Reciprocal requirement of Wnt signaling and SKN-1 underlies cryptic intraspecies 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 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 modifications of gene regulatory networks (GRNs). We report large cryptic variation in the requirement for two key inputs, SKN-1/Nrf2 and MOM-2/Wnt, into the *C. elegans* endoderm-determining GRN. Some natural variants show a nearly absolute requirement for SKN-1 and MOM-2, while in others, most of the embryos differentiate endoderm in their absence. GWAS and analysis of recombinant inbred lines reveal multiple genetic regions underlying this broad phenotypic variation. A striking reciprocal relationship is seen in which genomic variants, or debilitation of genes involved in endoderm formation, that result in high SKN-1 requirement show low MOM-2/Wnt requirement and *vice-versa*. Thus, 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 metazoans, initiating regulatory inputs are remarkable plastic even within a single species.

INTRODUCTION

While the core regulatory machinery that specifies embryonic germ layers and major organ identity in the ancestor of modern animals has been bequeathed to all extant animals, GRN architecture must be able to accommodate substantial plasticity to allow for evolutionary innovation in developmental strategies, changes in selective pressures, and genetic drift [1,2]. Genetic variation, often with neutral effects on fitness, provides for plasticity in GRN structure and implementation [2]. Although studies of laboratory strains of model organisms with a defined genetic background have been highly informative in identifying the key regulatory nodes in GRNs that specify developmental processes [3–5], these approaches do not reveal the evolutionary basis for plasticity in these networks. Which parameters of GRN architecture provide the greatest opportunity for genetically driven evolutionary change, and which are more rigidly fixed? The variation and incipient changes in GRN function and architecture can be discovered by analyzing phenotypic differences resulting from natural genetic variation present in distinct isolates of a single species [6–8].

The endoderm has been proposed to be the most ancient of the three embryonic germ layers in metazoans [9,10], having appeared prior to the advent of the bilateria about 600 Mya [11]. It follows, therefore, that the GRN for endoderm in extant animals has undergone substantial modifications over the long evolutionary time span since its emergence. However, the core transcriptional machinery for endoderm specification and differentiation appears to share common mechanisms across metazoan phylogeny. For example, cascades of GATA-type transcription factors function to promote endoderm development not only in triploblastic animals but in the most ancient creatures that possess endoderm [12–16]. Among the many observations supporting a common regulatory

63 mechanism for establishing the endoderm, it has been found that the endoderm-determining
64 GATA factor, END-1, in the nematode *C. elegans*, is sufficient to activate endoderm
65 development in cells that would otherwise become ectoderm in Xenopus [17]. This indicates
66 that the role of GATA factors in endoderm development has been preserved since the
67 nematodes and vertebrates diverged from a common ancestor that lived perhaps 600 Mya.

68 To assess the genetic basis for evolutionary plasticity and cryptic variation underlying
69 early embryonic germ layer specification, we have analyzed the well-described GRN for
70 endoderm specification in *C. elegans*. The E cell, which is produced in the very early *C. elegans*
71 embryo, is the progenitor of the entire endoderm, which subsequently gives rise exclusively
72 to the intestine. The EMS blastomere at the four-cell stage divides to produce the E founder
73 cell and its anterior sister, the MS founder cell, which is the progenitor for much of the
74 mesoderm [18]. Both E and MS fates are determined by maternally provided SKN-1, an
75 orthologue of the vertebrate Nrf2 bZIP transcription factor [19–21]. In the laboratory N2
76 strain, elimination of maternal SKN-1 function (through either knock down or knockout)
77 results in fully penetrant embryonic lethality as a result of misspecification of EMS cell
78 descendants. In these embryos, the fate of MS is transformed to that of its cousin, the
79 mesectodermal progenitor C cell. E cells similarly adopt a C cell-like fate in a majority, but not
80 all, of these embryos [19]. SKN-1 initiates mesendoderm development via the GRN in E and
81 MS cells in part by activating zygotic expression of the MED-1/2 divergent GATA transcription
82 factors [22,23]. This event mobilizes a cascade of GATA factors in the E cell lineage that
83 ultimately direct intestinal differentiation [21,24,25].

84 This differential requirement for SKN-1 in endoderm (E) and mesoderm (MS)
85 development is determined by its combinatorial action with triply redundant Wnt, MAPK, and

Src signaling systems, which act together to polarize EMS [26–29]. MOM-2/Wnt acts through the MOM-5/Frizzled receptor, mobilizing WRM-1/β-catenin, resulting in its cytoplasmic accumulation in the posterior side of EMS. WRM-1, together with LIT-1/NLK kinase, alters both the nucleocytoplasmic distribution and activity of the Wnt effector POP-1/Tcf [29–31], converting it from a repressor of endoderm in the MS cell lineage to an activator in the E cell lineage [32–37]. Loss of MOM-2 expression in the laboratory N2 strain results in a partial gutless phenotype, while removal of both MOM-2 and SKN-1, through either knockdown or knockout, leads to a completely penetrant loss of gut [29], revealing their genetically redundant roles.

The regulatory relationship between SKN-1 and POP-1, the effector of Wnt signaling, shows striking variation even in relatively closely related species, suggesting substantial evolutionary plasticity in this key node in the endoderm GRN. *C. elegans* embryos lacking maternal POP-1 always make gut, both in the normal E cell lineage and in the MS cell lineage. However, in embryos lacking both SKN-1 and POP-1, endoderm is virtually never made, implying that these two factors constitute a Boolean "OR" logic gate. In contrast, removal of either SKN-1 or POP-1 alone in *C. briggsae* causes >90% of embryos to lack gut, indicative of an "AND" logic gate (Fig. 1A, B) [38].

In this study, we sought to determine whether the aforementioned changes in regulatory logic of the two major inputs into endoderm development are evident within the radiation of a single species. The availability of many naturally inbred variants (isotypes) of *C. elegans* that show widespread genomic variation [39–41], provides a genetically rich resource for investigating potential quantitative variation in developmental GRNs. We report here that the requirement for activation of the endoderm GRN by SKN-1 or MOM-2, but not POP-1, is

profoundly variable between natural *C. elegans* isolates, and even between very closely related isotypes. Thus, the key regulatory inputs into this major embryonic decision is subject to exceedingly rapid evolutionary modification. Genome-wide association studies in isolates from the natural populations and targeted analysis of recombinant inbred lines (RILs), revealed that a multiplicity of loci and their interactions are responsible for the variation in the developmental requirement for SKN-1 and MOM-2. We identified a striking reciprocal requirement for SKN-1 and MOM-2: loci associated with a high requirement for SKN-1 show a lower requirement for MOM-2 and *vice-versa*. We further identified several other endoderm regulatory factors, including RICT-1, PLP-1, and MIG-5, that show similar reciprocal relationships between these two GRN inputs. These findings reveal that the activation of the GRN network in specifying a germ layer, one of the most critical and early developmental switches in embryos, is subject to remarkable genetic plasticity during the radiation of a species and that the dynamic and rapid change in network architecture reflects influences distributed across many genetic components that affect both SKN-1 and Wnt pathways.

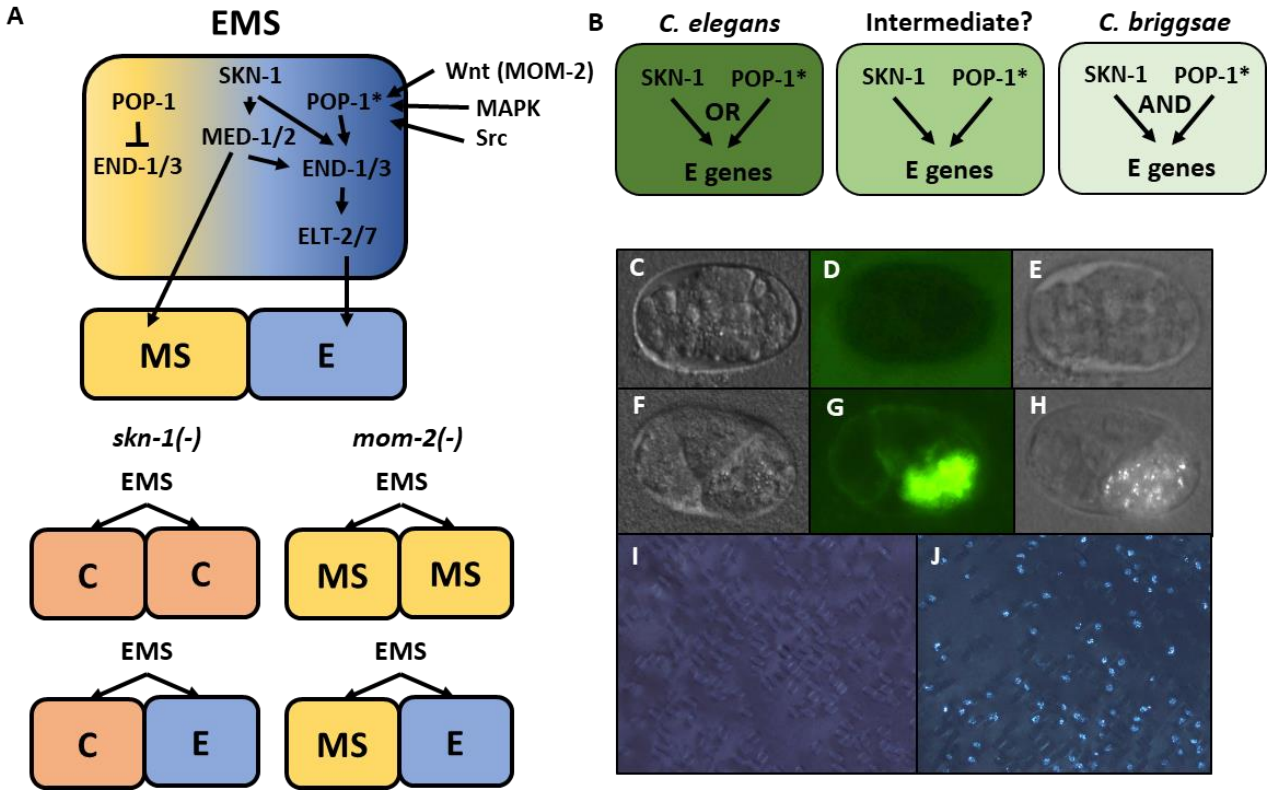


Fig 1. Endoderm regulatory pathway and scoring of gut differentiation.

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 embryonic arrest with a partially penetrant E→MS cell fate transformation, resulting in MS-like daughter cells. (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.

MATERIALS AND METHODS

C. elegans strains and maintenance

All wild isolates, each with a unique haplotype [40], were obtained from the Caenorhabditis Genetics Center (CGC) (see Supplemental file 1). Worm strains were maintained as described [42] and all experiments were performed at 20°C unless noted otherwise. The following mutant and transgenic strains were used in this study: JJ185 *dpy-13(e184) skn-1(zu67) IV; mDp1 (IV;f)* , JR3666 *(elt-2::GFP) X; (ifb-2::GFP) IV*, EU384 *dpy-11(e1180) mom-2(or42) V/nT1 [let-?(m435)] (IV;V)*, JJ1057 *pop-1(zu189) dpy-5(e61)/hT1 I; him-5(e1490)/hT1 V*, KQ1366 *(rict-1(ft7) II*, SU351 *mig-5(rh94)/mIn1 [dpy-10(e128) mIs14] II*, and RB1711 *plp-1(ok2155) IV*.

RNAi

Feeding-based RNAi experiments were performed as described [43]. RNAi clones were obtained from either the Vidal [44] or Ahringer libraries [45]. RNAi bacterial strains were grown at 37°C in LB containing 50 µg/ml ampicillin. The overnight culture was then diluted 1:10. After 4 hours of incubation at 37°C, 1 mM of IPTG was added and 60µl was seeded onto 35mm agar plates containing 1 mM IPTG and 25 µg/ml carbenicillin. Seeded plates were allowed to dry and used within five days. Five to 10 L4 animals were placed on RNAi plate. 24 hours later, they were transferred to another RNAi plate and allowed to lay eggs for four or 12 hours (12 hours for *skn-1* RNAi and four hours for the other RNAi). The adults were then removed, leaving the embryos to develop for an extra 7-9 hours. Embryos were quantified and imaged on agar pad using Nikon Ti-E inverted microscope.

Antibody staining

The embryonic gut cells and nuclei of all cells were stained with MH33 (mouse anti-IFB-2, deposited to the DSHB by Waterston, R.H.) and AHP418 (rabbit anti-acetylated histone H4, Serotec Bio-Rad) respectively. Fixation and permeabilization were carried out as described previously [46]. Goat anti-mouse Alexa Fluor® 594 and goat anti-rabbit Alexa Fluor® 488 secondary antibodies were used at 1:1000 dilution.

Quantification of endoderm specification

Gut was scored by presence of birefringent gut granule in arrested embryos [47,48]. For *skn-1(RNAi)*, the laboratory strain N2, which shows invariable ~30% of embryos with endoderm, was used as a control for all experiments.

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

To introgress *skn-1(zu67)* into wild isolates (WI), males from the wild isolate strains were crossed to JJ186 *dpy-13(e184) skn-1(zu67) IV; mDp1 (IV;f)* hermaphrodites. mDp1 is a free duplication that rescues the Dpy and lethal phenotypes of *dpy-13(e184)* and *skn-1(zu67)* respectively. Animals that have lost the free duplication will be Dpy and produce dead offspring. Wild type F1 hermaphrodites that have lost the free duplication as determined by presence of a ¼ Dpy progeny in the F2 were selected. 10 single non-Dpy F2 hermaphrodite descendants from F1 animals heterozygous for *skn-1(zu67)* (2/3 of which would be of the genotype *WI dpy-13(+) skn-1(+)/ dpy-13(e184) skn-1(zu67)* were backcrossed to their respective parental wild strain. 10 F3 hermaphrodites were picked to individual plates. Half of the F3 cross progeny would be heterozygous for *dpy-13(e184) skn-1(zu67)*, as evidenced by presence of F4 Dpy progeny that produced dead embryos. Non-Dpy siblings were used to continue the introgression as described. This was repeated for at least 5 rounds of

introgression. The embryonic gutless phenotype in the progeny of the Dpy animals was quantified.

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

Statistical Analyses: GWAS and EMMA

All data were analyzed and plotted using R software v 3.2.3 (https://www.r-project.org/). GWAS for both phenotypes was performed using *C. elegans* wild isolates and a previously published SNP map containing 4,690 SNPs [40] with the EMMA R package. P-values were calculated using mixed model analysis [49] (emma.REML.t() function) and IBS kinship matrix to account for population structure. For *skn-1* and *mom-2* RNAi phenotypic data, a genome-wide permutation-based FDR was also calculated for the EMMA results from 10,000 permuted values [50,51]. In addition, a linear model GWAS was performed with the same SNP map (but no kinship matrix) on both *mom-2* and *skn-1* datasets, with FDR calculations obtained from 10,000 permuted values. However, owing to the skewed nature of the *mom-2(RNAi)* data (Supplemental Fig. 4), genome-wide permutation-based FDR thresholds did not reveal any significant loci. The p-values for each individual SNP were then adjusted based on 1000 permutations at each locus. Significance thresholds were set at $p<0.01$ and 0.001.

Correlation Analysis

In order to test for the relationship between *mom-2 (RNAi)* and *skn-1 (RNAi)* phenotypic data, the difference between median phenotypic values for each SNP were calculated independently on both a genome-wide level (N = 4690) and at the SNPs most significantly associated with the *mom-2 (RNAi)* phenotype (N = 45, $p < 0.01$). Pearson's correlation test was used to calculate correlation between median phenotypic values for genome-wide analysis using a sliding window (N = 50 SNPs). Spearman's Rho was used to calculate the correlation using only SNPs most significantly associated with *mom-2* GWAS.

RIL construction and Genotype-By-Sequencing (GBS)

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

DNA was extracted using Blood and Tissue QIAGEN kit from worms from each of the RILs grown on four large NGM plates (90x15mm) with OP50 *E. coli* until starved (no more than a day). Samples were submitted in 96-well plate format at 10 ng/µl < n < 30 ng/µl. GBS libraries were constructed using digest products from ApeKI (GWCGC), using a protocol modified from [52]. After digestion, the barcoded adapters were ligated and fragments < 100bp were sequenced as single-end reads using an Illumina HiSeq 2000 lane (100 bp, single-end reads).

SNP calling was performed using the GBSversion3 pipeline in Trait Analysis by aSSociation, Evolution and Linkage (TASSEL) [53]. Briefly, fastq files were aligned to reference genome

WS252 using BWA v. 0.7.8-r455 and SNPs were filtered using vcftools [54]. Samples with greater than 90% missing data and SNPs with minor allele frequencies (mAF) of <1% were excluded from analysis, identifying 27,396 variants.

QTL mapping using R/qtl

Variants identified by GBS pipeline were filtered to match the SNPs present in the parental MY16 strain (using vcftools –recode command), and variants were converted to a 012 file (vcftools –012 command). Single-QTL analysis was performed in R/QTL [55] using 1770 variants and 95 RILs. Significant QTL were determined using Standard Interval Mapping (scanone() "em") and genome-wide significance thresholds were calculated by permuting the phenotype (N =1,000). Change in log-likelihood ratio score of 1.5 was used to calculate 95% confidence intervals and define QTL regions [56].

RESULTS

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

The relationship between SKN-1 and Wnt signaling through POP-1 in the endoderm GRN has undergone substantial divergence in the *Caenorhabditis* genus [38]. While neither input alone is absolutely required for endoderm specification in *C. elegans*, each is essential in *C. briggsae*, which has been estimated to have diverged from *C. elegans* ~20-40 Mya [57,58]. In contrast to the *C. elegans* N2 laboratory strain, removal of either SKN-1 or POP-1 alone results in fully penetrant conversion of the E founder cell fate into that of the mesectodermal C blastomere and of E to MS fate, respectively, in *C. briggsae* [38]. These findings revealed that the earliest inputs into the endoderm GRN are subject to substantial evolutionary differences between these two species (Fig. 1B). We sought to determine whether incipient evolutionary plasticity in this critical node at the earliest stages of endoderm development might be evident even within a single species of the *Caenorhabditis* genus by assessing their requirement in *C. elegans* wild isolates and testing whether the quantitative requirements of each input were correlated.

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

reproducible results: 100% of the embryos derived from *skn-1(RNAi)*-treated mothers arrest (n>100,000) and 32.0 ± 1.9% of the arrested embryos exhibited birefringent gut granules (Fig. 2A; Supplemental Fig. 1). We found that the LSJ1 laboratory strain, which is derived from the same original source as N2, but experienced very different selective pressures in the laboratory owing to its constant propagation in liquid culture over 40 years [59], gave virtually identical results to that of N2 (31.0% ± s.d 1.2%), implying that SKN-1-independent endoderm formation is a quantitatively stable trait. The low variability in this assay, and high number of embryos that can be readily examined (≥500 embryos per experiment), provides a sensitive and highly reliable system with which to analyze genetic variation in the endoderm GRN between independent *C. elegans* isolates.

To assess variation in SKN-1 requirement within the *C. elegans* species, we analyzed the outcome of knocking down SKN-1 by RNAi in 96 unique *C. elegans* wild isolates [40]. Owing to their propagation by self-fertilization, each of the isolates (isotypes) is a naturally inbred clonal population that is virtually homozygous and defines a unique haplotype. The reported estimated population mutation rate averages 8.3×10^{-4} per bp [40], and we found that a substantial fraction (29/97) of isotypes were quantitatively indistinguishable in phenotype between the N2 and LSJ1 laboratory strains (Fig. 2A). We found that all strains, with the exception of the RNAi-resistant Hawaiian CB4856 strain, invariably gave 100% embryonic lethality *with skn-1(RNAi)*, showing that on the basis of that criterion all strains are fully sensitive to RNAi. However, we observed dramatic variation in the fraction of embryos with differentiated gut across the complete set of strains, ranging from 0.9% to 60% (Fig. 2A). Repeated measurements with >500 embryos per replicate per strain revealed very high reproducibility (Supplemental Fig. 1), indicating that even small differences in the fraction of embryos generating endoderm could be reproducibly measured. Further, we found that some

293 wild isolates that were subsequently found to have identical genome sequences also gave
294 identical results. We note that these results contrast with those of Paaby *et al*. [60], who
295 found that RNAi in liquid culture of a set of 55 wild isolates resulted in much weaker effects,
296 both on lethality and on gut differentiation. This difference is likely attributable to variability
297 in RNAi efficacy in the latter study [61,62].

298 Although birefringent and autofluorescent rhabditin granules have been used as a
299 marker of gut specification and differentiation in many studies [47,48], it is conceivable that
300 the variation in fraction of embryos containing this marker that we observed might reflect
301 variations in gut granule formation rather than in gut differentiation *per se*. We note that
302 embryos from all strains showed a decisive "all-or-none" phenotype: i.e., they were either
303 strongly positive for gut differentiation or completely lacked gut granules, with virtually no
304 intermediate or ambiguous phenotypes. A threshold of gene activity in the GRN has been
305 shown to account for such an all-or-none switch in gut specification [22,63,64]. This
306 observation is inconsistent with possible variation in gut granule production: if SKN-1-
307 depleted embryos were defective in formation of the many granules present in each gut cell,
308 one might expect to observe gradations in numbers or signal intensity of these granules
309 between gut cells or across a set of embryos. Nonetheless, we extended our findings by
310 analyzing expression of the gut-specific intermediate filament IFB-2, a marker of late gut
311 differentiation, in selected strains representing the spectrum of phenotypes observed (Fig.
312 2B). As with gut granules, we found that embryos showed all-or-none expression of IFB-2. In
313 all cases, we found that the fraction of embryos containing immunoreactive IFB-2 was not
314 significantly different (Fisher's exact test, p-values > 0.05) from the fraction containing gut
315 granules, strongly suggesting that the strains vary in endoderm specification *per se* and
316 consistent with earlier studies of SKN-1 function [19,22].

Although we found that *skn-1(RNAi)* was 100% effective at inducing embryonic lethality in all strains (with the exception of the RNAi-defective Hawaiian strain, CB4856), it is conceivable that, at least for the strains that showed a weaker phenotype than for N2 (i.e., higher number of embryos specifying endoderm), the variation observed between strains might be attributable to differences in RNAi efficacy rather than in the endoderm GRN. To address this possibility, we introgressed the strong loss-of-function *skn-1(zu67)* chromosomal mutation into five wild isolates whose phenotypes spanned the spectrum observed (ranging from 2% of embryos with differentiated gut for MY16 to 50% for MY1) (Fig. 2C). In all cases, we found that introgression of the allele through five rounds of backcrosses resulted in a quantitative phenotype that was similar or identical to that of the effect observed with *skn-1(RNAi)*. The phenotypes of the introgressed allele were significantly different (p-values <0.01) from that of the parental N2 *skn-1(zu67)* strain, except for DL238, whose *skn-1(RNAi)* phenotype was indistinguishable from that of N2. The results obtained by introgression from four of the isotypes (CX11262, DL238, EG4724 and MY1), were not statistically different (Student t-test, p-values >0.05) from the corresponding RNAi knock down results (Fig. 2C) (i.e., the phenotype was suppressed or enhanced relative to N2 in these genetic backgrounds to the same extent as with *skn-1(RNAi)*). However, while the MY16 *skn-1(zu67)* strain shifted in the predicted direction (i.e., became stronger) when compared to the N2 strain, it was a weaker effect than was evident by RNAi knockdown, even following eight rounds of introgression. Nonetheless, diminished RNAi efficacy in MY16 cannot explain the large difference in *skn-1(RNAi)* phenotype between N2 and MY16, as the latter phenotype is much stronger, not weaker, than the former. As described below, we identified a modifier locus in the MY16 strain that is closely linked to the *skn-1* gene; it seems likely that the N2 chromosomal segment containing this modifier was carried with the *skn-1(zu67)* mutation

through the introgression crosses, thereby explaining the somewhat weaker phenotype of the introgressed allele compared to the RNAi effect in MY16. The results of introgression of the *skn-1(zu67)* chromosomal mutation confirm that the extreme variation in *skn-1(RNAi)* phenotype between the wild isolates results from *bona fide* cryptic variation in the endoderm GRN, rather than differences in RNAi efficacy.

We note that the strength of *skn-1(RNAi)* phenotype does not correlate with phylogenetic relatedness between the strains (Mantel test r = 0.21, NS). To illustrate, while some closely related strains (e.g., MY16 and MY23) showed a similar gut developmental phenotype, some very closely related strains (e.g., JU1491 and JU778) had phenotypes on the opposite ends of the phenotypic spectrum (Fig. 3A). We also did not observe any significant correlation between geographical distribution and *skn-1 (RNAi)* phenotype (Fig. 3B). These findings suggest that the endoderm GRN may be subject to rapid intraspecies evolutionary divergence and suggests that a small number of loci may underlie variation in the trait.

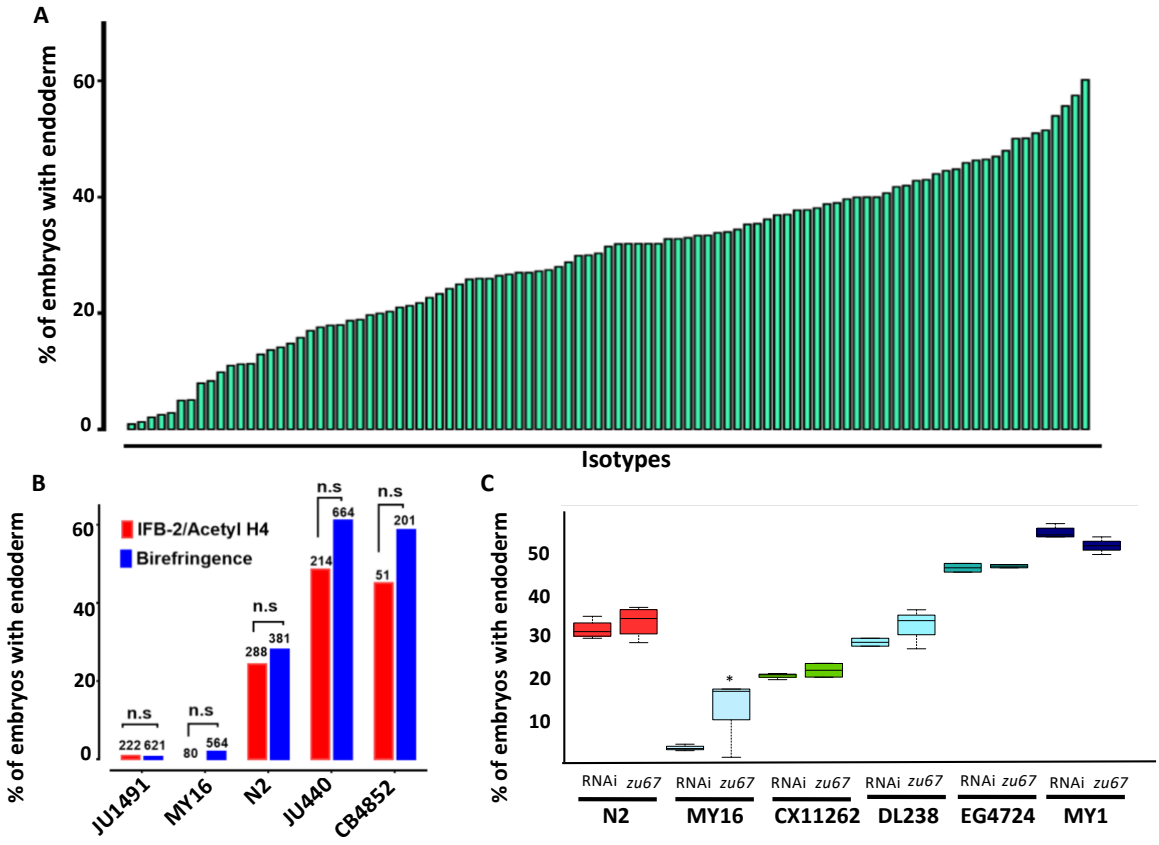


Fig. 2: Quantitative effects of loss of *skn-1* on endoderm formation.

(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 (5 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).

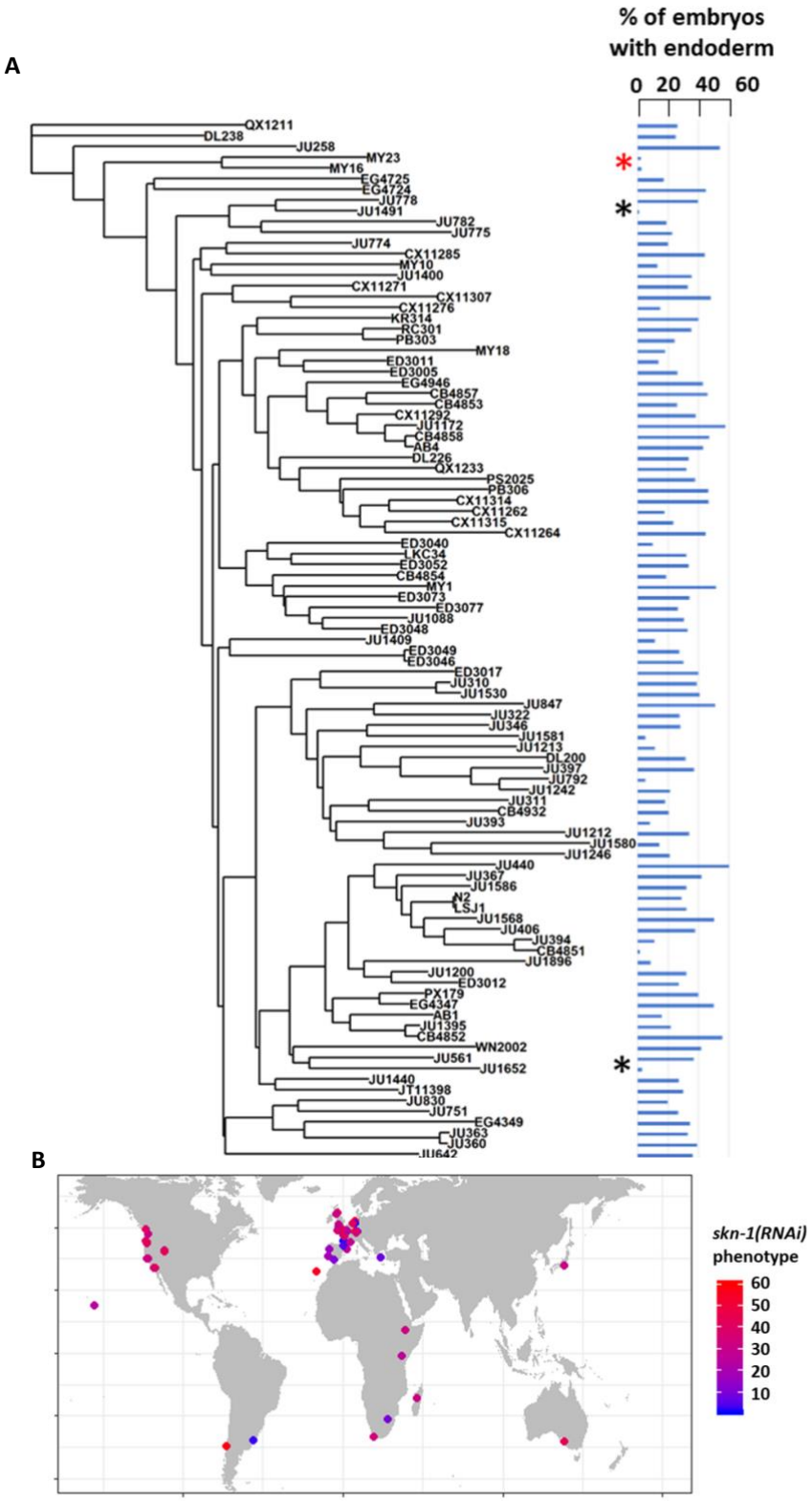


Fig. 3: SKN-1 requirement does not correlate with genotypic relatedness or geographical location. (A) *skn-1(RNAi)* phenotype of 97 isolates arranged with respect to the neighbor-joining tree constructed using 4,690 SNPs and pseudo-rooted to QX1211. Red asterisk indicates an example of closely related 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 (measured as Euclidean distance) are not significantly correlated (Mantel test, $r = 0.21$, NS). (B) Worldwide distribution of *skn-1(RNAi)* phenotype across 97 wild isolates. Each circle represents a single isotype.

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

The switch in the relationship of the SKN-1 and Wnt inputs between *C. elegans* ("OR" operator) and *C. briggsae* ("AND" operator) [38], and the extensive variation in the requirement for SKN-1 seen across *C. elegans* isolates, raised the possibility that the quantitative requirement for Wnt components might vary between unique isolates of *C. elegans*. It has been shown that signaling from Ras pathway varies in different *C. elegans* wild isolates and hyperactive Wnt signaling can compensate for reduced Ras activity in the vulva signaling network [6,65]. Given that removal of the maternal Wnt input also results in a partially penetrant gut defect (through either knock-out or knockdown of Wnt signaling components), it is conceivable that a compensatory relationship may exist between the SKN-1 and Wnt inputs. We investigated this possibility by examining the requirement for the MOM-2/Wnt ligand in the same wild isolates that were tested for the SKN-1 gut developmental requirement. Indeed, we observed broad variation in the requirement for MOM-2/Wnt in activation of the endoderm GRN between isotypes. *mom-2(RNAi)* of 94 isotypes resulted in embryonic arrest, indicating that, as with *skn-1(RNAi), mom-2(RNAi*) was effective at least by the criterion of lethality. Two isotypes, CB4853 and EG4349, did not exhibit *mom-2(RNAi)*-induced lethality and were omitted from further analyses. In the affected strains, the fraction of *mom-2(RNAi)* embryos with differentiated gut varied from ~40% to ~99% (Fig. 4A). As with *skn-1(RNAi)*, the *mom-2(RNAi)* phenotype of isotypes N2, JU440, and JU1213 was further confirmed by immunostaining with IFB-2 (Fig 4B), again demonstrating that birefringence of gut granules is a reliable proxy for endoderm formation for this analysis.

To assess whether the observed variation in the *mom-2(RNAi)* phenotype reflected differences in the GRN or RNAi efficacy, the *mom-2(or42)* allele was introgressed into three different genetic backgrounds chosen from the extreme ends of the phenotypic spectrum. *mom-2(RNAi)* of the laboratory N2 strain resulted in the developmental arrest of embryos. Of those, ~90% contained differentiated endoderm, a result that was highly reproducible. In contrast, the introgression of an apparent loss-of-function allele, *mom-2(or42)*, into the N2 strain results in a more extreme phenotype: only ~28% of embryos show endoderm differentiation (Fig. 4C) [29]. While this discrepancy can partly be explained by incomplete RNAi efficacy, it is notable that the penetrance of *mom-2* alleles vary widely [29]. We observed strain-specific variation in embryonic lethality response to RNAi of *mom-2* between the different isotypes. However, we found that the *mom-2(or42)* introgressed strains show qualitatively similar effects to those observed with *mom-2* RNAi. For example, the *mom-2(or42)* allele introgressed into the isotype JU1213 background resulted in a severe gutless phenotype (5.7% ± s.d 2.4%; n=2292) a similar but more extreme effect than was seen with RNAi (34.0% ± s.d 1.5%; n=1876). This is the strongest phenotype that has been reported for any known *mom-2* allele. On the other hand, introgression of the *mom-2* mutation gave rise to a significantly higher fraction of embryos with endoderm in isotypes DL226 (55.2% ± s.d 1.2%, n=1377) and PB303 (65.5% ± s.d 4.9%, n=2726), relative to the laboratory strain N2 (29.1% ± s.d 3.1%; n=1693), consistent with the RNAi phenotypes (Fig. 4C). These findings indicate that the differential requirement for MOM-2 is, at least in part, attributable to genetic modifiers in these strains.

As the MOM-2/Wnt signal is mediated through the POP-1 transcription factor, we sought to determine whether the requirement for POP-1 might also vary between isolates. We found that, while *pop-1(RNAi)* resulted in 100% embryonic lethality across all 96 RNAi-

425 sensitive isolates, 100% of the arrested embryos contained a differentiated gut (n>500 for
426 each isolate scored) (results not shown). Thus, all isolates behave similarly to the N2 strain
427 with respect to the requirement for POP-1. These results were confirmed by introgressing a
428 strong loss-of-function *pop-1(zu189)* allele into four wild isolates (N2, MY16, JU440, and
429 KR314) (Supplemental Fig. 2). The lack of variation in endoderm specification after loss of
430 POP-1 is not entirely unexpected. As has been observed in a *pop-1(-)* mutant strain,
431 elimination of the endoderm-repressive role of POP-1 in the MS lineage (which is not
432 influenced by the P2 signal) supersedes its endoderm activating role in the presence of SKN-
433 1. Indeed, the original observation that all *pop-1(-)* embryos in an N2 background contain gut
434 masked the activating function for POP-1, which is apparently only in the absence of SKN-1
435 [32,34,36]. It is likely that, as with the N2 strain, gut arises from both E and MS cells in all of
436 these strains; however, as we have scored only for presence or absence of gut, it is
437 conceivable that the E lineage is not properly specified in some strains, a possibility that
438 cannot be ruled out without higher resolution analysis.

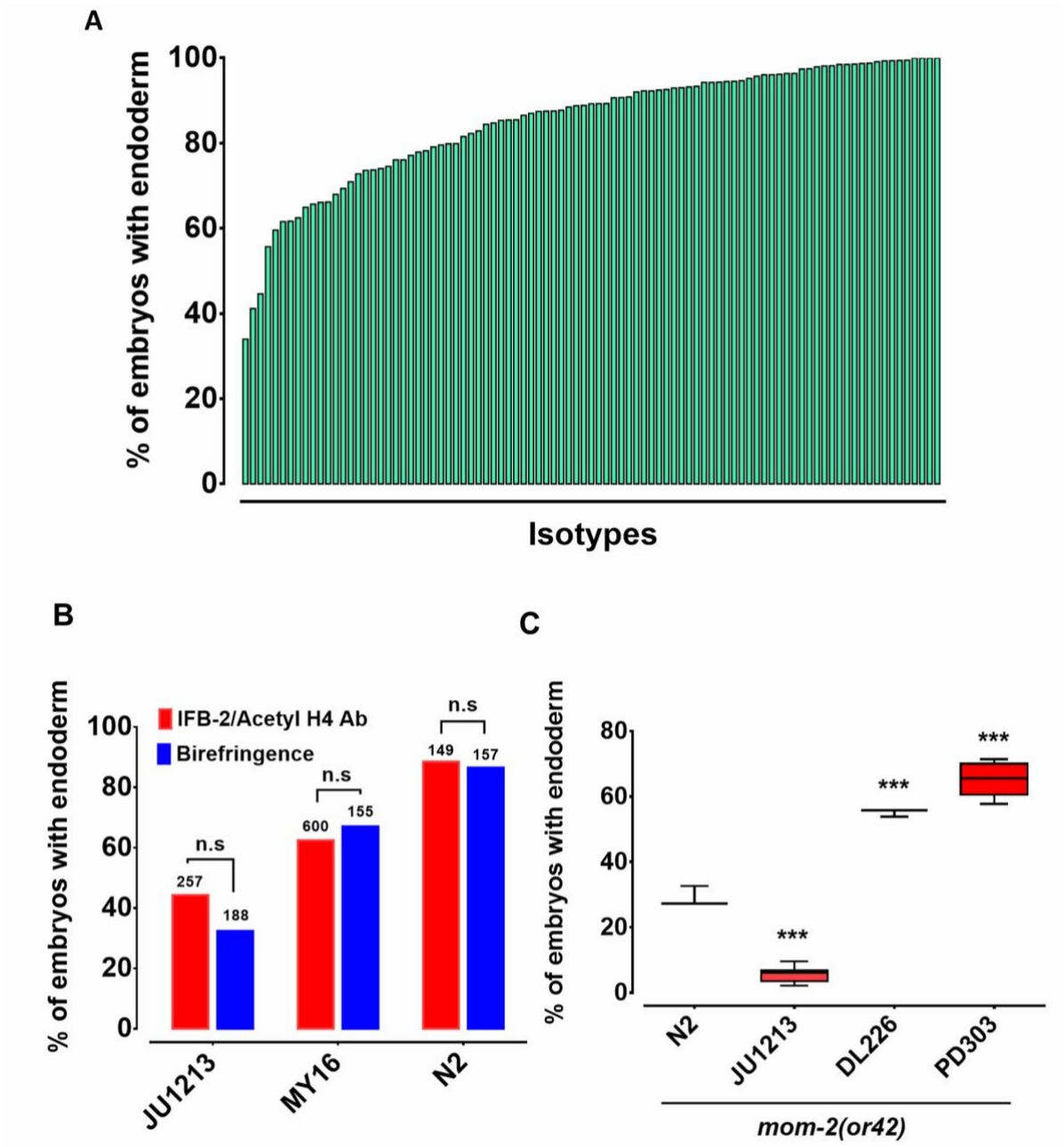


Fig. 4: Wide variation in the *mom-2(RNAi)* phenotype.

(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 staining 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*) shown by dashed line. Student t-test (*** p-value<0.001).

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

We sought to examine the genetic basis for the wide variation in SKN-1 and Wnt requirements across *C. elegans* isolates and to evaluate possible relationships in the variation seen with the SKN-1 and Wnt inputs by performing linear-model GWAS using the available SNP makers and map [40]. This analysis identified two highly significantly associated regions on chromosome IV and V (FDR < 1.5) that underlie the variation in SKN-1 requirement (Supplemental Fig. 3A). To ensure that these two QTLs were not an artifact of genetic relatedness between sets of strains, we applied the more stringent EMMA (Efficient Mixed-Model Analysis) algorithm, which adjusts for population structure (Fig. 5A) [49,66]. This approach also identified the same significant location on chromosome IV. The two mapping approaches show a moderate linear correlation (Spearman correlation coefficient = 0.43) (Supplemental Fig. 3B). In each case, the most statistically significant SNPs within each of the two identified QTLs are highly associated with the observed variance in SKN-1 requirement (Fig 5C; Supplemental Fig. 3C, D).

GWAS analysis on the *mom-2(RNAi)* phenotypic variation proved more challenging because this phenotype showed a highly skewed distribution (Shapiro-Wilk' test W =0.8682, p-value = $1.207X10^{-7}$) (Supplemental Fig. 4). Nevertheless, by applying a linear model GWAS and adjusting the individual p-values using a permutation-based approach (see Materials and Methods) and EMMA (Fig. 5B, Supplemental Fig. 5A), both GWAS and EMMA revealed highly correlated results (Pearson's R = 0.95, p-value < 2.2e-16) (Supplemental Fig. 5B). Although GWAS identified 45 significant SNPs distributed across the genome (GWAS adjusted p-values < 0.01), EMMA did not reveal any significant genomic regions for *mom-2(RNAi)* variation, suggesting that the MOM-2 requirement is a highly complex trait influenced by many loci.

However, when we compared the p-values of individual SNPs from *skn-1(RNAi)* and *mom-2(RNAi)* EMMA, a substantial overlap in the central region of chromosome of chromosome IV was observed (Supplemental Fig. 6). This genomic region showed striking reciprocality in phenotype compared to the SKN-1 results, as described below.

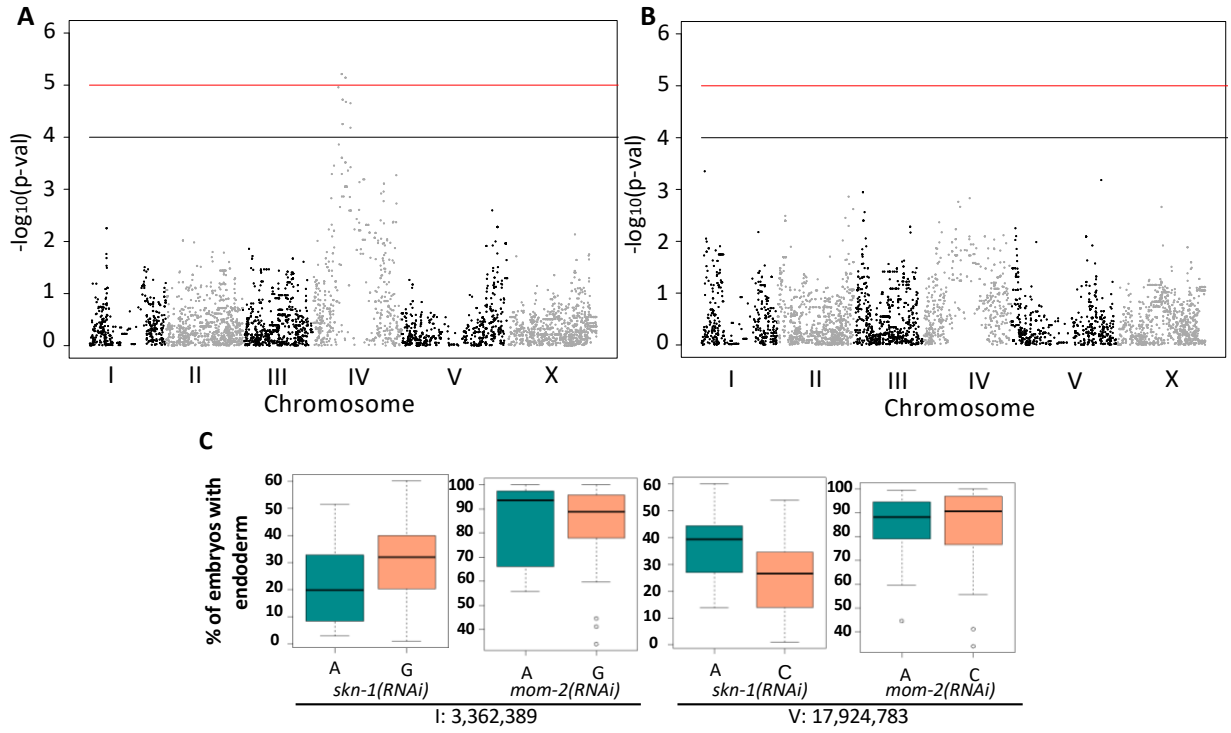


Fig. 5. Genome-Wide Association Studies of *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes.

(A) Manhattan plot of *skn-1(RNAi)* EMMA. 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 chromosomes are color-coded. 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 significant SNPs from *skn-1(RNAi)* GWAS at position 3,362,389 bp on chromosome I and position 17,924,783 bp on chromosome V (see Supplemental Fig. 3). 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 SKN-1 and MOM-2 (see Materials and Methods). In contrast to the very low variation

493 seen between multiple trials of each parental strain, analysis of the RNAi treated RIL strains
494 (>500 embryos/RIL) revealed a very broad distribution of phenotypes. We found that, while
495 some RILs gave phenotypes similar to that of the two parents, many showed intermediate
496 phenotypes and some were reproducibly more extreme than either parent, indicative of
497 transgressive segregation [67]. For *skn-1(RNAi)*, the phenotype varied widely across the RILs,
498 with 1 to 47% of embryos containing gut (Fig. 6A). This effect was even more striking with
499 *mom-2(RNAi)*, for which virtually the entire possible phenotypic spectrum was observed
500 across a selection of 31 RILs representing the span of *skn-1(RNAi)* phenotypes. The *mom-*
501 *2(RNAi)* phenotypes ranged from RILs showing 3% of embryos with gut to those showing 92%
502 (Fig. 6A). It is noteworthy that one RIL (JR3572, Supplemental File 2) showed a nearly
503 completely penetrant gutless phenotype, an effect that is much stronger than has been
504 previously observed for *mom-2(-)* [29]. These results indicate that a combination of natural
505 variants can nearly eliminate a requirement for MOM-2 altogether, while others make it
506 virtually essential for endoderm development. Collectively, these analyses reveal that
507 multiple quantitative trait loci (QTL) underlie SKN-1- and MOM-2-dependent endoderm
508 specification.

509 To identify QTLs from the recombinant population, we performed linkage mapping for
510 both phenotypes using both interval mapping and marker regression. For *skn-1(RNAi)*, two
511 major peaks were revealed on chromosomes II and IV (above 1% FDR estimated from 1,000
512 permutations). Two minor loci were found on chromosomes I and X (suggestive linkage,
513 above 20% FDR) (Fig 6B). For *mom-2(RNAi)*, two major independent QTL peaks were found
514 on Chromosomes I and II (above the 5% FDR estimated from 1,000 permutations). Although
515 the candidate peaks observed on Chromosome IV for *skn-1(RNAi)* (Fig. 6B) did not appear to
516 overlap with those for *mom-2(RNAi*), overlap was observed between the Chromosomes I and

II candidate regions for these two phenotypes. Moreover, many N2 and MY16 alleles in the QTL showed a reciprocal effect on SKN-1 and MOM-2 dependence (Fig. 6C-F).

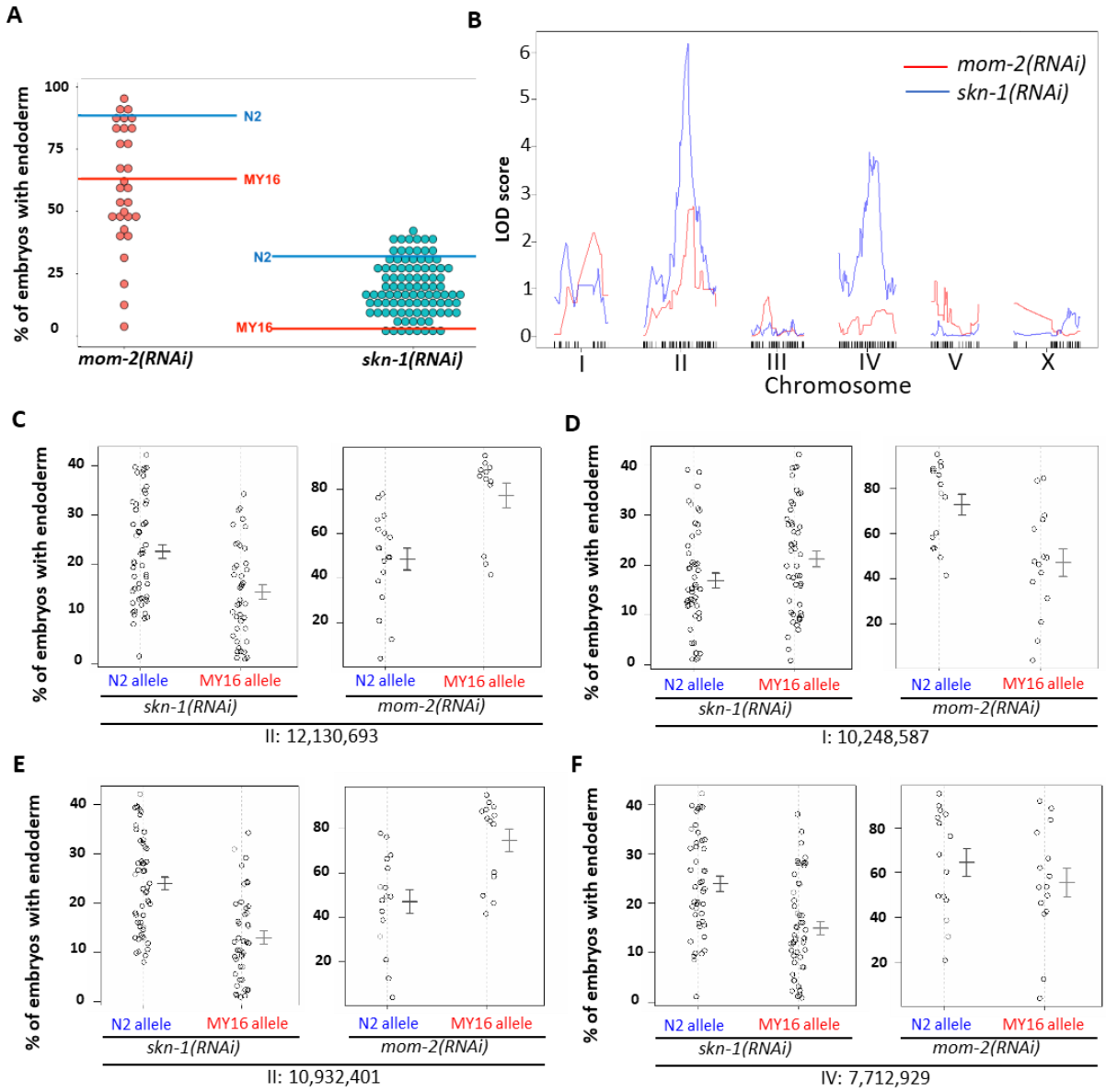


Fig 6. Quantitative genetic analysis of *mom-2(RNAi)* and *skn-1(RNAi)* phenotype in Recombinant Inbred Lines (RILs) between N2 and MY16.

(A) *mom-2(RNAi)* (left) and *skn-1(RNAi)* (right) phenotype of RILs. The phenotype of the parental strains, MY16 and N2 are shown by red and blue lines, respectively. (B) QTL analyses (interval mapping) of *skn-1(RNAi)* (blue line) and *mom-2(RNAi)* (red line) phenotype shown in (A). Genomic regions are shown on the x-axis and LOD score is shown on the y-axis. (C-F) Effect plots of significant SNPs from *mom-2(RNAi)* (C, D) and *skn-1(RNAi)* (E, F) QTL analyses of RILs. Each dot represents a RIL. The parental alleles are shown on the x-axis, and *skn-1(RNAi)* or *mom-2(RNAi)* phenotypes on the y-axis. Confidence intervals for the average phenotype in each genotype group are shown.

A cryptic compensatory relationship between the SKN-1 and Wnt regulatory inputs

As with *skn-1(RNAi)* findings, we found no correlation between the *mom-2(RNAi)* phenotype and phylogenetic relatedness or geographical distribution (Supplementary Fig. 7A, B), suggesting rapid intraspecies developmental system drift. This, together with the preceding findings, unveiled wide cryptic variation in the requirements for both SKN-1 and MOM-2/Wnt in the endoderm GRN and raised the possibility of a functional overlap in this variation. Comparisons of the GWAS and QTL mapping results for *skn-1* and *mom-2* showed an overlap in candidate QTL regions on chromosome I, II and IV (Fig. 5, Fig 6, Supplemental Fig. 6), suggesting a possible relationship between the genetic basis underlying these two traits. It is conceivable that some genetic backgrounds are generally more sensitive to loss of either input (e.g, the threshold for activating the GRN is higher) and others more robust to single-input loss. Alternatively, a higher requirement for one input might be associated with a relaxed requirement for the other, i.e., a reciprocal relationship.

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

We next sought to examine possible relationships between the two GRN inputs at higher resolution by comparing association of specific genetic regions with the quantitative requirement for each input. We took advantage of the available sequence data for all the

554 isotypes tested [40] and examined the impact of each allele on the *skn-1(RNAi)* and *mom-*
555 *2(RNAi)* phenotypes for the SNPs that were most highly associated with the variation in
556 requirement for SKN-1 by calculating the difference between the phenotypic medians for
557 each allele at each SNP. Comparisons of the GWAS analyses for variation in the requirement
558 for SKN-1 and MOM-2 showed a particularly strong overlap in candidate QTL regions for the
559 two phenotypes on chromosome IV (Fig. 5A, B, Supplemental Fig. 3A, 5, 6). To assess the
560 relationship between these and other significant SNPs, we analyzed the top 45 SNPs from the
561 *mom-2* GWAS data and found a strong negative correlation between the allelic effects for
562 SKN-1 and MOM-2 dependence. With very few exceptions, SNPs associated with a milder *skn-*
563 *1(RNAi)* phenotype (higher % with endoderm) showed a stronger *mom-2(RNAi)* phenotype
564 (low %) and *vice versa* (Fig. 7B), with an overall highly significant negative correlation
565 (Pearson's correlation R=-0.6099, p-value=0.0001).

566 The strong negative correlation we observed between the strength of the *skn-1(RNAi)*
567 and *mom-2(RNAi)* phenotypes for the SNPs that are most significantly associated with the
568 variation might be explained in part by the large blocks of linkage disequilibrium observed in
569 *C. elegans* [40]. Thus, in principle, relatively few genomic regions might, by chance, show the
570 reciprocal relationship, in which case all linked high-significance SNPs would similarly show
571 the negative correlation. It was therefore important to assess how widespread and consistent
572 this effect is across the entire genome. We dissected the relationship of the SKN-1 and MOM-
573 2 requirements across all chromosomes by analyzing the phenotypic strength in sliding
574 windows of 50 SNPs each across each chromosome, using all 4,690 SNPs. This analysis
575 revealed a striking overall trend: for all six chromosomes, regions associated with high SKN-1
576 requirement showed a tendency toward lower MOM-2/Wnt requirement and *vice-versa*. This
577 effect was most pronounced on chromosome I, which showed a very strong negative

correlation (R=-0.69). The effect was also clearly evident on chromosomes IV (R=-0.55), III (R=-0.45), and II (R=-0.43). Though weaker for chromosomes V and X (R=-0.23 for both), the correlation was nonetheless negative for these chromosomes as well (Fig. 7C-E; Supplemental Fig. 8A-E). Thus, the inverse relationship between the MOM-2 and SKN-1 requirement appears to be distributed across the entire genome. The sequences underlying the cryptic variation we observed might not be expected to be uniformly distributed throughout the genome and, indeed, we found that strength of the correlation varied widely between and even within chromosomes (Fig. 7C-E, Supplemental Fig. 8B-E).

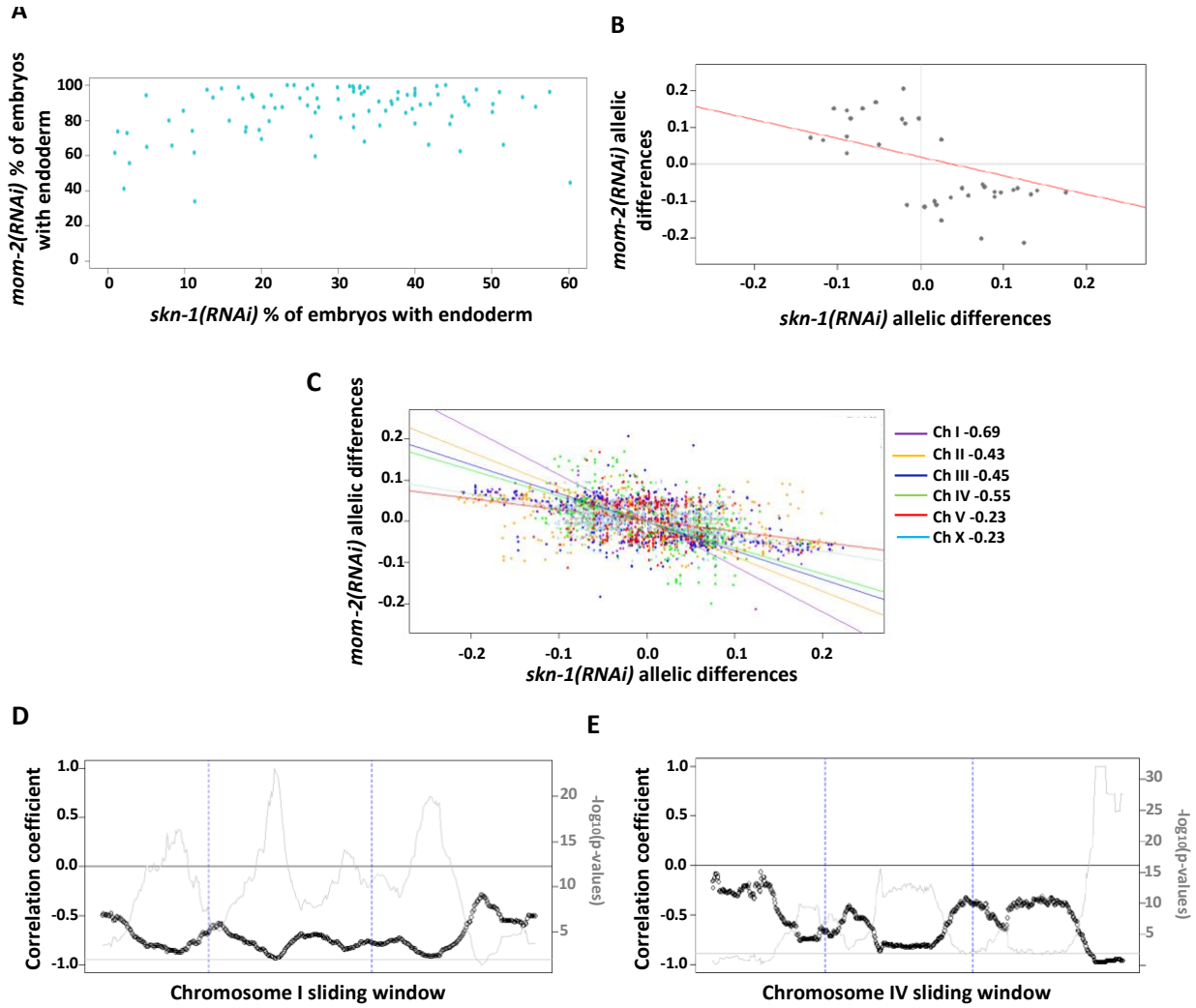


Fig. 7: Negative correlation of *skn-1(RNAi)* and *mom-2(RNAi)* allelic differences.

(A) Comparison of *skn-1(RNAi)* and *mom-2(RNAi)* phenotype in 94 strains 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) Negative correlation of *skn-1(RNAi)* and *mom-2(RNAi)* allelic differences at the top SNPs from the *mom-2(RNAi)* GWAS as calculated by subtracting the median (of the *skn-1(RNAi)* and *mom-2(RNAi)* phenotypes) of one allele to the median of the second. Pearson's correlation R=-0.6099, p-value=0.0001. (C) Genome-wide negative correlation in allelic effects. Each dot represents a SNP, and 4,690 SNPs in total were analyzed. The chromosomes and their corresponding regression lines are color-coded. The R-value of each chromosome is indicated. Correlation sliding window of (D) chromosome I and (E) chromosome IV. Windows of 50 SNPs were used to calculate the correlation coefficient and p-value. Black circles represent the correlation coefficient (R value) for each window (scale on the x-axis). Black line indicates the 0 threshold. Grey line represents the $-\log_{10}$ of the p-values for the corresponding correlation windows (scale on the y-axis). Grey horizontal line is the significance threshold set at p-value=0.01. Blue dotted lines divide chromosomal region into left, middle and right arms.

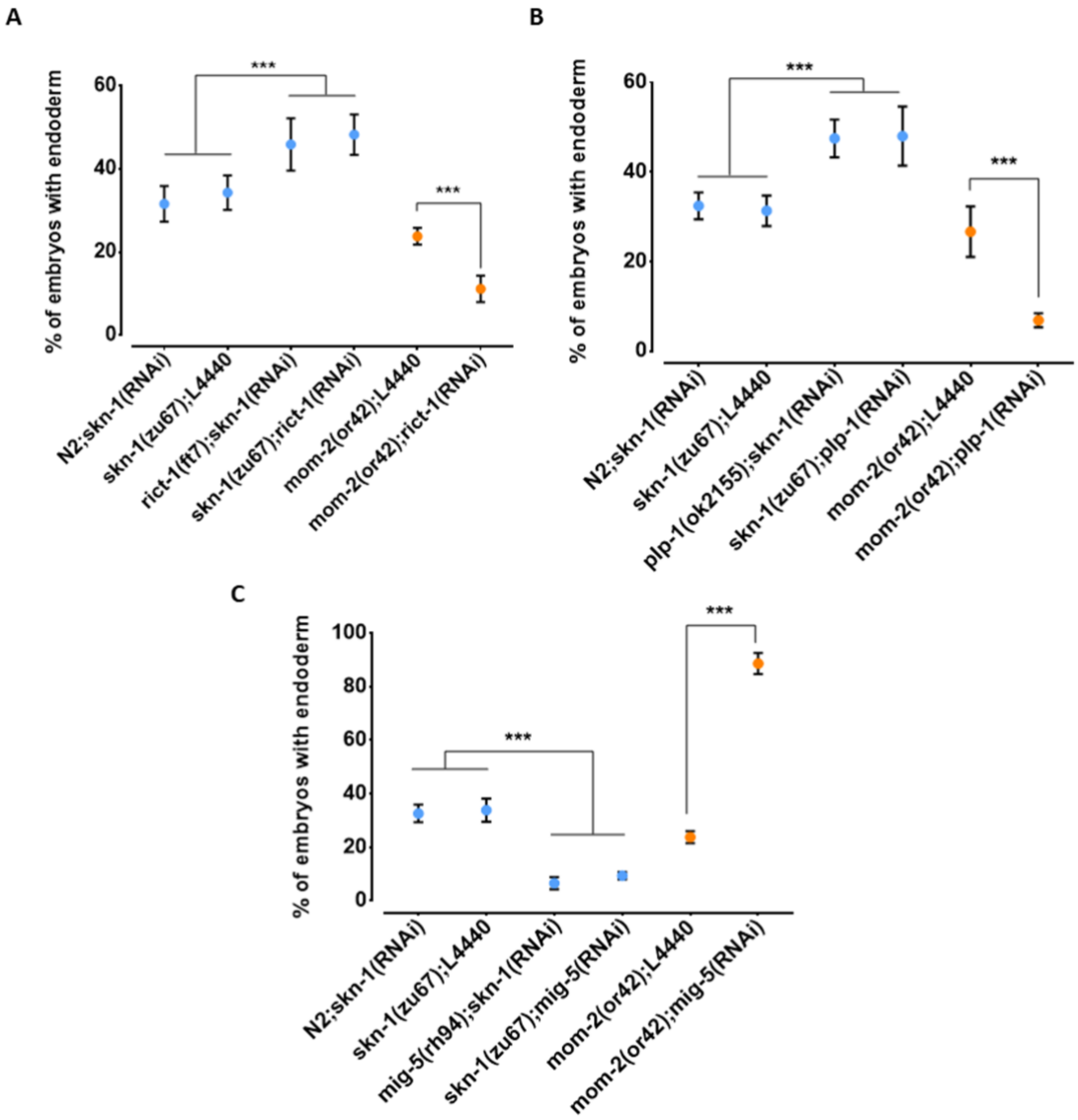
Multiple factors reciprocally regulate the requirement for SKN-1 and MOM-2/Wnt

We further explored this relationship between the requirement for SKN-1 and MOM-2 by testing other candidate genes implicated in endoderm development [68–70]. We found that loss of RICT-1, the *C. elegans* orthologue of the human RICTOR (Rapamycin-insensitive companion of mTOR; [71]), a component of the TORC2 complex, which has been shown to antagonize SKN-1 function [68], results in opposite effects on *skn-1(-)* and *mom-2(-)* mutants (Fig. 8A). Specifically, while *rict-1(RNAi)* suppresses the absence of gut in *skn-1(zu67)* embryos (*skn-1(zu67)*: 34.3% ± s.d 4.1% with gut *vs. skn-1(zu67)*; *rict-1(RNAi)*: 48.3% ± s.d 4.9%; p=<0.001), we found that it *enhances* this phenotype in *mom-2(or42)* mutants (*mom-2(or42)*: 23.8% ± s.d 2.0%; vs. *mom-2(or42)*; *rict-1(RNAi)*: 11.2% ± s.d 3.2%; p<0.001). Confirming this effect, a similar outcome was observed when SKN-1 was depleted by RNAi in *rict-1(ft7)* chromosomal mutants (*skn-1(RNAi)*: 31.6% ± s.d 4.3% with gut vs. *rict-1(ft7); skn-1(RNAi)*: 45.9% ± s.d 6.3%; p<0.05) (Fig. 8A). Similarly, RNAi depletion of PLP-1, the *C. elegans* homologue of the Pur alpha transcription factor that has been shown to bind to and regulate the *end-1* promoter [69], reciprocally affects the outcome of removing these two inputs in the same direction: loss of PLP-1 function suppresses the *skn-1(-)* phenotype (to 48.0% ± s.d 6.6%), and strongly enhances the *mom-2* phenotype (to 6.9% ± s.d 1.6%). Again, this result

618 was confirmed by RNAi of *skn-1* in a *plp-1(ok2156)* chromosomal mutant (Fig. 8B). Thus, as
619 observed with the effect across the genome with natural variants, we observed a striking
620 reciprocal effect of both of these genes on loss of SKN-1 and MOM-2.

621 We also observed a reciprocal effect on the SKN-1 and Wnt inputs with MIG-
622 5/*dishevelled*, a component of the Wnt pathway that acts downstream of the Wnt receptor
623 [70]; however, in this case the effect was in the opposite direction as seen for RICT-1 and PLP-
624 1. Loss of MIG-5 as a result of chromosomal mutation or RNAi leads to *enhancement* of the
625 *skn-1(-)* phenotype (*mig-5(rh94); skn-1(RNAi)*: 6.6% ± s.d 2.3%; *skn-1(zu67); mig-5(RNAi)*:
626 9.4% ± s.d 1.4%) and *suppression* of the *mom-2(-)* phenotype (88.6% ± s.d 4.0%) (Fig 8C).

627 Together, these findings reveal that, as observed with the natural variant alleles
628 (Fig.5C, Fig. 6C-F; Fig. 7B, C), RICT-1, PLP-1, and MIG-5 show opposite effects on the
629 phenotype of removing SKN-1 and MOM-2, suggesting a prevalence of genetic influences that
630 reciprocally influence the outcome in the absence of these two inputs.



631
632 **Fig 8. Reciprocal effects of RICT-1, PLP-1, and MIG-5 on *skn-1(-)* and *mom-2(-)* phenotypes**

633 (A, B) Loss of RICT-1 or PLP-1 enhances the *mom-2(or42)* loss-of-endoderm phenotype and suppresses
634 *skn-1(zu67)* and *skn-1(RNAi)* phenotype. (C) Loss of MIG-5 enhances the *skn-1(zu67)* and *skn-1(RNAi)*
635 phenotype and suppresses *mom-2(or42)* phenotype. At least three replicates were performed per
636 experiment. Student t-test (*** p-value<0.001). Data represented with Standard Deviations.

637

DISCUSSION

The remarkable variety of forms associated with the ~36 animal phyla [72] that emerged from a common metazoan ancestor >600 Mya is the product of numerous incremental changes in GRNs underlying the formation of the body plan and cell types. Here, we describe an unexpectedly broad divergence in the deployment of SKN-1/Nrf and MOM-2/Wnt signaling in generating the most ancient germ layer, the endoderm, within wild isolates of a single animal species, *C. elegans*. In this study, we report five major findings: 1) while the quantitative requirement for two distinct regulatory inputs that initiate expression of the endoderm GRN (SKN-1 and MOM-2) are highly reproducible in individual *C. elegans* isolates, there is wide cryptic variation between isolates. 2) Cryptic variation in the requirement for these regulatory factors shows substantial differences even between closely related strains, suggesting that these traits are subject to rapid evolutionary change in this species. 3) Quantitative genetic analyses of natural and recombinant populations revealed multiple loci underlying the variation in the requirement for SKN-1 and MOM-2 in endoderm specification. 4) The quantitative requirements for SKN-1 and MOM-2 in endoderm specification are negatively correlated across the genome, as shown by allelic effect analysis, implying a reciprocal requirement for the two inputs. 5)*, rict-1*, *plp-1*, *and mig-5* reciprocally influence the outcome of *skn-1(-)* and *mom-2(-)*, substantiating the reciprocal influences on the two GRN inputs. These findings reveal substantial plasticity and complexity underlying SKN-1 and MOM-2/Wnt regulatory inputs in mobilizing a conserved system for endoderm specification.

Together, these findings indicate that, while the core genetic toolkit for the development of the endoderm, the most ancient of the three germ layers, appears to have been preserved for well over half a billion years, the molecular regulatory inputs that initiate

its expression in *C. elegans* vary extremely rapidly over short evolutionary time scales within the species.

Evolutionary plasticity in maternal regulators of embryonic GRNs

The finding that the key regulatory inputs that initiate the endoderm GRN show dramatic plasticity is in accordance with the "hourglass" concept of embryonic development [73–75], in which divergent developmental mechanisms converge on a more constant state (i.e., a "phylotypic stage" at the molecular regulatory level). Indeed, it appears that a downstream GATA factor cascade that directs endoderm specification and differentiation is a highly conserved feature not only across *Caenorhabitis* species [38,76,77] but, in fact, across the broad spectrum of animal phyla [12–17]. These observations are also consistent with the notion that, while the late stages in organ differentiation involve activation of a very large number of target differentiation genes by a limited set of transcription factors, thereby restricting evolutionary divergence at that stage in the regulatory circuitry, the early stages involve the action of transcription factors on far fewer target genes, hence allowing for much greater evolutionary plasticity [21].

In Drosophila, early maternally acting genes show more rapid evolution than those expressed zygotically [78]. Moreover, maternal patterning systems that spatially regulate conserved patterning gene networks between broadly divergent insect species are highly divergent [79,80]. Further comparisons of early embryonic transcripts across many *Drosophila* species and *Aedes aegypti* revealed that maternal transcript pools that, like those of *C. elegans skn-1*, are present only transiently during early embryogenesis, and expression levels are highly variable across these species, spanning ~60 My of evolution [81]. What is particularly striking about our findings is that the varying requirement for key maternal

684 regulatory components is seen within the relatively recent radiation of a single species with
685 low genetic diversity [40]. Variation in gene expression predicts phenotypic severity of
686 mutations in different genetic backgrounds [82]. As quantitative transcriptional profiling of *C.*
687 *elegans* isotypes advances, it will be of interest to assess whether the highly evolvable
688 requirement for maternal regulatory inputs into the endoderm GRN similarly correlates with
689 rapid divergence in quantitative levels of maternal transcripts that are transiently deployed
690 in early embryos of this species.

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

692 GWAS and EMMA revealed several major candidate QTLs (Fig. 5, Supplemental Fig. 3,
693 5), implying that multigenic factors are causally responsible for the differences in requirement
694 for SKN-1 and MOM-2 between isotypes. This multigenic influence was also apparent from
695 analysis of RILs derived from N2 and MY16 parental strains, which identified several loci
696 associated with both traits. In addition, we found substantial epistasis between the different
697 genomic regions underlying this variation. Transgressive segregation of the requirement for
698 both SKN-1 and MOM-2 was seen in the RIL sets (Fig. 6A). For example, the MY16 strain which
699 shows an almost fully penetrant requirement for SKN-1 for gut development, appear to
700 harbor cryptic variants that suppress the requirement for SKN-1, allowing enhanced gut
701 development when combined with genetic factors in the N2 strain.

702 We observed substantial overlap on chromosome IV in the GWAS/EMMA analyses of
703 the *skn-1* and *mom-2* requirements in wild isotypes (Fig. 5, Supplemental Fig. 6) and on
704 chromosome II from analyses using RILs (Fig. 6B). This finding raises the possibility that some
705 QTLs may influence requirement for both inputs into the endoderm specification pathway: as
706 SKN-1 and Wnt converge to regulate expression of the *end-1/3* genes, it is conceivable that

common genetic variants might modulate the relative strength or outcome of both maternal inputs. However, our findings do not resolve whether these genetic variants act independently to influence the maternal regulatory inputs.

Genetic interactions are often neglected in large-scale genetic association studies [83] owing in part to the difficulty in confirming them [84]. Many studies [85–88], including ours here, showed that epistasis can strongly influence the behavior of certain variants upon genetic perturbation. In addition, selection on pleiotropically acting loci facilitates rapid developmental system drift [87,89,90]. Together, epistasis and selection on pleiotropic loci play important roles in the evolution of natural populations [89–92].

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

Although we did not observe a direct correlation between the *skn-1(-)* and *mom-2(-)* phenotypes across the isotypes studied here, a clear inverse correlation was observed when testing associated individual SNPs in significantly linked loci (Fig. 7). This reciprocal effect seen across large portions of the genome may be attributable in part to the large LD blocks present on all chromosomes (Fig. 7, Supplemental Fig. 8) [40]. However, our finding that this effect is seen across the entire genome raises the possibility that the SKN-1 and MOM-2/Wnt inputs might compensate for each other and that genetic variants that enhance the requirement for one of the inputs relaxes the requirement for the other. This reciprocality might reflect cross-regulatory interactions between these two maternal inputs or may be the result of evolutionary constraints imposed by selection on these genes, which act pleiotropically in a variety of processes.

We identified two genes, *rict-1* and *plp-1*, that show similar inverse effects on the requirements from *skn-1* and *mom*-2: debilitation of either gene enhances the phenotype of

730 *mom-2(-)* and suppresses that of *skn-1(-)*. RICT-1 function extends lifespan in *C. elegans*
731 through the action of SKN-1 [68], and loss of RICT-1 rescues the misspecification of the MS
732 and E blastomeres and lethality of *skn-1(-)* embryos [68], consistent with our finding.
733 However, the mechanism by which loss of *rict-1* synergizes with a defect in the Wnt pathway
734 is not clear. We previously showed that PLP-1, a homologue of the vertebrate transcription
735 factor pur alpha, binds to the *end-1* promoter and acts in parallel to the Wnt pathway and
736 downstream of the MAPK signal [69], thereby promoting gut formation. PLP-1 shows a similar
737 reciprocal relationship with SKN-1 and MOM-2 as with RICT-1 (Fig. 8). Given that PLP-1 binds
738 at a *cis* regulatory site in *end-1* near a putative POP-1 binding site [69], and that SKN-1 also
739 binds to the *end-1* regulatory region [64], it is conceivable that this reciprocality reflects
740 integration of information at the level of transcription factor binding sites. As the architecture
741 of the GRN is shaped by changes in *cis*-regulatory sequences [1,3], analyzing alterations in
742 SKN-1 and Wnt/POP-1 targets among *C. elegans* wild isolates may provide insights into how
743 genetic changes are accommodated without compromising the developmental output at
744 microevolutionary time scale.

745 MIG-5, a *dishevelled* orthologue, functions in the Wnt pathway in parallel to Src
746 signaling to regulate asymmetric cell division and endoderm induction [28,70]. We found that
747 the loss of *mig-5* function enhances the gut defect of *skn-1(-)* and suppresses that of the *mom-*
748 *2(-), the* opposite reciprocal relationship to that of *rict-1* and *plp-1,* and consistent with a
749 previous report (Fig. 8) [28]. These effects were not observed in embryos lacking function of
750 *dsh-2*, the orthologue of *mig-5* (data not shown), supporting a previous study that showed
751 overlapping but non-redundant roles of MIG-5 and DSH-2 in EMS spindle orientation and gut
752 specification [70]. Recent studies showed that Dishevelled can play both positive and negative
753 roles during axon guidance [93,94]. Dishevelled, upon Wnt-activation, promotes

754 hyperphosphorylation and inactivation of Frizzled receptor to fine-tune Wnt activity. It is
755 tempting to speculate that MIG-5 may perform similar function in EMS by downregulating
756 activating signals (Src or MAPK), in the absence of MOM-2.

757 We hypothesize that compensatory mechanisms may evolve to fine-tune the level of
758 gut-activating regulatory inputs. Successful developmental events depend on tight spatial and
759 temporal regulation of gene expression. For example, anterior-posterior patterning in the
760 *Drosophila* embryo is determined by the local concentrations of the Bicoid, Hunchback, and
761 Caudal transcription factors [95]. We postulate that SKN-1 and Wnt signaling is modulated so
762 that the downstream genes, *end-1/3*, which control specification and later differentiation of
763 endoderm progenitors, are expressed at optimal levels that ensure normal gut development.
764 Suboptimal END activity leads to poorly differentiated gut and both hypo- and hyperplasia in
765 the gut lineage [96–98]. Hyper- or hypo-activation of Wnt signaling has been implicated in
766 cancer development [99], bone diseases [100,101], and metabolic diseases [102,103],
767 demonstrating the importance of regulating the timing and dynamics of such developmental
768 signals within a quantitatively restricted window.

769 **Cryptic variation and evolvability of GRNs**

770 This study revealed substantial cryptic genetic modifications that alter the relative
771 importance of two partially redundant inputs into the *C. elegans* endoderm GRN, leading to
772 rapid change in the developmental network architecture (Fig. 9). Such modifications may
773 occur through transitional states that are apparent even within the radiation of this single
774 species. For example, the finding that POP-1 is not required for gut development even in a
775 wild isolate (e.g., MY16) that, like *C. briggsae*, shows a near-absolute requirement for SKN-1
776 may reflect a transitional state between the two species: *i.e.*, a nearly essential requirement

for SKN-1 but non-essential requirement for POP-1, an effect not previously seen in either species. In addition, duplicated GATA factors (the MEDs, ENDs, and ELTs) and partially redundant activating inputs (SKN-1, Wnt, Src, and MAPK) in endoderm GRN, provide an opportunity for genetic variation to accumulate and "experimentation" of new regulatory relationships without diminishing fitness [2,104,105].

Redundancy in the system may act to 'rescue' an initial mutation and allow for secondary mutations that might eventually lead to rewiring of the network. For example, loss of either MyoD or Myf5, two key regulators of muscle differentiation in metazoans, produces minimal defects in myogenesis as a result of compensatory relationship between the myogenic factors [106]. In vertebrates, gene duplication events have resulted in an expansion of Hox genes to a total of >200, resulting in prevalent redundancy [107–109]. This proliferation of redundant genes provides opportunities for evolutionary experimentation and subsequent specialization of new functions [109]. In *C. elegans*, loss of GAP-1 (a Ras inhibitor) or SLI-1 (a negative regulator of EGFR signaling) alone does not produce obvious defects, while double mutations lead to a multivulva phenotype [110]. Many other similar redundant relationships between redundant partners exist in the animal. Notably, the relative importance of Ras, Notch, and Wnt signals in vulva induction differ in various genetic backgrounds [6,65] and physiological conditions [111,112], resulting in flexibility in the system. While vulval development in *C. elegans,* when grown under standard laboratory conditions, predominantly favors utilization of the EGF/Ras signaling pathway [111], Wnt is the predominant signaling pathway in the related *Pristionchus pacificus*, which is ~250 MY divergent [113]. In addition, while *Cel-lin-17* functions positively to transduce the Wnt signal, *Ppa-lin-17/Fz* antagonizes Wnt signaling and instead the Wnt signal is transmitted by *Ppa-lin-18/Ryk*, which has acquired a novel SH3 domain not present in the *C. elegans* ortholog [114].

Thus, extensive rewiring of signaling networks and modularity of signaling motifs contribute to developmental systems drift.

The broad cryptic variation we have observed in this study may drive developmental system drift, giving rise to GRN architectures that differ in the relative strength of the network components. In the developmental hourglass model of evolvability of animal development, the early stages of embryonic development showed the least constraint in gene expression compared to either the phylotypic stage or post phylotypic stage. This is likely attributable either to positive selection during early embryonic and later larval stages or to developmental constraints. Analysis of developmental gene expression in mutation accumulation lines, which have evolved in the absence of any positive selection, showed similarity to the developmental hourglass model of evolvability, consistent with strong developmental constraints on the phylotypic stage [115]. However, they do not rule out the possibility that early and late stages of development might be more adaptive and therefore subject to positive selection. It will be of interest to learn the degree to which the divergence in network architecture might arise as a result of differences in the environment and selective pressures on different *C. elegans* isotypes.

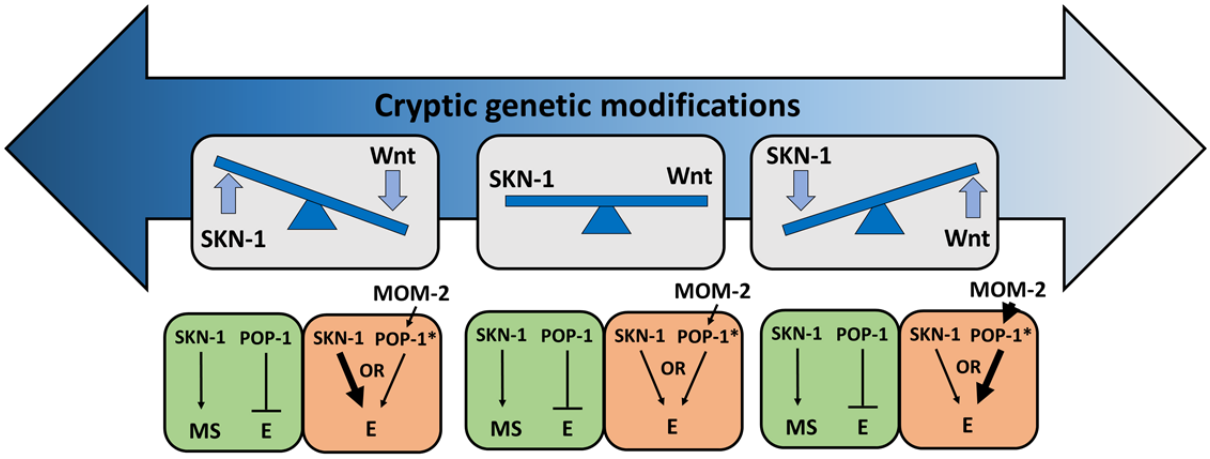


Fig. 9. Simplified models accounting for cryptic compensatory relationship between the SKN-1 and MOM-2/Wnt regulatory inputs in the endoderm GRN.

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.

825 **ACKNOWLEDGMENTS**

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

835 **COMPETING INTERESTS**

836 The authors declare no competing or financial interests.

837

838 **REFERENCES**

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

841 2. Félix M-A, Wagner A. Robustness and evolution: concepts, insights and challenges
842 from a developmental model system. Heredity (Edinb). Nature Publishing Group;
843 2008;100: 132–140. doi:10.1038/sj.hdy.6800915

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

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

850 5. Peter IS, Davidson EH. Assessing regulatory information in developmental gene
851 regulatory networks. Proc Natl Acad Sci U S A. National Academy of Sciences;
852 2017;114: 5862–5869. doi:10.1073/pnas.1610616114

853 6. Milloz J, Duveau F, Nuez I, Felix M-A. Intraspecific evolution of the intercellular
854 signaling network underlying a robust developmental system. Genes Dev. 2008;22:
855 3064–3075. doi:10.1101/gad.495308

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

859 8. Phinchongsakuldit J, MacArthur S, Brookfield JFY. Evolution of Developmental Genes:

Molecular Microevolution of Enhancer Sequences at the Ubx Locus in Drosophila and Its Impact on Developmental Phenotypes. Mol Biol Evol. 2003;21: 348–363. doi:10.1093/molbev/msh025

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

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

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

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

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

14. Boyle MJ, Seaver EC. Expression of FoxA and GATA transcription factors correlates with regionalized gut development in two lophotrochozoan marine worms: Chaetopterus (Annelida) and Themiste lageniformis (Sipuncula). Evodevo. 2010;1: 2. doi:10.1186/2041-9139-1-2

882 15. Gillis WJ, Bowerman B, Schneider SQ. Ectoderm- and endomesoderm-specific GATA
883 transcription factors in the marine annelid Platynereis dumerilli. Evol Dev. 2007;9:
884 39–50. doi:10.1111/j.1525-142X.2006.00136.x

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

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

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

896 19. Bowerman B, Eaton BA, Priess JR. skn-1, a maternally expressed gene required to
897 specify the fate of ventral blastomeres in the early C. elegans embryo. Cell. 1992;68:
898 1061–75. Available: http://www.ncbi.nlm.nih.gov/pubmed/1547503

899 20. Bowerman B, Draper BW, Mello CC, Priess JR. The maternal gene skn-1 encodes a
900 protein that is distributed unequally in early C. elegans embryos. Cell. Elsevier;
901 1993;74: 443–52. doi:10.1016/0092-8674(93)80046-H

902 21. Maduro MF, Rothman JH. Making Worm Guts: The Gene Regulatory Network of the
903 Caenorhabditis elegans Endoderm. Dev Biol. Academic Press; 2002;246: 68–85.

904 doi:10.1006/DBIO.2002.0655

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

908 23. Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH.
909 Restriction of mesendoderm to a single blastomere by the combined action of SKN-1
910 and a GSK-3beta homolog is mediated by MED-1 and -2 in C. elegans. Mol Cell.
911 2001;7: 475–85. Available: http://www.ncbi.nlm.nih.gov/pubmed/11463373

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

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

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

921 27. Shin TH, Yasuda J, Rocheleau CE, Lin R, Soto M, Bei Y, et al. MOM-4, a MAP kinase
922 kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce
923 anterior/posterior polarity signals in C. elegans. Mol Cell. 1999;4: 275–80. Available:
924 http://www.ncbi.nlm.nih.gov/pubmed/10488343

925 28. Bei Y, Hogan J, Berkowitz LA, Soto M, Rocheleau CE, Pang KM, et al. SRC-1 and Wnt

926 signaling act together to specify endoderm and to control cleavage orientation in
927 early C. elegans embryos. Dev Cell. 2002;3: 113–25. Available:
928 http://www.ncbi.nlm.nih.gov/pubmed/12110172

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

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

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

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

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

945 34. Maduro MF, Lin R, Rothman JH. Dynamics of a Developmental Switch: Recursive
946 Intracellular and Intranuclear Redistribution of Caenorhabditis elegans POP-1
947 Parallels Wnt-Inhibited Transcriptional Repression. Dev Biol. Academic Press;

2002;248: 128–142. doi:10.1006/DBIO.2002.0721

35. 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

36. 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

37. 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

38. 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

39. 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

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

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

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

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

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

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

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

984 47. Clokey G V, Jacobson LA. The autofluorescent “lipofuscin granules” in the intestinal
985 cells of Caenorhabditis elegans are secondary lysosomes. Mech Ageing Dev. 1986;35:
986 79–94. Available: http://www.ncbi.nlm.nih.gov/pubmed/3736133

987 48. Hermann GJ, Schroeder LK, Hieb CA, Kershner AM, Rabbitts BM, Fonarev P, et al.
988 Genetic Analysis of Lysosomal Trafficking in Caenorhabditis elegans. Mol Biol Cell.
989 2005;16: 3273–3288. doi:10.1091/mbc.E05

990 49. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient
991 Control of Population Structure in Model Organism Association Mapping. Genetics.

2008;178: 1709–1723. doi:10.1534/genetics.107.080101

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

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

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

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

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

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

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

1014 57. Zhao Z, Boyle TJ, Bao Z, Murray JI, Mericle B, Waterston RH. Comparative analysis of
1015 embryonic cell lineage between Caenorhabditis briggsae and Caenorhabditis elegans.
1016 Dev Biol. NIH Public Access; 2008;314: 93–9. doi:10.1016/j.ydbio.2007.11.015

1017 58. Cutter AD. Divergence Times in Caenorhabditis and Drosophila Inferred from Direct
1018 Estimates of the Neutral Mutation Rate. Mol Biol Evol. Oxford University Press;
1019 2008;25: 778–786. doi:10.1093/molbev/msn024

1020 59. Sterken MG, Snoek LB, Kammenga JE, Andersen EC. The laboratory domestication of
1021 Caenorhabditis elegans. Trends Genet. NIH Public Access; 2015;31: 224–31.
1022 doi:10.1016/j.tig.2015.02.009

1023 60. Paaby AB, White AG, Riccardi DD, Gunsalus KC, Piano F, Rockman M V. Wild worm
1024 embryogenesis harbors ubiquitous polygenic modifier variation. Elife. eLife Sciences
1025 Publications Limited; 2015;4: e09178. doi:10.7554/eLife.09178

1026 61. Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK, et al. Minimizing
1027 the risk of reporting false positives in large-scale RNAi screens. Nat Methods. 2006;3:
1028 777–779. doi:10.1038/nmeth1006-777

1029 62. Zhuang JJ, Hunter CP. RNA interference in Caenorhabditis elegans: Uptake,
1030 mechanism, and regulation. Parasitology. 2012;139: 560–573.
1031 doi:10.1017/S0031182011001788

1032 63. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Variability in gene expression
1033 underlies incomplete penetrance. Nature. Nature Publishing Group; 2010;463: 913–
1034 918. doi:10.1038/nature08781

1035 64. Zhu J, Hill RJ, Heid PJ, Fukuyama M, Sugimoto A, Priess JR, et al. end-1 encodes an

apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. Genes Dev. Cold Spring Harbor Laboratory Press; 1997;11: 2883–96. Available: http://www.ncbi.nlm.nih.gov/pubmed/9353257

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

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

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

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

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

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

1058 71. Tatebe H, Shiozaki K. Evolutionary Conservation of the Components in the TOR
1059 Signaling Pathways. Biomolecules. 2017;7: 77. doi:10.3390/biom7040077

1060 72. Adoutte A, Philippe H. The major lines of metazoan evolution: summary of traditional
1061 evidence and lessons from ribosomal RNA sequence analysis. EXS. 1993;63: 1–30.
1062 Available: http://www.ncbi.nlm.nih.gov/pubmed/8422536

1063 73. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. Gene
1064 expression divergence recapitulates the developmental hourglass model. Nature.
1065 Nature Publishing Group; 2010;468: 811–814. doi:10.1038/nature09634

1066 74. Raff RA. The shape of life : genes, development, and the evolution of animal form.
1067 University of Chicago Press; 1996.

1068 75. Domazet-Lošo T, Tautz D. A phylogenetically based transcriptome age index mirrors
1069 ontogenetic divergence patterns. Nature. Nature Publishing Group; 2010;468: 815–
1070 818. doi:10.1038/nature09632

1071 76. Coroian C, Broitman-Maduro G, Maduro MF. Med-type GATA factors and the
1072 evolution of mesendoderm specification in nematodes. Dev Biol. 2006;289: 444–455.
1073 doi:10.1016/j.ydbio.2005.10.024

1074 77. Maduro MF, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess JR, et al. Genetic
1075 redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol.
1076 2005;284: 509–522. doi:10.1016/j.ydbio.2005.05.016

1077 78. Mensch J, Serra F, Lavagnino NJ, Dopazo H, Hasson E. Positive Selection in
1078 Nucleoporins Challenges Constraints on Early Expressed Genes in Drosophila
1079 Development. Genome Biol Evol. 2013;5: 2231–2241. doi:10.1093/gbe/evt156

79. Davis GK, Patel NH. Short, Long, and Beyond: Molecular and Embryological Approaches to Insect Segmentation. Annu Rev Entomol. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA ; 2002;47: 669–699. doi:10.1146/annurev.ento.47.091201.145251

80. Lynch JA, El-Sherif E, Brown SJ. Comparisons of the embryonic development of Drosophila, Nasonia, and Tribolium. Wiley Interdiscip Rev Dev Biol. John Wiley & Sons, Ltd (10.1111); 2012;1: 16–39. doi:10.1002/wdev.3

81. Atallah J, Lott SE. Evolution of maternal and zygotic mRNA complements in the early Drosophila embryo. Dyer KA, editor. PLOS Genet. 2018;14: e1007838. doi:10.1371/journal.pgen.1007838

82. Vu V, Verster AJ, Schertzberg M, Chuluunbaatar T, Spensley M, Pajkic D, et al. Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. Cell. 2015;162: 391–402. doi:10.1016/j.cell.2015.06.037

83. Moore JH, Williams SM. Epistasis and Its Implications for Personal Genetics. Am J Hum Genet. Cell Press; 2009;85: 309–320. doi:10.1016/J.AJHG.2009.08.006

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

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

86. Volis S, Shulgina I, Zaretsky M, Koren O. Epistasis in natural populations of a

predominantly selfing plant. Heredity (Edinb). Nature Publishing Group; 2011;106: 300–9. doi:10.1038/hdy.2010.79

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

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

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

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

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

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

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

doi:10.1016/J.DEVCEL.2011.01.002

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

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

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

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

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

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

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

101. Baron R, Gori F. Targeting WNT signaling in the treatment of osteoporosis. Curr Opin

Pharmacol. Elsevier; 2018;40: 134–141. doi:10.1016/J.COPH.2018.04.011

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

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

104. Gibson G, Dworkin I. Uncovering cryptic genetic variation. Nat Rev Genet. Nature Publishing Group; 2004;5: 681–690. doi:10.1038/nrg1426

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

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

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

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

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

1167 110. Yoon CH, Chang C, Hopper NA, Lesa GM, Sternberg PW. Requirements of multiple
1168 domains of SLI-1, a Caenorhabditis elegans homologue of c-Cbl, and an inhibitory
1169 tyrosine in LET-23 in regulating vulval differentiation. Mol Biol Cell. American Society
1170 for Cell Biology; 2000;11: 4019–31. Available:
1171 http://www.ncbi.nlm.nih.gov/pubmed/11071924

1172 111. Braendle C, Félix M-A. Plasticity and Errors of a Robust Developmental System in
1173 Different Environments. Dev Cell. 2008;15: 714–724.
1174 doi:10.1016/j.devcel.2008.09.011

1175 112. Grimbert S, Vargas Velazquez AM, Braendle C. Physiological Starvation Promotes
1176 Caenorhabditis elegans Vulval Induction. G3 (Bethesda). Genetics Society of America;
1177 2018;8: 3069–3081. doi:10.1534/g3.118.200449

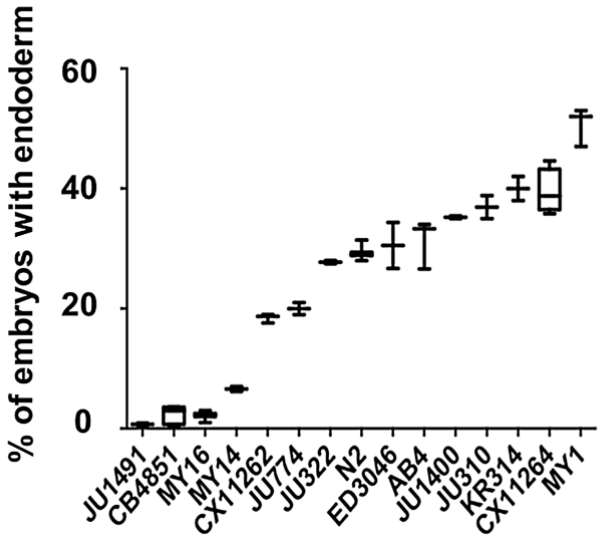
1178 113. Zheng M, Messerschmidt D, Jungblut B, Sommer RJ. Conservation and diversification
1179 of Wnt signaling function during the evolution of nematode vulva development. Nat
1180 Genet. 2005;37: 300–304. doi:10.1038/ng1512

1181 114. Wang X, Sommer RJ. Antagonism of LIN-17/Frizzled and LIN-18/Ryk in Nematode
1182 Vulva Induction Reveals Evolutionary Alterations in Core Developmental Pathways.
1183 Sternberg PW, editor. PLoS Biol. 2011;9: e1001110.
1184 doi:10.1371/journal.pbio.1001110

1185 115. Zalts H, Yanai I. Developmental constraints shape the evolution of the nematode mid-
1186 developmental transition. Nat Ecol Evol. Nature Publishing Group; 2017;1: 0113.
1187 doi:10.1038/s41559-017-0113

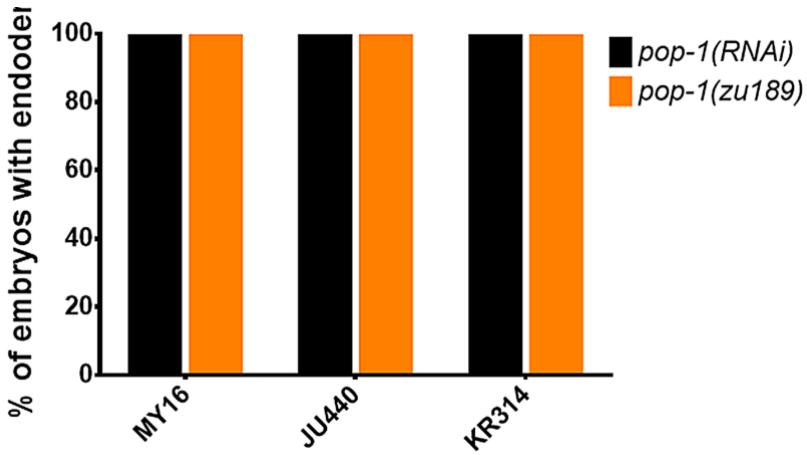
1188

Supplementary figures



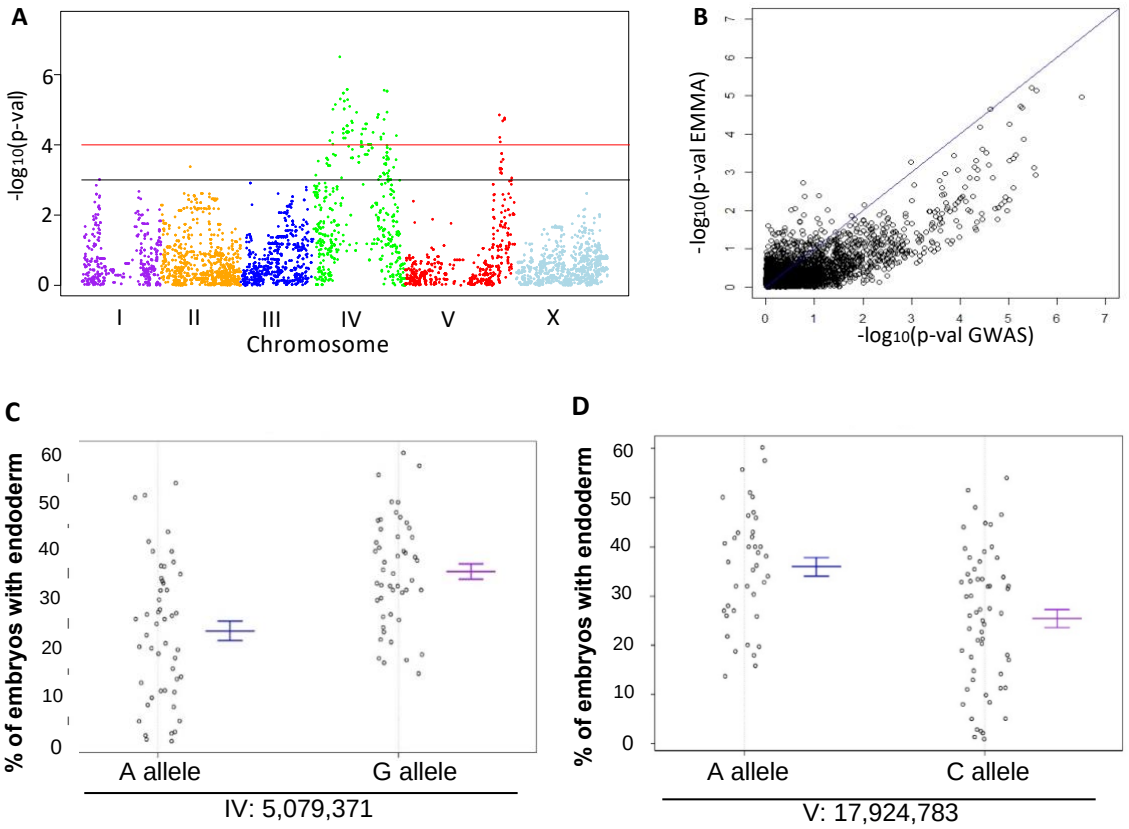
Supplemental Fig. 1: High reproducibility of* skn-1(RNAi) *phenotypes in various* C. elegans *isotypes.

A minimum of two replicates were obtained, with >500 embryos per replicate. Box-plot represents median with range bars showing upper and lower quartiles.



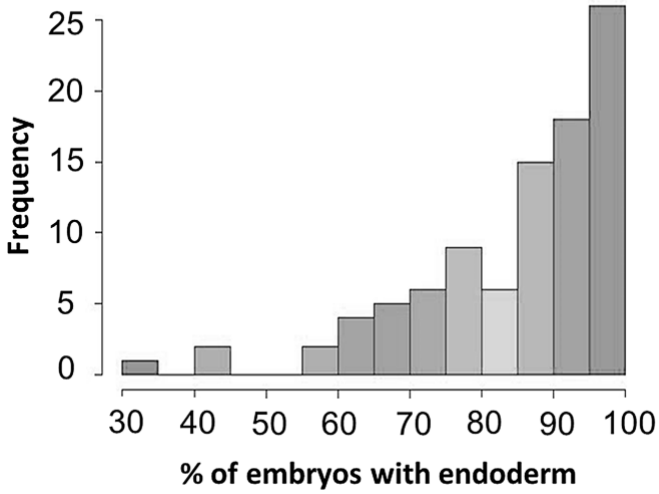
Supplemental Fig. 2: The requirement for POP-1 in endoderm formation does not vary in three introgressed strains.

Strains are shown on the x-axis and fraction of arrested embryos with endoderm are shown on the y-axis. Four introgressed lines were studied for each mutant strain. >200 embryos were scored per experiment.



Supplemental Fig. 3: linear model GWAS of skn-1 _embryonic phenotype highly correlates with the mixed-model analysis._

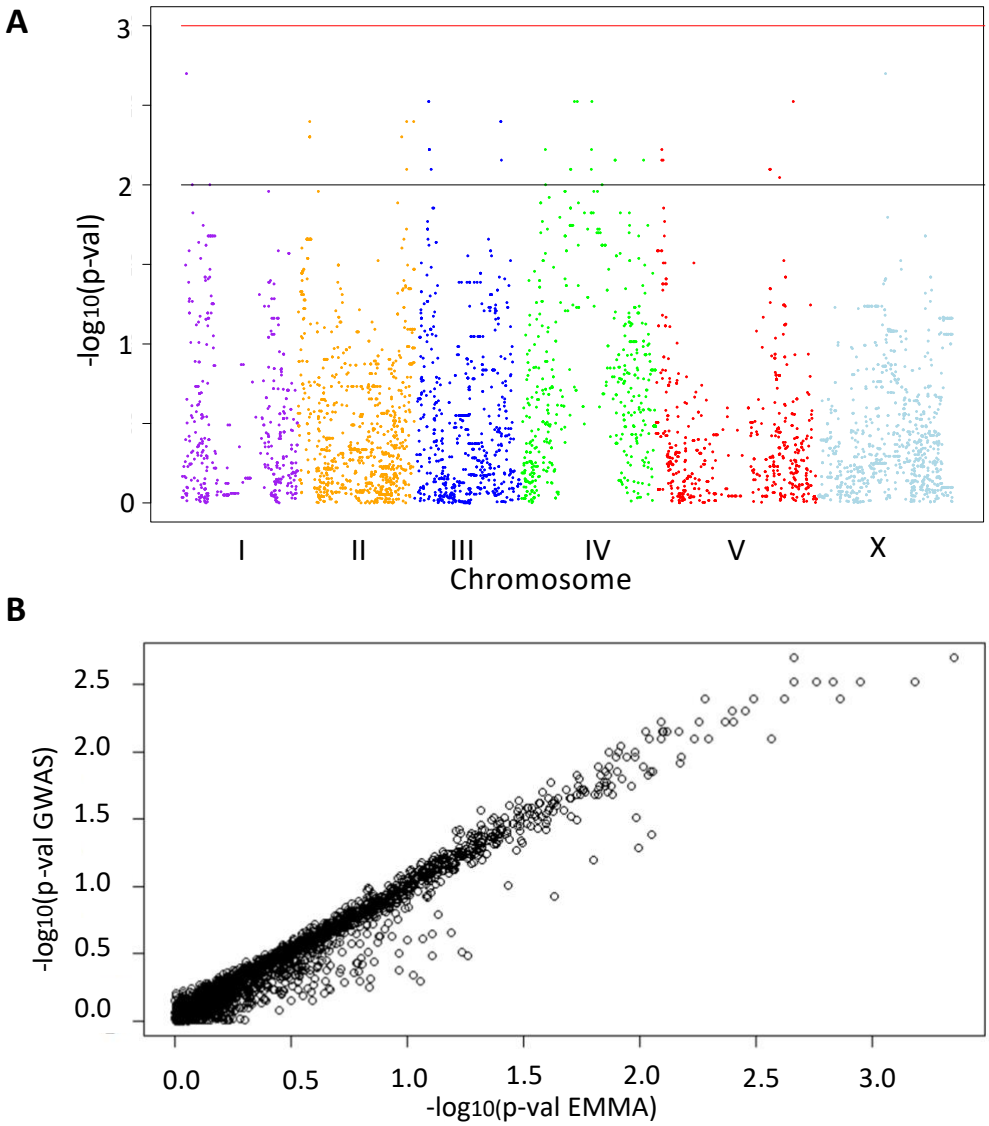
(A) Manhattan plot of skn-1(RNAi) GWAS. Red line represents 1.5% FDR (obtained from 10,000 permutated results) and black line represents 3.0% FDR. The chromosomes are color-coded. The y axis is the $-\log_{10}$ of p-value. (B) Correlation between skn-1(RNAi) GWAS and EMMA (see Fig. 5A). y-axis is the $-\log_{10}$ of the p-values from EMMA, x-axis is $-\log_{10}$ of the p-values from GWAS. A modest linear relationship is found (Spearman correlation coefficient = 0.43) as shown by the blue line. (C) Effect plot of the top SNP revealed by skn-1(RNAi) EMMA (see Fig. 5A). (D) Effect plot of the top SNP on Chromosome V revealed by skn-1(RNAi) GWAS. The variant position and genotype are shown on the x-axis, while the phenotype of strains carrying the alleles after skn-1(RNAi) treatment is shown on the y-axis. Confidence intervals for the average phenotype in each genotype group are shown.



1213

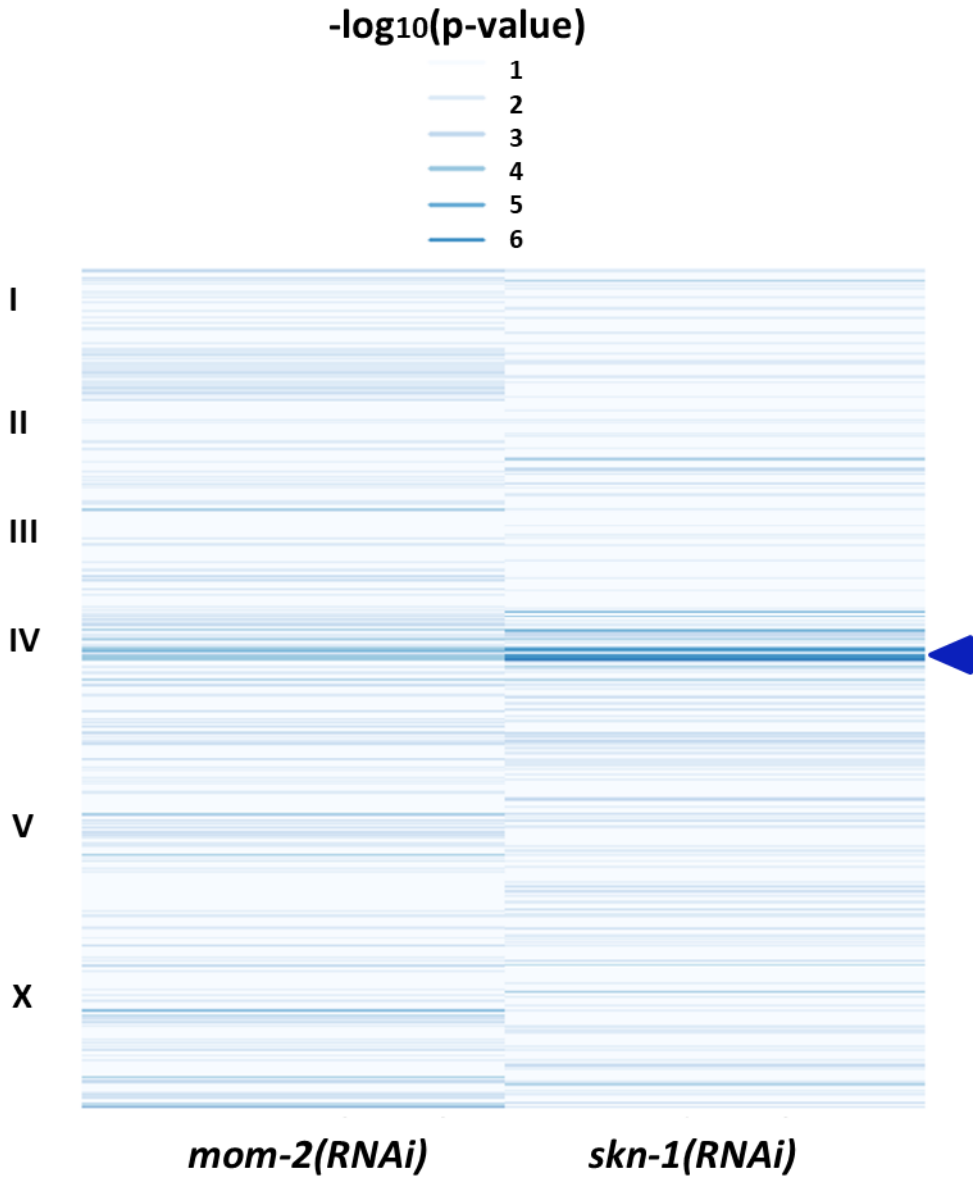
1214 ***Supplemental Fig. 4: Histogram of mom-2(RNAi) phenotype among the 94 wild isolates.***

1215 *A beta-distribution is observed (skewed to the right). Shapiro-Wilk normality test (W=0.8682, p-*
1216 *value=1.207X10^{-7}).*



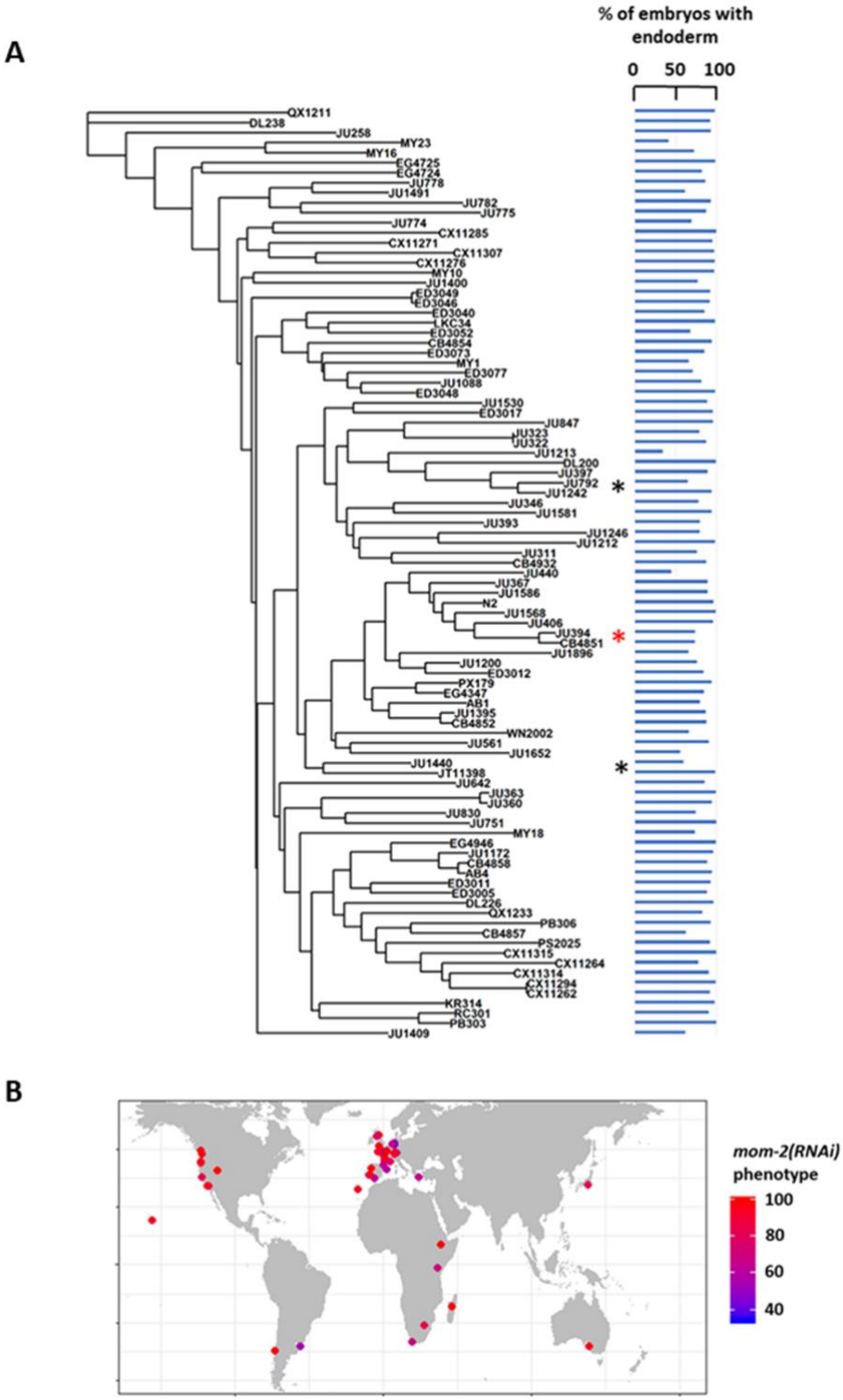
Supplemental Fig. 5: GWAS of *mom-2* embryonic phenotype is highly correlated with the mixed-model analysis.

(A) Manhattan plot of *mom-2(RNAi)* GWAS (permutation-adjusted p-values). Black line represents p-value<0.01, while red line represents p-value<0.001. Genomic regions are shown on the x-axis. The chromosomes are color-coded. The y axis is the –log10 of p-value in the linear model. (B) Correlation between *mom-2(RNAi*) GWAS and EMMA. The y-axis represents the $-\log_{10}$ of the p-values from the GWAS approach, while the x-axis represents $-\log_{10}$ of the p-values from the EMMA approach. A strong linear relationship is found (Pearson's correlation R = 0.95, p-value < 2.2e-16).



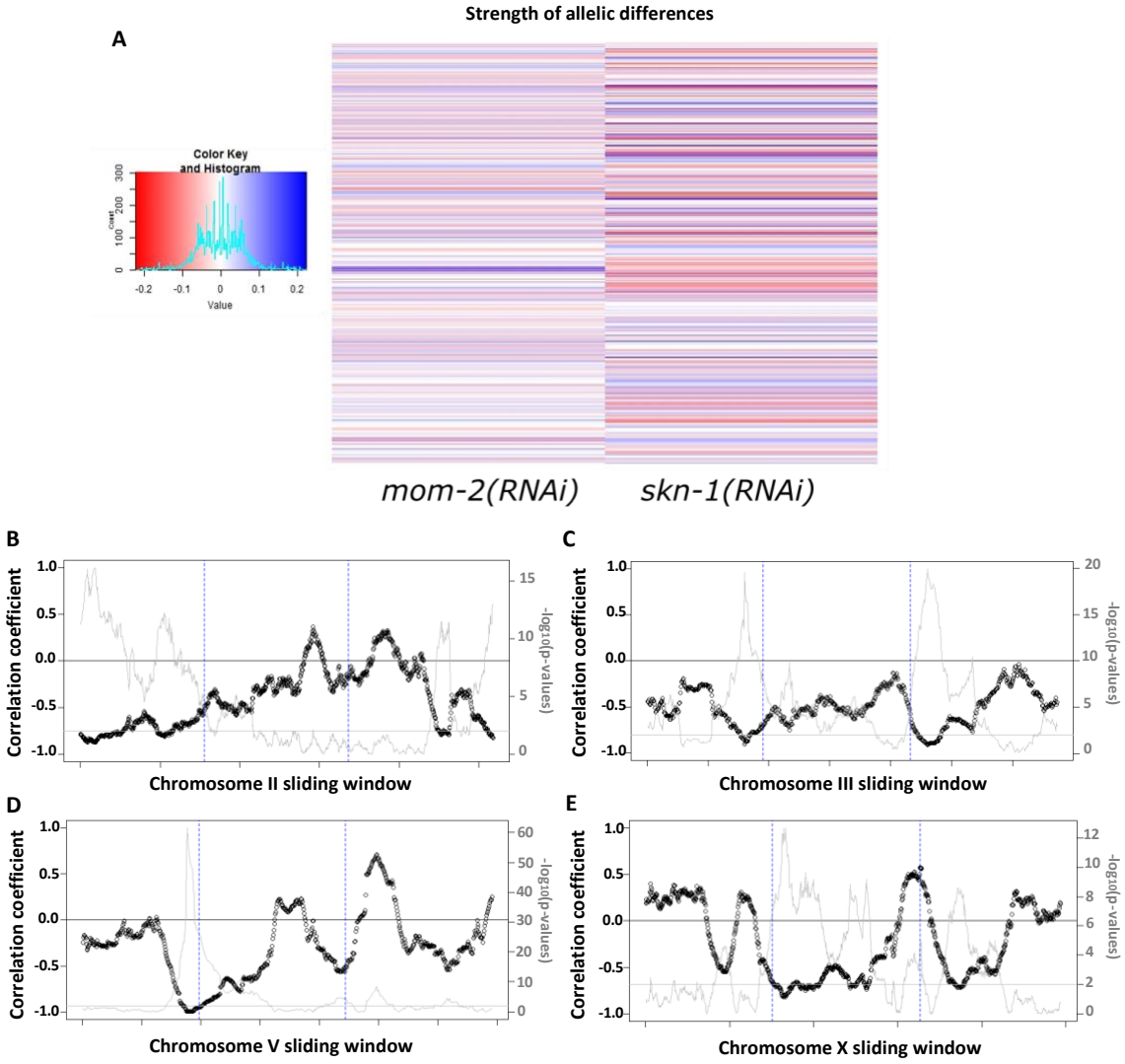
Supplemental Fig. 6: Comparison of EMMA p-values for both* mom-2 *and* skn-1 *RNAi phenotypes.

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



Supplemental Fig. 7: 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. (B) Worldwide distribution of* mom-2(RNAi) *phenotype across 94 isolates. Each circle represents a single isolate.*



1241

1242 ***Supplemental Fig. 8: Negative correlation of skn-1(RNAi) and mom-2(RNAi) allelic differences.***

1243 *(A) Heat map of allelic differences per SNP for skn-1(RNAi) and mom-2(RNAi), as calculated by the*
1244 *phenotypic median differences per allele at each SNP. Each line represents a color-coded result of a single*
1245 *locus, covering the entire genome. Correlation sliding window of (B) chromosome II, (C) chromosome III, (D)*
1246 *chromosome V and (E) chromosome X. Windows of 50 SNPs were used to calculate the correlation*
1247 *coefficient and p-value. Black circles represent the correlation coefficient (R value) for each window (scale*
1248 *on the x-axis). Black line indicates the 0 threshold. Grey line represents the $-\log_{10}$ of the p-values for the*
1249 *corresponding correlation windows (scale on the y-axis). Grey horizontal line is the significance threshold*
1250 *set at p-value=0.01. Dotted blue lines divide chromosomal region into left, middle and right arms.*