

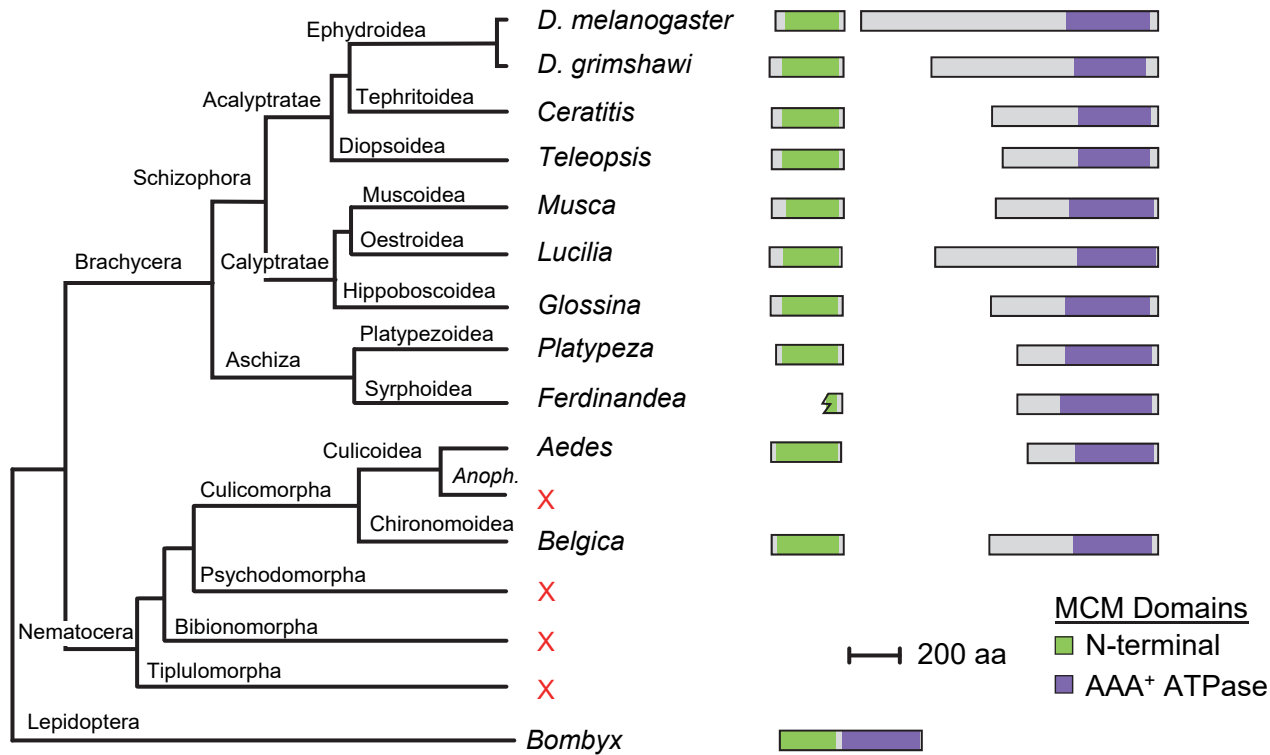
Figure S1. Occurrence of Msh4, Msh5, MCM8, MCM9, MEI-217, and MEI-218 in Diptera. The dendrogram on the left illustrates relationships among Dipteran taxa for which sufficient genome or transcriptome sequence is available to determine with reasonable confidence the presence or absence of genes encoding proteins relevant to this work. Circles to the right indicate presence (filled) or absence (open) of each gene/protein. For the suborder Brachycera, major superfamilies within Schizophora and the sister taxon Aschiza are shown, as well as the superfamily Asiloidea. For the suborder Nematocera, only infraorders are shown, except for Culicomorpha, where both superfamilies are indicated. Within the superfamily Culicoidea (mosquitoes), MEI-217 and MEI-218 are found in *Culex* and *Aedes* but are missing from all of the 20 *Anopheles* species whose genomes are sequenced.

It is hypothesized that the mei-MCM complex functionally replaces Msh 4/5 in *Drosophila* (Kohl, Jones, and Sekelsky 2012). We do not find orthologs of Msh4, Msh5, or Msh9 in species in the Dipteran sub-order Brachycera, suggesting that the structure and function of the *Drosophila* mei-MCM complex may have its origins in the ancestral founder of this lineage. Interestingly, Asiloidea appear to have retained an ortholog of MCM9. It may be informative to examine these species more thoroughly when additional sequences become available.

Figure S2

Hartmann *et al.*

A



B

Sequence alignment of MCM domains across various insect species. The alignment shows the N-terminal (green) and AAA+ ATPase (purple) regions. The sequences are as follows:

D. melanogaster: Q L E A C H L R I N *
CAACTGGAAGCATGTCATCTACGAATAAACTAGAGAAAAAG
M S S T N K L E K K

Ceratitis capitata: I F Q A W S V M T H *
ATATTTCAAGCATGGAGCGTAATGACACATTAGAAACCACG
M E R N D T L E T T

Musca domestica: E F Q A W N L T T I *
GAGTTTCAAGCATGGAACCTAACAACAATATAAGTACAGAT
M E P N N N I S T D

Platypeza anthrax: E F E S W H L T R T Q V *
GAATTTGAATCATGGCATTGACAAGGACTCAAGTTTAGAT
M A F D K D S S L D

Aedes aegypti: W A F E E L *
TGGGCTTCGAGGAATTGtaggatgatgATGGAGTCTACGCTA
M E S T L

Belgica antarctica: A T V *
GCCACCGTTtaattcattgccagtcactgtccatcATGTCGTTTT
M S F

Bombyx mori: K S I I I W S L E K I E Q C K A P M T H
AAATCCATTATAATATGGAGTTTGGAAAAGATTGAACAATGTAAAGCACCAATGACCCAT

Figure S2

Hartmann *et al.*

C

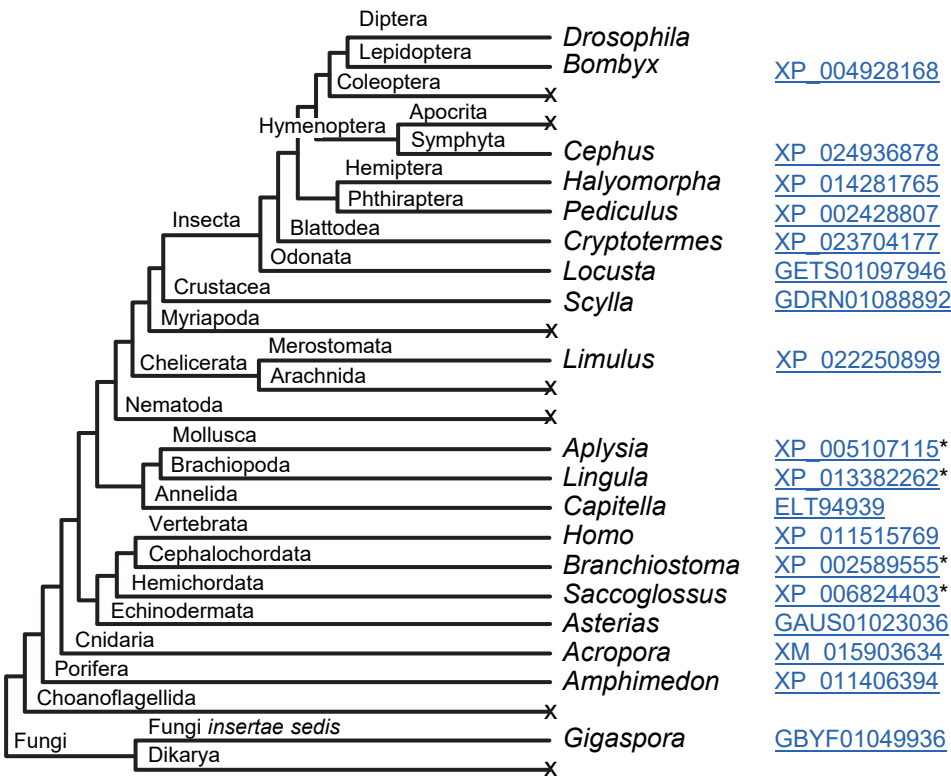


Figure S2. Structures of MEI-217 and MEI-218 in Diptera. (A) The dendrogram is the same as in Figure S1, with additional species to illustrate the variation in domain architectures. Domain architectures for representative species are to the right (the jagged N-terminal end in *Ferdiandea cuprea* indicates incomplete sequence). Domains were determined by PHYRE2 alignment to Protein Data Bank entry c5udb7 (a cryo-electron microscopy structure of *S. cerevisiae* MCM7). Accession numbers for the sequences included are listed below. Accession numbers that start with J are from Ensemble Metazoa genomic assemblies (found at <http://metazoan.ensemble.org>). The *Aedes*, *Musca*, and *Glossina* sequences are genomic contigs from Vectorbase (<http://vectorbase.org>). All other sequences are from NCBI (<http://ncbi.nih.nlm.gov>); those starting with a G are from the transcriptome shotgun assembly (TSA) database.

Species	Accession
<i>Drosophila melanogaster</i>	NM_167557.3
<i>Drosophila grimshawi</i>	XM_001992187.1
<i>Ceratitis capitata</i>	GAMC01014250.1
<i>Teleopsis whitei</i>	GBBQ01026862.1
<i>Musca domestica</i>	scf7180000644883
<i>Lucilia cuprina</i>	JRES01000755:1975-11359
<i>Glossina morsitans</i>	scf7180000644883
<i>Platypeza anthrax</i>	GCGU01008763.1
<i>Ferdiandea cuprea</i>	GCHQ01011487.1
<i>Aedes aegypti</i>	AAGE02016621.1
<i>Belgica Antarctica</i>	JPYR01000187:32247-35225
<i>Bombyx mori</i>	AK381112.1

(B) Junctions between open reading frames (ORFs) for the N-terminal and AAA+ ATPase domains are shown. At the top are three species from Shizophora and one from Aschiza, showing overlapping ORFs. Amino acids at the end of the N-terminal domain are shown in green above the DNA sequence (the position of the stop codon is highlighted in green); amino acids for the beginning of the AAA+ ATPase domain are in purple below the DNA sequence (the position of the start codon is highlighted in purple). Below that are two Nematocera species, showing separate but non-overlapping ORFs. Non-coding sequence between the ORFs is in lowercase text. At the bottom is a non-Diptera representative, the moth *Bombyx mori*. In this case, as in all other non-Dipteran species with an Mcmdc2 ortholog and in replicative MCM proteins, the two domains are on the same polypeptide, separated by a short linker.

(C) Distribution of Mcmdc2 in Opisthokonta. Dendrogram shows phyla, some sub-phyla, and several orders within the sub-phylum Insecta in which we can find clear orthologs of Mcmdc2 or in which there are sufficient genome or transcriptome sequences to suggest loss of Mcmdc2 with reasonable confidence. For those clades with an ortholog, a representative genus is listed, along with an accession number. Taxa in which we could not find any orthologs are indicated with an x. We have not found Mcmdc2 orthologs outside of Opisthokonta.

Figure S3

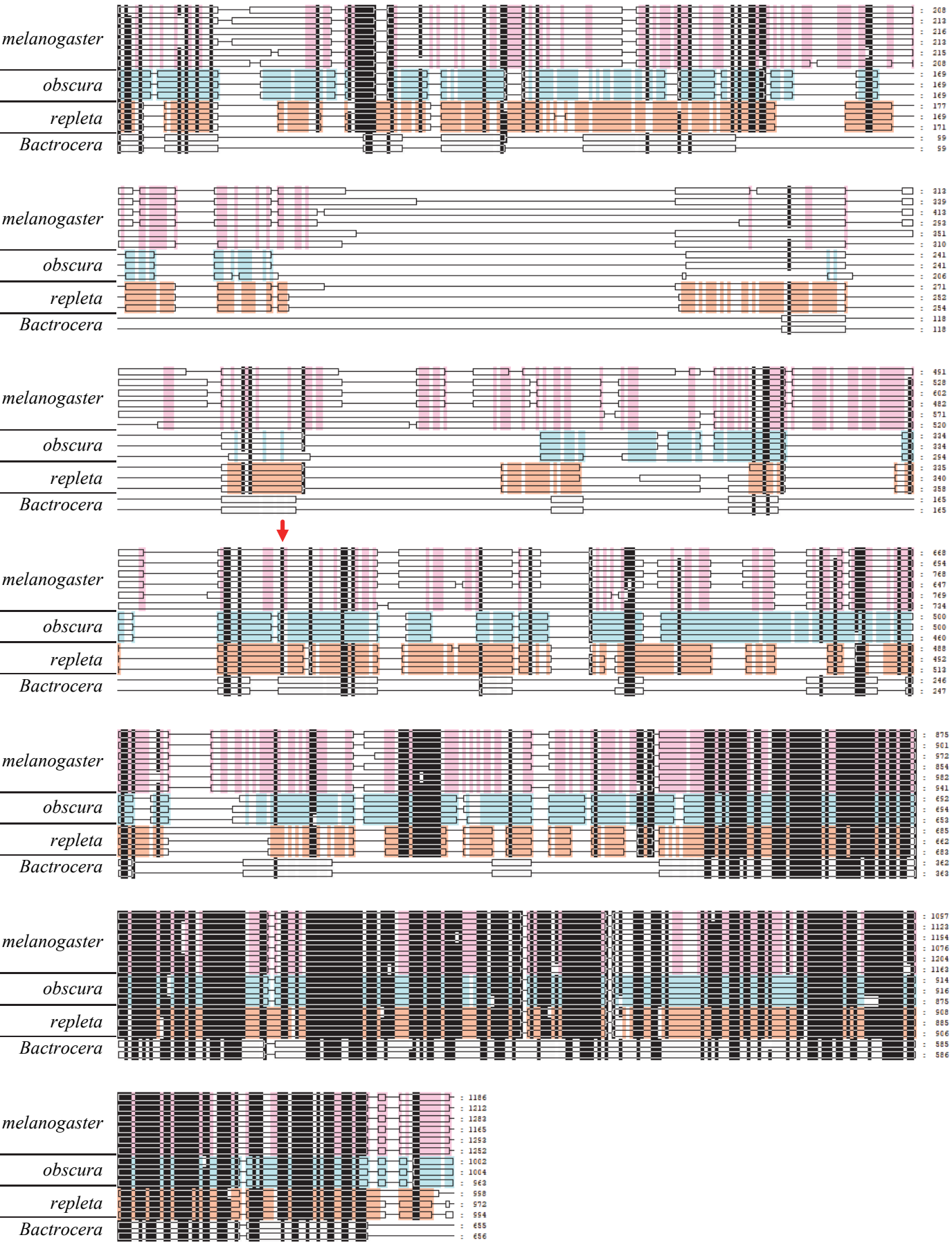


Figure S3. Sequence conservation and divergence in MEI-218. MEI-218 orthologs from 12 *Drosophila* species and two *Bactrocera* species (Tephritid fruit flies, also in the Acalyptratae subsection of Schizophora) are depicted. The red arrow indicates the start of the N-terminally-truncated MEI-218 analyzed in this work.

Species are, from top to bottom and sorted by species group as indicated on the figure:

<i>melanogaster:</i>	<i>Drosophila melanogaster</i>
	<i>Drosophila sechelia</i>
	<i>Drosophila simulans</i>
	<i>Drosophila mauritiana</i>
	<i>Drosophila yakuba</i>
	<i>Drosophila erecta</i>
<i>obscura:</i>	<i>Drosophila pseudoobscura pseudoobscura</i>
	<i>Drosophila persimilis</i>
	<i>Drosophila miranda</i>
<i>repleta:</i>	<i>Drosophila navojoa</i>
	<i>Drosophila mojavensis</i>
	<i>Drosophila arizonae</i>
<i>Bactrocera:</i>	<i>Bactrocera latifrons</i>
	<i>Bactrocera dorsalis</i>

Thin horizontal lines denote gaps introduced in the alignment process. Vertical lines indicate amino acid identity or similarity, using the Dayhoff PAM 200 matrix. Black is conserved among at least 11 sequences (e.g., one mismatch in the 12 *Drosophila* species). Pink indicates conservation within the *melanogaster* species group, aqua within the *obscura* group, and orange within the *repleta* group. The red arrow denotes the start codon for the truncated MEI-218 described in the text.

Sequences were aligned in MEGA (v. 10.0.4) using the MUSCLE algorithm. Manual adjustments were done in GeneDoc v. 2.7.000. This visualization is the summary view produced by GeneDoc, with species groups with conservation mode shading enabled.

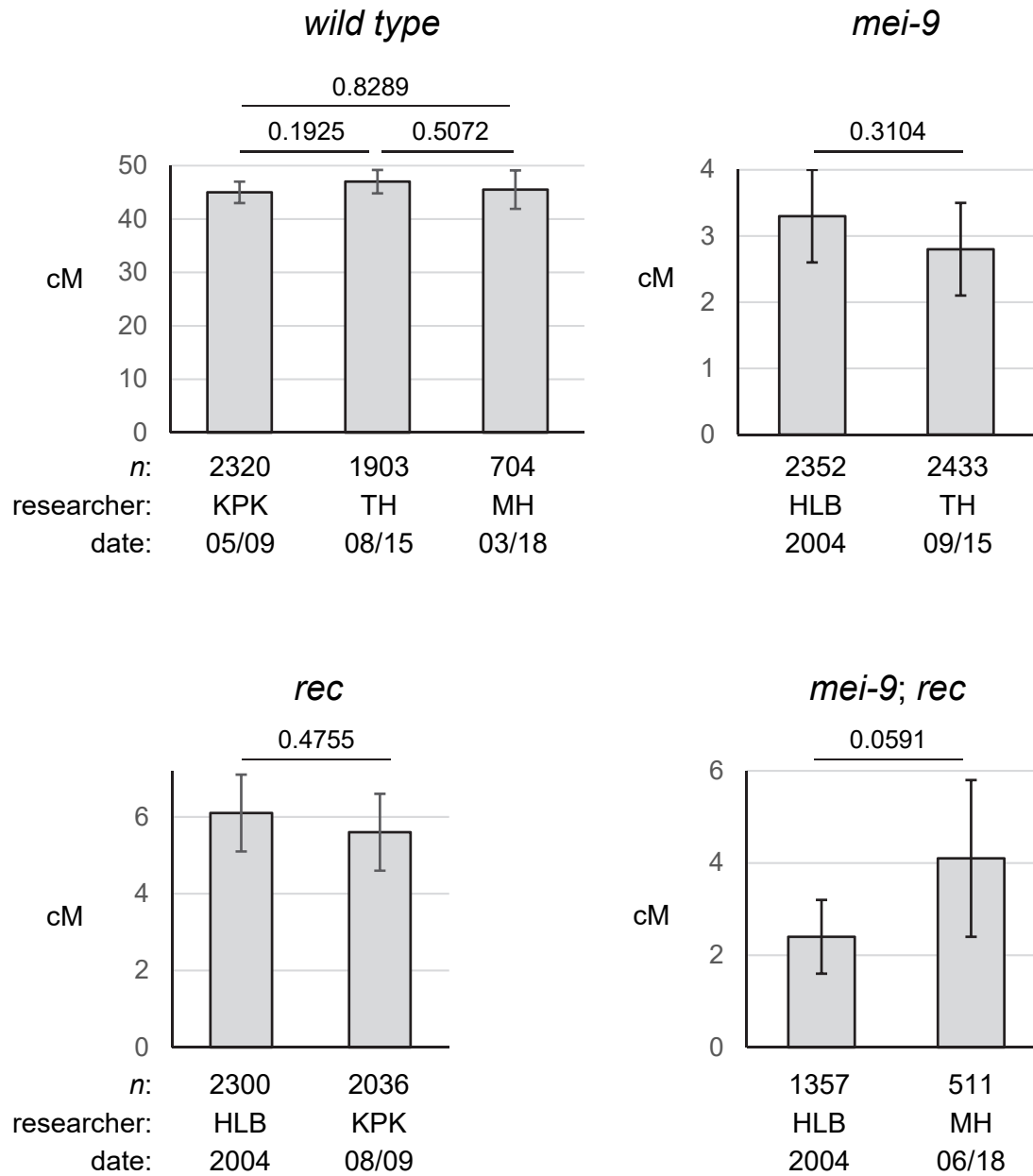


Figure S4. Reproducibility of recombination assays. Graphs show the total number of cM for the *net* – *cn* intervals (*al* – *cn* for HLB assays; *al* is 0.2 cM to the right of *net*) for recombination assays in the same genotype scored by different researchers at different times. Error bars are 95% confidence intervals. *P* values are from Fisher's exact test on 2x2 contingency tables with total crossovers and parental flies for each genotype. Below each bar is the total number of flies scored (*n*), the researcher doing the scoring, and the date of the experiment. MH, KPK, and TH are authors of this work, with some data published in Kohl *et al.* 2012 and Hatkevich *et al.* 2017. HLB refers to data published in Blanton *et al.* 2005. The *mei-9; rec* MH experiment had six double crossovers and one triple crossover. Because all of these were *b pr⁺ cn* and no multiple crossover events were observed in the 2.7x larger HLB dataset, we suspect these were from some non-meiotic event(s); these flies were removed from this analysis.

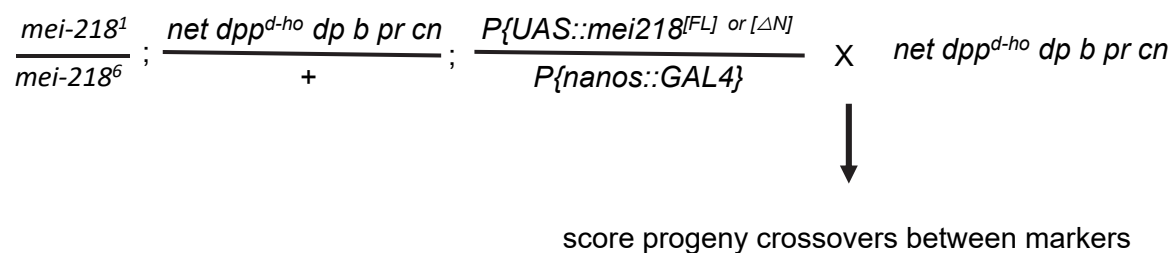


Figure S5. Cross scheme of *mei-218* overexpression. *mei-218^{FL}* or *mei-218^{ΔN}* are expressed from *UAS::Gal4* system on the third chromosome in the background of a *mei-218* null mutation. Recessive markers on chromosome 2 are used for crossover scoring. Virgin female progeny are crossed back to flies homozygous for the recessive marker chromosome.