1 and Wnt inputs between *C. elegans* ("OR"
410 operator) and *C. briggsae* ("AND" operator) [39], and the extensive variation in the
411 requirement for SKN-1 seen across *C. elegans* isolates, raised the possibility that the
412 quantitative requirement for Wnt components might vary between unique isolates of *C.*
413 *elegans*. It has been shown that signaling from Ras pathway varies in different *C. elegans* wild
414 isolates and hyperactive Wnt signaling can compensate for reduced Ras activity in the vulva
415 signaling network [6,71]. Given that removal of the maternal Wnt input also results in a
416 partially penetrant gut defect (through either knock-out or knockdown of Wnt signaling
417 components), it is conceivable that a compensatory relationship may exist between the SKN-
418 1 and Wnt inputs. We investigated this possibility by examining the requirement for the
419 MOM-2/Wnt ligand in the same wild isolates that were tested for the SKN-1 gut
420 developmental requirement. Indeed, we observed broad variation in the requirement for
421 MOM-2/Wnt in activation of the endoderm GRN between isotypes. *mom-2(RNAi)* of 94
422 isotypes resulted in embryonic arrest, indicating that, as with *skn-1(RNAi), mom-2(RNAi*) was
423 effective at least by the criterion of lethality. Two isotypes, CB4853 and EG4349, did not
424 exhibit *mom-2(RNAi)*-induced lethality and were omitted from further analyses. In the
425 affected strains, the fraction of *mom-2(RNAi)* embryos with differentiated gut varied from
426 ~40% to ~99% (Fig. 4A, Supplementary File 1). As with *skn-1(RNAi)*, the *mom-2(RNAi)*
427 phenotype of isotypes N2, JU440, and JU1213 was further confirmed by immunostaining with
428 IFB-2 (Fig 4B), again demonstrating that birefringence of gut granules is a reliable proxy for
429 endoderm formation for this analysis.

430 To assess whether the observed variation in the *mom-2(RNAi)* phenotype reflected
431 differences in the GRN or RNAi efficacy, the *mom-2(or42)* allele was introgressed into three
432 different genetic backgrounds chosen from the extreme ends of the phenotypic spectrum.
433 *mom-2(RNAi)* of the laboratory N2 strain resulted in the developmental arrest of embryos. Of
434 those, ~90% contained differentiated endoderm, a result that was highly reproducible. In
435 contrast, the introgression of an apparent loss-of-function allele, *mom-2(or42)*, into the N2
436 strain results in a more extreme phenotype: only ~28% of embryos show endoderm
437 differentiation (Fig. 4C) [30]. While this discrepancy can partly be explained by incomplete
438 RNAi efficacy, it is notable that the penetrance of *mom-2* alleles vary widely [30]. We
439 observed strain-specific variation in embryonic lethality response to RNAi of *mom-2* between
440 the different isotypes. However, we found that the *mom-2(or42)* introgressed strains show
441 qualitatively similar effects to those observed with *mom-2* RNAi. For example, the *mom-*
442 *2(or42)* allele introgressed into the isotype JU1213 background resulted in a low fraction of
443 arrested embryos with gut (5.7% ± s.d 2.4%; n=2292), a more extreme effect than was seen
444 with RNAi (34.0% ± s.d 1.5%; n=1876). This is the strongest phenotype that has been reported
445 for any known *mom-2* allele. On the other hand, introgression of the *mom-2* mutation gave
446 rise to a significantly higher fraction of embryos with endoderm in isotypes DL226 (55.2% ±
447 s.d 1.2%, n=1377) and PB303 (65.5% ± s.d 4.9%, n=2726), relative to the laboratory strain N2
448 (29.1% ± s.d 3.1%; n=1693), consistent with the RNAi phenotypes (Fig. 4C). These findings
449 indicate that the differential requirement for MOM-2 is, at least in part, attributable to
450 genetic modifiers in these strains. As with *skn-1(RNAi),* we found no correlation between the
451 *mom-2(RNAi)* phenotype and phylogenetic relatedness or geographical distribution
452 (Supplementary Fig. 5), suggesting rapid intraspecies developmental system drift.

453 As the MOM-2/Wnt signal is mediated through the POP-1 transcription factor, we
454 sought to determine whether the requirement for POP-1 might also vary between isolates.
455 We found that, while *pop-1(RNAi)* resulted in 100% embryonic lethality across all 96 RNAi-
456 sensitive isolates, 100% of the arrested embryos contained a differentiated gut (n>500 for
457 each isolate scored) (results not shown). Thus, all isolates behave similarly to the N2 strain
458 with respect to the requirement for POP-1. These results were confirmed by introgressing a
459 strong loss-of-function *pop-1(zu189)* allele into four wild isolates (N2, MY16, JU440, and
460 KR314) (Supplementary Fig. 2). The lack of variation in endoderm specification after loss of
461 POP-1 is not entirely unexpected. As has been observed in a *pop-1(-)* mutant strain,
462 elimination of the endoderm-repressive role of POP-1 in the MS lineage (which is not
463 influenced by the P2 signal) supersedes its endoderm activating role in the presence of SKN-
464 1. Indeed, the original observation that all *pop-1(-)* embryos in an N2 background contain gut
465 masked the activating function for POP-1, which is apparently only in the absence of SKN-1
466 [33,35,37]. It is likely that, as with the N2 strain, gut arises from both E and MS cells in all of
467 these strains; however, as we have scored only for presence or absence of gut, it is
468 conceivable that the E lineage is not properly specified in some strains, a possibility that
469 cannot be ruled out without higher resolution analysis.

470 Our results contrast with those of Paaby *et al.* [72], who reported that RNAi of 29
471 maternal-effect genes across a set of 55 wild isolated in liquid culture resulted in generally
472 weaker effects on lethality than we observed. This difference is likely attributable to
473 diminished and variable RNAi efficacy in the latter study owing to the different culture
474 methods used (see MATERIALS AND METHODS) [48,73]. To assess this possibility further, we
475 compared our results with those of Paaby *et al.* (2015) and found no correlation between the
476 variation in fraction of embryos with gut and the lethality observed in that report with both

477 *mom-2(RNAi)* and *skn-1(RNAi)* (Pearson's R = 0.19, p = 0.23; Pearson's R = 0.22, p = 0.17,
478 respectively). In addition, Paaby *et al.* (2015) found that the genetically divergent strain
479 QX1211 consistently showed weak penetrance across all targeted genes, while under our
480 experimental conditions, QX1211 exhibited a slightly stronger skn-1(RNAi)phenotype (25.2%
481 *vs.* 32.0%, Fisher's exact test p-value = 0.03) and a similar mom-2(RNAi) phenotype (90% *vs*.
482 90%, Fisher's exact test p-value = 0.9) compared to the N2 strains with fully penetrant
483 lethality in all cases.

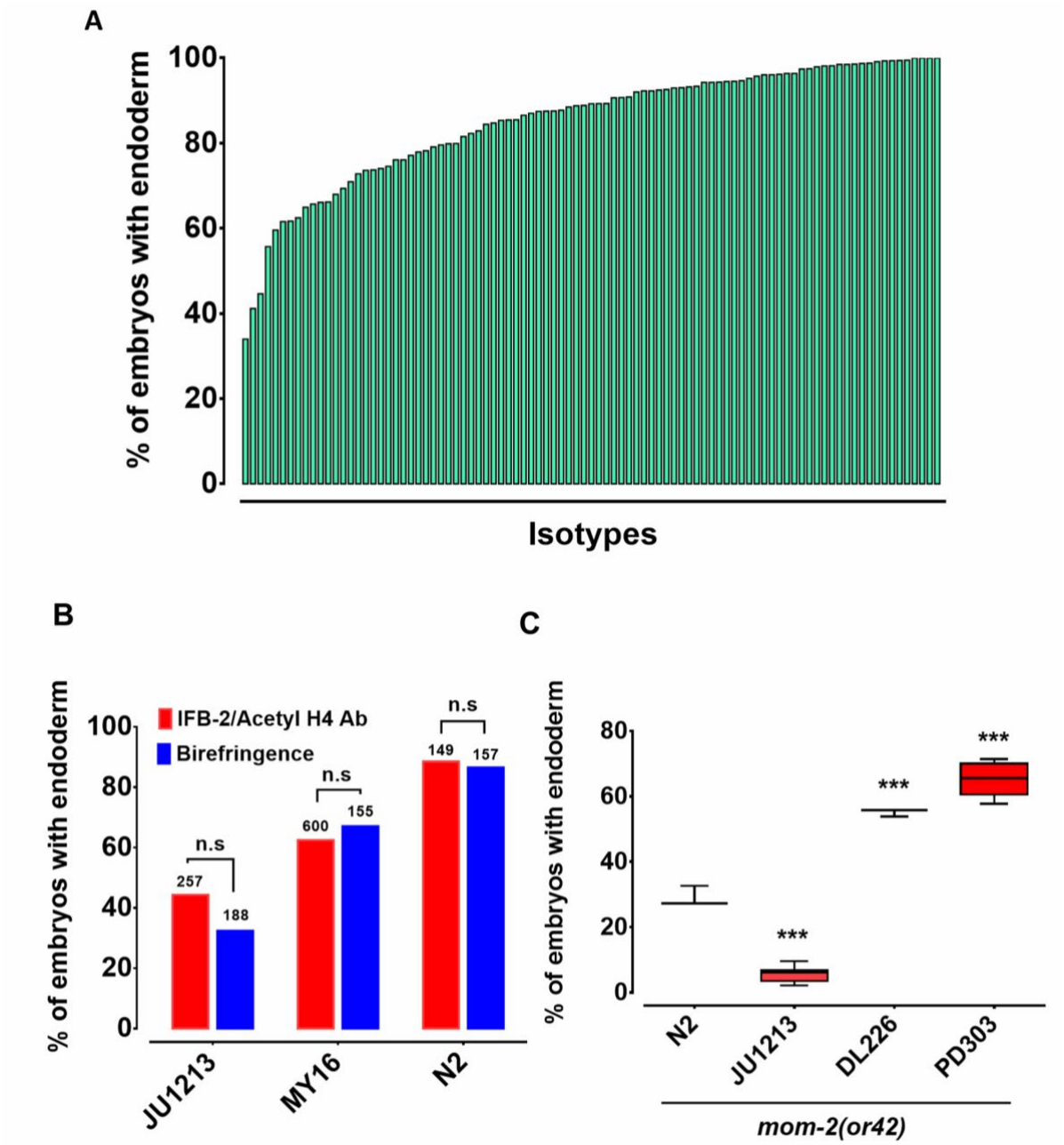


Fig. 4: Widespread variation in the *mom-2(RNAi)* phenotype.

(A) Spectrum of *mom-2(RNAi)* effects across the *C. elegans* isolates. The effects of *mom-2(RNAi)* are quantified as the average percentage of arrested embryos with endoderm (y-axis). Each column represents the mean for each wild isolate (n >500 embryos were scored for each experiment with at least two replicates per isotype). (B) Comparison of *mom-2(RNAi)* phenotype using two different gut markers (birefringent gut granules and MH33 immunostaining of IFB-2) in three different genetic backgrounds. In all cases, no significant statistical difference was found between the two quantitative methods. Fisher's exact test (NS p-value > 0.05). (C) Comparison of the effect of *mom-2(or42)* on endoderm development after introgression into four different genetic backgrounds. At least three independent introgressed lines were studied for each wild isotype. The results were compared *to N2; mom-2(or42*). Student t-test (*** p-value < 0.001).

496 **Genome-wide association studies (GWAS) and analysis of RILs identify multiple genomic**
497 **regions underlying variation in the two major endoderm GRN inputs**

498 We sought to examine the genetic basis for the wide variation in SKN-1 and Wnt
499 requirements across *C. elegans* isolates and to evaluate possible relationships in the variation
500 seen with the SKN-1 and Wnt inputs by performing GWAS using the available SNP markers
501 and map [42], adjusting for population structure by using Efficient Mixed-Model Analysis
502 (EMMA) (Fig. 5A, B) [52,74]. This approach identified two significant closely-located positions
503 on chromosome IV that underlie the variation in SKN-1 requirement (Fig. 5A, Supplementary
504 Table 1).

505 GWAS of the *mom-2(RNAi)* variation proved more challenging because this phenotype
506 showed a highly skewed distribution (Shapiro-Wilk' test W =0.8682, p-value = $1.207X10^{-7}$)
507 (Supplementary Fig. 3). While GWAS did not reveal any genomic regions for the *mom-2(RNAi)*
508 variation that exceeded an FDR of 5%, we found that the most strongly associated loci for the
509 *mom-2(RNAi)* phenotype also showed large effects for *skn-1(RNAi)* (Fig. 5C). In particular, we
510 observed substantial overlap in the p-values for individual SNPs from *skn-1(RNAi)* and *mom-*
511 *2(RNAi)* in the central region of chromosome IV (Supplementary Fig. 4), raising the possibility
512 that common genetic factors might underlie these phenotypes.

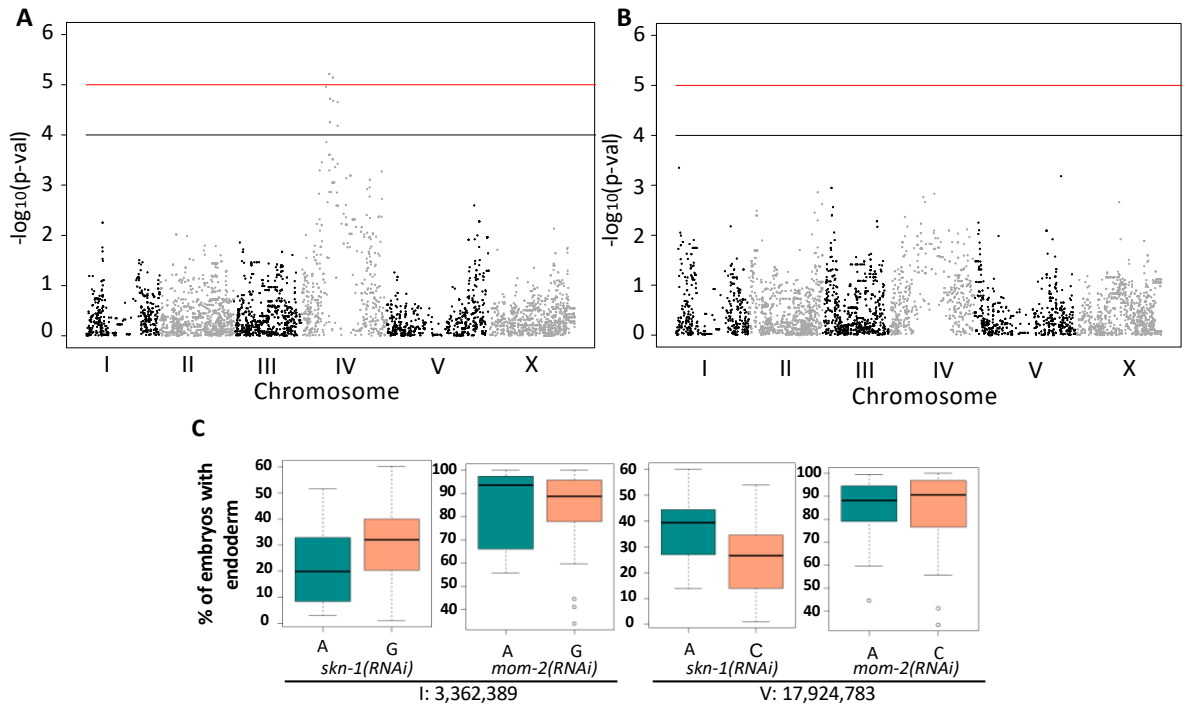


Fig. 5. Genome-Wide Association Studies of *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes.

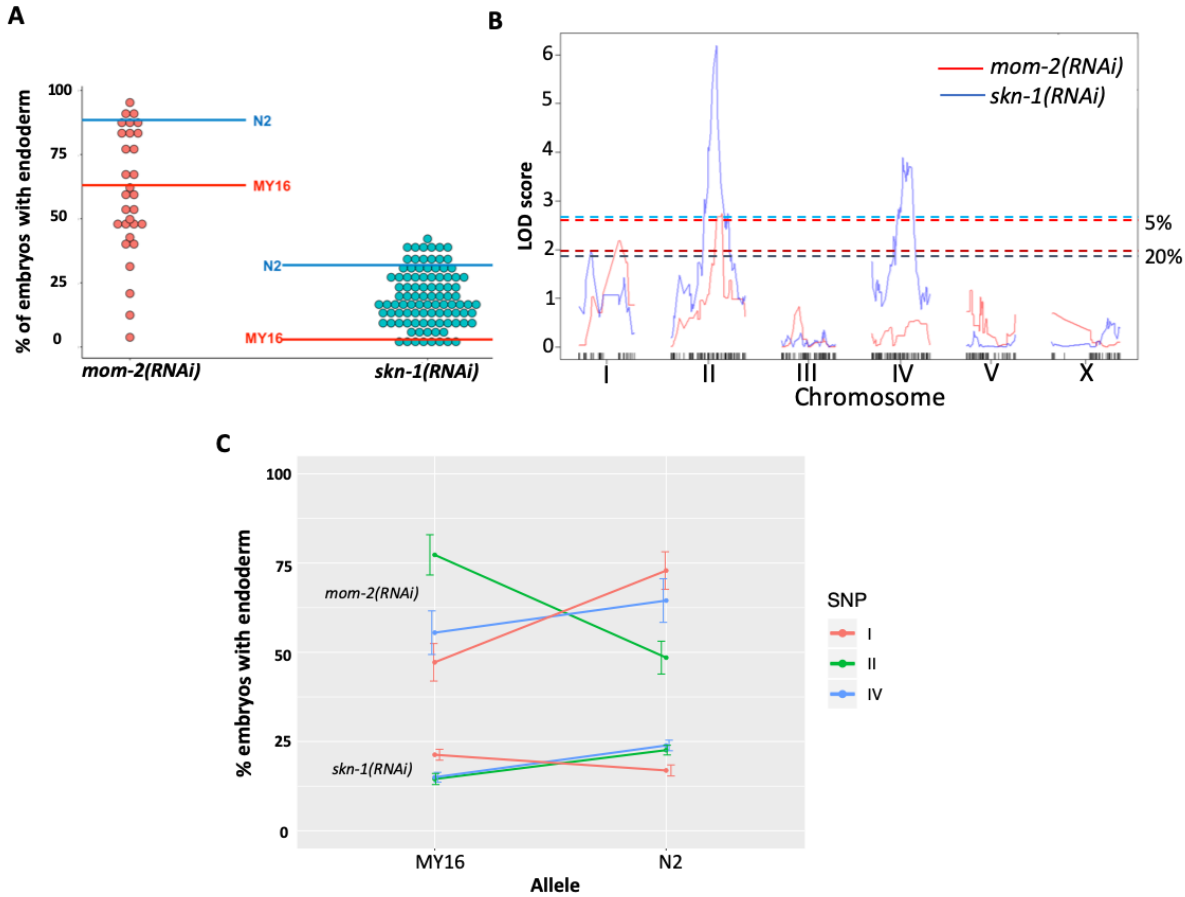
(A) Manhattan plot of *skn-1(RNAi)* GWAS. The red line indicates a genome-wide 1.5% FDR (permutation-based FDR, from 10,000 permutated results). Black line represents 3.0% FDR. The y axis is the $-\log_{10}$ of p-value. (B) Manhattan plot of *mom-2 (RNAi)* EMMA. The y axis is the –log10 of p-value. Genomic regions are shown on the x-axis. (C) Effect plots of the most strongly-linked SNPs from *mom-2(RNAi)* GWAS at position 3,362,389 bp on chromosome I and position 17,924,783 bp on chromosome V. Horizontal lines within each box represent the median, and the boxes represent 25th–75th percentile.

In an effort to narrow in on causal loci underlying the *skn-1(-)* and *mom-2(-)* phenotypic variation, and to assess possible relationships between these two GRN inputs, we prepared and analyzed 95 recombinant inbred lines (RILs) between two *C. elegans* isotypes, N2 and MY16. These strains were chosen for their widely varying differences in requirement for both inputs (see MATERIALS AND METHODS). In contrast to the very low variation seen between multiple trials of each parental strain, analysis of the RNAi-treated RIL strains (>500 embryos/RIL) revealed a very broad distribution of phenotypes. We found that, while some RILs showed phenotypes similar to that of the two parents, many showed intermediate phenotypes and some were reproducibly more extreme than either parent, indicative of transgressive segregation [75]. For *skn-1(RNAi)*, the phenotype varied widely across the RILs,

531 with 1 to 47% of embryos containing gut (Fig. 6A, Supplementary File 2). This effect was even
532 stronger with *mom-2(RNAi)*, for which virtually the entire possible phenotypic spectrum was
533 observed across a selection of 31 RILs representing the span of *skn-1(RNAi)* phenotypes. The
534 *mom-2(RNAi)* phenotypes ranged from RILs showing 3% of embryos with gut to those
535 showing 92% (Fig. 6A). In all RILs, *skn-1(RNAi)* and *mom-2(RNAi)* resulted in 100% lethality. It
536 is noteworthy that one RIL (JR3572, Supplementary File 2) showed a nearly completely
537 penetrant gutless phenotype, an effect that is much stronger than has been previously
538 observed for *mom-2(-)* [30]. These results indicate that a combination of natural variants can
539 nearly eliminate a requirement for MOM-2 altogether, while others make it virtually essential
540 for endoderm development. Collectively, these analyses reveal that multiple quantitative
541 trait loci (QTL) underlie SKN-1- and MOM-2-dependent endoderm specification.

542 To identify QTLs from the recombinant population, we performed linkage mapping for
543 both phenotypes using interval mapping (see MATERIALS AND METHODS). For *skn-1(RNAi)*,
544 two major peaks were revealed on chromosomes II and IV (above 1% FDR estimated from
545 1,000 permutations). Two minor loci were found on chromosomes I and X (suggestive linkage,
546 above 20% FDR) (Fig 6B). For *mom-2(RNAi)*, two major independent QTL peaks were found
547 on chromosomes I and II (above the 5% FDR estimated from 1,000 permutations). Although
548 the candidate peaks observed on chromosome IV for *skn-1(RNAi)* (Fig. 6B) did not appear to
549 overlap with those for *mom-2(RNAi*), overlap was observed between the chromosomes I and
550 II candidate regions for these two phenotypes. These QTLs show large individual effects on
551 both phenotypes (Fig. 6C),

552



553

554 **Fig 6. Quantitative genetic analysis of *mom-2(RNAi)* and *skn-1(RNAi)* phenotype in Recombinant Inbred**
555 **Lines (RILs) between N2 and MY16.**

556 (A) *mom-2(RNAi)* (left) and *skn-1(RNAi)* (right) phenotype of RILs. The phenotype of the parental strains,
557 MY16 and N2 are shown by red and blue lines, respectively. (B) QTL analyses (interval mapping) of *skn-*
558 *1(RNAi)* (blue line) and *mom-2(RNAi)* (red line) phenotype shown in (A). Genomic regions are shown on the
559 x-axis and LOD score is shown on the y-axis. Significance thresholds for *mom-2(RNAi)* and *skn-1(RNAi)* at
560 5% and 20% linkage represented in red and blue dashed lines, respectively. (C) Effect plots of significant
561 SNPs from *mom-2(RNAi)* and *skn-1(RNAi)*, indicated by chromosome number and color, showing the
562 direction of the allelic effects. Confidence intervals for the average phenotype in each genotype group are
563 shown.

564

565 **Potential cryptic relationships between SKN-1 and MOM-2 inputs**
566
567 The preceding findings unveiled wide cryptic variation in the requirements for both
568 SKN-1 and MOM-2/Wnt in the endoderm GRN, raising the possibility that the variation
569 affecting the two inputs might be related. Indeed, comparisons of the GWAS and QTL

570 mapping results for *skn-1* and *mom-2* showed an overlap in candidate QTL regions on
571 chromosome I, II and IV (Fig. 5, Fig 6, Supplementary Fig. 4), suggesting a possible connection
572 between the genetic basis underlying these two traits. It is conceivable that some genetic
573 backgrounds are generally more sensitive to loss of either input (e.g., the threshold for
574 activating the GRN is higher) and others more robust to single-input loss. Alternatively, a
575 higher requirement for one input might be associated with a relaxed requirement for the
576 other, i.e., a reciprocal relationship.

577 As an initial assessment of these alternatives, we examined whether the requirements
578 for SKN-1 and MOM-2 across the strains were significantly correlated. This analysis revealed
579 no strong relationship between the cryptic variation in the requirement for these inputs seen
580 across all the strains (Spearman correlation R=0.18, p-value=0.07) (Fig. 7A). This apparent lack
581 of correlation at the level of strains is not unexpected, as many factors likely contribute to the
582 cryptic variation and the comparison reflects the collective effect of all causal loci in the
583 genome of each strain (Fig. 5, 6).

584 We next sought to examine possible relationships between the two GRN inputs at
585 higher resolution by comparing association of specific genetic regions with the quantitative
586 requirement for each input. We used the available sequencing data for all isotypes tested [42]
587 and examined the impact of each allele on the *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes,
588 correcting for outliers and using a pruned SNP map (see MATERIALS AND METHODS). We
589 found a weak positive correlation (Pearson's R = 0.21, p= p-value < 2.2e-16, Supplementary
590 Fig. 6) between the allelic effects. One possible explanation for this observation might be that
591 variants across the set of wild isolates may generally influence the threshold for activating the
592 positive feedback loops that lock down gut development [49,67], thereby altering the

593 sensitivity for regulatory inputs into the endoderm pathway. Alternatively, although evidence
594 for variation in germline RNAi sensitivity among *C. elegans* wild isolates is lacking (except for
595 CB4856, which has been omitted from our study) [69,76], and we have shown above that
596 variation in SKN-1 and MOM-2 requirement reflects in large part cryptic genetic differences
597 in the endoderm GRN, it remains possible that a fraction of the variation found in the two
598 phenotypes tested is attributable to varying RNAi penetrance, which may underlie the minor
599 positive correlation between *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes.

600 In contrast, analysis of the N2/MY16 RILs uncovered a potential reciprocal relationship
601 between the requirements for SKN-1 and MOM-2: we observed a negative correlation
602 between the *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes across the genome (Fig 7B, genome-
603 wide Pearson's R = -0.35, p = 0.001, correcting for LD and outliers as above; correlation
604 without chromosome IV R = -0.59, p < 0.001). This finding suggested that at least some
605 quantitative variants result in opposing effects on the requirement for SKN-1 and MOM-2.

606 While a reciprocal relationship was observed generally across the genome spanning
607 five of the chromosomes, we observed the opposite correlation on chromosome IV (Pearson's
608 R = 0.83, p-value = 1.695×10^{-5}). No correlation was observed for chromosome IV with the wild
609 isolates (Pearson's R = 0.08, NS, Supplementary Fig. 6). As there is a major QTL on
610 chromosome IV for the SKN-1 requirement and there is substantial overlap in the same region
611 with the GWAS analysis of the *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes, we sought to
612 dissect further the relationship between the requirement for MOM-2 and SKN-1 in this
613 region. We created six near-isogenic lines (NILs) in which the QTL region for the *skn-1(RNAi)*
614 phenotype on chromosome IV from N2 was introgressed into the MY16 background, and *vice-*
615 *versa* (Supplementary Fig. 7). Control lines were created from the same crosses at the same

616 generation by selecting the original parental region (e.g., selecting for the N2 region in an N2
617 background and MY16 in MY16 background). We found that the region affects the *skn-*
618 *1(RNAi)* phenotype as expected: the N2 region increased the fraction of gut in an MY16
619 background, and the MY16 regions decreased this fraction in an N2 background. However,
620 for *mom-2(RNAi),* while introgressing the N2 region in MY16 dramatically changed the
621 phenotype (Fig. 7C), we found that the MY16 region was not sufficient to alter the phenotype
622 in an N2 background. We created segregant NILs in which one of the genetic markers was lost
623 (see MATERIALS AND METHODS) and found that replacing the N2 region with the
624 corresponding MY16 region in all cases results in a stronger *mom-2(RNAi)* phenotype.
625 However, for the *skn-1(RNAi)* phenotype six of nine segregants showed the opposite effect:
626 i.e., a weaker phenotype (Fig. 7D), revealing that when contributing variants were separated
627 by recombination, a reciprocal effect was frequently seen. These observations suggest that
628 complex genetic interactions between variants on chromosome IV might mask the potential
629 reciprocal effects that were observed on the other chromosomes.

630

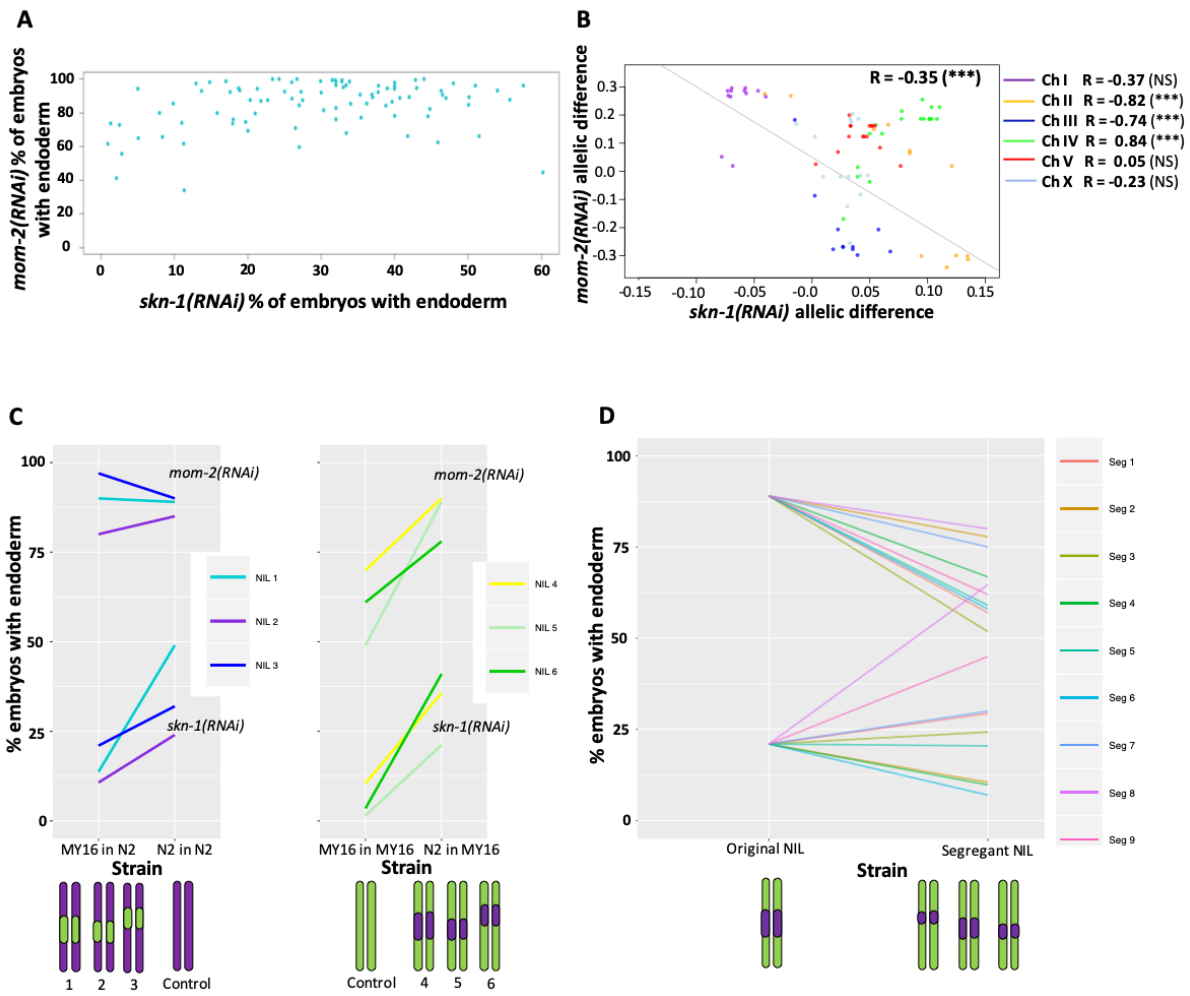


Fig. 7: Correlation of *skn-1(RNAi)* and *mom-2(RNAi)* allelic differences.

(A) Comparison of *skn-1(RNAi)* and *mom-2(RNAi)* phenotype in 94 natural isolates tested. No correlation was found (Spearman correlation R=0.1844, p-value=0.07). Each dot corresponds to a wild isolate. Y-axis, *skn-1(RNAi)* phenotype, x-axis, *mom-2(RNAi)* phenotype. (B) Genome-wide correlation of *skn-1(RNAi)* and *mom-2(RNAi)* allelic differences in the N2xMY16 RILs. Each dot represents a SNP. Chromosomes are color-coded with their Pearson's R values represented (NS = Not Significant, *** = p-value < 0.001). Regression line in grey. (C) Six different NILs were created for chromosome IV, each of which was compared with a control NIL from the same cross (e.g., MY16 in MY16 as control for N2 in MY16). A schematic of the introgressed regions is represented below the plots. Percentage of *skn-1(RNAi)* or *mom-2(RNAi)* embryos with gut is represented. (D) Changes in phenotype for both *skn-1(RNAi)* and *mom-2(RNAi)* following recombination in segregant NILs, which a schematic representation of the segregant NILs below the plot.

Multiple factors reciprocally regulate the requirement for SKN-1 and MOM-2/Wnt

While the above findings revealed that the relationship between the requirement for SKN-1 and MOM-2 may be complicated by genetic interactions, our results raised the

647 possibility of compensatory relationships between them. To further assess this possibility, we
648 tested other candidate genes that reside in the QTL regions and that have been implicated in
649 endoderm development [77–79]. We found that loss of RICT-1, the *C. elegans* orthologue of
650 the human RICTOR (Rapamycin-insensitive companion of mTOR; [80]), a component of the
651 TORC2 complex, which has been shown to antagonize SKN-1 function [77], results in opposite
652 effects on *skn-1(-)* and *mom-2(-)* mutants (Fig. 8A). Specifically, while *rict-1(RNAi)* suppresses
653 the absence of gut in *skn-1(zu67)* embryos (*skn-1(zu67*): 34.3% ± s.d 4.1% with gut *vs. skn-*
654 *1(zu67)*; *rict-1(RNAi)*: 48.3% ± s.d 4.9%; p=<0.001), we found that it *enhances* this phenotype
655 in *mom-2(or42)* mutants (*mom-2(or42)*: 23.8% ± s.d 2.0%; vs. *mom-2(or42)*; *rict-1(RNAi)*:
656 11.2% ± s.d 3.2%; p<0.001). Confirming this effect, a similar outcome was observed when
657 SKN-1 was depleted by RNAi in *rict-1(ft7)* chromosomal mutants (*skn-1(RNAi)*: 31.6% ± s.d
658 4.3% with gut vs. *rict-1(ft7); skn-1(RNAi)*: 45.9% ± s.d 6.3%; p<0.05) (Fig. 8A). Similarly, RNAi
659 depletion of PLP-1, the *C. elegans* homologue of the Pur alpha transcription factor that has
660 been shown to bind to and regulate the *end-1* promoter [78], reciprocally affects the outcome
661 of removing these two inputs in the same direction: loss of PLP-1 function suppresses the *skn-*
662 *1(-)* phenotype (to 48.0% ± s.d 6.6%), and strongly enhances the *mom-2* phenotype (to 6.9%
663 ± s.d 1.6%). Again, this result was confirmed by RNAi of *skn-1* in a *plp-1(ok2156)* chromosomal
664 mutant (Fig. 8B). Thus, as observed with the effect across the genome with natural variants,
665 we observed a substantial reciprocal effect of both of these genes on loss of SKN-1 and MOM-
666 2.

667 We also observed a reciprocal effect on the SKN-1 and Wnt inputs with MIG-
668 5/*dishevelled*, a component of the Wnt pathway that acts downstream of the Wnt receptor
669 [79]; however, in this case the effect was in the opposite direction as seen for RICT-1 and PLP-
670 1. Loss of MIG-5 as a result of chromosomal mutation or RNAi leads to *enhancement* of the

671 *skn-1(-)* phenotype (*mig-5(rh94); skn-1(RNAi)*: 6.6% ± s.d 2.3%; *skn-1(zu67); mig-5(RNAi)*:

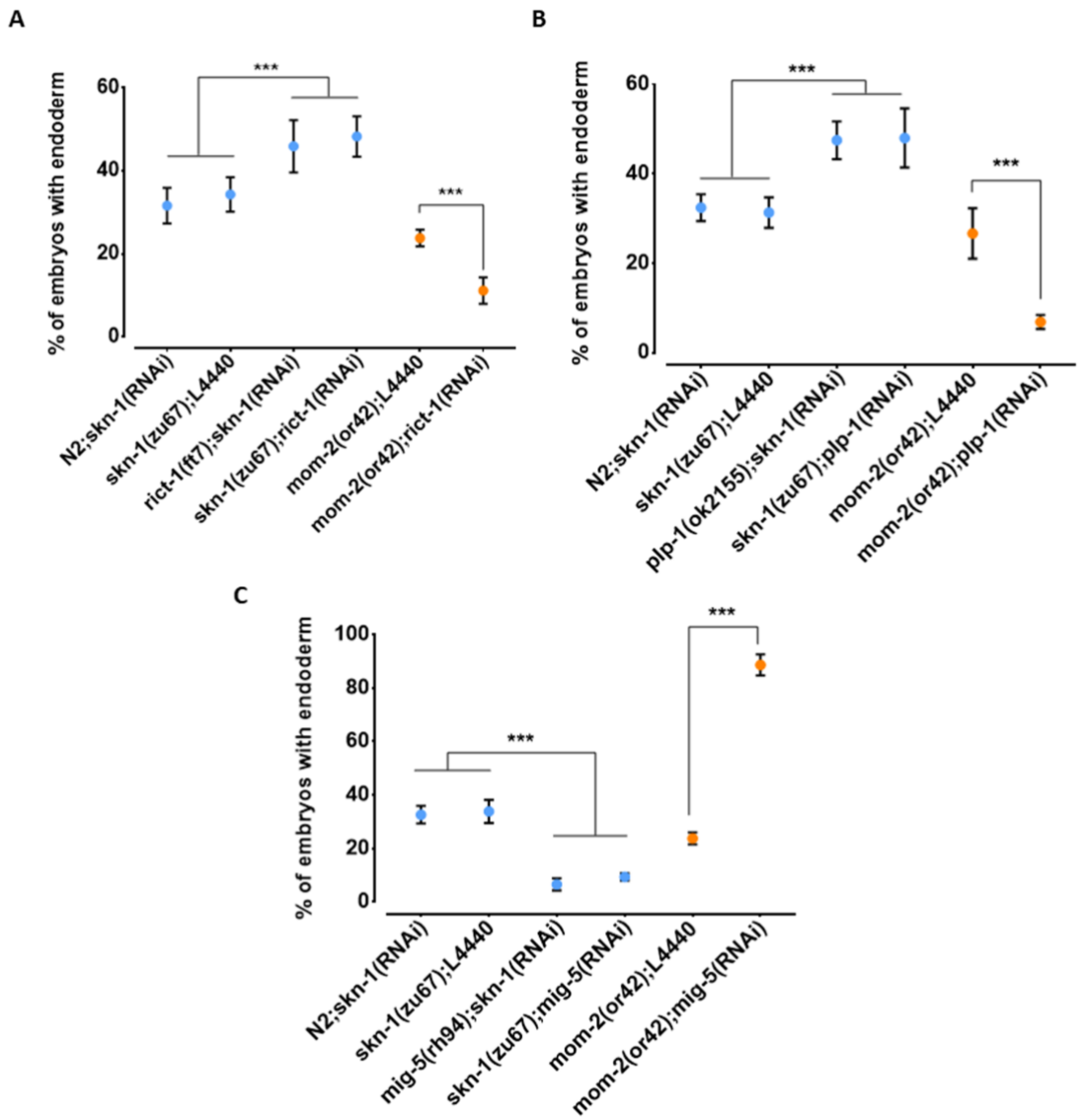
672 9.4% ± s.d 1.4%) and *suppression* of the *mom-2(-)* phenotype (88.6% ± s.d 4.0%) (Fig 8C).

673 Together, these findings reveal that, as observed with many of the N2/MY16 RILs

674 variants across most of the genome, RICT-1, PLP-1, and MIG-5 show opposite effects on the

675 phenotype of removing SKN-1 and MOM-2, suggesting a trend toward genetic influences that

676 reciprocally influence the outcome in the absence of these two inputs.



677

678 **Fig 8. Reciprocal effects of RICT-1, PLP-1, and MIG-5 on *skn-1(-)* and *mom-2(-)* phenotypes**

679 (A, B) Loss of RICT-1 or PLP-1 enhances the *mom-2(or42)* loss-of-endoderm phenotype and suppresses
680 *skn-1(zu67)* and *skn-1(RNAi)* phenotype. (C) Loss of MIG-5 enhances the *skn-1(zu67)* and *skn-1(RNAi)*
681 phenotype and suppresses *mom-2(or42)* phenotype. At least three replicates were performed per

682 experiment and >200 embryos per experiment. Student t-test (*** p-value<0.001). Data represented
683 with Standard Deviations.

684

685 **DISCUSSION**

686 The remarkable variety of forms associated with the ~36 animal phyla [81] that
687 emerged from a common metazoan ancestor >600 Mya is the product of numerous
688 incremental changes in GRNs underlying the formation of the body plan and cell types [1,82].
689 Here, we describe an unexpectedly broad divergence in the deployment of SKN-1/Nrf and
690 MOM-2/Wnt signaling in generating the most ancient germ layer, the endoderm, within wild
691 isolates of a single animal species, *C. elegans*. In this study, we report five major findings: 1)
692 while the quantitative requirement for two distinct regulatory inputs that initiate expression
693 of the endoderm GRN (SKN-1 and MOM-2) are highly reproducible in individual *C. elegans*
694 isolates, there is wide cryptic variation between isolates. 2) Cryptic variation in the
695 requirement for these regulatory factors shows substantial differences even between closely
696 related strains, suggesting that these traits are subject to rapid evolutionary change in this
697 species. 3) Quantitative genetic analyses of natural and recombinant populations revealed
698 multiple loci underlying the variation in the requirement for SKN-1 and MOM-2 in endoderm
699 specification. 4) The requirements for SKN-1 and MOM-2 in endoderm specification is
700 frequently reciprocal in their relation to other genetic factors. 5) *rict-1*, *plp-1*, *and mig-5*
701 reciprocally influence the outcome of *skn-1(-)* and *mom-2(-)*, substantiating the reciprocal
702 influences on the two GRN inputs. These findings reveal prevalent plasticity and complexity
703 underlying SKN-1 and MOM-2/Wnt regulatory inputs in mobilizing a conserved system for
704 endoderm specification. Thus, while the core genetic toolkit for the development of the
705 endoderm, the most ancient of the three germ layers, appears to have been preserved for
706 well over half a billion years, the molecular regulatory inputs that initiate its expression in *C.*
707 *elegans* vary extremely rapidly over short evolutionary time scales within the species.

708

709 **Multigenic variation in the requirement for SKN-1 and MOM-2**

710 Quantitative analyses of the wild isolates and RILs revealed that multigenic factors are
711 responsible for the difference in requirement for SKN-1 and MOM-2 between isotypes.
712 Notably, we observed substantial overlap on chromosome IV in the GWAS analyses of the
713 *skn-1* and *mom-2* requirements in wild isotypes (Fig. 5, Supplementary Fig. 4) and on
714 chromosome II from analyses using RILs (Fig. 6B). This finding raises the possibility that some
715 QTLs may influence requirement for both inputs into the endoderm specification pathway: as
716 SKN-1 and Wnt converge to regulate expression of the *end-1/3* genes, it is conceivable that
717 common genetic variants might modulate the relative strength or outcome of both maternal
718 inputs. However, our findings do not resolve whether these genetic variants act
719 independently to influence the maternal regulatory inputs. Genetic interactions are often
720 neglected in large-scale genetic association studies [83] owing in part to the difficulty in
721 confirming them [83]. Many studies [84–87], including ours here, showed that epistasis can
722 strongly influence the behavior of certain variants upon genetic perturbation. In addition,
723 selection on pleiotropically acting loci facilitates rapid developmental system drift. Together,
724 epistasis and selection on pleiotropic loci play important roles in the evolution of natural
725 populations [88–91].

726 **Potential compensatory relationships between SKN-1 and MOM-2/Wnt**

727 Although we did not observe a direct correlation between the *skn-1(-)* and *mom-2(-)*
728 phenotypes across the isotypes studied here, we found a negative correlation across much of
729 the genome for the N2 X MY16 RILs (Fig. 6, 7). Further, while GWAS and QTL analysis of natural
730 and inbred lines, respectively, did not reveal a causal region in chromosome IV for *mom-*

731 *2(RNAi)* variation, analysis of NILs results did uncover at least one QTL affecting this
732 phenotype. Moreover, while broad regions of the chromosome showed a positive correlation
733 between the SKN-1 and MOM-2 requirements, isolation of variants in NILs revealed an inverse
734 requirement for these inputs for at least some regions on this chromosome. These results
735 reflect the limitations of genome-wide studies of complex genetic traits: in the case of
736 chromosome IV, several closely linked loci appear to influence both the SKN-1 and MOM-2
737 requirements.

738 Our findings raise the possibility that the SKN-1 and MOM-2/Wnt inputs might
739 compensate for each other and that genetic variants that enhance the requirement for one
740 of the inputs may often relax the requirement for the other. Such reciprocality could reflect
741 cross-regulatory interactions between these two maternal inputs or could be the result of
742 evolutionary constraints imposed by selection on these genes, which act pleiotropically in a
743 variety of processes. Further supporting this possibility, we identified two genes, *rict-1* and
744 *plp-1*, that show similar inverse effects on the requirements from *skn-1* and *mom*-2:
745 debilitation of either gene enhances the phenotype of *mom-2(-)* and suppresses that of *skn-*
746 *1(-)*. RICT-1 function extends lifespan in *C. elegans* through the action of SKN-1[77], and loss
747 of RICT-1 rescues the misspecification of the MS and E blastomeres and lethality of *skn-1(-)*
748 embryos [77], consistent with our finding. We previously reported that PLP-1, a homologue
749 of the vertebrate transcription factor pur alpha, binds to the *end-1* promoter and acts in
750 parallel to the Wnt pathway and downstream of the MAPK signal [78], thereby promoting gut
751 formation. PLP-1 shows a similar reciprocal relationship with SKN-1 and MOM-2 as with RICT-
752 1 (Fig. 8). Given that PLP-1 binds at a *cis* regulatory site in *end-1* near a putative POP-1 binding
753 site [78], and that SKN-1 also binds to the *end-1* regulatory region [68], it is conceivable that
754 this reciprocality might reflect integration of information at the level of transcription factor

755 binding sites. As the architecture of the GRN is shaped by changes in *cis*-regulatory sequences
756 [1,3], analyzing alterations in SKN-1 and Wnt/POP-1 targets among *C. elegans* wild isolates
757 may provide insights into how genetic changes are accommodated without compromising the
758 developmental output at microevolutionary time scale.

759 MIG-5, a *dishevelled* orthologue, functions in the Wnt pathway in parallel to Src
760 signaling to regulate asymmetric cell division and endoderm induction [29,79]. We found that
761 the loss of *mig-5* function enhances the gut defect of *skn-1(-)* and suppresses that of the *mom-*
762 *2(-), the* opposite reciprocal relationship to that of *rict-1* and *plp-1,* and consistent with a
763 previous report (Fig. 8) [29]. These effects were not observed in embryos lacking function of
764 *dsh-2*, the orthologue of *mig-5* (data not shown), supporting a previous study that showed
765 overlapping but non-redundant roles of MIG-5 and DSH-2 in EMS spindle orientation and gut
766 specification [79]. Recent studies showed that Dishevelled can play both positive and negative
767 roles during axon guidance [92,93]. Dishevelled, upon Wnt-activation, promotes
768 hyperphosphorylation and inactivation of Frizzled receptor to fine-tune Wnt activity. It is
769 tempting to speculate that MIG-5 may perform similar function in EMS by downregulating
770 activating signals (Src or MAPK), in the absence of MOM-2.

771 We hypothesize that compensatory mechanisms might evolve to fine-tune the level
772 of gut-activating regulatory inputs. Successful developmental events depend on tight spatial
773 and temporal regulation of gene expression. For example, anterior-posterior patterning in the
774 *Drosophila* embryo is determined by the local concentrations of the Bicoid, Hunchback, and
775 Caudal transcription factors [94]. We postulate that SKN-1 and Wnt signaling is modulated so
776 that the downstream genes, *end-1/3*, which control specification and later differentiation of
777 endoderm progenitors, are expressed at optimal levels that ensure normal gut development.

778 Suboptimal END activity leads to poorly differentiated gut and both hypo- and hyperplasia in
779 the gut lineage [95–97]. Hyper- or hypo-activation of Wnt signaling has been implicated in
780 cancer development [98], bone diseases [99,100], and metabolic diseases [101,102],
781 demonstrating the importance of regulating the timing and dynamics of such developmental
782 signals within a quantitatively restricted window.

783 **Cryptic variation and evolvability of GRNs**

784 This study revealed substantial cryptic genetic modifications that alter the relative
785 importance of two partially redundant inputs into the *C. elegans* endoderm GRN, leading to
786 rapid change in the developmental network architecture (Fig. 9). Such modifications may
787 occur through transitional states that are apparent even within this single species. For
788 example, the finding that POP-1 is not required for gut development even in a wild isolate
789 (e.g., MY16) that, like *C. briggsae*, shows a near-absolute requirement for SKN-1 may reflect
790 a transitional state between the two species: *i.e.*, a nearly essential requirement for SKN-1
791 but non-essential requirement for POP-1, an effect not previously seen in either species. In
792 addition, duplicated GATA factors (the MEDs, ENDs, and ELTs) and partially redundant
793 activating inputs (SKN-1, Wnt, Src, and MAPK) in endoderm GRN, provide an opportunity for
794 genetic variation to accumulate and "experimentation" of new regulatory relationships
795 without diminishing fitness [2,102,103].

796 Redundancy in the regulatory inputs may act to 'rescue' an initial mutation and allow
797 for secondary mutations that might eventually lead to rewiring of the network. For example,
798 loss of either MyoD or Myf5, two key regulators of muscle differentiation in metazoans,
799 produces minimal defects in myogenesis as a result of compensatory relationship between
800 the myogenic factors [104]. In vertebrates, gene duplication events have resulted in an

801 expansion of Hox genes to a total of >200, resulting in prevalent redundancy [105–107]. This
802 proliferation of redundant genes provides opportunities for evolutionary experimentation,
803 subsequent specialization of new functions, and developmental system drift [107,108]. In *C.*
804 *elegans*, loss of GAP-1 (a Ras inhibitor) or SLI-1 (a negative regulator of EGFR signaling) alone
805 does not produce obvious defects, while double mutations lead to a multivulva phenotype
806 [109]. Similar redundant relationships between redundant partners exist in many other
807 contexts in the animal. Notably, the relative importance of Ras, Notch, and Wnt signals in
808 vulva induction differ in various genetic backgrounds [6,71] and physiological conditions
809 [110,111], resulting in flexibility in the system. While vulval development in *C. elegans,* when
810 grown under standard laboratory conditions, predominantly favors utilization of the EGF/Ras
811 signaling pathway [110], Wnt is the predominant signaling pathway in the related *Pristionchus*
812 *pacificus*, which is ~250 MY divergent [112,113]. In addition, while *Cel-lin-17* functions
813 positively to transduce the Wnt signal, *Ppa-lin-17/Fz* antagonizes Wnt signaling and instead
814 the Wnt signal is transmitted by *Ppa-lin-18/Ryk*, which has acquired a novel SH3 domain not
815 present in the *C. elegans* ortholog [114]. Thus, extensive rewiring of signaling networks and
816 modularity of signaling motifs contribute to developmental systems drift [108,115].

817 The broad cryptic variation may drive developmental system drift, giving rise to GRN
818 architectures that differ in the relative strength of the network components. Our finding that
819 the key regulatory inputs that initiate the endoderm GRN show dramatic plasticity is
820 consistent with comparative transcriptomic studies that demonstrate high gene expression
821 variability and divergence during early embryonic stages in fly [116–118], worm [117–120],
822 Xenopus [118,121,122], zebrafish [118,122], and mouse [118,122]. Therefore, early
823 developmental events may be highly evolvable, in part due to weak purifying selection on
824 maternal-effect genes [123,124]. This is in accordance with the “hourglass” concept of

825 embryonic development [125–127], in which divergent developmental mechanisms during
826 early embryogenesis converge on a more constant state (i.e., a “phylotypic stage” at the
827 molecular regulatory level). Indeed, unlike the terminal differentiation factor ELT-2, upstream
828 MEDs and ENDs genes are present only in closely related Caenorhabditis species
829 [15,97,128,129]. This is likely attributable either to positive selection during early embryonic
830 and later larval stages or to developmental constraints. Analysis of developmental gene
831 expression in mutation accumulation lines, which have evolved in the absence of any positive
832 selection, showed similarity to the developmental hourglass model of evolvability, consistent
833 with strong developmental constraints on the phylotypic stage [120]. However, they do not
834 rule out the possibility that early and late stages of development might be more adaptive and
835 therefore subject to positive selection. It will be of interest to learn the degree to which the
836 divergence in network architecture might arise as a result of differences in the environment
837 and selective pressures on different *C. elegans* isotypes.

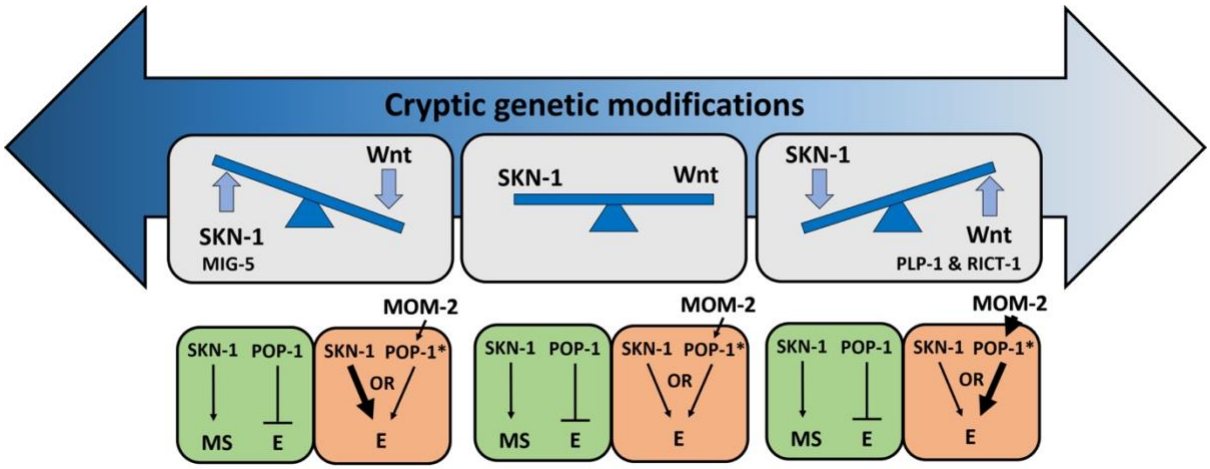


Fig. 9. Simplified models for potential cryptic compensatory relationship between the SKN-1 and MOM-2/Wnt regulatory inputs in the endoderm GRN.

Accumulation of cryptic genetic modifications drives rapid rewiring of the GRN, causing broad variation of SKN-1 and MOM-2/Wnt dependence in endoderm (E) specification among *C. elegans* isotypes. Wnt-signaled POP-1 (indicated by *) acts as an E activator, while unmodified POP-1 in the MS blastomere acts as a repressor of E fate in all *C. elegans* isotypes. The relative strength of the inputs is indicated by the thickness of the arrow. RICT-1, PLP-1 and MIG-5 reciprocally influence the outcome in the absence of the two inputs

848 **ACKNOWLEDGEMENTS**

849 We thank members of the Rothman lab, especially Sagen Flowers and Kristoffer C. Mellingen
850 for experimental assistance, and Snell labs, particularly Dr. Kien Ly, for helpful advice and
851 feedback. We thank Dr. Kathy Ruggiero (University of Auckland, New Zealand) for helpful
852 advice on GWAS methodology and Dr. James McGhee (University of Calgary, Canada) for
853 providing the MH33 antibody. Nematode strains used in this work were provided by the
854 Caenorhabditis Genetics Center, which is funded by the National Institutes of Health - Office
855 of Research Infrastructure Programs (P40 OD010440). Y.N.T.C was supported during part of
856 this work by a University of Auckland Doctoral Scholarship. This work was supported by grants
857 from the NIH (#1R01HD082347 and # 1R01HD081266) to J.H.R.

858 **COMPETING INTERESTS**

859 The authors declare no competing or financial interests.

860 **REFERENCES**

861 1. Peter IS, Davidson EH. Evolution of gene regulatory networks controlling body plan
862 development. Cell. Elsevier; 2011;144: 970–85. doi:10.1016/j.cell.2011.02.017

863 2. Félix M-A, Wagner A. Robustness and evolution: concepts, insights and challenges from a
864 developmental model system. Heredity (Edinb). Nature Publishing Group; 2008;100: 132–
865 140. doi:10.1038/sj.hdy.6800915

866 3. Davidson EH, Levine MS. Properties of developmental gene regulatory networks. Proc Natl
867 Acad Sci U S A. National Academy of Sciences; 2008;105: 20063–6.
868 doi:10.1073/pnas.0806007105

869 4. Oliveri P, Tu Q, Davidson EH. Global regulatory logic for specification of an embryonic cell
870 lineage. Proc Natl Acad Sci U S A. National Academy of Sciences; 2008;105: 5955–62.
871 doi:10.1073/pnas.0711220105

872 5. Peter IS, Davidson EH. Assessing regulatory information in developmental gene regulatory
873 networks. Proc Natl Acad Sci U S A. National Academy of Sciences; 2017;114: 5862–5869.
874 doi:10.1073/pnas.1610616114

875 6. Milloz J, Duveau F, Nuez I, Felix M-A. Intraspecific evolution of the intercellular signaling
876 network underlying a robust developmental system. Genes Dev. 2008;22: 3064–3075.
877 doi:10.1101/gad.495308

878 7. Nunes MDS, Arif S, Schlötterer C, McGregor AP. A perspective on micro-evo-devo: progress
879 and potential. Genetics. Genetics; 2013;195: 625–34. doi:10.1534/genetics.113.156463

880 8. Phinchongsakuldit J, MacArthur S, Brookfield JFY. Evolution of Developmental Genes:
881 Molecular Microevolution of Enhancer Sequences at the Ubx Locus in Drosophila and Its
882 Impact on Developmental Phenotypes. Mol Biol Evol. 2003;21: 348–363.
883 doi:10.1093/molbev/msh025

884 9. Hashimshony T, Feder M, Levin M, Hall BK, Yanai I. Spatiotemporal transcriptomics reveals
885 the evolutionary history of the endoderm germ layer. Nature. Nature Publishing Group;
886 2015;519: 219–222. doi:10.1038/nature13996

887 10. Rodaway A, Patient R. Mesendoderm: An Ancient Germ Layer? Cell. Cell Press; 2001;105:
888 169–172. doi:10.1016/S0092-8674(01)00307-5

889 11. Peterson KJ, Lyons JB, Nowak KS, Takacs CM, Wargo MJ, McPeek MA. Estimating metazoan
890 divergence times with a molecular clock. Proc Natl Acad Sci U S A. National Academy of
891 Sciences; 2004;101: 6536–41. doi:10.1073/pnas.0401670101

892 12. Martindale MQ, Pang K, Finnerty JR. Investigating the origins of triploblasty: “mesodermal”
893 gene expression in a diploblastic animal, the sea anemone Nematostella vectensis (phylum,
894 Cnidaria; class, Anthozoa). Development. The Company of Biologists Ltd; 2004;131: 2463–74.
895 doi:10.1242/dev.01119

896 13. Boyle MJ, Seaver EC. Developmental expression of foxA and gata genes during gut formation
897 in the polychaete annelid, Capitella sp. I. Evol Dev. 2008;10: 89–105. doi:10.1111/j.1525-
898 142X.2007.00216.x

899 14. Boyle MJ, Seaver EC. Expression of FoxA and GATA transcription factors correlates with
900 regionalized gut development in two lophotrochozoan marine worms: Chaetopterus
901 (Annelida) and Themiste lageniformis (Sipuncula). Evodevo. 2010;1: 2. doi:10.1186/2041-

902 9139-1-2

903 15. Gillis WJ, Bowerman B, Schneider SQ. Ectoderm- and endomesoderm-specific GATA
904 transcription factors in the marine annelid Platynereis dumerilli. Evol Dev. 2007;9: 39–50.
905 doi:10.1111/j.1525-142X.2006.00136.x

906 16. Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh C-H, et al. A Provisional Regulatory
907 Gene Network for Specification of Endomesoderm in the Sea Urchin Embryo. Dev Biol.
908 2002;246: 162–190. doi:10.1006/dbio.2002.0635

909 17. Shoichet SA, Malik TH, Rothman JH, Shivdasani RA. Action of the Caenorhabditis elegans
910 GATA factor END-1 in Xenopus suggests that similar mechanisms initiate endoderm
911 development in ecdysozoa and vertebrates. Proc Natl Acad Sci U S A. National Academy of
912 Sciences; 2000;97: 4076–81. Available: http://www.ncbi.nlm.nih.gov/pubmed/10760276

913 18. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the
914 nematode Caenorhabditis elegans. Dev Biol. 1983;100: 64–119. Available:
915 http://www.ncbi.nlm.nih.gov/pubmed/6684600

916 19. Bowerman B, Eaton BA, Priess JR. skn-1, a maternally expressed gene required to specify the
917 fate of ventral blastomeres in the early C. elegans embryo. Cell. 1992;68: 1061–75. Available:
918 http://www.ncbi.nlm.nih.gov/pubmed/1547503

919 20. Bowerman B, Draper BW, Mello CC, Priess JR. The maternal gene skn-1 encodes a protein
920 that is distributed unequally in early C. elegans embryos. Cell. Elsevier; 1993;74: 443–52.
921 doi:10.1016/0092-8674(93)80046-H

922 21. Maduro MF, Rothman JH. Making Worm Guts: The Gene Regulatory Network of the
923 Caenorhabditis elegans Endoderm. Dev Biol. Academic Press; 2002;246: 68–85.
924 doi:10.1006/DBIO.2002.0655

925 22. Maduro MF, Broitman-Maduro G, Mengarelli I, Rothman JH. Maternal deployment of the
926 embryonic SKN-1 → MED-1,2 cell specification pathway in C. elegans. Dev Biol. Academic
927 Press; 2007;301: 590–601. doi:10.1016/J.YDBIO.2006.08.029

928 23. Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH. Restriction of
929 mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3beta
930 homolog is mediated by MED-1 and -2 in C. elegans. Mol Cell. 2001;7: 475–85. Available:
931 http://www.ncbi.nlm.nih.gov/pubmed/11463373

932 24. Maduro MF. Gut development in C. elegans. Semin Cell Dev Biol. 2017;66: 3–11.
933 doi:10.1016/j.semcdb.2017.01.001

934 25. Wiesenfahrt T, Osborne Nishimura E, Berg JY, McGhee JD. Probing and rearranging the
935 transcription factor network controlling the *C. elegans* endoderm. Worm. 2016;5: e1198869.
936 doi:10.1080/21624054.2016.1198869

937 26. McGhee J. The C. elegans intestine. WormBook. 2007; 1–36. doi:10.1895/wormbook.1.133.1

938 27. Meneghini MD, Ishitani T, Carter JC, Hisamoto N, Ninomiya-Tsuji J, Thorpe CJ, et al. MAP
939 kinase and Wnt pathways converge to downregulate an HMG-domain repressor in
940 Caenorhabditis elegans. Nature. Nature Publishing Group; 1999;399: 793–797.
941 doi:10.1038/21666

942 28. Shin TH, Yasuda J, Rocheleau CE, Lin R, Soto M, Bei Y, et al. MOM-4, a MAP kinase kinase
943 kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity
944 signals in C. elegans. Mol Cell. 1999;4: 275–80. Available:

http://www.ncbi.nlm.nih.gov/pubmed/10488343

29. Bei Y, Hogan J, Berkowitz LA, Soto M, Rocheleau CE, Pang KM, et al. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early C. elegans embryos. Dev Cell. 2002;3: 113–25. Available: http://www.ncbi.nlm.nih.gov/pubmed/12110172

30. Thorpe CJ, Schlesinger A, Carter JC, Bowerman B. Wnt signaling polarizes an early C. elegans blastomere to distinguish endoderm from mesoderm. Cell. 1997;90: 695–705. Available: http://www.ncbi.nlm.nih.gov/pubmed/9288749

31. Nakamura K, Kim S, Ishidate T, Bei Y, Pang K, Shirayama M, et al. Wnt signaling drives WRM-1/beta-catenin asymmetries in early C. elegans embryos. Genes Dev. Cold Spring Harbor Laboratory Press; 2005;19: 1749–54. doi:10.1101/gad.1323705

32. Rocheleau CE, Yasuda J, Shin TH, Lin R, Sawa H, Okano H, et al. WRM-1 Activates the LIT-1 Protein Kinase to Transduce Anterior/Posterior Polarity Signals in C. elegans. Cell. Cell Press; 1999;97: 717–726. doi:10.1016/S0092-8674(00)80784-9

33. Owraghi M, Broitman-Maduro G, Luu T, Roberson H, Maduro MF. Roles of the Wnt effector POP-1/TCF in the C. elegans endomesoderm specification gene network. Dev Biol. NIH Public Access; 2010;340: 209–21. doi:10.1016/j.ydbio.2009.09.042

34. Huang S, Shetty P, Robertson SM, Lin R. Binary cell fate specification during C. elegans embryogenesis driven by reiterated reciprocal asymmetry of TCF POP-1 and its coactivator -catenin SYS-1. Development. 2007;134: 2685–2695. doi:10.1242/dev.008268

35. Maduro MF, Lin R, Rothman JH. Dynamics of a Developmental Switch: Recursive Intracellular and Intranuclear Redistribution of Caenorhabditis elegans POP-1 Parallels Wnt-Inhibited Transcriptional Repression. Dev Biol. Academic Press; 2002;248: 128–142. doi:10.1006/DBIO.2002.0721

36. Phillips BT, Kidd AR, King R, Hardin J, Kimble J. Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. Proc Natl Acad Sci. 2007;104: 3231–3236. doi:10.1073/pnas.0611507104

37. Maduro MF, Kasmir JJ, Zhu J, Rothman JH. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Dev Biol. 2005;285: 510–523. doi:10.1016/j.ydbio.2005.06.022

38. Shetty P, Lo M-C, Robertson SM, Lin R. C. elegans TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. Dev Biol. 2005;285: 584–592. doi:10.1016/j.ydbio.2005.07.008

39. Lin KT-H, Broitman-Maduro G, Hung WWK, Cervantes S, Maduro MF. Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in C. briggsae as compared with C. elegans. Dev Biol. 2009;325: 296–306. doi:10.1016/j.ydbio.2008.10.001

40. Félix M-A, Braendle C. The natural history of Caenorhabditis elegans. Curr Biol. Elsevier; 2010;20: R965-9. doi:10.1016/j.cub.2010.09.050

41. Cook DE, Zdraljevic S, Roberts JP, Andersen EC. CeNDR, the *Caenorhabditis elegans* natural diversity resource. Nucleic Acids Res. 2017;45: D650–D657. doi:10.1093/nar/gkw893

42. 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. NIH Public

988 Access; 2012;44: 285–90. doi:10.1038/ng.1050

989 43. Kontani K, Moskowitz IPG, Rothman JH. Repression of Cell-Cell Fusion by Components of the
990 C. elegans Vacuolar ATPase Complex. Dev Cell. 2005;8: 787–794.
991 doi:10.1016/j.devcel.2005.02.018

992 44. Brenner S. The genetics of Caenorhabditis elegans. Genetics. 1974;77: 71–94. Available:
993 http://www.ncbi.nlm.nih.gov/pubmed/4366476

994 45. Kamath RS, Ahringer J. Genome-wide RNAi screening in Caenorhabditis elegans. Methods.
995 Academic Press; 2003;30: 313–321. doi:10.1016/S1046-2023(03)00050-1

996 46. Rual J-F, Ceron J, Koreth J, Hao T, Nicot A-S, Hirozane-Kishikawa T, et al. Toward Improving
997 Caenorhabditis elegans Phenome Mapping With an ORFeome-Based RNAi Library. Genome
998 Res. 2004;14: 2162–2168. doi:10.1101/gr.2505604

999 47. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al. Systematic functional
1000 analysis of the Caenorhabditis elegans genome using RNAi. Nature. Nature Publishing Group;
1001 2003;421: 231–237. doi:10.1038/nature01278

1002 48. Gomez-Amaro RL, Valentine ER, Carretero M, LeBoeuf SE, Rangaraju S, Broaddus CD, et al.
1003 Measuring Food Intake and Nutrient Absorption in Caenorhabditis elegans. Genetics.
1004 Genetics Society of America; 2015;200: 443–54. doi:10.1534/genetics.115.175851

1005 49. Sommermann EM, Strohmaier KR, Maduro MF, Rothman JH. Endoderm development in
1006 Caenorhabditis elegans: The synergistic action of ELT-2 and -7 mediates the
1007 specification→differentiation transition. Dev Biol. 2010;347: 154–166.
1008 doi:10.1016/j.ydbio.2010.08.020

1009 50. Clokey G V, Jacobson LA. The autofluorescent “lipofuscin granules” in the intestinal cells of
1010 Caenorhabditis elegans are secondary lysosomes. Mech Ageing Dev. 1986;35: 79–94.
1011 Available: http://www.ncbi.nlm.nih.gov/pubmed/3736133

1012 51. Hermann GJ, Schroeder LK, Hieb CA, Kershner AM, Rabbitts BM, Fonarev P, et al. Genetic
1013 Analysis of Lysosomal Trafficking in Caenorhabditis elegans. Mol Biol Cell. 2005;16: 3273–
1014 3288. doi:10.1091/mbc.E05

1015 52. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient Control of
1016 Population Structure in Model Organism Association Mapping. Genetics. 2008;178: 1709–
1017 1723. doi:10.1534/genetics.107.080101

1018 53. Millstein J, Volfson D. Computationally efficient permutation-based confidence interval
1019 estimation for tail-area FDR. Front Genet. Frontiers Media SA; 2013;4: 179.
1020 doi:10.3389/fgene.2013.00179

1021 54. Hansen E, Kerr KF. A Comparison of Two Classes of Methods for Estimating False Discovery
1022 Rates in Microarray Studies. Scientifica (Cairo). 2012;2012: 1–9. doi:10.6064/2012/519394

1023 55. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language.
1024 Bioinformatics. Narnia; 2004;20: 289–290. doi:10.1093/bioinformatics/btg412

1025 56. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things).
1026 Methods Ecol Evol. John Wiley & Sons, Ltd (10.1111); 2012;3: 217–223. doi:10.1111/j.2041-
1027 210X.2011.00169.x

1028 57. Ives AR, Midford PE, Garland T. Within-Species Variation and Measurement Error in
1029 Phylogenetic Comparative Methods. Oakley T, editor. Syst Biol. 2007;56: 252–270.
1030 doi:10.1080/10635150701313830

1031 58. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set
1032 for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet.
1033 2007;81: 559–575. doi:10.1086/519795

1034 59. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A Robust, Simple
1035 Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. Orban L, editor. PLoS
1036 One. Public Library of Science; 2011;6: e19379. doi:10.1371/journal.pone.0019379

1037 60. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for
1038 association mapping of complex traits in diverse samples. Bioinformatics. 2007;23: 2633–
1039 2635. doi:10.1093/bioinformatics/btm308

1040 61. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call
1041 format and VCFtools. Bioinformatics. Oxford University Press; 2011;27: 2156–8.
1042 doi:10.1093/bioinformatics/btr330

1043 62. Broman KW, Sen S. A Guide to QTL Mapping with R/qtl. New York, NY: Springer New York;
1044 2009; 1–20. doi:10.1007/978-0-387-92125-9

1045 63. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses.
1046 Bioinformatics. 2003;19: 889–90. Available: http://www.ncbi.nlm.nih.gov/pubmed/12724300

1047 64. Zhao Z, Boyle TJ, Bao Z, Murray JI, Mericle B, Waterston RH. Comparative analysis of
1048 embryonic cell lineage between Caenorhabditis briggsae and Caenorhabditis elegans. Dev
1049 Biol. NIH Public Access; 2008;314: 93–9. doi:10.1016/j.ydbio.2007.11.015

1050 65. Cutter AD. Divergence Times in Caenorhabditis and Drosophila Inferred from Direct Estimates
1051 of the Neutral Mutation Rate. Mol Biol Evol. Oxford University Press; 2008;25: 778–786.
1052 doi:10.1093/molbev/msn024

1053 66. Sterken MG, Snoek LB, Kammenga JE, Andersen EC. The laboratory domestication of
1054 Caenorhabditis elegans. Trends Genet. NIH Public Access; 2015;31: 224–31.
1055 doi:10.1016/j.tig.2015.02.009

1056 67. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Variability in gene expression underlies
1057 incomplete penetrance. Nature. Nature Publishing Group; 2010;463: 913–918.
1058 doi:10.1038/nature08781

1059 68. Zhu J, Hill RJ, Heid PJ, Fukuyama M, Sugimoto A, Priess JR, et al. end-1 encodes an apparent
1060 GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos.
1061 Genes Dev. Cold Spring Harbor Laboratory Press; 1997;11: 2883–96. Available:
1062 http://www.ncbi.nlm.nih.gov/pubmed/9353257

1063 69. Tijsterman M, Okihara KL, Thijssen K, Plasterk RH. PPW-1, a PAZ/PIWI Protein Required for
1064 Efficient Germline RNAi, Is Defective in a Natural Isolate of C. elegans. Curr Biol. Cell Press;
1065 2002;12: 1535–1540. doi:10.1016/S0960-9822(02)01110-7

1066 70. Pollard DA, Rockman M V. Resistance to Germline RNA Interference in a Caenorhabditis
1067 elegans Wild Isolate Exhibits Complexity and Nonadditivity. G3 Genes, Genomes, Genet. G3:
1068 Genes, Genomes, Genetics; 2013;3: 941–947. doi:10.1534/G3.113.005785

1069 71. Gleason JE, Korswagen HC, Eisenmann DM. Activation of Wnt signaling bypasses the
1070 requirement for RTK/Ras signaling during C. elegans vulval induction. Genes Dev. Cold Spring
1071 Harbor Laboratory Press; 2002;16: 1281–90. doi:10.1101/gad.981602

1072 72. Paaby AB, White AG, Riccardi DD, Gunsalus KC, Piano F, Rockman M V. Wild worm
1073 embryogenesis harbors ubiquitous polygenic modifier variation. Elife. eLife Sciences

1074 Publications Limited; 2015;4: e09178. doi:10.7554/eLife.09178

1075 73. Çelen İ, Doh JH, Sabanayagam CR. Effects of liquid cultivation on gene expression and
1076 phenotype of C. elegans. BMC Genomics. BioMed Central; 2018;19: 562. doi:10.1186/s12864-
1077 018-4948-7

1078 74. Wang J, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. An estimator for
1079 pairwise relatedness using molecular markers. Genetics. Genetics; 2002;160: 1203–15.
1080 doi:10.1534/genetics.167.1.531

1081 75. Rieseberg LH, Widmer A, Arntz AM, Burke B. The genetic architecture necessary for
1082 transgressive segregation is common in both natural and domesticated populations. Philos
1083 Trans R Soc B Biol Sci. 2003;358: 1141–1147. doi:10.1098/rstb.2003.1283

1084 76. Félix M-A, Ashe A, Piffaretti J, Wu G, Nuez I, Bélicard T, et al. Natural and Experimental
1085 Infection of Caenorhabditis Nematodes by Novel Viruses Related to Nodaviruses. Hodgkin J,
1086 editor. PLoS Biol. Public Library of Science; 2011;9: e1000586.
1087 doi:10.1371/journal.pbio.1000586

1088 77. Ruf V, Holzem C, Peyman T, Walz G, Blackwell TK, Neumann-Haefelin E. TORC2 signaling
1089 antagonizes SKN-1 to induce C. elegans mesendodermal embryonic development. Dev Biol.
1090 Academic Press; 2013;384: 214–227. doi:10.1016/J.YDBIO.2013.08.011

1091 78. Witze ES, Field ED, Hunt DF, Rothman JH. C. elegans pur alpha, an activator of end-1,
1092 synergizes with the Wnt pathway to specify endoderm. Dev Biol. 2009;327: 12–23.
1093 doi:10.1016/j.ydbio.2008.11.015

1094 79. Walston T, Tuskey C, Edgar L, Hawkins N, Ellis G, Bowerman B, et al. Multiple Wnt Signaling
1095 Pathways Converge to Orient the Mitotic Spindle in Early C. elegans Embryos. Dev Cell. Cell
1096 Press; 2004;7: 831–841. doi:10.1016/J.DEVCEL.2004.10.008

1097 80. Tatebe H, Shiozaki K. Evolutionary Conservation of the Components in the TOR Signaling
1098 Pathways. Biomolecules. 2017;7: 77. doi:10.3390/biom7040077

1099 81. Adoutte A, Philippe H. The major lines of metazoan evolution: summary of traditional
1100 evidence and lessons from ribosomal RNA sequence analysis. EXS. 1993;63: 1–30. Available:
1101 http://www.ncbi.nlm.nih.gov/pubmed/8422536

1102 82. Carroll SB. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of
1103 Morphological Evolution. Cell. 2008;134: 25–36. doi:10.1016/j.cell.2008.06.030

1104 83. Page GP, George V, Go RC, Page PZ, Allison DB. “Are We There Yet?”: Deciding When One Has
1105 Demonstrated Specific Genetic Causation in Complex Diseases and Quantitative Traits. Am J
1106 Hum Genet. 2003;73: 711–719. doi:10.1086/378900

1107 84. Mackay TFC. Epistasis and quantitative traits: using model organisms to study gene-gene
1108 interactions. Nat Rev Genet. NIH Public Access; 2014;15: 22–33. doi:10.1038/nrg3627

1109 85. Volis S, Shulgina I, Zaretsky M, Koren O. Epistasis in natural populations of a predominantly
1110 selfing plant. Heredity (Edinb). Nature Publishing Group; 2011;106: 300–9.
1111 doi:10.1038/hdy.2010.79

1112 86. Félix M-A. Cryptic Quantitative Evolution of the Vulva Intercellular Signaling Network in
1113 Caenorhabditis. Curr Biol. 2007;17: 103–114. doi:10.1016/j.cub.2006.12.024

1114 87. Barkoulas M, van Zon JS, Milloz J, van Oudenaarden A, Félix M-A. Robustness and Epistasis in
1115 the C. elegans Vulval Signaling Network Revealed by Pathway Dosage Modulation. Dev Cell.
1116 Cell Press; 2013;24: 64–75. doi:10.1016/J.DEVCEL.2012.12.001

1117 88. Duveau F, Félix M-A. Role of Pleiotropy in the Evolution of a Cryptic Developmental Variation
1118 in Caenorhabditis elegans. Noor MAF, editor. PLoS Biol. Public Library of Science; 2012;10:
1119 e1001230. doi:10.1371/journal.pbio.1001230

1120 89. Johnson NA, Porter AH. Evolution of branched regulatory genetic pathways: directional
1121 selection on pleiotropic loci accelerates developmental system drift. Genetica. Springer
1122 Netherlands; 2006;129: 57–70. doi:10.1007/s10709-006-0033-2

1123 90. Phillips PC. Epistasis--the essential role of gene interactions in the structure and evolution of
1124 genetic systems. Nat Rev Genet. NIH Public Access; 2008;9: 855–67. doi:10.1038/nrg2452

1125 91. Wagner GP, Zhang J. The pleiotropic structure of the genotype–phenotype map: the
1126 evolvability of complex organisms. Nat Rev Genet. Nature Publishing Group; 2011;12: 204–
1127 213. doi:10.1038/nrg2949

1128 92. Shafer B, Onishi K, Lo C, Colakoglu G, Zou Y. Vangl2 Promotes Wnt/Planar Cell Polarity-like
1129 Signaling by Antagonizing Dvl1-Mediated Feedback Inhibition in Growth Cone Guidance. Dev
1130 Cell. Cell Press; 2011;20: 177–191. doi:10.1016/J.DEVCEL.2011.01.002

1131 93. Zheng C, Diaz-Cuadros M, Chalfie M. Dishevelled attenuates the repelling activity of Wnt
1132 signaling during neurite outgrowth in Caenorhabditis elegans. Proc Natl Acad Sci U S A.
1133 National Academy of Sciences; 2015;112: 13243–8. doi:10.1073/pnas.1518686112

1134 94. Rivera-Pomar R, Jãckle H. From gradients to stripes in Drosophila embryogenesis: filling in the
1135 gaps. Trends Genet. Elsevier Current Trends; 1996;12: 478–483. doi:10.1016/0168-
1136 9525(96)10044-5

1137 95. Maduro MF, Broitman-Maduro G, Choi H, Carranza F, Wu AC-Y, Rifkin SA. MED GATA factors
1138 promote robust development of the C. elegans endoderm. Dev Biol. Academic Press;
1139 2015;404: 66–79. doi:10.1016/J.YDBIO.2015.04.025

1140 96. Choi H, Broitman-Maduro G, Maduro MF. Partially compromised specification causes
1141 stochastic effects on gut development in C. elegans. Dev Biol. Academic Press; 2017;427: 49–
1142 60. doi:10.1016/J.YDBIO.2017.05.007

1143 97. Maduro MF. Developmental robustness in the *Caenorhabditis elegans* embryo. Mol Reprod
1144 Dev. 2015;82: 918–931. doi:10.1002/mrd.22582

1145 98. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. Oncogene. Nature Publishing Group;
1146 2017;36: 1461–1473. doi:10.1038/onc.2016.304

1147 99. Jenkins ZA, van Kogelenberg M, Morgan T, Jeffs A, Fukuzawa R, Pearl E, et al. Germline
1148 mutations in WTX cause a sclerosing skeletal dysplasia but do not predispose to
1149 tumorigenesis. Nat Genet. 2009;41: 95–100. doi:10.1038/ng.270

1150 100. Baron R, Gori F. Targeting WNT signaling in the treatment of osteoporosis. Curr Opin
1151 Pharmacol. Elsevier; 2018;40: 134–141. doi:10.1016/J.COPH.2018.04.011

1152 101. Chen N, Wang J. Wnt/β-Catenin Signaling and Obesity. Front Physiol. 2018;9: 792.
1153 doi:10.3389/fphys.2018.00792

1154 102. Schinner S. Wnt-signalling and the Metabolic Syndrome. Horm Metab Res. 2009;41: 159–163.
1155 doi:10.1055/s-0028-1119408

1156 103. Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. Phenotypic robustness conferred by
1157 apparently redundant transcriptional enhancers. Nature. 2010;466: 490–493.
1158 doi:10.1038/nature09158

1159 104. Mohun T. Muscle differentiation. Curr Opin Cell Biol. 1992;4: 923–8. Available:
1160 http://www.ncbi.nlm.nih.gov/pubmed/1485959

1161 105. Imai Y, Gates MA, Melby AE, Kimelman D, Schier AF, Talbot WS. The homeobox genes vox
1162 and vent are redundant repressors of dorsal fates in zebrafish. Development. 2001;128:
1163 2407–20. Available: http://www.ncbi.nlm.nih.gov/pubmed/11493559

1164 106. Manley NR, Capecchi MR. Hox Group 3 Paralogous Genes Act Synergistically in the Formation
1165 of Somitic and Neural Crest-Derived Structures. Dev Biol. Academic Press; 1997;192: 274–
1166 288. doi:10.1006/DBIO.1997.8765

1167 107. Nam J, Nei M. Evolutionary change of the numbers of homeobox genes in bilateral animals.
1168 Mol Biol Evol. NIH Public Access; 2005;22: 2386–94. doi:10.1093/molbev/msi229

1169 108. True JR, Haag ES. Developmental system drift and flexibility in evolutionary trajectories. Evol
1170 Dev. John Wiley & Sons, Ltd (10.1111); 2001;3: 109–119. doi:10.1046/j.1525-
1171 142x.2001.003002109.x

1172 109. Yoon CH, Chang C, Hopper NA, Lesa GM, Sternberg PW. Requirements of multiple domains of
1173 SLI-1, a Caenorhabditis elegans homologue of c-Cbl, and an inhibitory tyrosine in LET-23 in
1174 regulating vulval differentiation. Mol Biol Cell. American Society for Cell Biology; 2000;11:
1175 4019–31. Available: http://www.ncbi.nlm.nih.gov/pubmed/11071924

1176 110. Braendle C, Félix M-A. Plasticity and Errors of a Robust Developmental System in Different
1177 Environments. Dev Cell. 2008;15: 714–724. doi:10.1016/j.devcel.2008.09.011

1178 111. Grimbert S, Vargas Velazquez AM, Braendle C. Physiological Starvation Promotes
1179 Caenorhabditis elegans Vulval Induction. G3 (Bethesda). Genetics Society of America; 2018;8:
1180 3069–3081. doi:10.1534/g3.118.200449

1181 112. Zheng M, Messerschmidt D, Jungblut B, Sommer RJ. Conservation and diversification of Wnt
1182 signaling function during the evolution of nematode vulva development. Nat Genet. 2005;37:
1183 300–304. doi:10.1038/ng1512

1184 113. Tian H, Schlager B, Xiao H, Sommer RJ. Wnt Signaling Induces Vulva Development in the
1185 Nematode Pristionchus pacificus. Curr Biol. 2008;18: 142–146. doi:10.1016/j.cub.2007.12.048

1186 114. Wang X, Sommer RJ. Antagonism of LIN-17/Frizzled and LIN-18/Ryk in Nematode Vulva
1187 Induction Reveals Evolutionary Alterations in Core Developmental Pathways. Sternberg PW,
1188 editor. PLoS Biol. 2011;9: e1001110. doi:10.1371/journal.pbio.1001110

1189 115. Haag ES, Fitch DHA, Delattre M. From "the Worm" to "the Worms"
1190 and Back Again: The Evolutionary Developmental Biology of Nematodes. Genetics. Genetics
1191 Society of America; 2018;210: 397–433. doi:10.1534/genetics.118.300243

1192 116. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. Gene expression
1193 divergence recapitulates the developmental hourglass model. Nature. Nature Publishing
1194 Group; 2010;468: 811–814. doi:10.1038/nature09634

1195 117. Gerstein MB, Rozowsky J, Yan K-K, Wang D, Cheng C, Brown JB, et al. Comparative analysis of
1196 the transcriptome across distant species. Nature. Nature Publishing Group; 2014;512: 445–
1197 448. doi:10.1038/nature13424

1198 118. Levin M, Anavy L, Cole AG, Winter E, Mostov N, Khair S, et al. The mid-developmental
1199 transition and the evolution of animal body plans. Nature. Nature Publishing Group;
1200 2016;531: 637–641. doi:10.1038/nature16994

1201 119. Levin M, Hashimshony T, Wagner F, Yanai I. Developmental Milestones Punctuate Gene

1202 Expression in the Caenorhabditis Embryo. Dev Cell. 2012;22: 1101–1108.
1203 doi:10.1016/j.devcel.2012.04.004

1204 120. Zalts H, Yanai I. Developmental constraints shape the evolution of the nematode mid-
1205 developmental transition. Nat Ecol Evol. Nature Publishing Group; 2017;1: 0113.
1206 doi:10.1038/s41559-017-0113

1207 121. Yanai I, Peshkin L, Jorgensen P, Kirschner MW. Mapping Gene Expression in Two Xenopus
1208 Species: Evolutionary Constraints and Developmental Flexibility. Dev Cell. 2011;20: 483–496.
1209 doi:10.1016/j.devcel.2011.03.015

1210 122. Irie N, Kuratani S. Comparative transcriptome analysis reveals vertebrate phylotypic period
1211 during organogenesis. Nat Commun. Nature Publishing Group; 2011;2: 248.
1212 doi:10.1038/ncomms1248

1213 123. Cruickshank T, Wade MJ. Microevolutionary support for a developmental hourglass: gene
1214 expression patterns shape sequence variation and divergence in *Drosophila*. Evol Dev. John
1215 Wiley & Sons, Ltd (10.1111); 2008;10: 583–590. doi:10.1111/j.1525-142X.2008.00273.x

1216 124. Cutter AD, Garrett RH, Mark S, Wang W, Sun L. Molecular evolution across developmental
1217 time reveals rapid divergence in early embryogenesis. Evol Lett. John Wiley & Sons, Ltd; 2019;
1218 doi:10.1002/evl3.122

1219 125. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. Gene expression
1220 divergence recapitulates the developmental hourglass model. Nature. Nature Publishing
1221 Group; 2010;468: 811–814. doi:10.1038/nature09634

1222 126. Raff RA. The shape of life : genes, development, and the evolution of animal form. University
1223 of Chicago Press; 1996.

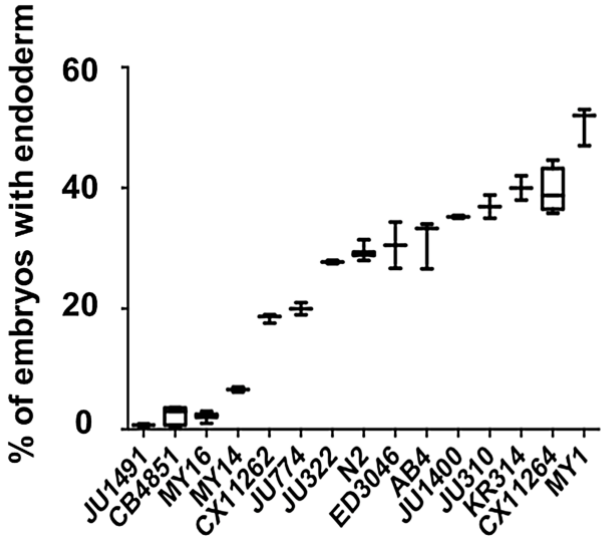
1224 127. Domazet-Lošo T, Tautz D. A phylogenetically based transcriptome age index mirrors
1225 ontogenetic divergence patterns. Nature. Nature Publishing Group; 2010;468: 815–818.
1226 doi:10.1038/nature09632

1227 128. Coroian C, Broitman-Maduro G, Maduro MF. Med-type GATA factors and the evolution of
1228 mesendoderm specification in nematodes. Dev Biol. 2006;289: 444–455.
1229 doi:10.1016/j.ydbio.2005.10.024

1230 129. Maduro MF, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess JR, et al. Genetic redundancy in
1231 endoderm specification within the genus Caenorhabditis. Dev Biol. 2005;284: 509–522.
1232 doi:10.1016/j.ydbio.2005.05.016

1233

1234 ## Supplementary figures

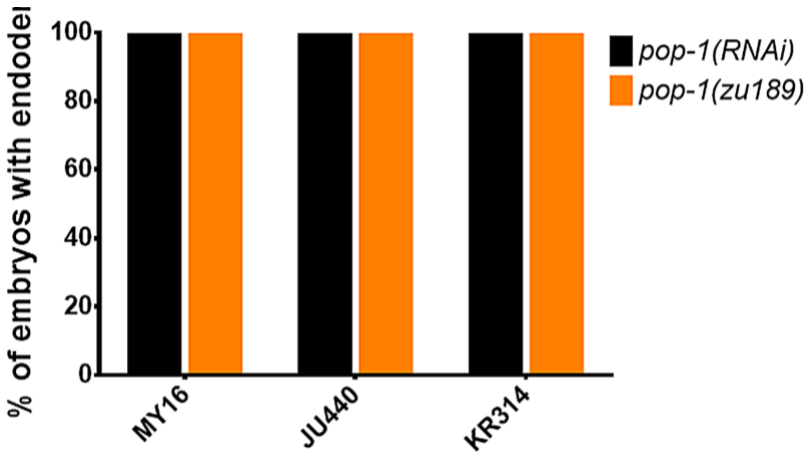


1235

1236 ***Supplementary Fig. 1: High reproducibility of* skn-1(RNAi) *phenotypes in various* C. elegans *isotypes.***

1237 *A minimum of two replicates were obtained, with >500 embryos per replicate (N = 3). Box-plot represents*
1238 *median with range bars showing upper and lower quartiles.*

1239



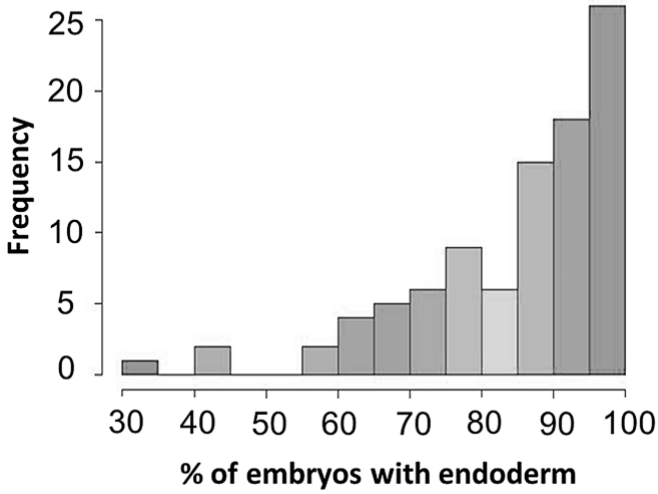
1240

1241 ***Supplementary Fig. 2: The requirement for POP-1 in endoderm formation does not vary in three***
1242 ***introgressed strains.***

1243 *Strains are shown on the x-axis and fraction of arrested embryos with endoderm are shown on the y-axis.*
1244 *The orange bars represent the results from mutant lines: MY16;pop-1(zu189), JU440;pop-1(zu189), and*
1245 *KR314;pop-1(zu189). Four introgressed lines were created for each new mutant strain. Black bars represent*
1246 *pop-1(RNAi) results on the wild isolates indicated. >200 embryos were scored per experiment.*

1247

1248

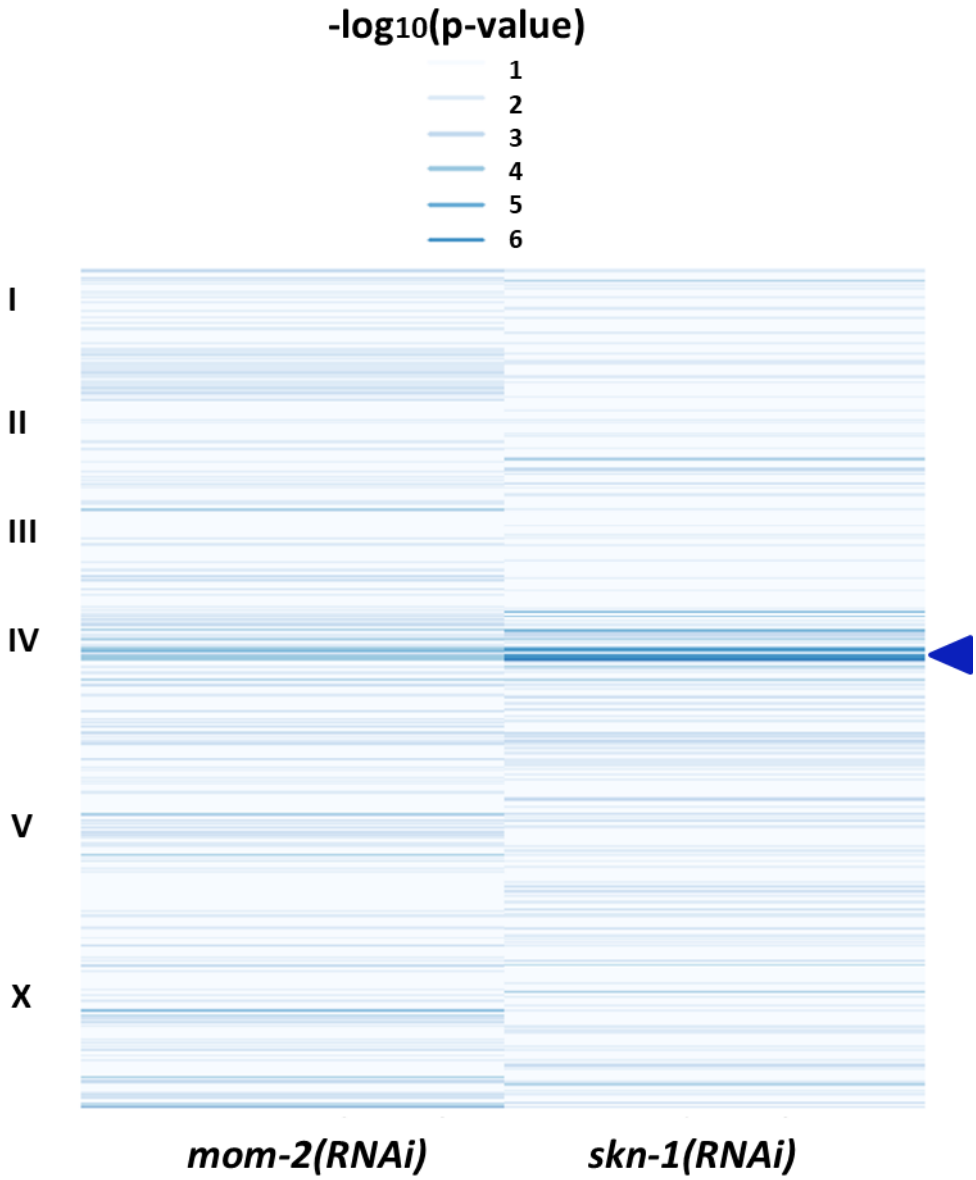


1249

1250 ***Supplementary Fig. 3: Histogram of mom-2(RNAi) phenotype among the 94 wild isolates.***

1251 *A beta-distribution is observed (skewed to the right). Shapiro-Wilk normality test (W=0.8682, p-*
1252 *value=1.207X10^{-7}).*

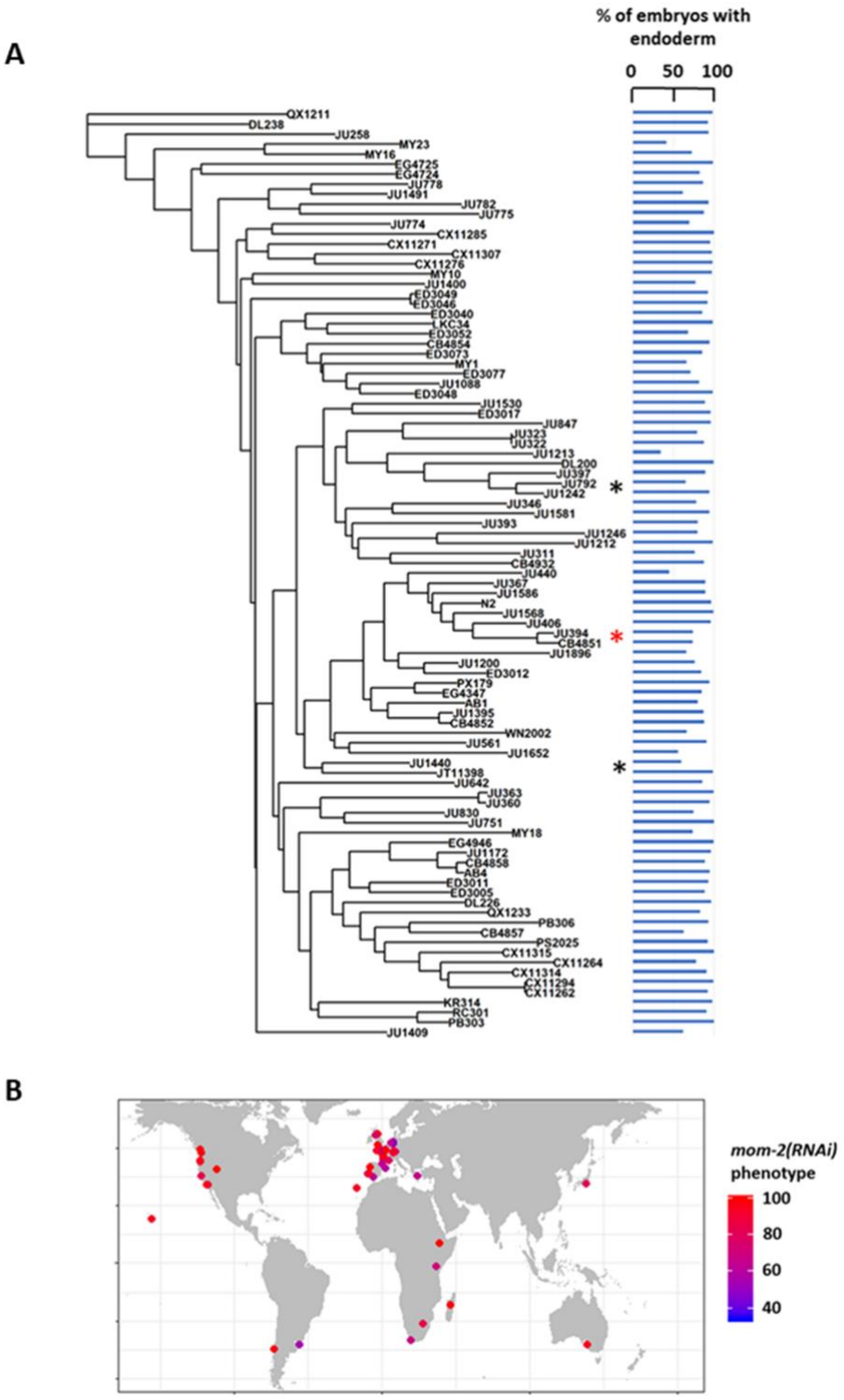
1253



1254

1255 ***Supplementary Fig. 4: Comparison of EMMA p-values for both* mom-2 *and* skn-1 *RNAi phenotypes.***

1256 *Heatmap of p-values for* mom-2(RNAi) *(left) and* skn-1(RNAi) *(right) as calculated in the GWAS analyses*
1257 *(see Fig. 5A, B). Strength of association between genotype and endoderm formation phenotypes is*
1258 *represented as –log$_{10}$(p-value), here depicted as a heatmap (lighter colors – weaker association, darker*
1259 *colors – stronger association). An overlap (indicated by arrow head) is found in a small region of*
1260 *chromosome IV, but no further correlations are observed. Significant SNPs for skn-1(RNAi) are shown in*
1261 *Table S1.*

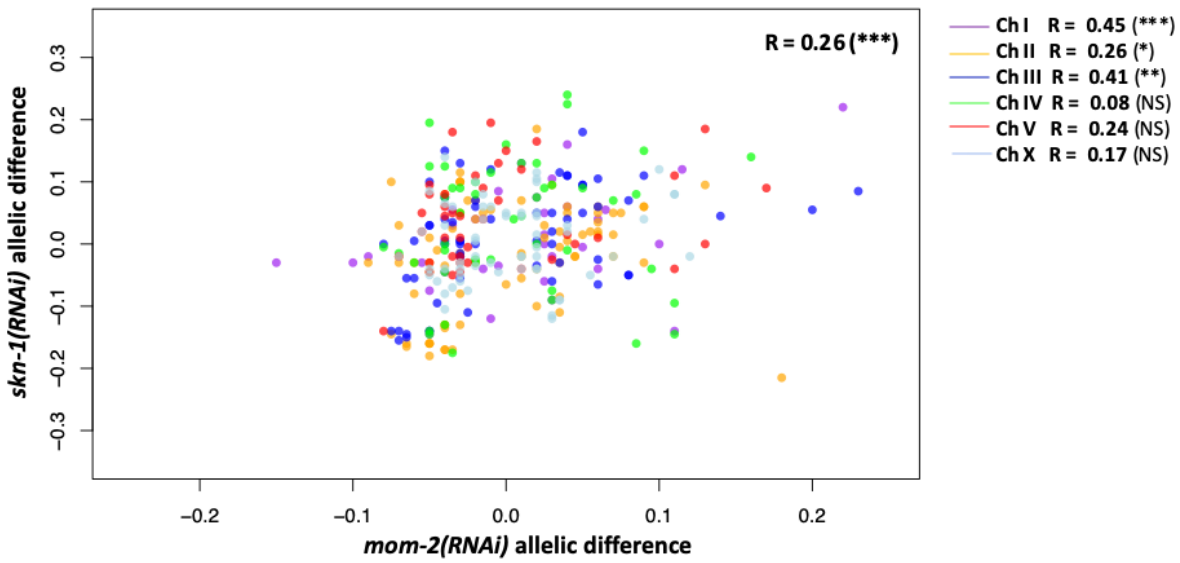


Supplementary Fig. 5: MOM-2 requirement does not correlate with genotypic relatedness or geographical location.

(A) mom-2*(RNAi) phenotype of 94 isolates arranged with respect to the neighbor-joining tree constructed using 4,690 SNPs and pseudo-rooted to QX1211. Red asterisk indicates an example of closely related strains (JU394 and CB4851) with similar phenotypes, while black asterisks indicate examples sister strains (JU792 and JU1242; JU1440 and JT11398) with distinct phenotypes (λ = 6.94X10^-05, p-value = 1). (B) Worldwide distribution of* mom-2(RNAi) *phenotype across 94 isolates. Each circle represents a single isolate.*

1270

1271

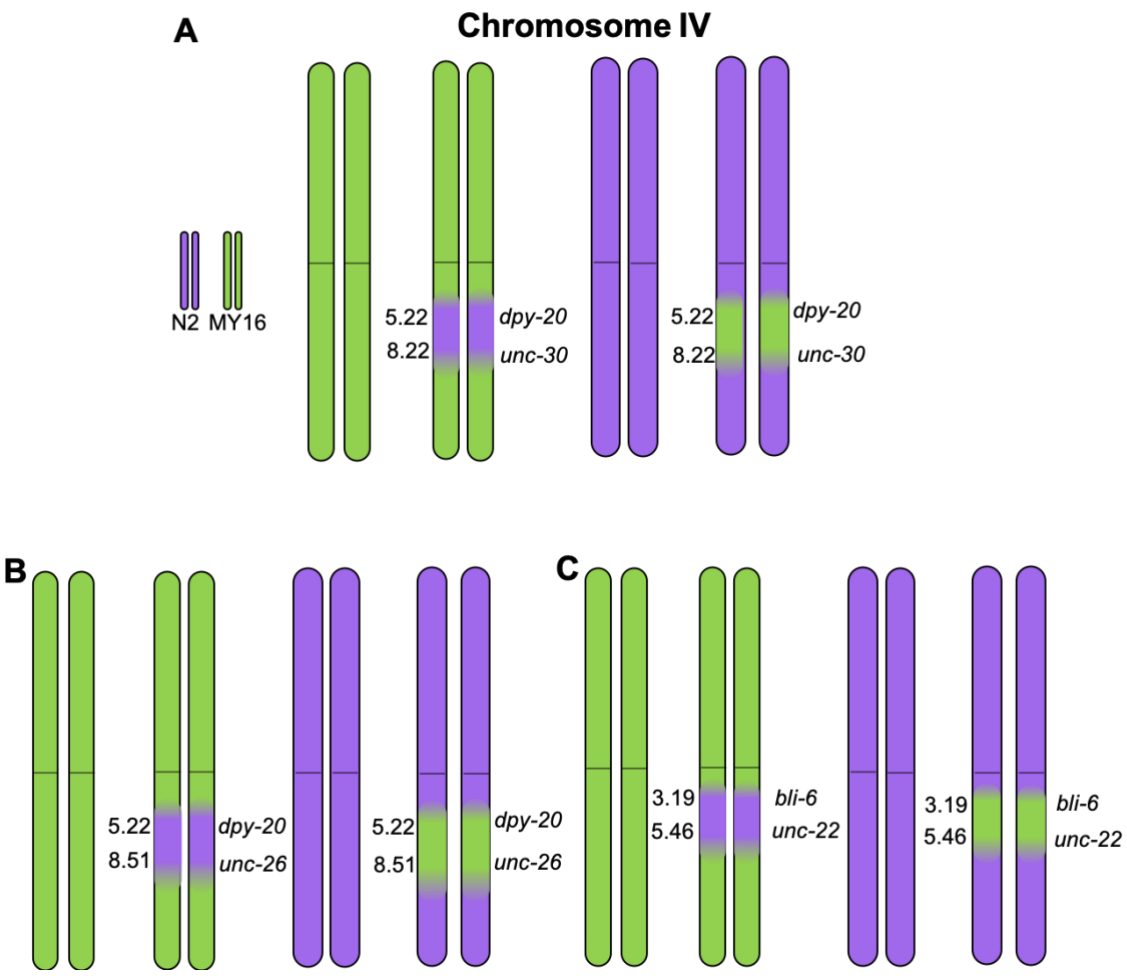


1272

1273 ***Supplementary Fig. 6: Correlation between the skn-1(RNAi) and mom-2(RNAi) allelic differences.***

1274 *Each dot represents a SNP, colour-coded by chromosome. A small genome-wide correlation is observed, but*
1275 *significant for only three of the chromosomes individually. Represented here is a set of pruned SNPs (N =*
1276 *321) to cover the whole genome, and corrected for LD (also used for calculations). Z-score is used to*
1277 *calculate median of each allelic group, correcting for outliers. Strength of correlation (Pearson's R)*
1278 *represented. Significance levels: non-significant, p-value < 0.05 (*), p-value < 0.01 (**), p-value < 0.001*
1279 *(***).*

1280



Supplementary Fig. 7: Introgression of N2 and MY16 regions in chromosome IV.

Near-Isogenic Lines (NILs) were built between strains N2 (purple) and MY16 (green), using visual mutations as markers, with their locations in cM indicated in each graph. (A) Introgression of chromosome IV region flanked by dpy-20 and unc-30. (B) Introgression of chromosome IV region flanked by dpy-20 and unc-26. (C) Introgression of chromosome IV region flanked by bli-6 and unc-22.

Supplementary Table S1: Significantly linked SNPs for skn-1(RNAi) GWAS.

SNP	***EMMA -log(p)***	***N2 allele***	***Variant allele***
IV: 5,079,371 bp	*4.957651645*	*A*	*G*
IV: 5,725,367 bp	*5.211140897*	*C*	*T*
IV: 5,761,153 bp	*5.211140897*	*G*	*A*
IV: 5,891,378 bp	*4.252324884*	*G*	*A*
IV: 5,920,597 bp	*4.720037892*	*T*	*A*
IV: 5,921,302 bp	*4.720037892*	*T*	*G*
IV: 5,921,510 bp	*4.252324884*	*C*	*T*
IV: 6,453,892 bp	*5.142174312*	*T*	*A*

IV: 6,511,989 bp	*5.142174312*	*C*	*A*
IV: 6,563,740 bp	*4.678423021*	*C*	*T*
IV: 7,453,945 bp	*4.652004517*	*G*	*A*
IV: 7,453,143 bp	*4.181579989*	*A*	*G*

1288