G3: Genes|Genomes|Genetics Early Online, published on August 15, 2014 as doi:10.1534/g3.114.013151

# Widespread genomic incompatibilities in *Caenorhabditis elegans*

L. Basten Snoek[1,*], Helen E. Orbidans[2,*], Jana J. Stastna[2], Aafke Aartse[1], Miriam Rodriguez[1], Joost A.G. Riksen[1], Jan E. Kammenga[1,3] and Simon C. Harvey[2,4]

[1] Laboratory of Nematology, Wageningen University, 6708 PB Wageningen, The Netherlands.

[2] Biomolecular Research Group, School of Human and Life Sciences, Canterbury Christ Church University, North Holmes Road, Canterbury, CT1 1QU, UK.

[3] jan.kammenga@wur.nl

[4] simon.harvey@canterbury.ac.uk

* These authors contributed equally to this work.

Running Title: Incompatibilities in *C. elegans*.

Key words: C. elegans, Bateson-Dobzhansky-Muller, reproductive isolation, negative epistasis, QTL, introgression line.

Word count: 5692

Table count: 2

Figure count: 6

Data is archived in WormQTL (http://www.wormqtl.org).

ABSTRACT In the Bateson-Dobzhansky-Muller (BDM) model of speciation, incompatibilities emerge from the deleterious interactions between alleles that are neutral or advantageous in the original genetic backgrounds, *i.e.* negative epistatic effects. Within species such interactions are responsible for outbreeding depression and F2 (hybrid) breakdown. We sought to identify BDM incompatibilities in the nematode *Caenorhabditis elegans* by looking for genomic regions that disrupt egg laying; a complex, highly regulated and coordinated phenotype. Investigation of introgression lines and recombinant inbred lines derived from the isolates CB4856 and N2 uncovered multiple incompatibility quantitative trait loci (QTL). These QTL produce a synthetic egg-laying defective phenotype not seen in CB4856 and N2 nor in other wild isolates. For two of the QTL regions, results are inconsistent with a model of pairwise interaction between two loci, suggesting that the incompatibilities are a consequence of complex interactions between multiple loci. Analysis of additional life history traits indicates that the QTL regions identified in these screens are associated with effects on other traits such as lifespan and reproduction, suggesting that the incompatibilities are likely to be deleterious. Taken together, these results indicate that numerous BDM incompatibilities that could contribute to reproductive isolation can be detected and mapped within *C. elegans*.

In order to understand the mechanisms that lead to speciation, insight is required into the genetic basis of reproductive isolation. The most widely accepted explanation for the genetic basis of intrinsic, post-zygotic reproductive isolation between species is the Bateson-Dobzhansky-Muller (BDM) model (Bateson 1909; Dobzhansky 1936; Muller 1942). This relies on negative epistasis between alleles, and normally considers the case of alleles that have been fixed in different lineages. In hybrids, negative epistasis between alleles that have not been tested together by natural selection result in reduced hybrid fitness (Phillips 2008). Such epistatic interactions have been shown to be involved in, for instance, hybrid male sterility in *Drosophila* (Perez and Wu 1995; Orr and Irving 2001; Tao *et al.* 2003) and are also important in human disease and in complex traits more generally (see Phillips 2008 and Mackay 2014 for review). In recent years, the causal polymorphisms underlying BDM incompatibilities have been identified in a limited number of species, with divergence in both coding sequence and in regulatory elements producing incompatibilities (see Presgraves 2010 for review).

BDM incompatibilities will however also arise within species (see Cutter 2012 for review) and theoretical analyses suggest that interactions between synthetic deleterious loci are common (Phillips and Johnson 1998; Lachance et al. 2011). This is supported by the widespread observation of outbreeding depression in hybrids between divergent populations (e.g. Templeton 1986; Edmands 1999; Dolgin *et al.* 2007; Drury *et al.* 2013; Gimond *et al.* 2013). A small number of BDM incompatibilities have now been identified within species, mostly producing major effects (e.g. Seidel *et al.* 2008; Bikard *et al.* 2009; Drury *et al.* 2011; Baird and Stonesifer 2012). More recently, a genome wide screen in *D. melanogaster* RILs identified many epistatic interactions, two of which were shown to have major effects on fecundity (Corbett-Detig *et al.* 2013). It is however likely that the alleles and regions that have been found to date represent only a subset of the polymorphic

incompatibilities within a species, *i.e.* the major effects identified to date represent those that are easy to detect (see Rockman 2012 for a general discussion of this issue).

As outbreeding depression has been documented between isolates of the free-living nematode *Caenorhabditis elegans* (Dolgin *et al.* 2007) it is likely that a range of potential incompatibilities exists between isolates. We therefore sought to identify small-effect incompatibilities between the isolates CB4856 and N2. We sought these by looking at the disruption of a complex, highly regulated and coordinated, phenotype, egg-laying, and undertook screens for genomic regions that disrupt this process. At 20°C, *C. elegans* N2 eggs are normally laid about three hours after fertilization at around the 30-cell stage (Hirsh *et al.* 1976), with hatching occurring approximately fourteen hours later (Sulston *et al.* 1983). Disruption of the egg-laying process produces an *egl* (*eg*g *l*aying abnormal) phenotype, with one class of *egl* mutation characterized by an increase in the number of fertilized eggs retained within the body and eggs being laid at a much later stage of development. Mutations producing this *egl* phenotype have been identified in genes that affect chemosensation, muscle development, the cell lineage, sex determination and dauer larvae development (Greenwald and Horvitz 1980; Horvitz and Sulston 1980; Waterston *et al.* 1980; Trent *et al.* 1983; WormBase). We therefore considered that this phenotype represented a suitably large target for the development of incompatibilities. Screens were undertaken using *C. elegans* recombinant inbred lines (RILs) and introgression lines (ILs) produced from the isolates CB4856 and N2 (see Methods for details of these lines) and identified multiple quantitative trait loci (QTL) that result in a synthetic *egl* phenotype. For two of the QTL regions identified, analysis of the ILs indicates that the incompatibilities are a consequence of complex interactions between multiple loci. Incompatibility regions identified in these screens are also shown to be associated with negative effects on lifespan and on

reproduction, suggesting that the incompatibilities are likely to be deleterious. In combination, these results indicate that numerous BDM incompatibilities that could lead to reproductive isolation can be detected within *C. elegans*.

# MATERIALS AND METHODS

*Worms*

Experiments were performed using the N2 (Bristol) isolate (obtained from the *Caenorhabditis* genetics centre), wild isolates of *C. elegans* (obtained from Marie-Anne Félix, IBENS, Paris, France, and from the CGC), RILs produced from crosses between CB4856 and N2 (see, for details, Li *et al.* 2006; Kammenga *et al.* 2007; Kammenga *et al.* 2008; Li *et al.* 2010; Viñuela *et al.* 2010; Elvin *et al.* 2011; Rodriguez *et al.* 2012; Viñuela *et al.* 2012), and a panel of CB4856/N2 ILs derived from these RILs in which regions of the CB4846 genome have been introgressed into an N2 background (see, for details, Doroszuk *et al.* 2009; Green *et al.* 2013). Briefly, the RILs were created from crosses between N2 and CB4856, with the F1 progeny subsequently inbred, by transfer of single animals at each generation, for 20 generations. RILs were then genotyped at 121 markers across the genome (20 each on chromosomes *I*, *II*, *III*, *IV and X*, and *21* on *V*). The ILs were produced from specific RILs, chosen based on the CB4856 regions they contain, these RILs were back-crossed to N2, genotyped, further back-crossed as appropriate, and then genotyped at the same markers as the RILs and at two additional markers on chromosome *IV* (for a total of 123 markers). This resulted in the production of a panel of ILs, each containing a single segment of the CB4856 genome in an N2 background.

Worms were maintained using standard methods and fed on the OP50 strain of *Escherichia coli* (Stiernagle 2006). All experiments were undertaken at 20°C and were initiated from synchronized populations of L1s produced by allowing eggs isolated from hypochlorite treated adults (Stiernagle 2006) to hatch on plates without food and to develop for 24 hours. Within assays, genotypes were randomized and plates blind coded, with plates that became infected by fungi excluded from

134 analyses.

135

136 *Embryo stage analysis in the RILs and ILs*

137 The various stages of embryo morphogenesis are well defined in *C. elegans* (Von
138 Ehrenstein and Schierenberg 1980) and can be identified with a dissecting
139 microscope. Most screens undertaken for mutations producing an *egl* phenotype
140 relied on screening worms early in the reproductive period to identify hermaphrodites
141 that had died by internal hatching of progeny (bagging) or that were bloated with late-
142 stage eggs (Greenwald and Horvitz 1980; Horvitz and Sulston 1980; Waterston *et al.*
143 1980; Trent *et al.* 1983). Subsequent analysis of these mutants showed that most
144 worms capable of releasing eggs tended to lay them at a much later stage of
145 development than the wild-type (Trent *et al.* 1983). As we aimed to identify genomic
146 regions that, when in a different genetic background, produced an *egl* phenotype, we
147 determined the stages of eggs laid by worms late in the reproductive period. Our
148 reasoning for screening late in the reproductive life is that this would allow the
149 identification of differences reliant on age-related loss-of-function. Preliminary
150 experiments (data not shown, but see Figs 4 and 5) indicated that both N2 and
151 CB4856 continue to lay almost all eggs at very early stages of development (Fig S4)
152 throughout the reproductive period. We therefore considered that laying eggs at a
153 late stage of development could be considered a consequence of an incompatibility
154 between N2 and CB4856 alleles.

155

156 For embryo stage analysis, we classified progeny into four stages: stage I from
157 fertilization to the end of gastrulation; stage II from ‘lima bean’ to ‘comma’ stage
158 embryos; stage III ‘tadpole’ to ‘pretzel’ stage; and L1 (stages as described by Von
159 Ehrenstein and Schierenberg 1980), see also (Trent *et al.* 1983). Unless otherwise
160 noted, embryo stages were assayed on the third day of reproduction, six days after
161 recovery from L1 arrest, with adults transferred to fresh NGM plates five days after

feeding to allow progeny to be discarded. On the day of assay, for each genotype, five to ten worms were moved to a fresh NGM plate for 2 hours and then discarded. Eggs laid within this two hour window were then observed and the developmental stage classified. For the RIL and IL assays, lines were randomized across experimental blocks and N2 and CB4856 wild types were included as controls in each block. Other assays were conducted in the same manner. Analysis of embryo staging for each experimental block took less than an hour and rescoring of plates during initial experiments indicated that this time did not affect embryo stage data.

All analyses were conducted in custom written scripts in "R" version 2.13.1 x 64. To analyze these data, the effect of genotype on the stage at which the eggs were deposited was tested by ANOVA, with all the individual egg stage scores used as input "egg-stage~genotype+e". This was only used to determine the effect of the genotype on the variation in egg-stage. For the IL and RIL data, the mean square of the trait and the residuals were then used to determine the heritability of the trait in each panel. To find genomic regions associated with the control of egg stage we used QTL mapping. For QTL mapping in the RILs we used a single marker model, with the percentage of total progeny at a certain stage used as a phenotype. In the RILs, the percentage of progeny at > stage II was also mapped. Genome-wide thresholds were determined by 1000 permutations. In each permutation round the phenotypic scores were randomly distributed over the RILs after which genome wide QTL were mapped. The most significant linkage was recorded for each permutation round. The 95% highest –log10(p) value was taken as the 0.05 genome-wide threshold. A similar method was used to determine the threshold for multiple QTL mapping (MQM).

*Bin mapping*

Bin-mapping in the ILs was done as described by Doroszuk *et al.* (2009) and Green *et al.* (2013), with the exception that a chi-square test was used as a statistical test. The percentages of eggs per stage of N2 was used as expected distribution and tested against the distribution per bin. Threshold was determined by 10,000 permutations. Each permutation picked the egg-stage scores of two groups of three randomly selected dishes. These two groups were than used in a chi-square test. The 95% highest –log10(p) value was taken as the 0.05 FDR threshold. This method was also used to determine the threshold in IL vs IL mapping.

*MQM method*

A forward marker selection was used as MQM method. The mapping was initiated by single marker mapping. The marker with most significant linkage was added to the mapping model as a cofactor. The cofactor was excluded from the model when markers closer than 5 markers from the cofactor were considered or when the significance of the cofactor was > 0.05. This process was repeated until no new QTL/cofactors could be added.

*Fixed locus mapping*

To investigate the effect of the major QTL of the left of chr IV on QTL mapping we fixed the locus by splitting the RILs into two groups. One group with an N2 allele at the left of chr IV and one group with a CB4856 allele at the left of chr IV. Single marker mapping using linear regression was subsequently used to find QTL in these two groups of RILs.

*Sub-IL generation*

In order to further investigate the effects of introgressions on chromosome IV on the control of egg stage, we also analyzed an additional set of Sub-ILs (ewIR4001-4011). These were generated by crossing ewIR052 with N2 and selecting for new

recombinants in the F4 offspring. F4 offspring were obtained by single worm decent. RFLP Markers described in Li *et al.* (2006) and Doroszuk *et al.* (2009) spanning the original ewIR052 introgression were used for recombination detection.

*IL vs. IL mapping*

To test if the egg-stage distribution between each IL-pair were different a chi-square test was used. The percentages of eggs per stage of one IL was used as expected distribution and tested against the distribution of the other IL. Pairs were then compared as described by Shao *et al.* (2010) and Green *et al.* (2013) to find QTL.

*Embryo stage analysis in wild isolates*

Preliminary experiments and the RIL and IL analyses indicated that N2 and CB4856 lay the majority of their eggs at very early stages of development. To investigate natural variation in this trait more broadly, we assayed, as described above, a range of wild isolates. The IL ewIR51, which contains a CB4856 introgression on chromosome IV that results in the production of large numbers of late stage progeny (see Figs 2 and S2), was included in these assays as a control.

*Analysis of the chromosome IV QTL*

To determine how the embryo stage of progeny changed across the reproductive period, we compared ewIR51, ewIR52 (another IL containing the chromosome IV QTL), CB4856 and N2. Here the embryo stages of progeny were determined, as described above, daily for the first three days of the reproductive period. To determine if the production of large numbers of late stage embryos was associated with an increase in the number of fertilized eggs *in utero*, as seen in many *egl* mutants (Trent *et al.* 1983), we compared ewIR51, ewIR52 and N2. To do this, individual hermaphrodites were transferred to a drop of hypochlorite solution (Stiernagle 2006) on an NGM plate with food. Plates were then incubated at 20°C for

two days when the number of progeny that had developed was determined. Again, these assays were undertaken daily for the first three days of the reproductive period.

*Relationship to other traits*

To determine how variation in other life history traits relates to the synthetic *egl* effects observed in the RIL and IL lines, all ILs containing introgressions on chromosomes II and IV were assayed for body length, lifetime fecundity and lifespan. These analyses also identified any animals that died by bagging. These assays used standard methods for the analysis of reproductive traits in *C. elegans* (Hodgkin and Doniach 1997). Body length was determined as described by Harvey and Orbidans (2011) for worms two days after recovery from L1 arrest, with individuals photographed using a Moticam 2000 video camera (Motic, Wetzlar, Germany) and the length from the mouth to the base of the tail, determined in ImageJ (Rasband 1997-2009). Worms were considered to have died if they were not moving and failed to respond to touch.

*Data storage*

All data was stored in WormQTL (www.wormqtl.org; Snoek *et al.* 2013; Snoek *et al.* 2014b; van der Velde *et al.* 2014).

# RESULTS

Analyses of the RILs (101 lines) and the ILs (87 lines) indicated that genotype significantly affected the stage at which eggs were laid ($p < 1e\text{-}15$ in both cases). The heritability of the egg stage was also very high (estimated as 96.1% in the RILs and 92.9% in the ILs based on individual egg measurements and 74.8% on multiple population averages per genotype in the ILs), although variability between replicates suggests that the heritability based on the individual egg measurements is most likely an over-estimation. In both sets of lines, the N2 and CB4856 controls are not significantly different (chi-square test; $p \sim 1$), with both lines laying mostly stage I eggs (~96% and ~90% respectively for N2 and CB4856).

The phenotypic distribution in both the RILs and ILs shows a one sided transgression, with many genotypes laying large proportions of their eggs at much later stages (Figs 1 and S1) than either of the parental isolates. About half of the RILs laid 50% or more eggs at stage III or later, with about 20% of the ILs displaying such extreme phenotypes (Fig. 1). These lines therefore phenocopy mild *egl* mutations, *i.e.* they would be classified as M/E, most/early, with all or most progeny released, a few early-stage eggs, and many late-stage eggs observed on the plate (Trent *et al.* 1983). The observed transgression therefore provides evidence that the stage at which an egg is deposited is a polygenic trait. Moreover it suggests that either N2 and CB4856 each carry positive and negative allele(s) of the genes involved that are acting additively, or that the observed effects are a consequence of incompatibilities between diverged N2 and CB4856 alleles at different loci, i.e. negative epistatic effects, or a combination of both of these. That more RILs than ILs show an *egl* phenotype, suggests that multiple regions of the genome and interactions between those contribute to the laying of late stage eggs (comparison between Fig. 1 A and B).

*QTL mapping in the RILs and ILs*

QTL mapping in the RILs identified one highly significant locus at the left of chromosome IV (Fig. 2A). This locus can be found for the percentage of progeny at stage I, stage III, L1 and > stage II, with the CB4856 allele at this locus increased the proportions of late stage progeny. These analyses also identified minor QTL for the proportion of L1s on both chromosomes I and II. MQM analysis indicated that additional QTL can be detected on chromosomes I, III and IV (Table 1, Fig. S2). A two locus scan for epistatic interactions suggested that there were interactions between many of these QTLs, but, due to limited power, these were not significant after correction for multiple testing.

Bin mapping in the ILs using the data from the initial genome wide screen identified a total of 8 QTL where the CB4856 introgression increased the production of late-stage eggs (Fig. 2B). ILs with introgressions harboring one of these QTL were retested in a separate experiment and this analysis resulted in the confirmation of 4 out of the 8 QTL (Table 1, Fig. 2B), with 3 of these QTL overlapping the major QTL and minor QTL identified in the RILs. The IL analysis also suggests the presence of additional QTL on chromosome V and on the X chromosome. In combination, the RIL and IL analyses therefore reproducibly identify regions of chromosomes I, II and IV where introgression of the CB4856 region into an N2 background results in an increased production of late-stage eggs.

The QTL identified by bin mapping span very large regions of the genome, up to almost a whole chromosome in the case of chromosomes I and IV (Table 1, Fig. 2B). Because of this, we investigated the individual ILs for clues on the number of alleles/QTL present. This was done by using a chi-square test to test for a difference in stage numbers between N2 and the individual ILs (Table 1, Fig. S3). These analyses detect and confirm the stage increasing CB4856 QTL on chromosomes I, II,

III and IV. Given that ~90% of progeny in N2 and CB4856 are stage I eggs, comparison of the ILs and N2 will only detect CB4856 alleles that increase progeny stage. Such analyses suggest that many regions of the genome disrupt the normal process of egg-laying. For example, on chromosome I this suggests the presence of at least 3 separate QTL as three non-overlapping ILs are different from N2 (Table 1, Fig. S3). In contrast to such analyses, comparison of overlapping ILs allows the identification of regions that contain CB4856 alleles that decrease progeny stage (Table 1, Fig. 3). These comparisons support the conclusions that the QTL detected here can be separated into multiple factors.

*Embryo stage analysis in wild isolates*

To determine if late stage egg production was seen in wild isolates of *C. elegans*, the embryo stage of hermaphrodites from a range of wild isolates on the third day of reproduction was tested. These analyses indicated that there are differences between lines, but that wild isolates all lay eggs at a predominantly early stage of development (Fig. 4). This further supports our classification of the late-stage embryo production trait as an incompatibility.

*Analysis of the chromosome IV QTL*

Analysis of embryo stage across the reproductive period indicates that the trait is age-related, such that the proportion of embryos laid at later stages of development increases throughout the reproductive period (Fig. 5A). This suggests that it may represent a change in the rates at which the worms are senescing. Previously identified differences in developmental speed between RILs derived from crosses between the isolates N2 and CB4856 only span a few hours (Francesconi and Lehner 2014; Snoek *et al*. 2014a) and cannot therefore cause the (large) differences in egg-stages between lines.

Many *egl* mutations cause worms to retain large numbers of eggs *in utero*, with young adults displaying a slightly bloated phenotype and older worms often containing many times the normal number of fertilized embryos. Comparison of two ILs containing the major chromosome IV QTL to N2 (Fig. 5B) indicated that the number of eggs *in utero* is slightly increased during the first two days of reproduction, but that there is no increase seen on the third day of reproduction.

*Relationship to other traits*

Analysis of all ILs containing introgressions on chromosomes II and IV indicated that all traits were variable (Fig. 6A and B), with these analyses defining QTL for all traits (Table 2 and S2). Comparison of these QTL to those found in previous analyses indicates that many QTL are found in multiple studies. For instance, variation in body size between N2 and BO has previously been mapped to chromosome IV (Knight *et al.* 2001) and one of the chromosome IV body size QTL identified here (Table 2) contains *tra-3*, a gene polymorphic between CB4856 and N2 that affects how body size changes across temperatures (Kammenga *et al.* 2007). Similarly, previous comparisons using CB4856 and N2 RILs identified a fecundity QTL on chromosome IV (Gutteling *et al.* 2007), although this was found at 12°C and not at 24°C. The patterns of variation identified here do however indicate that the control of these traits is complex, with chromosome IV containing five separate QTL affecting body size (Table 2).

There was no overall correlation between the traits assayed, showing that multiple independent functional allelic differences exist between N2 and CB4856 (on chromosomes II and IV). These analyses do however indicate that QTL affecting bagging, lifetime fecundity and lifespan can be identified in regions associated with the production of late stage progeny (Fig. 6). These data also provide direct evidence for epistatic interactions affecting both lifespan and fecundity on chromosome *II*, with

376 the IL vs IL analyses of ewlR021-23 indicating an epistatic interaction between the

377 CB4856 region in ewlR021 and the region in ewlR023 (Fig. 6 and Table 2).

# DISCUSSION

Within the *Caenorhabditis* species there is a continuum between distinct, reproductively isolated, species and species where isolates are at the very earliest stages of speciation (Baird and Stonesifer 2012; Kozlowska *et al.* 2012; Gimond *et al.* 2013). The polymorphisms that result in outbreeding depression and hybrid breakdown within species underlie developmental transitions that can ultimately lead to speciation. Our analyses of ILs and RILs derived from the isolates CB4856 and N2 indicate that many of these lines phenocopy mild *egl* mutations, laying progeny at an advanced stage of development (Fig. 1). Genetic analyses of these data revealed multiple QTL affecting egg-laying (Fig. 2 and Table 1). These data indicate that the stage at which an egg is deposited is a polygenic trait. However, it is not clear from this analysis if this is a consequence of the additive action of positive and negative allele(s) from CB4856, of epistatic interactions between loci, or a combination of both. The observation that all of the wild isolates lay very early stage eggs (Fig. 4) and that the QTL are associated with increased bagging does however argue that laying late stage eggs is deleterious and therefore that selection will be acting to minimize this.

The other phenotypes linked to these *egl* effects involve fitness traits (Table 2). The clearest association is with bagging, with ILs underlying the QTL on both chromosomes showing increased bagging (Fig. 6, Table 2). This association between laying late stage progeny and an increased rate of bagging is unsurprising given that this is a common phenotype in *egl* mutants (Trent *et al.* 1983). The patterns of bagging observed on both chromosome II and IV indicates that these do not represent simple interactions between two loci. For example, comparison of ILs ewIR21-23 (Fig. 6) suggest the presence of interactions with other loci on the same chromosome (e.g. between alleles present in ewLR21 and those in ewlR23). As this

trait is, like production of late-stage embryo trait, based on the proportion of the population showing the trait, it is not possible to use these comparisons to distinguish between QTL acting additively and those that are a function of epistatic interactions. This is not the case for the lifespan and fecundity QTL that we detect in the two incompatibility regions (Table 2), as positive effect QTL would be detected in comparisons between ILs and N2. Here, both regions support the interpretation of the QTL as epistatic interactions. For instance, comparisons between ILs on chromosome II define two positive effect QTL for both fecundity and lifespan (Table 2), but the introgressions in this region are not consistent with this, as it would imply two positive effect QTL in ewlR22 and one each in ewlR21 and 23 (Figure 6). As ewlR21 and 23 are not different to N2, a more parsimonious explanation would be that the increased lifespan and fecundity seen in ewlR22 is a consequence of an interaction between CB4856 alleles that are separated in ewlR21 and 23. In this context, it noteworthy that ewlR21 has a slightly reduced lifespan in this assay and has been previously shown, using these ILs, to contain a CB4856 allele that reduces lifespan (Doroszuk *et al.* 2009). A similar case for a complex interaction can be made for the lifespan QTL identified on chromosome IV (Table 2), a QTL also found by (Doroszuk *et al.* 2009). Given that fecundity QTL are detected at both ends of chromosome IV (Fig. 6 and Table 2) it is not clear if a model of additive QTL is more consistent with these data than one reliant on epistatic interactions.

Given the detrimental effects of the QTL we have detected, it is likely that they would represent weak post zygotic barriers. Conceptually, the effects we have detected can be viewed in a number of differing ways. They could be the consequence of transgressive segregation, although in this case this is unlikely as the trait mapped is essentially synthetic and not seen in either parent or in other wild isolates. Alternatively, the trait could be the result of a disruption of canalization and represent the exposure of cryptic genetic variation. In general, canalization acts to limit trait

433 sensitivity to changes in the environment and/or the genetic background
434 (Waddington 1942; Schalhausen 1949; Lerner 1954). Within species, such
435 incompatibilities will appear similar to cryptic variation, a situation where genetic or
436 environmental perturbation is required to reveal otherwise hidden genetic variation
437 (Gibson and Dworkin 2004; Li *et al.* 2006; Masel and Siegal 2009; Snoek *et al.* 2012;
438 Paaby and Rockman 2014). Here, the origin of cryptic variation may represent the
439 evolution of epistatic correction of deleterious effects of a particular mutation (that
440 may or may not also produce adaptive changes). Such changes would be analogous
441 to the local compensatory mutations that occur both between and within species to
442 correct structural changes in proteins (Long *et al.* 2013)
443
444 The life history of *C. elegans* may facilitate the build-up of such deleterious
445 mutations. For example, fixation within a line of adaptive mutations that have
446 pleiotropic deleterious effects, or mildly deleterious mutations (as aided by the
447 extensive selfing and the bottlenecking resulting from the *C. elegans* life-history)
448 would allow the subsequent selection for compensatory mutations. As compensatory
449 mutations appear commonly in *C. elegans*, as shown by experiments that have re-
450 imposed selection on mutation accumulation lines (Estes and Lynch 2003; Denver *et*
451 *al.* 2010; Estes *et al.* 2011), this could result in a negative interaction between the
452 compensatory mutation and the original allele. This would produce a situation where
453 local adaptation (first mutation advantageous) or cryptic genetic variation (first
454 mutation deleterious and now associated with a compensatory mutation) would
455 produce, at least, a pair of co-adapted genes. In making the RILs and the ILs the
456 links between co-adapted genes might be broken up and cryptic genetic variation
457 that only exists to correct otherwise deleterious polymorphisms is revealed. It is clear
458 that there is significant genotypic and phenotypic variation between *C. elegans* wild
459 isolates (Hodgkin and Doniach 1997; Viney *et al.* 2003; Barrière and Félix 2005;
460 Barrière and Félix 2007; Harvey et al. 2008; Harvey 2009; Maydan *et al.* 2010;

461 Andersen *et al.* 2012; Green *et al.* 2013; Thompson *et al.* 2013; Volkers *et al.* 2013;
462 Snoek *et al.* 2014b). Large scale analysis of *C. elegans* isolates reveals little
463 grouping by isolation environment or by country of origin on a global scale (Andersen
464 *et al.* 2012), although there is evidence at smaller scales that suggests local
465 adaptation (Volkers *et al.* 2013). Hence, there is much potential for local adaptation
466 to produce the kinds of interactions proposed here.
467
468 The mapping resolution of the QTL identified here precludes a detailed search for
469 candidate genes. However, comparison of the locations of the QTL identified here to
470 the results of eQTL studies of lines produced from crosses between N2 and CB4856
471 (Li *et al.* 2006; Rockman *et al.* 2010; Viñuela *et al.* 2010; Viñuela *et al.* 2012; Snoek
472 *et al.* 2013; van der Velde *et al.* 2014) suggest that a number of the genome hotspots
473 for trans acting eQTL do co-localize with incompatibility QTL. This is particularly the
474 case with the incompatibility QTL on the top of chromosome IV (Fig. 2), where a very
475 strong eQTL hotspot has been identified under a range of conditions (Rockman *et al.*
476 2010; Viñuela *et al.* 2010; Viñuela *et al.* 2012). This part of chromosome IV also
477 contains multiple QTL affecting dauer larvae development in growing populations
478 (Green *et al.* 2013) and a large number of separate QTL affecting olfactory
479 preference between *Serratia marcescens*, a bacterium pathogenic to *C. elegans*, and
480 *E. coli* (Glater *et al.* 2014). The large number of phenotypes now known to be linked
481 this region and the observed complexity of their regulation, as implied by the number
482 of separable QTL in the region (Green et al. 2013; Glater et al. 2014) (Table 1; Fig.
483 5), mean that determining how these variants are related will be interesting for their
484 potential role in speciation. More generally, given the extensive lab adaptation
485 observed in the N2 isolate (McGrath *et al.* 2009; Weber *et al.* 2010; Duveau and
486 Félix 2012) it would be informative to investigate the role of these changes in the
487 incompatibilities observed here as such alleles are known to be of recent origin. This

would therefore demonstrate that short periods of strong selection can rapidly produce incompatibilities.

To date, the mechanisms that isolate four *Caenorhabditis* species, *C. elegans*, *C. briggsae*, *C. remanei*, and *C. sp.* strain CB5161, now named *C. brenneri* (Sudhaus and Kiontke 2007), have been described (Baird *et al.* 1992). Work on more recently isolated *Caenorhabditis* species, which can form viable, and in some cases fertile, hybrids, has also started to address the genetic bases of speciation in this group (Baird and Stonesifer 2012; Kozlowska *et al.* 2012; Gimond *et al.* 2013). As outbreeding depression is also observed in the other predominantly self-fertilizing *Caenorhabditis* species (Ross *et al.* 2011; Baird and Stonesifer 2012; Kozlowska *et al.* 2012; Gimond *et al.* 2013) it is likely that BDM incompatibilities will also be detectable within these species. Over the longer term, the identification of the causative loci for the QTL identified here would allow comparison with the changes that produce more extreme reproductive isolation and the alleles involved in the very early stages of speciation that have been detected in other *Caenorhabditis* species (Dey *et al*. 2012; Kozlowska *et al.* 2012). This suggests that the *Caenorhabditis* species have the potential to be hugely informative about the genetics of speciation and more generally about the role of epistatic interactions in the control of complex traits.

## ACKNOWLEDGMENTS

Strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Elements of the work reported here were undertaken by HEO while supported by a Nuffield Foundation Undergraduate Research Bursary (URB/39836). SCH was supported by a Research Grant from the Royal Society, LBS and JK were funded by the ERASysbio-plus ZonMW project

GRAPPLE - Iterative modelling of gene regulatory interactions underlying stress, disease and ageing in *C. elegans* (project nr. 90201066), and NEMADAPT, NWO-ALW (project 855.01.151). We thank Wormbase (www.wormbase.org) for being a rich and versatile source of information. We thank Morris Swertz and Joeri van der Velde for their help with making the data accessible through WormQTL.

# LITERATURE CITED


Andersen, E. C., J. P. Gerke, J. A. Shapiro, J. R. Crissman, R. Ghosh et al., 2012 Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. Nat. Genet. 44: 285-290.

Baird, S. E., and R. Stonesifer, 2012 Reproductive isolation in *Caenorhabditis briggsae*: dysgenic interactions between maternal- and zygotic-effect loci result in a delayed development phenotype. Worm 1: 189-195.

Baird, S. E., M. E. Sutherlin and S. W. Emmons, 1992 Reproductive isolation in Rhabditidae (Nematoda, Secernentea) - mechanisms that isolate 6 species of 3 genera. Evolution 46: 585-594.

Barrière, A., and M. A. Félix, 2005 High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. Curr. Biol. 15: 1176-1184.

Barrière, A., and M. A. Félix, 2007 Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. Genetics 176: 999-1011.

Bateson, W., 1909 Heredity and variation in modern lights. In Darwin and Modern Science (Seward, A.C., ed.), Cambridge University Press: pp. 85-101.

Bikard, D., D. Patel, C. Le Mette, V. Giorgi, C. Camilleri et al., 2009 Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. Science 323: 623-626.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl and J. F. Ayroles, 2013 Genetic incompatibilities are widespread within species. Nature 504: 135-137.

Cutter, A. D., 2012 The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. Trends Ecol. Evol. 27: 209-218.

Denver, D. R., D. K. Howe, L. J. Wilhelm, C. A. Palmer, J. L. Anderson et al., 2010 Selective sweeps and parallel mutation in the adaptive recovery from deleterious mutation in *Caenorhabditis elegans*. Genome Res. 20: 1663-1671.

549 Dey, A., Y. Jeon, G. X. Wang and A. D. Cutter, 2012 Global population genetic
550 structure of *Caenorhabditis remanei* reveals incipient speciation. Genetics 191:
551 1257-1269.
552 Dobzhansky, T., 1936 Studies on hybrid sterility. II. Localization of sterility factors in
553 *Drosophila pseudoobscura* hybrids. Genetics 21: 113-135.
554 Dolgin, E. S., B. Charlesworth, S. E. Baird and A. D. Cutter, 2007 Inbreeding and
555 outbreeding depression in *Caenorhabditis* nematodes. Evolution 61: 1339-1352.
556 Doroszuk, A., L. B. Snoek, E. Fradin, J. Riksen and J. Kammenga, 2009 A genome-
557 wide library of CB4856/N2 introgression lines of *Caenorhabditis elegans*. Nucleic
558 Acids Res. 37: e110.
559 Drury D. W., R. C. Ehmke, V. N. Jideonwo, and M. J. Wade, 2013 Developmental
560 trajectories and breakdown in F1 interpopulation hybrids of *Tribolium castaneum*.
561 Ecol. Evol. 3: 1992-2001
562 Drury D. W., V. N. Jideonwo, R. C. Ehmke, and M. J. Wade, 2011 An unusual barrier
563 to gene flow: perpetually immature larvae from inter-population crosses in the flour
564 beetle, *Tribolium castaneum*. J. Evol. Biol. 24: 2678-2686
565 Duveau, F., and M. A. Félix, 2012 Role of pleiotropy in the evolution of a cryptic
566 developmental variation in *Caenorhabditis elegans*. PLoS Biol. 10: e1001230.
567 Edmands, S., 1999 Heterosis and outbreeding depression in interpopulation crosses
568 spanning a wide range of divergence. Evolution 53: 1757-1768.
569 Elvin, M., L. B. Snoek, M. Frejno, U. Klemstein, J. E. Kammenga et al., 2011 A
570 fitness assay for comparing RNAi effects across multiple *C. elegans* genotypes.
571 BMC Genomics 12: 510.
572 Estes, S., and M. Lynch, 2003 Rapid fitness recovery in mutationally degraded lines
573 of *Caenorhabditis elegans*. Evolution 57: 1022-1030.
574 Estes, S., P. C. Phillips and D. R. Denver, 2011 Fitness recovery and compensatory
575 evolution in natural mutant lines of *C. elegans*. Evolution 65: 2335-2344.

576 Francesconi, M., and B. Lehner, 2014 The effects of genetic variation on gene
577 expression dynamics during development. Nature 505: 208-211.
578 Gibson, G., and I. Dworkin, 2004 Uncovering cryptic genetic variation. Nat. Rev.
579 Genet. 5: 681-690.
580 Gimond, C., R. Jovelin, S. Han, C. Ferrari, A. D. Cutter et al., 2013 Outbreeding
581 depression with low genetic variation in selfing *Caenorhabditis* nematodes.
582 Evolution 67: 3087-3101.
583 Glater, E. E., M. V. Rockman and C. I. Bargmann, 2014 Multigenic natural variation
584 underlies *Caenorhabditis elegans* olfactory preference for the bacterial pathogen
585 *Serratia marcescens*. G3 (Bethesda) 4: 265-276.
586 Green, J. W., L. B. Snoek, J. E. Kammenga and S. C. Harvey, 2013 Genetic
587 mapping of variation in dauer larvae development in growing populations of
588 *Caenorhabditis elegans*. Heredity (Edinb) 111: 306-313.
589 Greenwald, I. S., and H. R. Horvitz, 1980 unc-93(e1500): A behavioral mutant of
590 *Caenorhabditis elegans* that defines a gene with a wild-type null phenotype.
591 Genetics 96: 147-164.
592 Gutteling, E. W., J. A. Riksen, J. Bakker and J. E. Kammenga, 2007 Mapping
593 phenotypic plasticity and genotype-environment interactions affecting life-history
594 traits in *Caenorhabditis elegans*. Heredity (Edinb) 98: 28-37.
595 Harvey, S. C., 2009 Non-dauer larval dispersal in *Caenorhabditis elegans*. J. Exp.
596 Zool. B Mol. Dev. Evol. 312B: 224-230.
597 Harvey, S. C., and H. E. Orbidans, 2011 All eggs are not equal: the maternal
598 environment affects progeny reproduction and developmental fate in
599 *Caenorhabditis elegans*. PLoS One 6: e25840.
600 Harvey, S. C., A. Shorto and M. E. Viney, 2008 Quantitative genetic analysis of life-
601 history traits of *Caenorhabditis elegans* in stressful environments. BMC Evol. Biol.
602 8: 15.

603 Hirsh, D., D. Oppenheim and M. Klass, 1976 Development of the reproductive
604 system of *Caenorhabditis elegans*. Dev. Biol. 49: 200-219.
605 Hodgkin, J., and T. Doniach, 1997 Natural variation and copulatory plug formation in
606 *Caenorhabditis elegans*. Genetics 146: 149-164.
607 Horvitz, H. R., and J. E. Sulston, 1980 Isolation and genetic characterization of cell-
608 lineage mutants of the nematode *Caenorhabditis elegans*. Genetics 96: 435-454.
609 Kammenga, J. E., A. Doroszuk, J. A. Riksen, E. Hazendonk, L. Spiridon et al., 2007
610 A *Caenorhabditis elegans* wild type defies the temperature-size rule owing to a
611 single nucleotide polymorphism in tra-3. PLoS Genet. 3: e34.
612 Kammenga, J. E., P. C. Phillips, M. De Bono and A. Doroszuk, 2008 Beyond induced
613 mutants: using worms to study natural variation in genetic pathways. Trends
614 Genet. 24: 178-185.
615 Knight, C. G., R. B. Azevedo and A. M. Leroi, 2001 Testing life-history pleiotropy in
616 *Caenorhabditis elegans*. Evolution 55: 1795-1804.
617 Kozlowska, J. L., A. R. Ahmad, E. Jahesh and A. D. Cutter, 2012 Genetic variation
618 for postzygotic reproductive isolation between *Caenorhabditis briggsae* and
619 *Caenorhabditis Sp 9*. Evolution 66: 1180-1195.
620 Lachance, J.L., N.A. Johnson and J.R. True, 2011 The population genetics of X-
621 autosome synthetic lethals and steriles. Genetics 189: 1011-1027.
622 Lerner, I. M., 1954 Genetic homeostasis. Wiley, New York.
623 Li, Y., O. A. Alvarez, E. W. Gutteling, M. Tijsterman, J. Fu et al., 2006 Mapping
624 determinants of gene expression plasticity by genetical genomics in *C. elegans*.
625 PLoS Genet. 2: e222.
626 Li, Y., R. Breitling, L. B. Snoek, K. J. van der Velde, M. A. Swertz et al., 2010 Global
627 genetic robustness of the alternative splicing machinery in *Caenorhabditis*
628 *elegans*. Genetics 186: 405-410.
629 Long Q., F. A. Rabanal, D. Meng, C. D. Huber, A. Farlow, A. Platzer, Q. Zhang, B. J.
630 Vilhjálmsson, A. Korte, V. Nizhynska, V. Voronin, P. Korte, L. Sedman, T.

631 Mandáková, M. A. Lysak, Ü. Seren, I. Hellmann and M. Nordborg 2013 Massive
632 genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden.
633 Nat. Genet. 45: 884-890. Mackay, T. F., 2014 Epistasis and quantitative traits:
634 using model organisms to study gene-gene interactions. Nat. Rev. Genet. 15: 22-
635 33.
636 Masel, J., and M. L. Siegal, 2009 Robustness: mechanisms and consequences.
637 Trends Genet. 25: 395-403.
638 Maydan, J. S., A. Lorch, M. L. Edgley, S. Flibotte and D. G. Moerman, 2010 Copy
639 number variation in the genomes of twelve natural isolates of *Caenorhabditis*
640 *elegans*. BMC Genomics 11: 62.
641 McGrath, P. T., M. V. Rockman, M. Zimmer, H. Jang, E. Z. Macosko et al., 2009
642 Quantitative mapping of a digenic behavioral trait implicates globin variation in *C.*
643 *elegans* sensory behaviors. Neuron 61: 692-699.
644 Muller, H. J., 1942 Isolating mechanisms, evolution and temperature. Temperature,
645 Evolution, Development, ed. T. Dobzhansky, vol 6, Biological Symposiua: A
646 Series of Volumes Devoted to Current Symposia in the Field of Biology
647 (Lancaster, PA: Jaques Cattell Press), 6: 71-125.
648 Orr, H. A., and S. Irving, 2001 Complex epistasis and the genetic basis of hybrid
649 sterility in the *Drosophila pseudoobscura* Bogota-USA hybridization. Genetics
650 158: 1089-1100.
651 Paaby, A. B., and M. V. Rockman, 2014 Cryptic genetic variation: evolution's hidden
652 substrate. Nature Reviews Genetics 15: 247-58
653 Perez, D. E., and C. I. Wu, 1995 Further characterization of the Odysseus locus of
654 hybrid sterility in *Drosophila*: one gene is not enough. Genetics 140: 201-206.
655 Phillips, P. C., 2008 Epistasis - the essential role of gene interactions in the structure
656 and evolution of genetic systems. Nat. Rev. Genet. 9: 855-867.
657 Phillips, P. C., and N. A. Johnson, 1998 The population genetics of synthetic lethals.
658 Genetics 150: 449-458.

659 Presgraves, D. C., 2010 The molecular evolutionary basis of species formation. Nat
660 Rev. Genet. 11: 175-180.
661 Rockman M. V. 2012 The QTN program and the alleles that matter for evolution: all
662 that's gold does not glitter. Evolution 66: 1-17.
663 Rockman, M. V., S. S. Skrovanek and L. Kruglyak, 2010 Selection at linked sites
664 shapes heritable phenotypic variation in *C. elegans*. Science 330: 372-376.
665 Rodriguez, M., L. B. Snoek, J. A. Riksen, R. P. Bevers and J. E. Kammenga, 2012
666 Genetic variation for stress-response hormesis in *C. elegans* lifespan. Exp.
667 Gerontol. 47: 581-587.
668 Ross, J. A., D. C. Koboldt, J. E. Staisch, H. M. Chamberlin, B. P. Gupta et al., 2011
669 Caenorhabditis briggsae recombinant inbred line genotypes reveal inter-strain
670 incompatibility and the evolution of recombination. PLoS Genet. 7: e1002174.
671 Schmalhausen, I. I., 1949 Factors of evolution: the theory of stabilizing selection. .
672 Blakiston, Philadelphia.
673 Seidel, H. S., M. V. Rockman and L. Kruglyak, 2008 Widespread genetic
674 incompatibility in C. elegans maintained by balancing selection. Science 319: 589-
675 594.
676 Shao, H., D. S. Sinasac, L. C. Burrage, C. A. Hodges, P. J. Supelak et al., 2010
677 Analyzing complex traits with congenic strains. Mamm. Genome 21: 276-286.
678 Snoek, L. B., I. R. Terpstra, R. Dekter, G. Van den Ackerveken and A. J. Peeters,
679 2012 Genetical genomics reveals large scale genotype-by-environment
680 interactions in *Arabidopsis thaliana*. Front. Genet. 3: 317.
681 Snoek, L. B., K. J. Van der Velde, D. Arends, Y. Li, A. Beyer et al., 2013 WormQTL--
682 public archive and analysis web portal for natural variation data in Caenorhabditis
683 spp. Nucleic Acids Res. 41: D738-743.
684 Snoek, L. B., M. G. Sterken, R. J. Volkers, M. Klatter, K. J. Bosman et al., 2014a A
685 rapid and massive gene expression shift marking adolescent transition in C.
686 elegans. Sci. Rep. 4: 3912.

687 Snoek, L. B., K. J. van der Velde, Y. Li, R. C. Jansen, M. A. Swertz et al., 2014b
688 Worm variation made accessible: take your shopping cart to store, link, and
689 investigate! Worm 3: e28357.
690 Stiernagle, T., 2006 Maintenance of *C. elegans*. WormBook: 1-11.
691 Sudhaus, W., and K. Kiontke, 2007 Comparison of the cryptic nematode species
692 *Caenorhabditis brenneri sp n.* and *C remanei* (Nematoda : Rhabditidae) with the
693 stem species pattern of the *Caenorhabditis elegans* group. Zootaxa: 45-62.
694 Sulston, J. E., E. Schierenberg, J. G. White and J. N. Thomson, 1983 The
695 embryonic-cell lineage of the nematode *Caenorhabditis elegans*. Dev. Biol. 100:
696 64-119.
697 Tao, Y., Z. B. Zeng, J. Li, D. L. Hartl and C. C. Laurie, 2003 Genetic dissection of
698 hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II.
699 Mapping hybrid male sterility loci on the third chromosome. Genetics 164: 1399-
700 1418.
701 Templeton, A. R., 1986 Coadaptation and outbreeding depression. M. Soulé, ed.
702 Conservation biology: the science of scarcity and diversity. Sinauer Associates,
703 Sunderland, MA.: 33-59.
704 Thompson, O., M. Edgley, P. Strasbourger, S. Flibotte, B. Ewing et al., 2013 The
705 million mutation project: a new approach to genetics in *Caenorhabditis elegans*.
706 Genome Res. 23: 1749-1762.
707 Trent, C., N. Tsuing and H. R. Horvitz, 1983 Egg-laying defective mutants of the
708 nematode *Caenorhabditis elegans*. Genetics 104: 619-647.
709 van der Velde, K. J., M. de Haan, K. Zych, D. Arends, L. B. Snoek et al., 2014
710 WormQTLHD--a web database for linking human disease to natural variation data
711 in *C. elegans*. Nucleic Acids Res. 42: D794-801.
712 Viney, M. E., M. P. Gardner and J. A. Jackson, 2003 Variation in *Caenorhabditis*
713 *elegans* dauer larva formation. Dev. Growth Diff. 45: 389-396.

714 Viñuela, A., L. B. Snoek, J. A. Riksen and J. E. Kammenga, 2010 Genome-wide
715 gene expression regulation as a function of genotype and age in *C. elegans*.
716 Genome Res. 20: 929-937.
717 Viñuela, A., L. B. Snoek, J. A. Riksen and J. E. Kammenga, 2012 Aging uncouples
718 heritability and expression-QTL in *Caenorhabditis elegans*. G3 (Bethesda) 2: 597-
719 605.
720 Volkers, R. J., L. B. Snoek, C. J. Hubar, R. Coopman, W. Chen et al., 2013 Gene-
721 environment and protein-degradation signatures characterize genomic and
722 phenotypic diversity in wild *Caenorhabditis elegans* populations. BMC Biol. 11: 93.
723 Von Ehrenstein, G., and E. Schierenberg, 1980 Cell lineages and development of
724 *Caenorhabditis elegans* and other nematodes. Nematodes as Biological Models,
725 Edited by B.M. Zuckerman. Academic Press, New York.: 1-71.
726 Waddington, C. H., 1942 The canalization of development and the inheritance of
727 acquired characters. Nature 150: 563-565.
728 Waterston, R. H., J. N. Thomson and S. Brenner, 1980 Mutants with Altered Muscle
729 Structure in *Caenorhabditis elegans*. Dev. Biol. 77: 271-302.
730 Weber, K. P., S. De, I. Kozarewa, D. J. Turner, M. M. Babu et al., 2010 Whole
731 genome sequencing highlights genetic changes associated with laboratory
732 domestication of *C. elegans*. PLoS One 5: e13922.
733

734 **Table 1: Locations and effect of QTL detected for egg-stages.** The column label
735 Chr show the chromosome on which the QTL was found. The N2L, CBL, CBR, N2R
736 show the position of the left N2, left CB, right CB and right N2 boundaries of the QTL.
737 The “Detected by” indicates the methods by which the QTL were found/supported.

| Chr | N2L | CBL | CBR | N2R | CB effect | Detected by |
|---|---|---|---|---|---|---|
| I | 2818974 | 3502476 | 3502476 | 4338254 | + | MQM, (BIN), Single IL, ILvsIL |
| I | 9569913 | 10259909 | 10259909 | 11085295 | + | Single IL, ILvsIL |
| I | 11085295 | 11085295 | 11085295 | 11760179 | - | ILvsIL |
| II | 2755074 | 3403575 | 4147051 | 4800868 | + | (BIN), Single IL, ILvsIL |
| II | 4147051 | 4800868 | 10414073 | 11180836 | - | ILvsIL |
| III | 5925983 | 6847169 | 7998164 | 8318553 | - | ILvsIL |
| III | 10027496 | 10613119 | 10613119 | 11341120 | + | MQM, Single IL, ILvsIL |
| III | 10613119 | 11341120 | 11341120 | 12301725 | - | ILvsIL |
| IV | Not applicable | 151889 | 1381409 | 2288742 | + | SM, MQM, Single IL, ILvsIL |
| IV | 2288742 | 3067374 | 3067374 | 3920366 | + | SM, MQM, Single IL, ILvsIL |
| IV | 10122930 | 10909560 | 10909560 | 11668242 | - | ILvsIL |
| IV | 10909560 | 11668242 | 11668242 | 12748880 | + | SM, MQM, Single IL, ILvsIL |
| V | 10368660 | 10912994 | 16008404 | 17377158 | + | Single IL, ILvsIL |
| V | 17377158 | 18574593 | 18574593 | 19525561 | - | ILvsIL |
| V | 18574593 | 19525561 | 20758352 | 20893784 | + | Single IL, ILvsIL |
| X | 5010049 | 5770179 | 5770179 | 7067019 | - | ILvsIL |
| X | 5770179 | 7067019 | 7982354 | 8691677 | + | Single IL, ILvsIL |

738

739

**Table 2: Locations and effect of QTL detected for body length, lifetime fecundity, lifespan and bagging.** QTL limits are shown by the locations of the flanking markers with N2 genotype and the adjacent markers with a CB4856 genotype. QTL marked as Single IL were detected in comparisons between ILs and N2, those marked IL vs IL were detected in comparisons between ILs. Only ILs on chromosome II and IV were tested.

| Trait | Chr | N2L | CBL | CBR | N2R | CB effect | Detected by |
|---|---|---|---|---|---|---|---|
| Size | IV | 766649 | 1381409 | 3067374 | 3920366 | - | Single IL |
| | IV | 5819735 | 6599685 | 12748880 | 13667267 | - | IL vs IL |
| | IV | 8397264 | 9102404 | 9102404 | 10122930 | + | IL vs IL |
| | IV | 11668242 | 12748880 | 12748880 | 13667267 | + | Single IL |
| | IV | 12748880 | 13667267 | 16371991 | 17084259 | - | Single IL |
| Lifespan | II | Not applicable | 176721 | 2755074 | 3403575 | + | IL vs IL |
| | II | 4147051 | 4800868 | 10414073 | 11180836 | + | IL vs IL |
| | IV | 3920366 | 4991858 | 5819735 | 6599685 | - | IL vs IL |
| Fecundity | II | Not applicable | 176721 | 2755074 | 3403575 | + | IL vs IL |
| | II | 4147051 | 4800868 | 10414073 | 11180836 | + | IL vs IL |
| | IV | 12748880 | 13667267 | 16371991 | 17084259 | - | Single IL, IL vs IL |
| Bagging | II | Not applicable | 176721 | 2755074 | 3403575 | + | IL vs IL |
| | II | 4147051 | 4800868 | 10414073 | 11180836 | + | IL vs IL |
| | IV | 3067374 | 3920366 | 3920366 | 4991858 | + | IL vs IL |

## Figure legends

**Figure 1: Embryo stage distribution in the RILs and ILs.** The cumulative percentage of embryo stages per RIL (A) and IL (B). Lines were sorted by the percentage of embryos > stage II.

**Figure 2: QTL mapping in the RILs and ILs.** (A) Mapping of embryo stage in the RILs, with the significance (-log10(p)) multiplied by the sign of the effect of the N2 allele plotted against the marker positions in mega basepairs for the percentage of total eggs in stage I eggs (black solid line), stage II eggs (black dashed line), stage III eggs (black dotted line), stage IV eggs (grey solid line), L1s (grey dashed line), and the proportion of progeny > stage II (grey dotted line). (B) Genome wide bin mapping of late stage embryo production (proportion > stage II), showing the significance (-log10(*p*)) by chi-square test of ILs sharing a certain genomic part against N2.

**Figure 3: Comparison between ILs of chromosome IV.** The CB4856 introgression per IL is shown by the coloured rectangle. Triangles join adjacent CB4856 and N2 markers. Embryo stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white). QTL are indicated on the X axis by red (+) or blue (-) boxes (denoting that the CB4856 allele increases or decreases the proportion of late stage embryos, respectively).

**Figure 4: Embryo stages of wild-isolates.** Embryo stage distribution shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white). CB4856 scores from different experiments (n = 281). (N = IL51 (ewIR51): 52; JU393: 85; JU1401: 76; JU1411: 83; MY2: 98; JU345: 84; PX174: 30; MY1: 94; JU1494: 91; CB4856: 103; N2: 112; CB4853: 60; JU262: 18). For > stage II eggs all the wild isolates are significantly different from ewIR51 ($p < 0.01$, two-sided t-test on

plate averages). None of the > stage II differences between the wild-isolated were significantly different (p> 0.05, two-sided t-test on plate averages).

**Figure 5: Variation across the reproductive period.** (A) Embryo stage distribution across the reproductive period shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white). (B) Number of eggs *in utero* across the reproductive period.

**Figure 6: Incompatibility QTL are associated with variation in other traits.** Average lifespan, lifetime fecundity, body size at L4 and proportion of worms that die by bagging for ILs containing introgressions on chromosome II (A) and IV (B). The CB4856 introgression per IL is shown by the colored rectangle. Triangles join adjacent CB4856 and N2 markers. Error bars represent ±1 S.E., dashed lines and shaded bars represent trait values in N2 and ILs significantly different from N2 ($p < 0.05$) are shown in black.

## Supplemental material

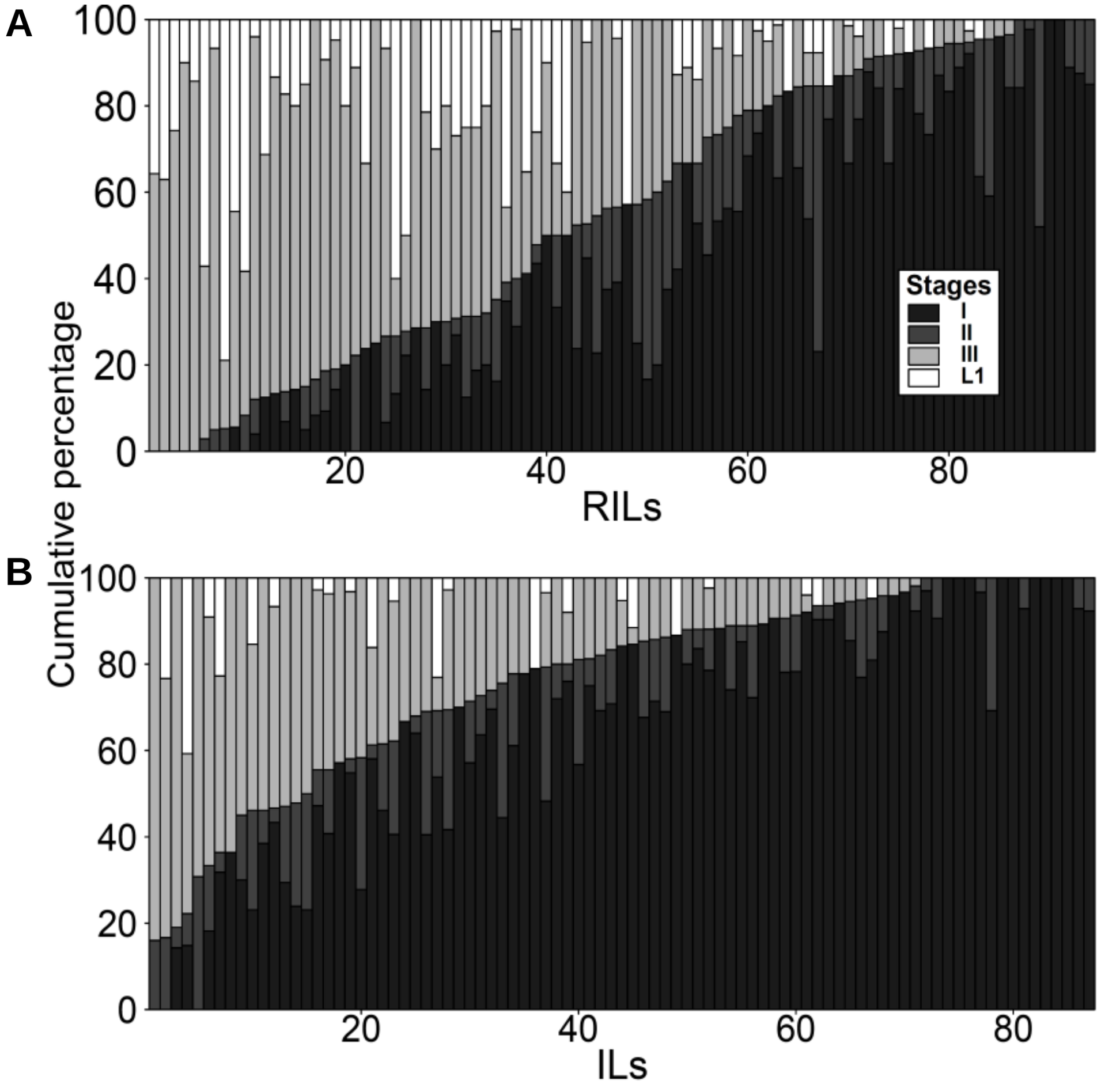
**Figure S1: Embryo stage distribution across the RILs and ILs.** The frequency distribution of eggs laid at a certain stage as a percentage of all progeny per genotype in the RILs (A) in grey and ILs (B) in yellow. The parental phenotypes are indicated by the vertical lines, N2 solid, CB4856 dashed. Progeny stage is indicated above each panel.

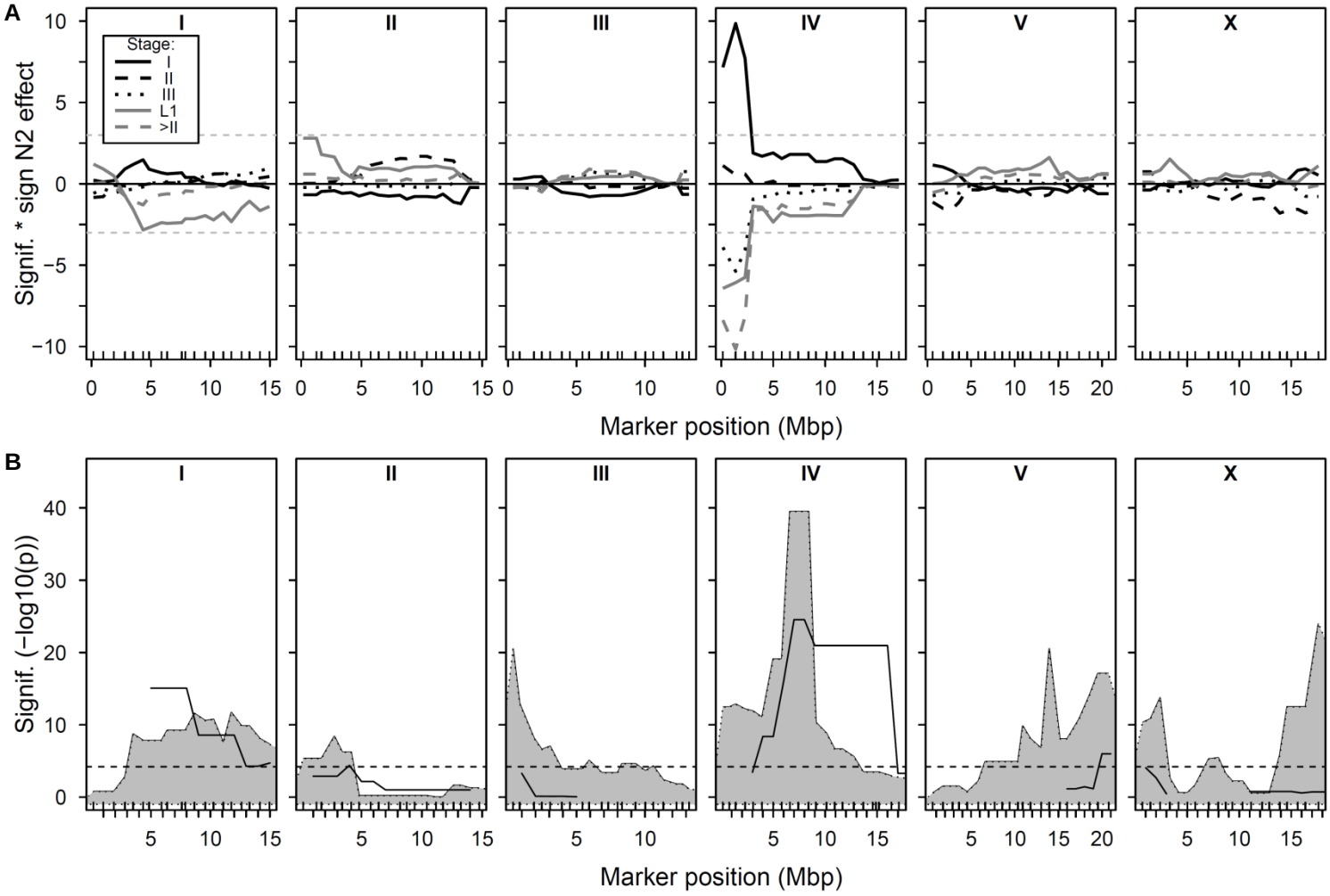
**Figure S2: MQM analysis of the RILs.** Percentage of stage I and II eggs per RIL was used as input for MQM QTL mapping. We used both forward mapping and backward mapping indicated in the figure titles. For forward mapping cofactor markers were added until no significant new QTL were found, for backward mapping 5 equally spaced co-factor marker were used in the starting model and each round the least significant co-factor was removed until no co-factors with $p > 0.1$ remained. For both forward and backward MQM mapping a window-size of 5 markers was used in calculating the final QTL profiles. We also fixed the major QTL at the top of chromosome IV to observe the QTL profiles of a N2 or CB fixed major QTL.

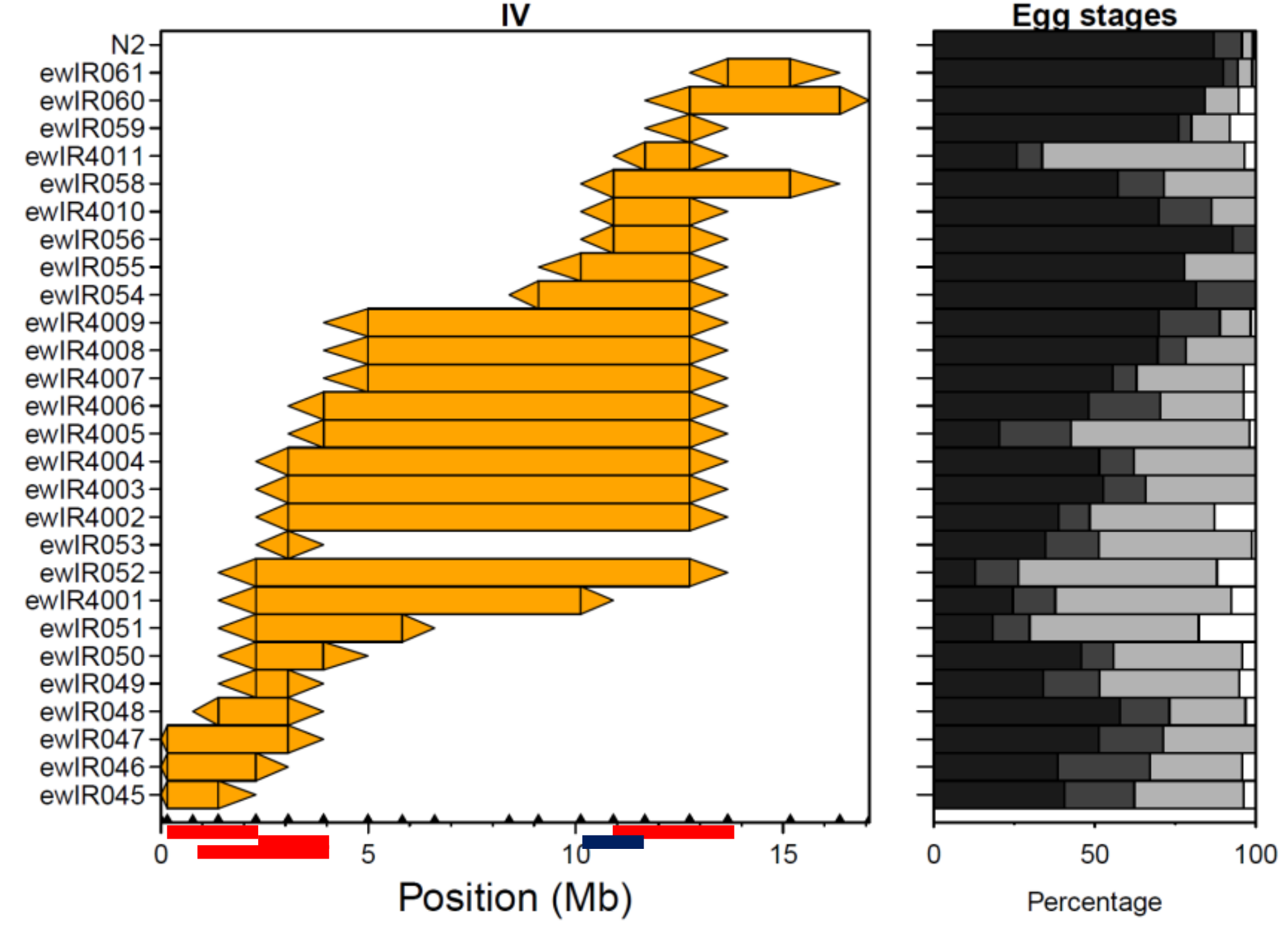
**Figure S3: Comparison between ILs in the genomewide screen.** The CB4856 introgression per IL is shown by the coloured rectangle. Triangles join adjacent CB4856 and N2 markers. Embryo stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white).

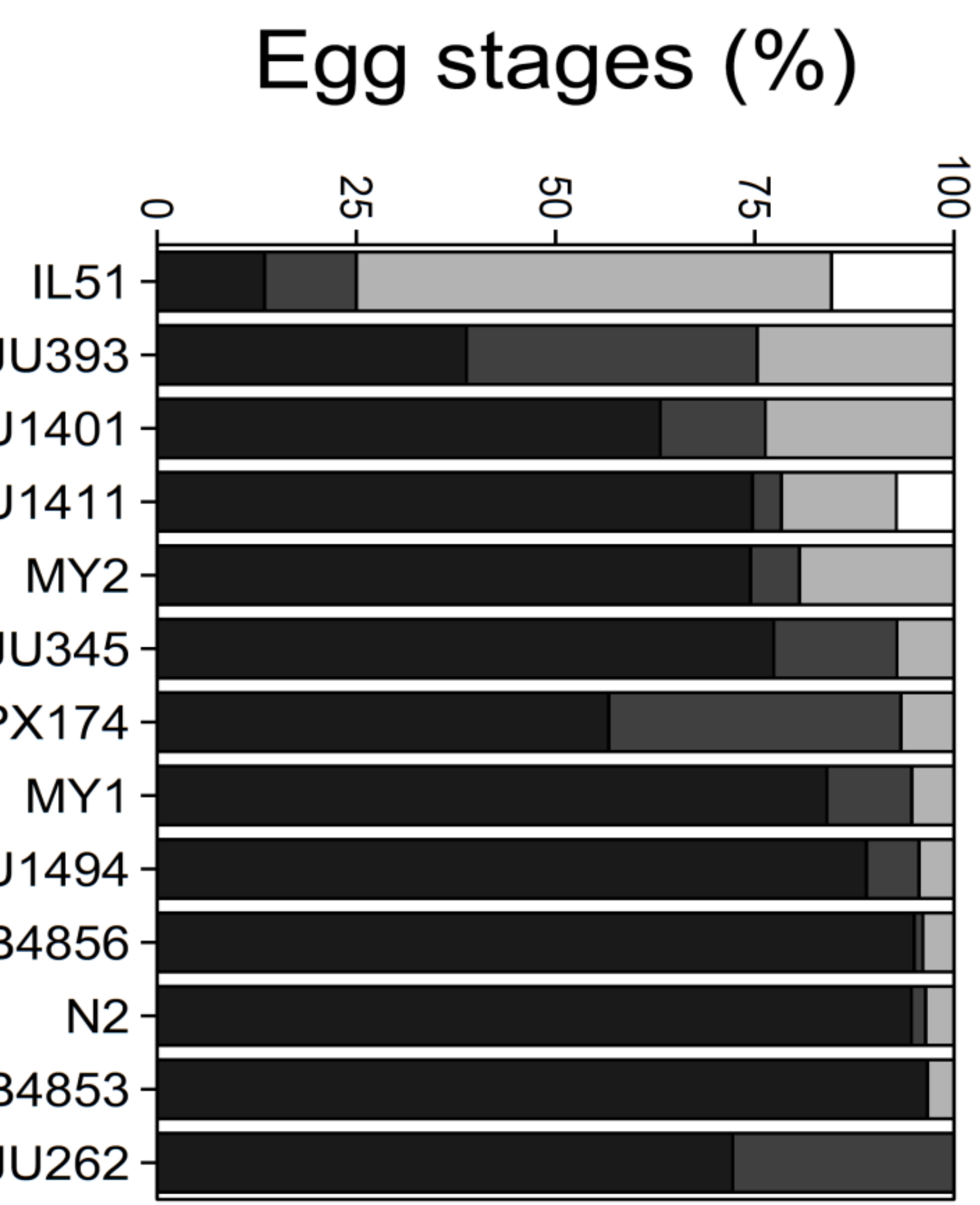
**Figure S4: Egg-stages in single N2 dishes used in this paper.** The number of eggs measured is indicated in red.

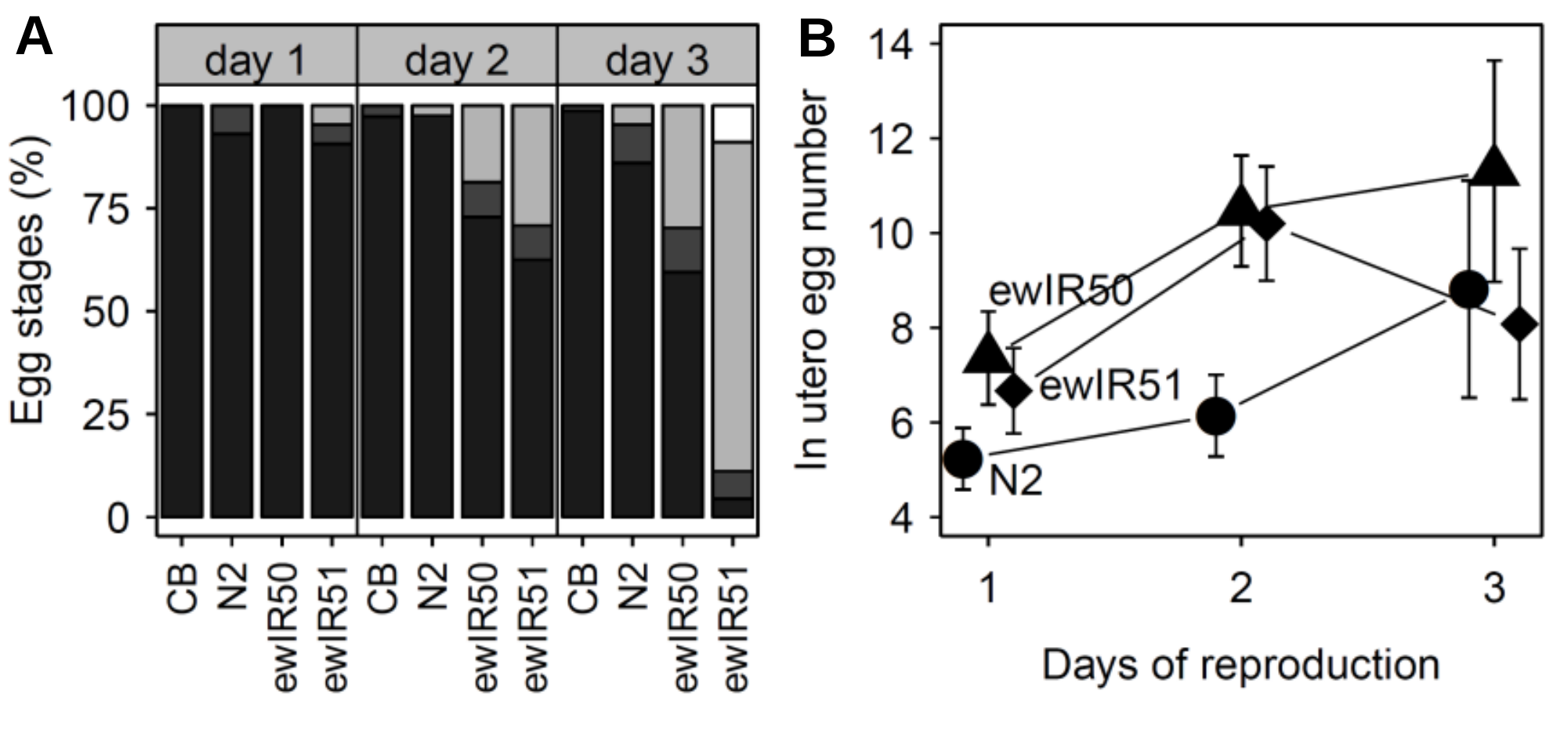
**Table S1: IL vs IL comparison for other traits.** QTL intervals and all IL by IL comparisons.

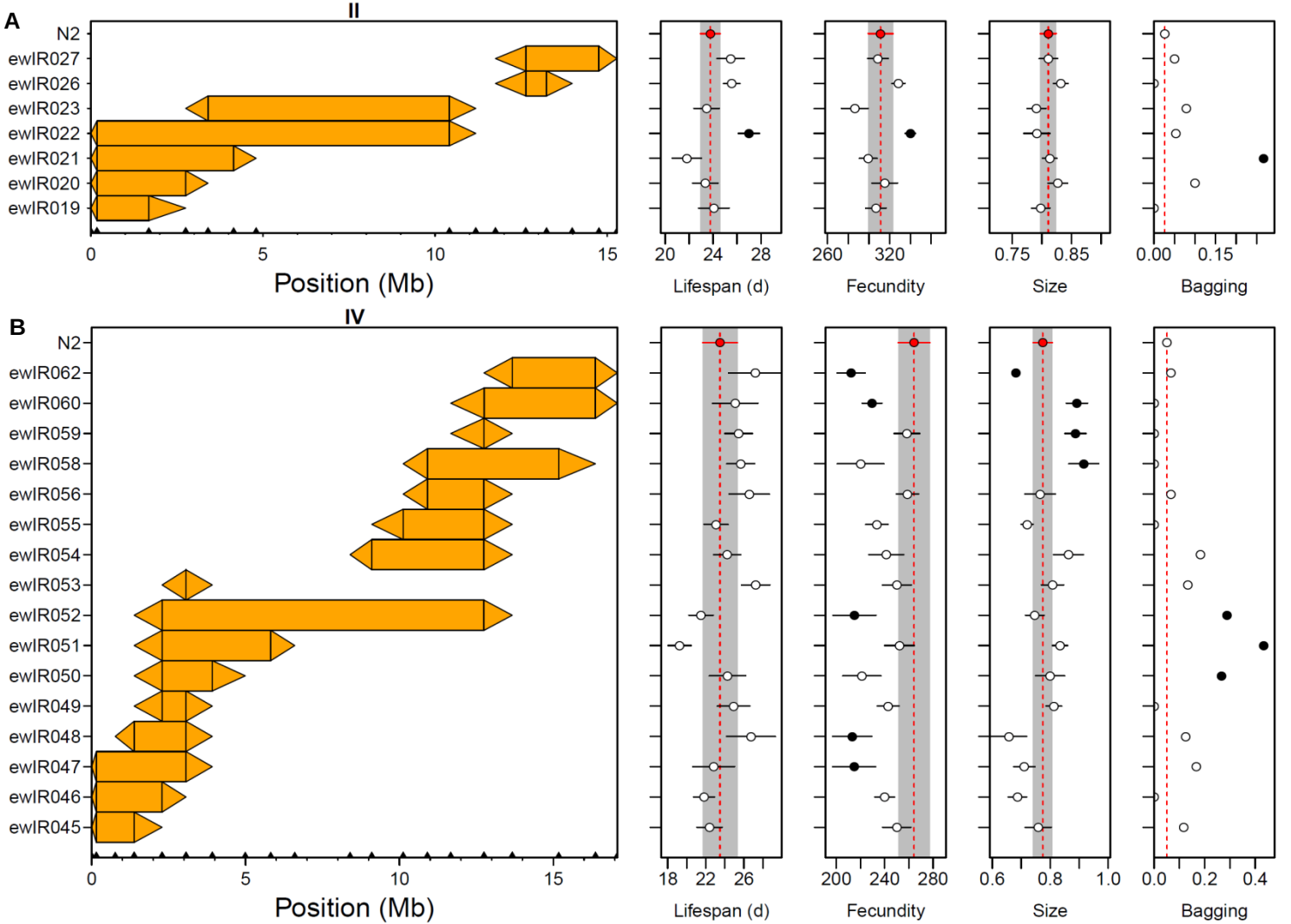

A
100
80
60
40
20
0
Cumulative percentage
20
40
60
80
RILs
Stages
I
II
III
L1
B
100
80
60
40
20
0
Cumulative percentage
20
40
60
80
ILs


A
I
II
III
IV
V
X
Stage:
I
II
III
L1
>II
Signif. * sign N2 effect
10
5
0
−5
−10
0
5
10
15
20
Marker position (Mbp)
B
Signif. (−log10(p))
40
30
20
10
0
Marker position (Mbp)


IV
Egg stages
N2
ewIR061
ewIR060
ewIR059
ewIR4011
ewIR058
ewIR4010
ewIR056
ewIR055
ewIR054
ewIR4009
ewIR4008
ewIR4007
ewIR4006
ewIR4005
ewIR4004
ewIR4003
ewIR4002
ewIR053
ewIR052
ewIR4001
ewIR051
ewIR050
ewIR049
ewIR048
ewIR047
ewIR046
ewIR045
0
5
10
15
Position (Mb)
0
50
100
Percentage


Egg stages (%)
0
25
50
75
100
IL51
JU393
JU1401
JU1411
MY2
JU345
PX174
MY1
JU1494
34856
N2
34853
JU262


A
day 1
day 2
day 3
Egg stages (%)
100
75
50
25
0
CB
N2
ewIR50
ewIR51
B
ln utero egg number
14
12
10
8
6
4
ewIR50
ewIR51
N2
1
2
3
Days of reproduction


A
II
N2
ewIR027
ewIR026
ewIR023
ewIR022
ewIR021
ewIR020
ewIR019
0
5
10
15
Position (Mb)
20
24
28
Lifespan (d)
260
320
Fecundity
0.75
0.85
Size
0.00
0.15
Bagging
B
IV
N2
ewIR062
ewIR060
ewIR059
ewIR058
ewIR056
ewIR055
ewIR054
ewIR053
ewIR052
ewIR051
ewIR050
ewIR049
ewIR048
ewIR047
ewIR046
ewIR045
0
5
10
15
Position (Mb)
18
22
26
Lifespan (d)
200
240
280
Fecundity
0.6
0.8
1.0
Size
0.0
0.2
0.4
Bagging

# Supplementary figures

## Widespread genomic incompatibilities in *Caenorhabditis elegans*

L. Basten Snoek[1,*], Helen E. Orbidans[2,*], Jana J. Stastna[2], Aafke Aartse[1], Miriam Rodriguez[1], Joost A.G. Riksen[1], Jan E. Kammenga[1,3] and Simon C. Harvey[2,4]

[1] Laboratory of Nematology, Wageningen University, 6708 PB Wageningen, The Netherlands.

[2] Biomolecular Research Group, Department of Geographical and Life Sciences, Canterbury Christ Church University, North Holmes Road, Canterbury, CT1 1QU, UK.

[3] jan.kammenga@wur.nl

[4] simon.harvey@canterbury.ac.uk

* These authors contributed equally to this work.

**Page 3** **Figure S1: Progeny stage distribution across the RILs and ILs.** The frequency distribution of progeny laid at a certain stage as a percentage of all progeny per genotype in the RILs (A) in grey and ILs (B) in yellow. The parental phenotypes are indicated by the vertical lines, N2 solid, CB4856 dashed. Progeny stage is indicated above each panel.

**Page 4 – 7** **Figure S2: MQM analysis of the RILs.** Percentage of stage I and II eggs per RIL was used as input for MQM QTL mapping. We used both forward mapping and backward mapping indicated in the figure titles. For forward mapping cofactor markers were added until no significant new QTL were found, for backward mapping 5 equally spaced co-factor marker were used in the starting model and each round the least significant co-factor was removed until no co-factors with $p > 0.1$ remained. For both forward and backward MQM mapping a window-size of 5 markers was used in calculating the final QTL profiles. We also fixed the major QTL at the top of chromosome IV to observe the QTL profiles of a N2 or CB4856 fixed major QTL.

**Page 8 – 14** **Figure S3: Comparison between ILs in the genomewide screen.** The CB4856 introgression per IL is shown by the coloured rectangle. Triangles join adjacent CB4856 and N2 markers. Progeny stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white).

**Page 15** **Figure S4: Egg-stages in single N2 dishes used in this paper.** The number of eggs measured is indicated in red. Progeny stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white).

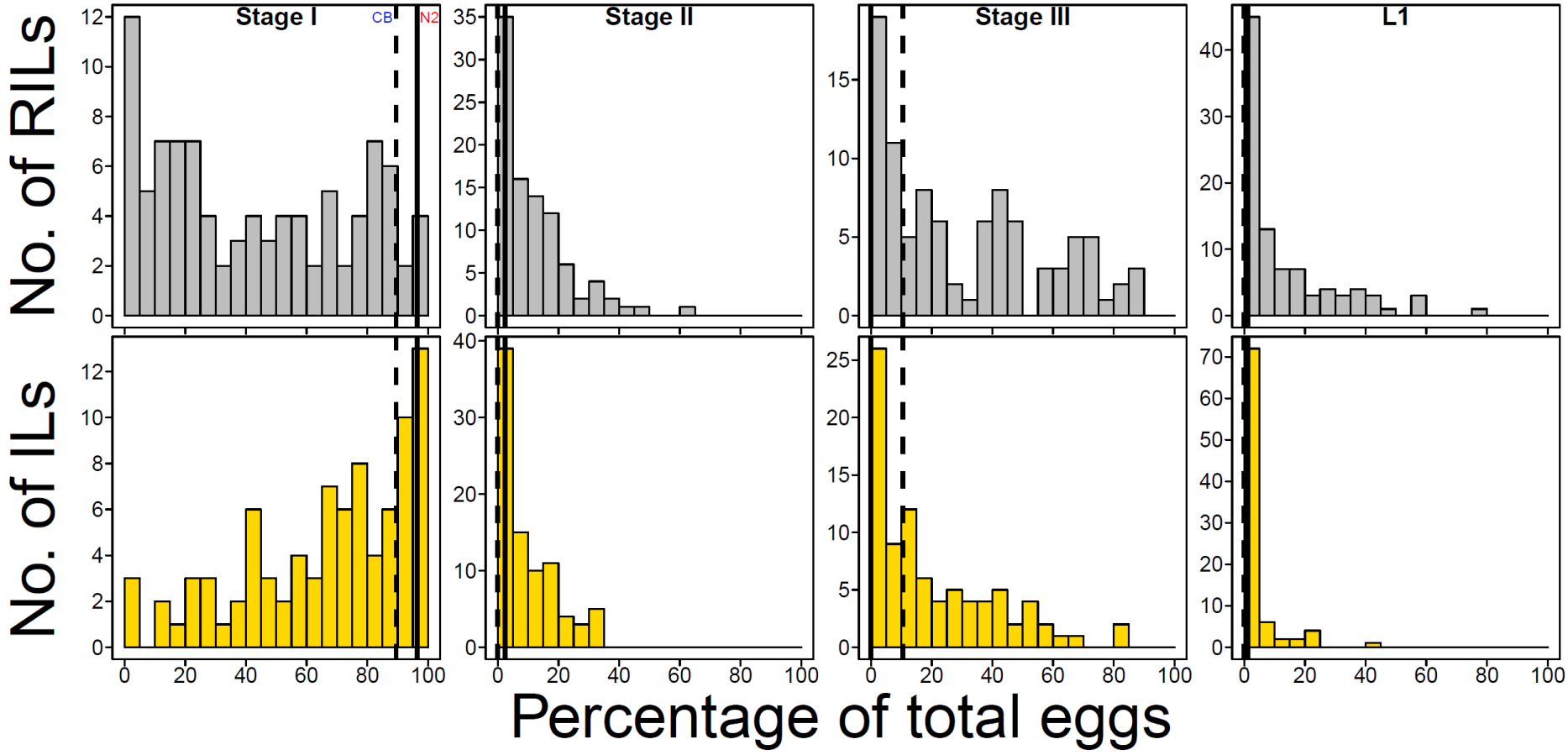
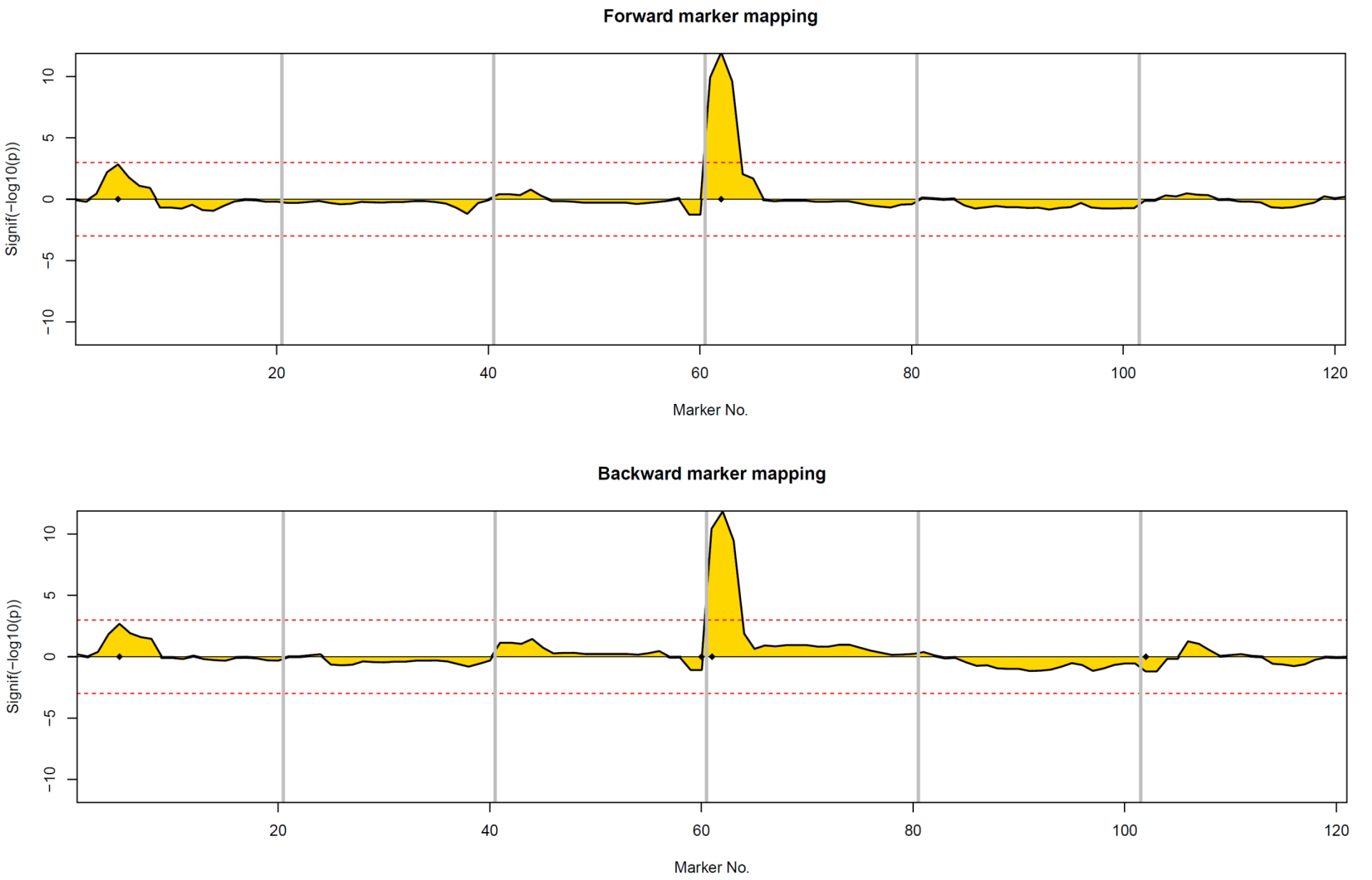
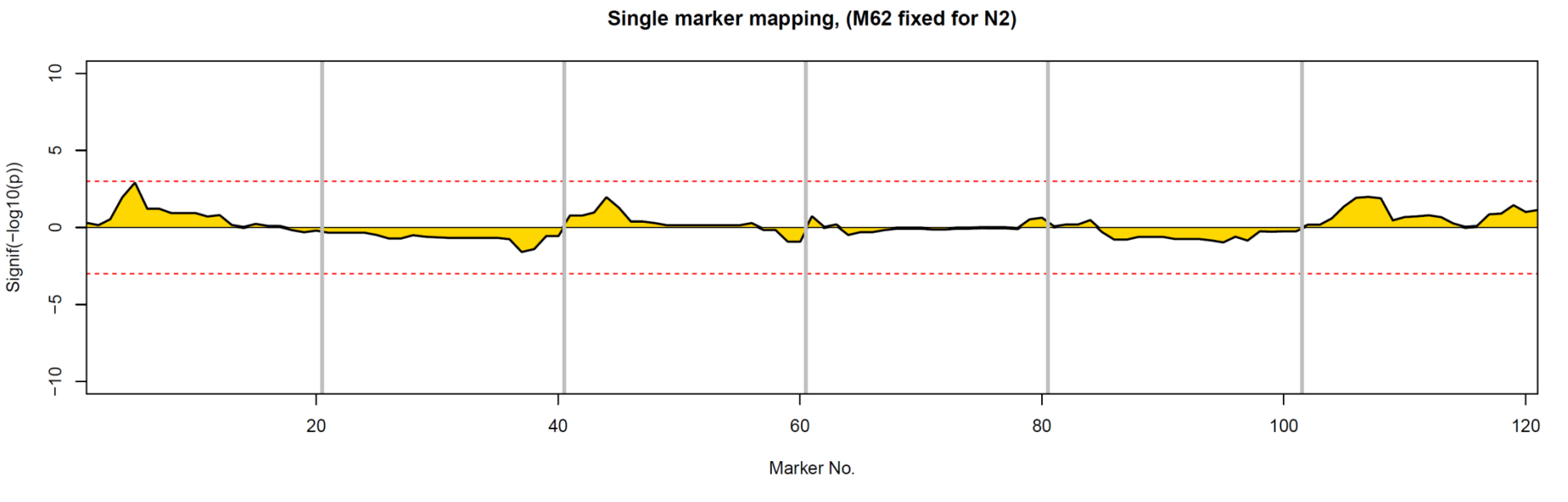
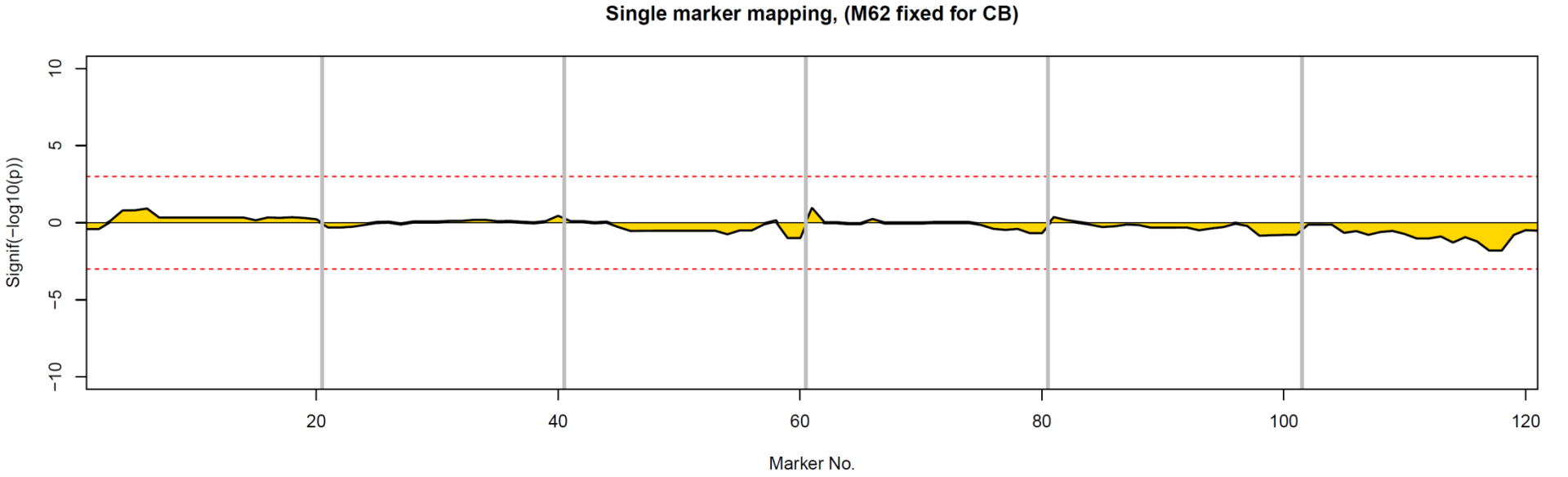


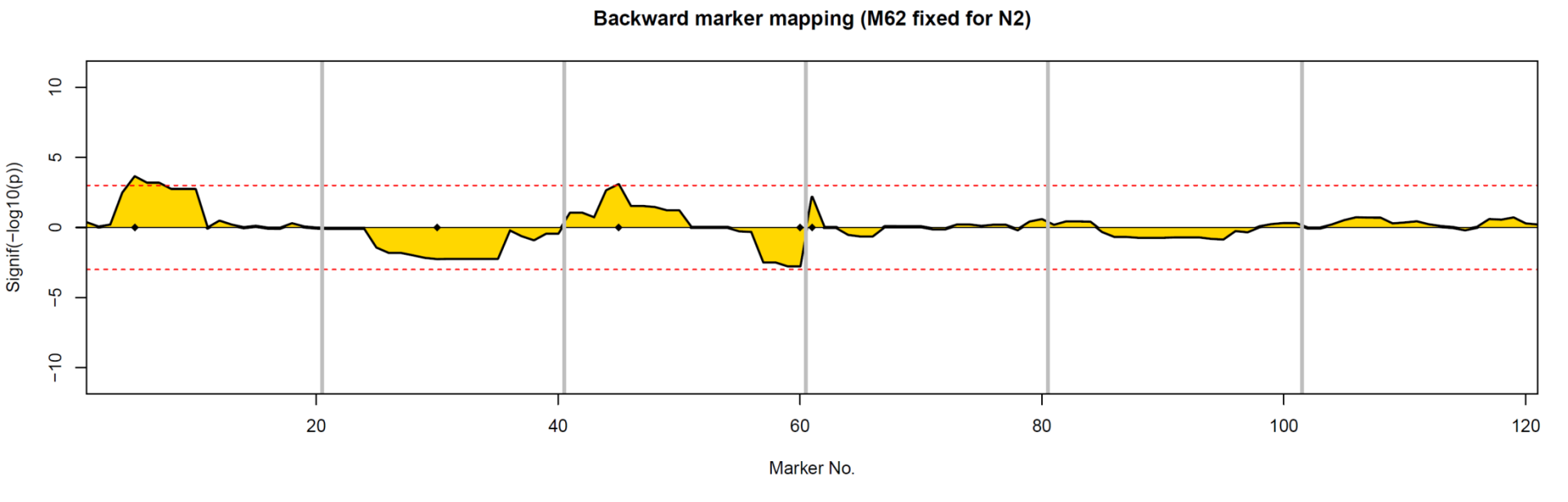

**Figure S1: Progeny stage distribution across the RILs and ILs.** The frequency distribution of progeny laid at a certain stage as a percentage of all progeny per genotype in the RILs (A) in grey and ILs (B) in yellow. The parental phenotypes are indicated by the vertical lines, N2 solid, CB4856 dashed. Progeny stage is indicated above each panel.

**Figure S2: MQM analysis of the RILs.** Percentage of stage I and II eggs per RIL was used as input for MQM QTL mapping. We used both forward mapping and backward mapping indicated in the figure titles. For forward mapping cofactor markers were added until no significant new QTL were found, for backward mapping 5 equally spaced co-factor marker were used in the starting model and each round the least significant co-factor was removed until no co-factors with $p > 0.1$ remained. For both forward and backward MQM mapping a window-size of 5 markers was used in calculating the final QTL profiles. We also fixed the major QTL at the top of chromosome IV to observe the QTL profiles of a N2 or CB fixed major QTL.


Forward marker mapping
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.
Backward marker mapping
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.


Single marker mapping, (M62 fixed for N2)
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.


Single marker mapping, (M62 fixed for CB)
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.


Backward marker mapping (M62 fixed for N2)
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.

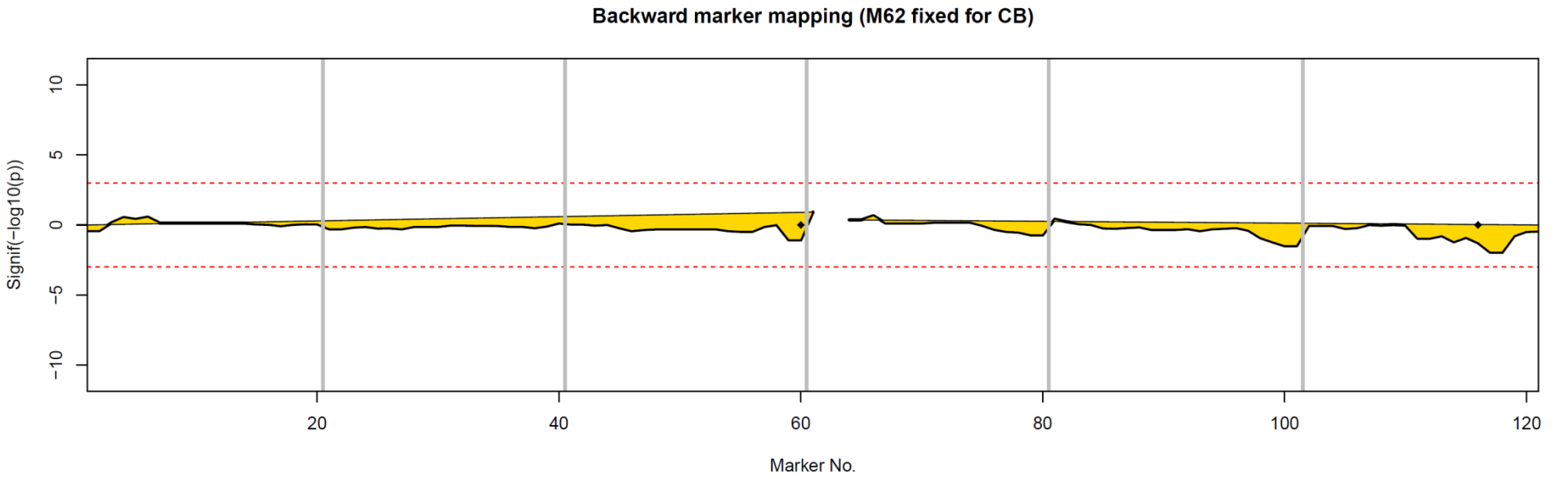
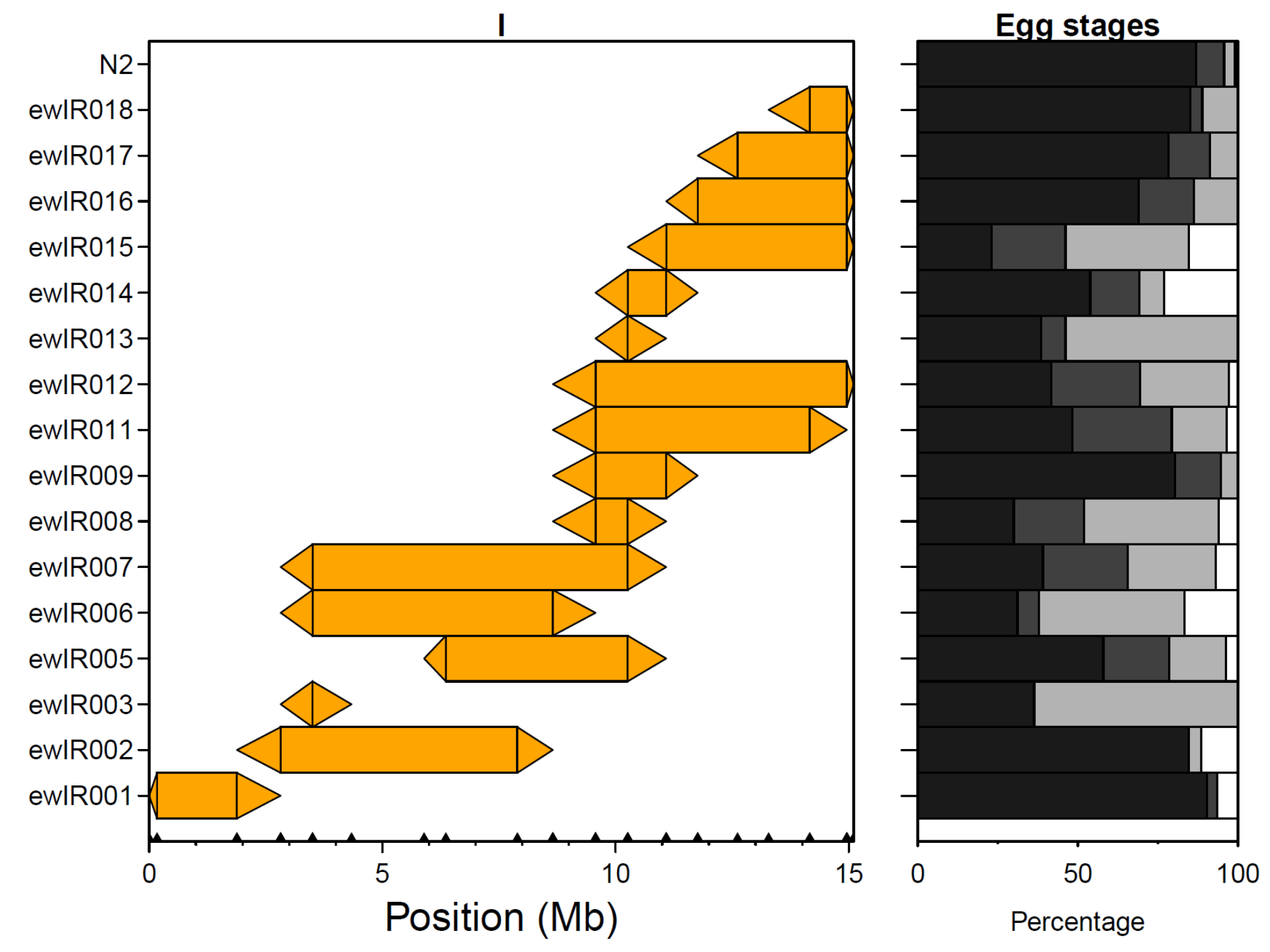
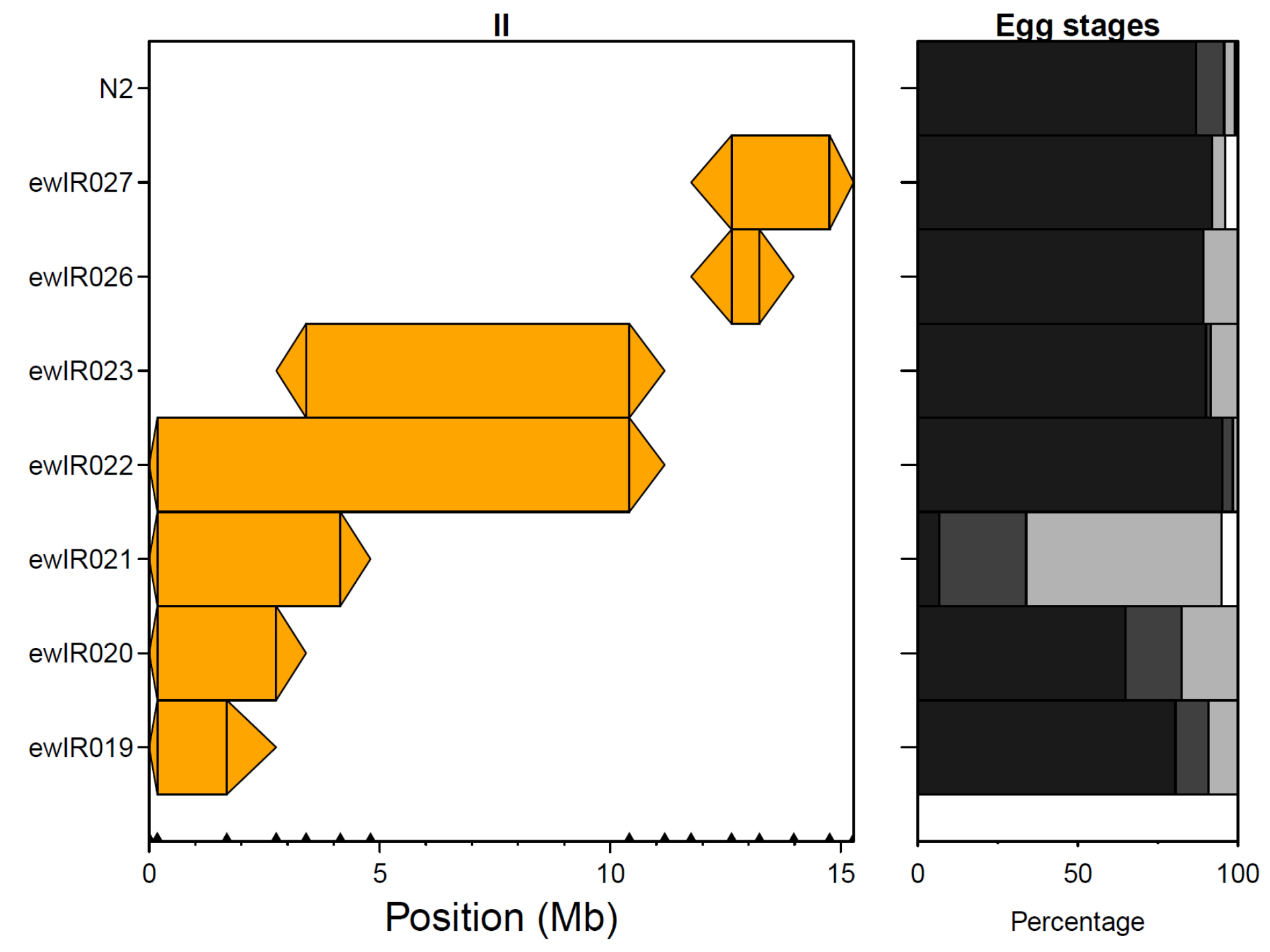
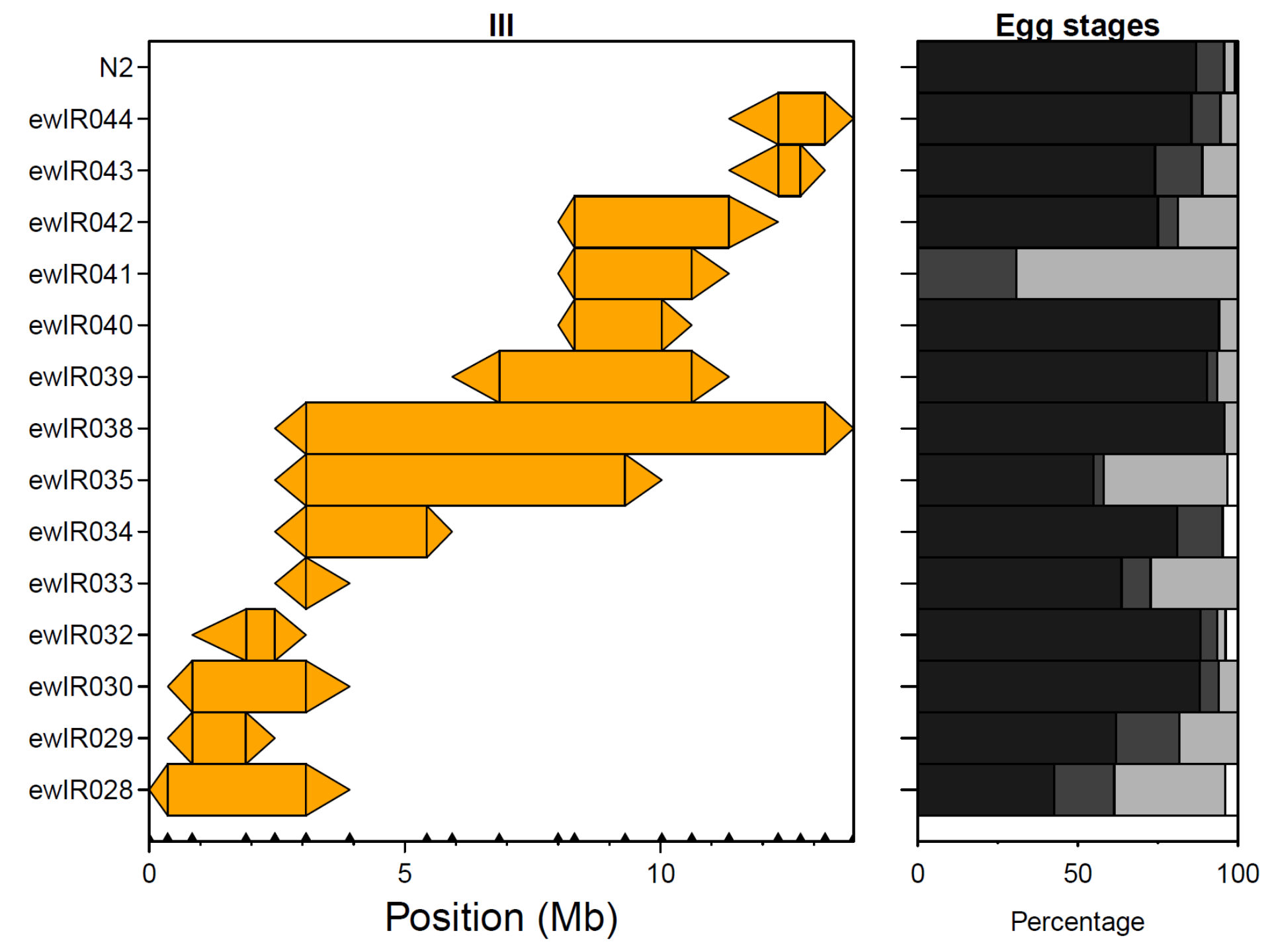
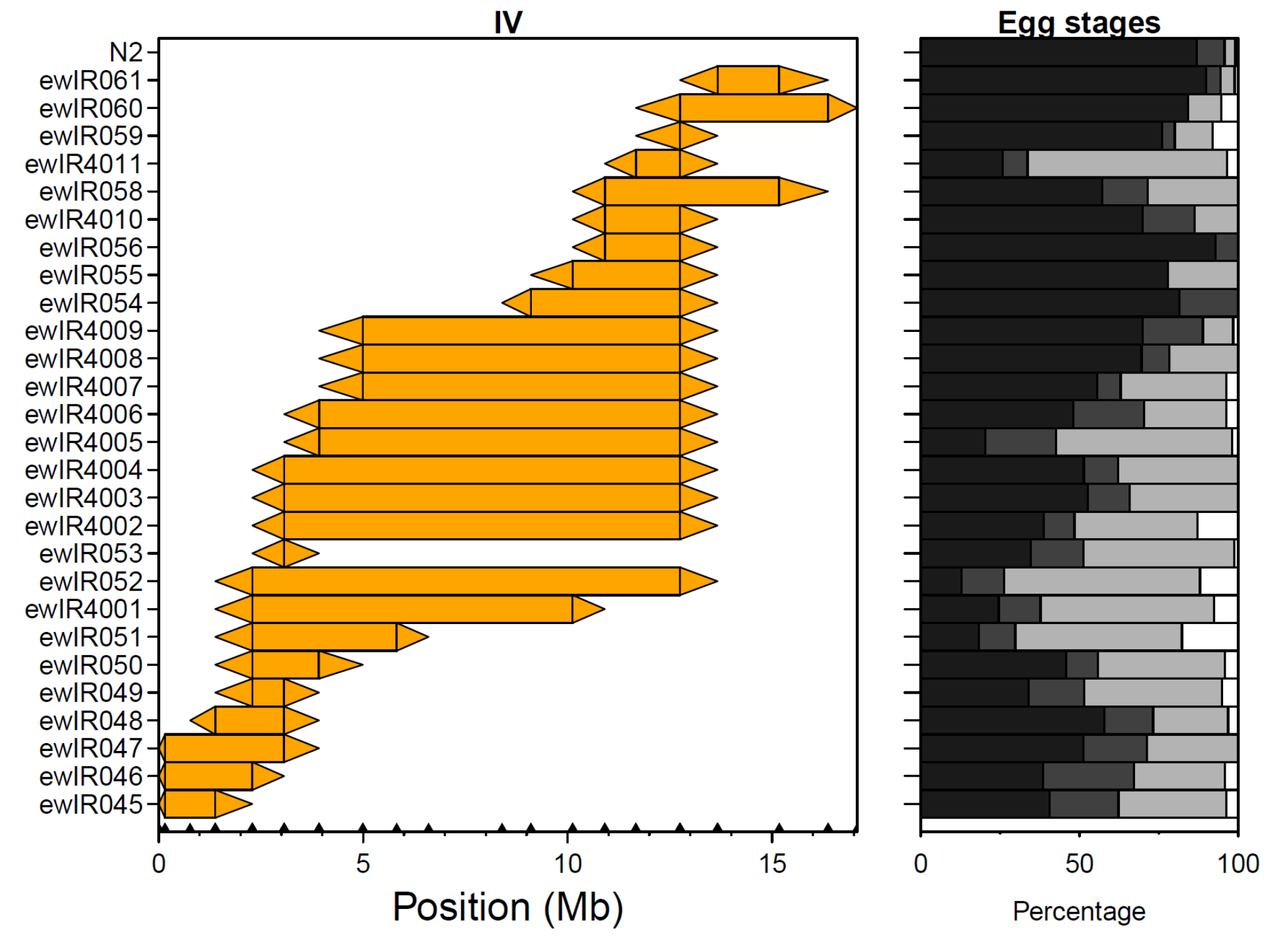

Backward marker mapping (M62 fixed for CB)
Signif(−log10(p))
10
5
0
−5
−10
20
40
60
80
100
120
Marker No.

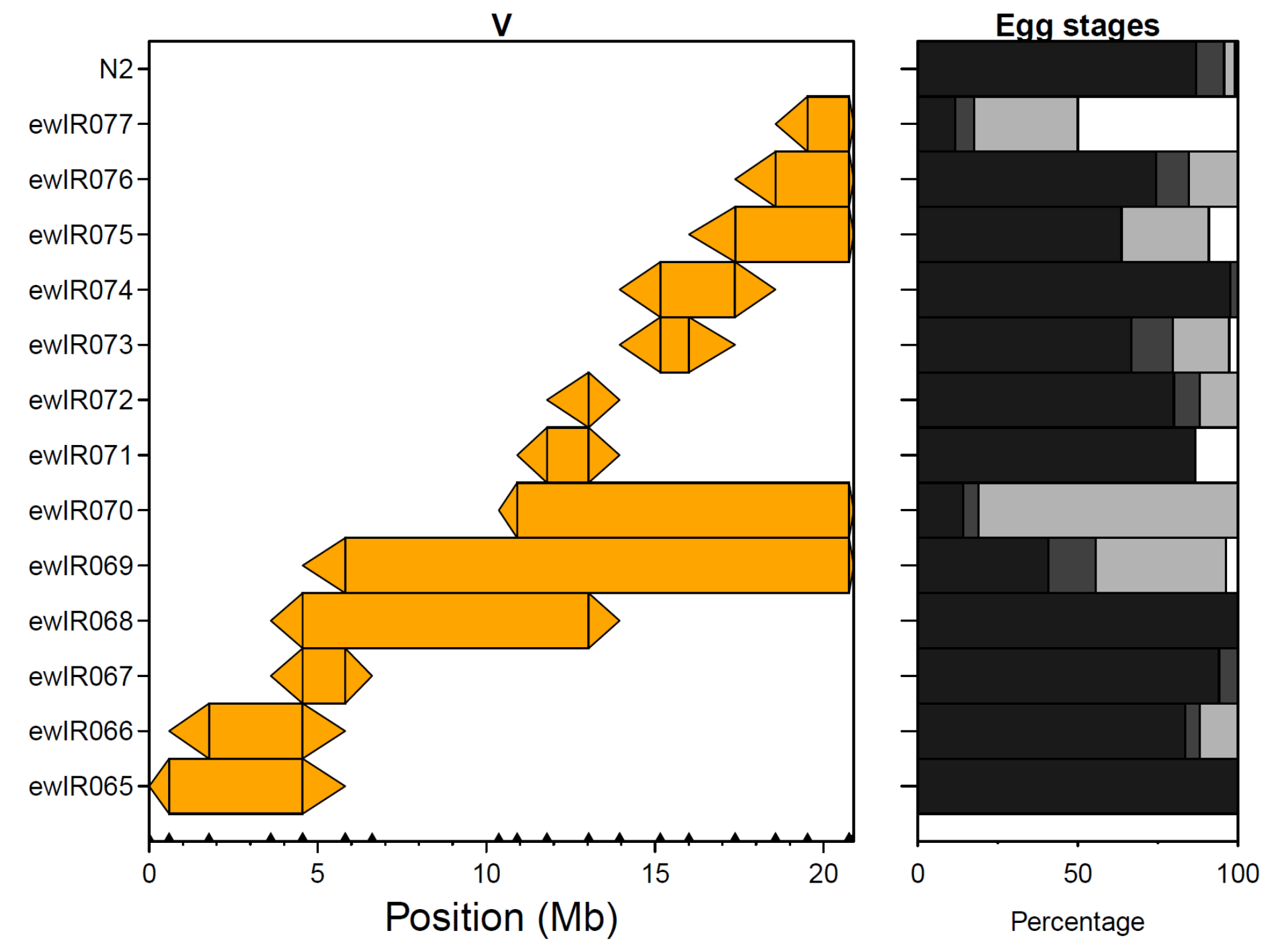
**Figure S3: Comparison between ILs in the genomewide screen.** The CB4856 introgression per IL is shown by the coloured rectangle. Triangles join adjacent CB4856 and N2 markers. Progeny stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white).

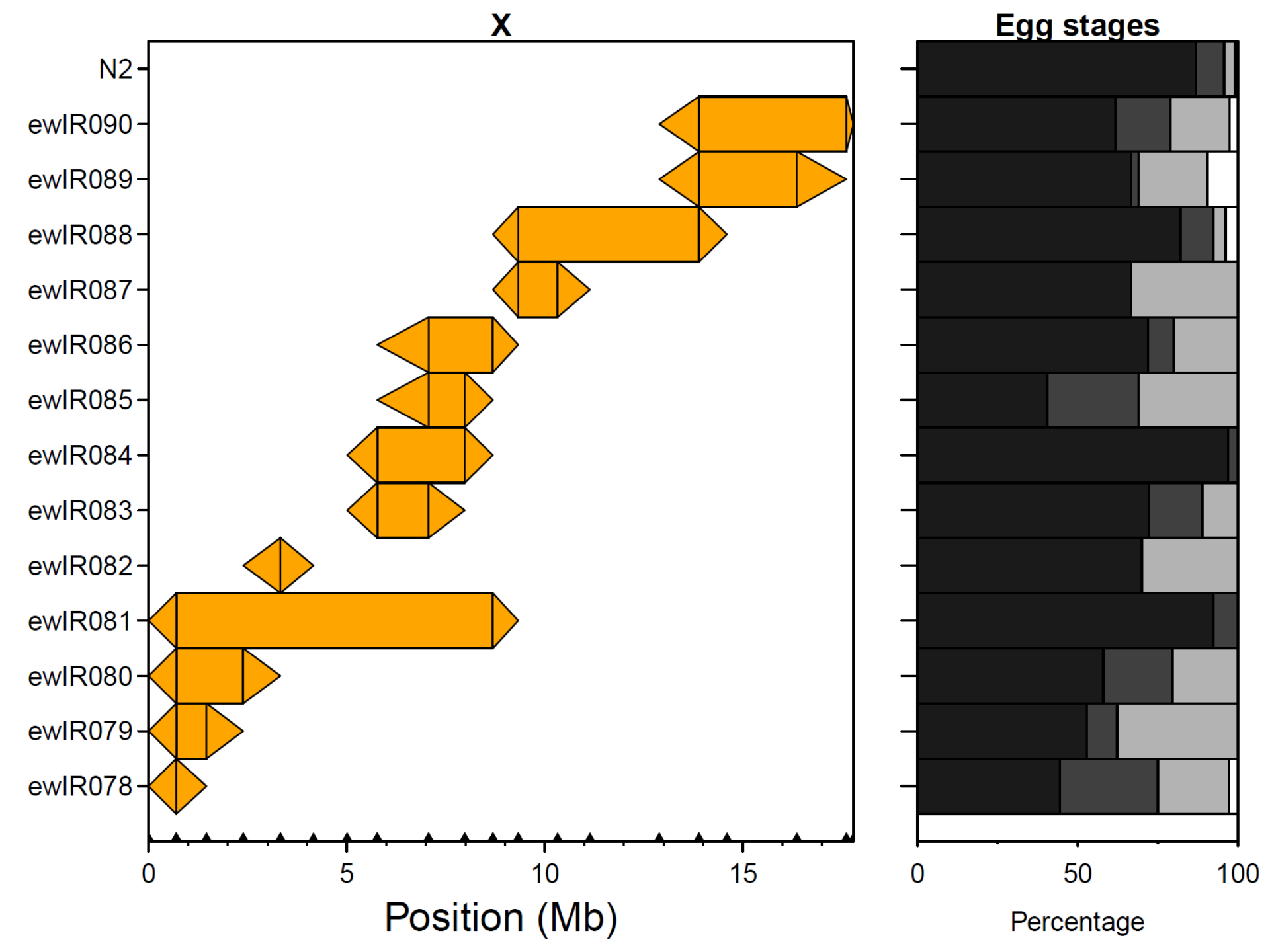

I
Egg stages
N2
ewIR018
ewIR017
ewIR016
ewIR015
ewIR014
ewIR013
ewIR012
ewIR011
ewIR009
ewIR008
ewIR007
ewIR006
ewIR005
ewIR003
ewIR002
ewIR001
0
5
10
15
Position (Mb)
0
50
100
Percentage


II
Egg stages
N2
ewIR027
ewIR026
ewIR023
ewIR022
ewIR021
ewIR020
ewIR019
0
5
10
15
Position (Mb)
0
50
100
Percentage


III
Egg stages
N2
ewIR044
ewIR043
ewIR042
ewIR041
ewIR040
ewIR039
ewIR038
ewIR035
ewIR034
ewIR033
ewIR032
ewIR030
ewIR029
ewIR028
0
5
10
Position (Mb)
0
50
100
Percentage


IV
Egg stages
N2
ewIR061
ewIR060
ewIR059
ewIR4011
ewIR058
ewIR4010
ewIR056
ewIR055
ewIR054
ewIR4009
ewIR4008
ewIR4007
ewIR4006
ewIR4005
ewIR4004
ewIR4003
ewIR4002
ewIR053
ewIR052
ewIR4001
ewIR051
ewIR050
ewIR049
ewIR048
ewIR047
ewIR046
ewIR045
0
5
10
15
Position (Mb)
0
50
100
Percentage


V
Egg stages
N2
ewIR077
ewIR076
ewIR075
ewIR074
ewIR073
ewIR072
ewIR071
ewIR070
ewIR069
ewIR068
ewIR067
ewIR066
ewIR065
0
5
10
15
20
Position (Mb)
0
50
100
Percentage


X
Egg stages
N2
ewIR090
ewIR089
ewIR088
ewIR087
ewIR086
ewIR085
ewIR084
ewIR083
ewIR082
ewIR081
ewIR080
ewIR079
ewIR078
0
5
10
15
Position (Mb)
0
50
100
Percentage

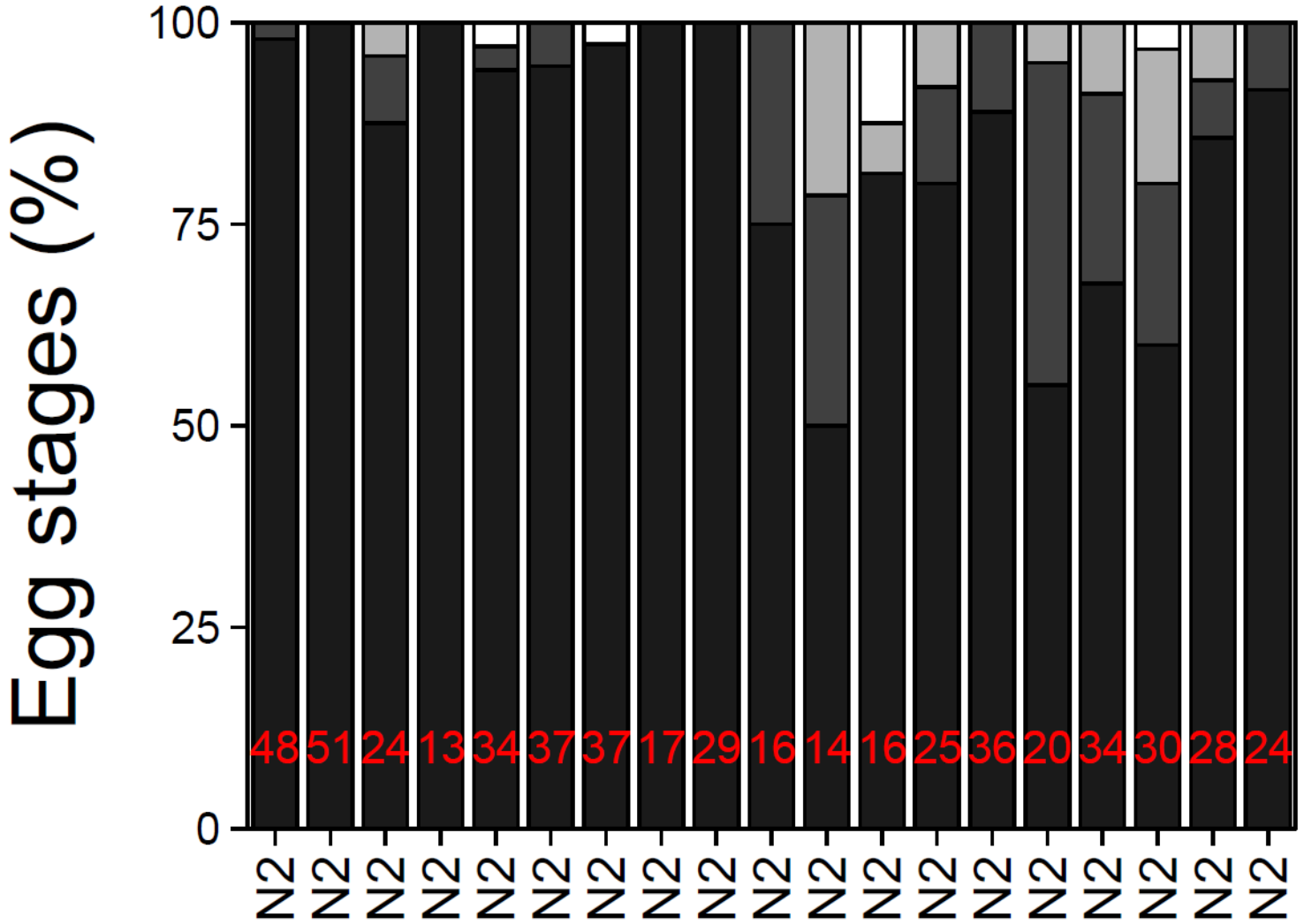



**Figure S4: Egg-stages in single N2 dishes used in this paper.** The number of eggs measured is indicated in red. Progeny stage distribution is shown as cumulative percentage of total progeny. From dark to light: Stage I, II, III and L1 (in white).