Supporting Information for Furman et al., Experimental exchange of paralogous domains in the MLH family provides evidence of sub-functionalization after gene duplication.

Figure S1 Models for Mlh1-Pms1 functioning in MMR and Mlh1-Mlh3 in meiotic crossing over. A. In MMR in baker's yeast, MSH complexes (Msh2-Msh3, Msh2-Msh6) bind to DNA mismatches and interact primarily with Mlh1-Pms1 in steps that result in the incision and resynthesis of the newly replicated DNA strand. Mlh1-Mlh3 plays a minor role in MMR and is recruited by Msh2-Msh3 to repair insertion/deletion mismatches. B. Cartoon of the *lys2-A*$_{14}$ reversion assay used to measure frameshift mutations (primarily -1) that result in a reversion to a Lys$^+$ phenotype. Shown are the reversion rates for *pms1*Δ (Table 2) and *mlh3*Δ (Table 1) strains relative to wild-type. C. In baker's yeast induced to undergo meiosis, the Spo11 complex catalyzes double-strand breaks (DSBs) which then undergo 5' to 3' resection. 3' single stranded tails invade the homologous chromosome to form a D-loop intermediate which, in the major pathway for crossover formation, can be further stabilized by ZMM proteins such as Zip3 and Msh4-Msh5 to enable DNA synthesis and branch migration. A double Holliday junction is then formed that is asymmetrically cleaved in an Mlh1-Mlh3-dependent step to yield primarily crossover products (Manhart and Alani 2016; Zakharyevich et al. 2012). D. Spore autonomous fluorescence assay used to measure single meiotic crossover events (tetratypes) in the *CEN8-THR1* interval of chromosome VIII (Thacker et al. 2011).

Figure S2 Maps and DNA sequences of plasmids pEAA238, pEAI254, and pEAM168.

Figure S3 Functional domains of yeast Mlh3. A. Cartoon depictions of the Mlh1-Mlh3, Mlh1-Mlh2, and Mlh1-Pms1 complexes, with the N-terminal ATP binding and C-terminal endonuclease/Mlh1 interaction domains separated by intrinsically disordered linker domains. PONDR (Predictor of Natural Disorder VSL2 disorder prediction algorithm; Obradovic et al. 2003) analysis was used by Furman et al. (2021) to propose locations of the intrinsically

disordered linker domains in *S. cerevisiae* Mlh1, Mlh2, Mlh3 and Pms1. B. Junction points for the Mlh3 and Pms1 linker sequences are shown with the flanking amino acid sequences. C. Ribbon diagram (left panel) of the C-terminal domain of *B. subtilis* MutL, with the regulatory (QEMIVP) and endonuclease motifs highlighted (Pillon et al. 2010; PDB 3KDK, RCSB Protein Data Bank, modeled using PyMOL). D. Sequence alignments of the endonuclease (green) and PCNA binding motifs (brown) found in a subset of MLH proteins. Sequences for the yeast *mlh3-PIP* mutations analyzed in this study are also shown. Bacterial MutL protein sequences are indicated as follows: Bs, *Bacillus subtilis*; Sp, *Streptococcus pneumoniae*; Ec, *E. coli*; St, *Salmonella typhimurium*. The eukaryotic MLH1, PMS1 (PMS2 in humans) and MLH3 amino acid sequences are indicated with an h for human, m for mouse, and y for baker's yeast.

Figure S4 Phylogenetic analysis of Mlh1, Mlh2, Pms1 and Mlh3 homologs from 34 fungal (Ascomycete) species. A. Mlh1 (red branches), Mlh3 (green), Mlh2 (orange), and Pms1 (purple) homologs from 34 Ascomycetes species were exported for phylogenic analysis (Materials and Methods). *E. coli* MutS and *B. subtilis* MutL were used to root the tree. The scale bar indicates the number of changes per amino acid site. Abbreviations for the fungal species that the Mlh1, Mlh3, Mlh2, and Pms1 homologs were derived from are as follows: *Saccharomyces cerevisiae (Scer), Ashbya gossypii (Agos), Candida albicans (Calb), Candida dubliniensis (Cdub), Candida glabrata (Cgla), Candida guilliermondii (Cgui), Candida lusitaniae (Clus), Candida tropoicalis (Ctro), Debaryomyces hansenii (Dhan), Kluyveromyces lactis (Klac), Kluyveromyces polysporus (Kpol), Kluyveromyces thermotolerans (Kthe), Kluyveromyces waltii (Kwal), Saccharomyces bayanus (Sbay), Saccharomyces castellii (Scas), Saccharomyces kluyveri (Sklu), Saccharomyces mikatae (Smik), Saccharomyces paradoxus (Spar), Lodderomyces elongisporus (Lelo), Pichia stipitis (Psti), Yarrowia lipomyces (Ylip), Aspergillus fumigatus (Afum), Neosartorya fischeri (Nfis), Aspergillus clavatus (Acla), Aspergillus terreus (Ater), Coccidioides immitis (Cimm), Neurospora crassa (Ncra), Podospora anserina (Pans), Fusarium graminearum (Fgra), Fusarium verticillioides (Fver), Trichoderma*

reesei (Tree), Sclerotinia sclerotiorum (Sscl), Stagonospora nodorom (Snod), and *Schizosaccharomyces pombe (Spom*). A gene missing from a particular species does not necessarily indicate gene loss in that species but could also reflect incomplete genome sequencing or other bioinformatic difficulty in locating the orthologous sequence. Campbell et al. (2014) found Mlh2 homologs in many but not all Ascomycetes; for example, they did not identify them in earlier diverging Ascomycete subphylum, or in Basidiomycetes. However, they identified Mlh2 homologs in other basal fungi, suggesting that Mlh2 was lost in Basidiomycetes, instead of being gained in Ascomycetes. Note species with non-canonical (or lacking) meiosis (*C. alb, C. dub, C. tro, C. gui and D. han*; Clark et al. 2013). B. Mlh3, Pms1, Mlh1 phylogenetic analysis using GBlocks. A select subset of more conserved and confidently aligned amino acid sites. Specifically, GBlocks (Castresana 2000) was used to select more conserved blocks of the alignment used to create the phylogenetic analysis presented in panel A. The conserved blocks that were used lacked insertion and deletion events, with the phylogenetic tree inferred from 212 sites. The scale bar indicates the number of changes per amino acid site.

Figure S5 Alignments of Block 1 (A), Block 2 (B), Blocks 3-4 (C) and Blocks 5-6 (D) mutations analyzed in this study. Amino acid sequences of Mlh3 and Pms1. Homologs were obtained from the list presented in the legend to Figure S4A.

Figure S6 EMBOSS Needle alignment (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) of the N-terminal, linker, and C-terminal domains of baker's yeast (S288c background) Mlh3 and Pms1.

Table S1 Plasmids used in this study. Unless indicated, *MLH3* and *PMS1* genes were derived from the SK1 strain background and all plasmids with the pEA plasmid designation were built in the Alani laboratory (Materials and Methods). The designation of the *M* (*MLH3*-derived)

and *P* (*PMS1*-derived) abbreviations for the *MLH3-PMS1* chimeras, and the detailed description of the block mutations can be found in the Materials and Methods and Figures 2.

Table S2 Strains used in this study. *lys2-A$_{14}$* reversion assays were performed on the SK1 strains EAY3252 (*wild-type*), EAY3255 (*mlh3Δ*), and EAY3255 derivatives containing integrated *mlh3$_{SK1}$::KanMX* alleles (Table 1). Meiotic crossover assays were performed on the following diploids: EAY3252/EAY3486 (*wild-type*), EAY3255/EAY3486 (*mlh3Δ*), and EAY3255::*mlh3* alleles/EAY3486 (Table 1). *lys2-A$_{14}$* reversion assays were performed on the S288c strains EAY1269, EAY4595::*mlh3Δ,* and EAY3097 transformed with plasmids as indicated (Tables 1 and 2).

Table S3 Pearson's Chi Squared Contingency Test for meiotic crossing over data presented in Table 1. P-values were determined using a Pearson's Chi-Squared contingency test, with a Benjamini-Hochberg correction for 32 comparisons. A $p<0.002$ cut off was used for statistical significance.

Table S4 Uniprot, NCBI, and RefSeq identifications for the genes used to perform the phylogenetic analysis in Figure 1B.

Literature Cited

Campbell, C. S., H. Hombauer, A. Srivastan, N. Bowen, K. Gries, *et al*., 2014 Mlh2 is an accessory factor for DNA mismatch repair in *Saccharomyces cerevisiae*. PLoS Genetics 10: e1004327

Castresana, J., 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17: 540-552. DOI: 10.1093/oxfordjournals.molbev.a026334

Christianson, T. W., R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter, 1992 Multifunctional yeast high-copy-number shuttle vectors. Gene 110: 119–122.

Clark, N. L., E. Alani, and C. F. Aquadro, 2013 Evolutionary rate covariation in meiotic proteins results from fluctuating evolutionary pressure in yeasts and mammals. Genetics 193: 529–538.

Furman, C. M., R. Elbashir, and E. Alani, 2021 Expanded roles for the MutL family of DNA mismatch repair proteins. Yeast 38: 39-53. doi: 10.1002/yea.3512.

Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac, C. J. Brown, and A. K. Dunker, 2003. Predicting intrinsic disorder from amino acid sequence. Proteins 53: 566–572. https://doi.org/10.1002/prot.10532

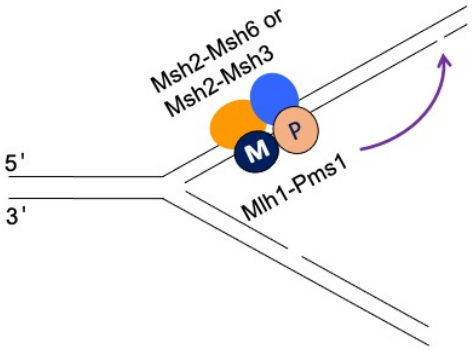
Manhart, C. M., and E. Alani, 2016 Roles for mismatch repair family proteins in promoting meiotic crossing over. DNA Repair 38: 84-93.

Pillon, M. C., J. J. Lorenowicz, M. Uckelmann, A. D. Klocko, R. R. Mitchell, *et al*., 2010 Structure of the endonuclease domain of MutL: unlicensed to cut. Mol. Cell 39: 145–151.

Thacker, D., I. Lam, M. Knop, and S. Keeney, 2011 Exploiting Spore-Autonomous Fluorescent protein expression to quantify meiotic chromosome behaviors in *Saccharomyces cerevisiae*. Genetics 189: 423–439.

Zakharyevich, K., S. Tang, Y. Ma, and N. Hunter, 2012 Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. Cell 149: 334-334.

A. DNA mismatch repair



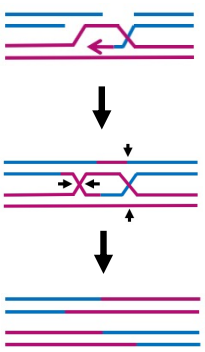
B. *lys2* reversion assay

***lys2-A*$_{14}$** → ***lys2-A*$_{13}$**

Lys$^-$ **Lys$^+$**

wild-type	1X
pms1Δ	~6,000X
mlh3Δ	~6X

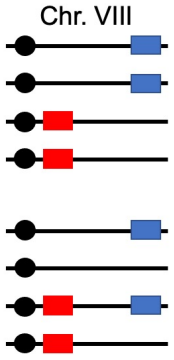
C. Meiotic crossing over



ZMM factors:
stabilize single-end
invasion intermediates

Mlh1-Mlh3, Exo1:
biased resolution of
double Holliday junctions

D. Meiotic crossover assay



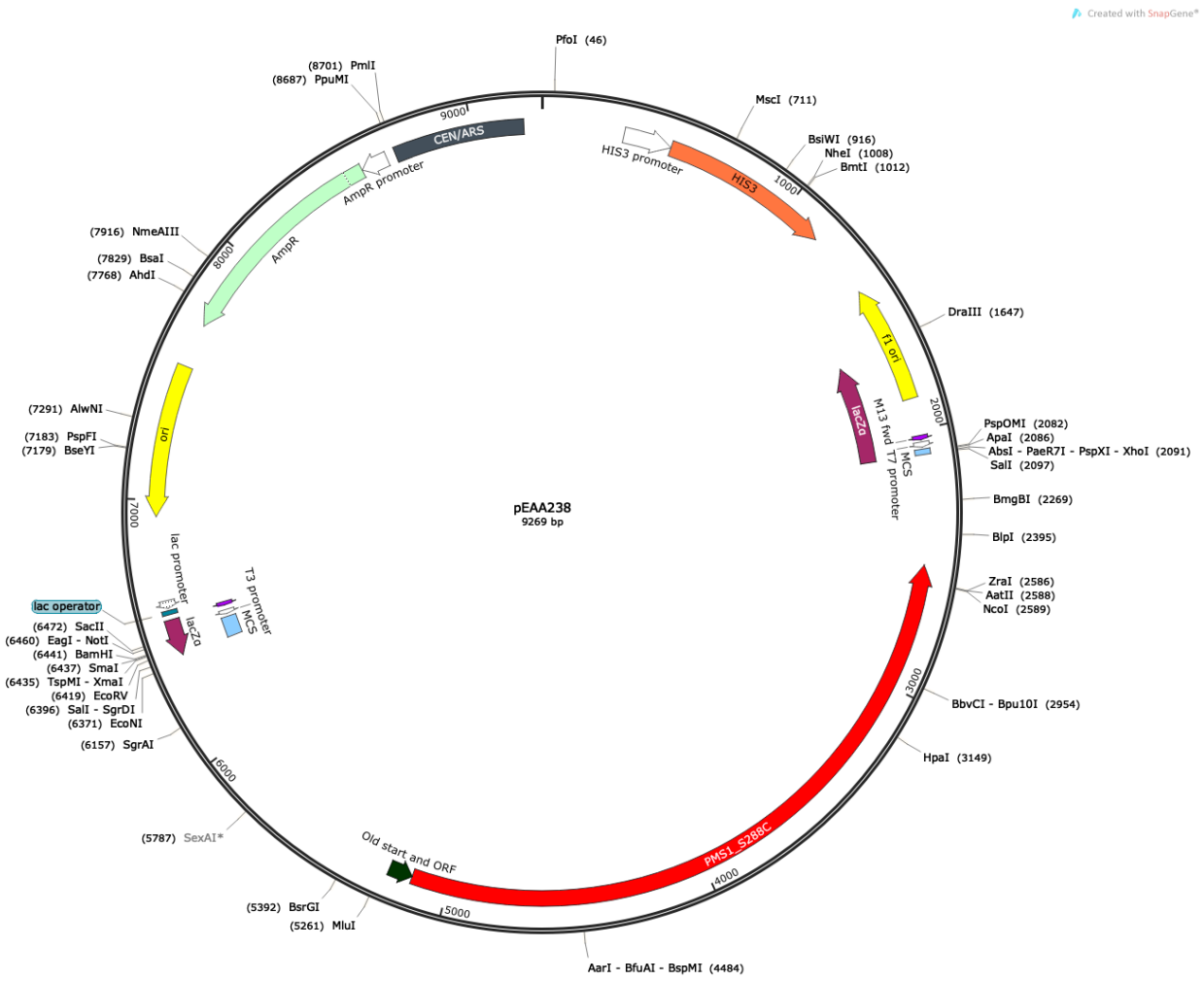
Parental ditype

Tetratype (single CO)
37% wild-type
18% *mlh3Δ*

Figure S1

Figure S2. Maps and DNA sequences of pEAA238, pEAI254, pEAM168

pEAA238

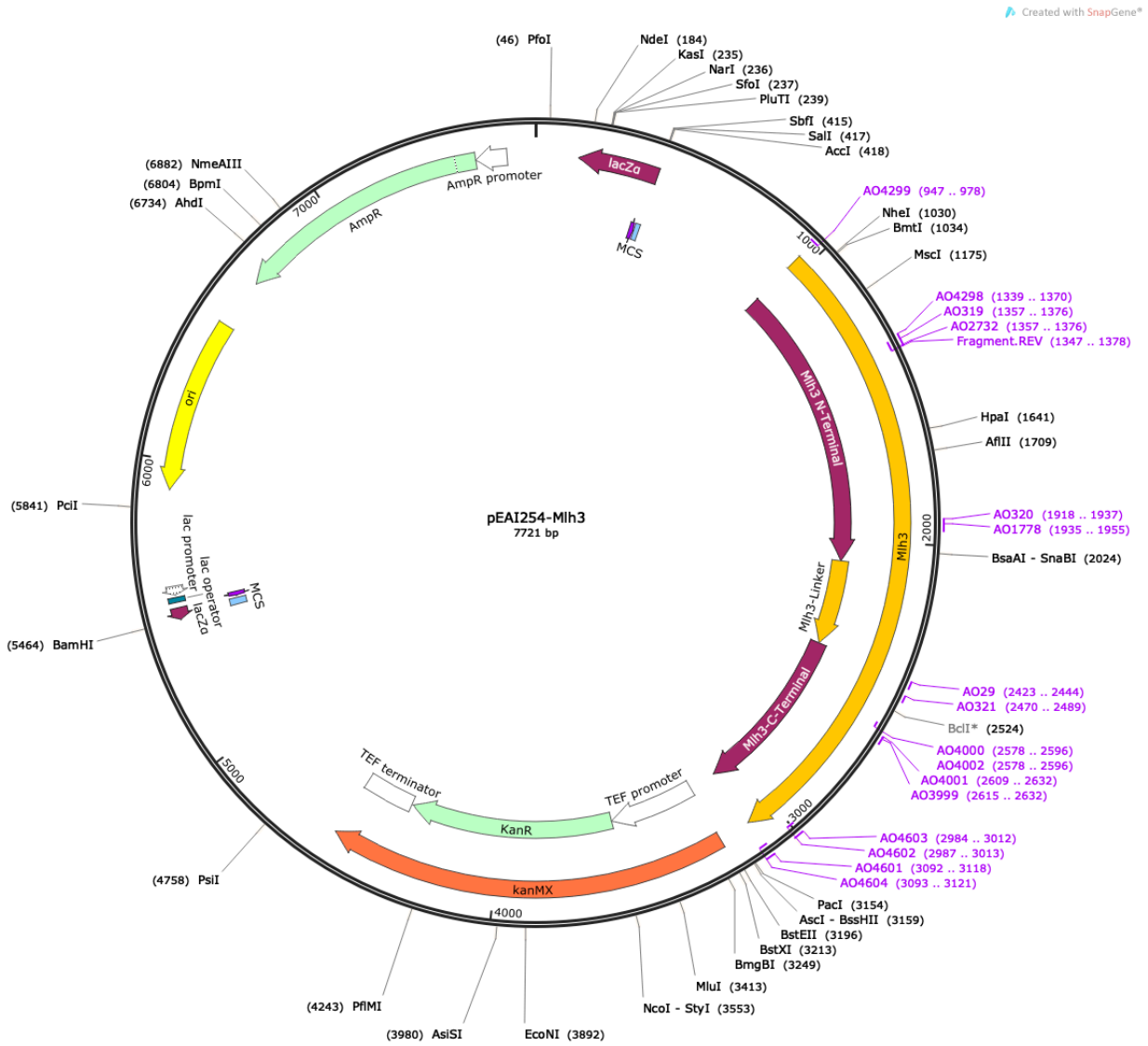


DNA sequence of pEAA238

5’TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCG
CGTCAGCGCGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAAATTCCCGTTTTAAGAGCTTGGTGAGCGCTAG
GAGTCACTGCCAGGTATCGTTTGAACACGGCATTAGTCAGGGAAGTCATAACACAGTCCTTTCCCGCAATTTTCTTTTTCTATTACTCTTGGCCTCCTCTAGTACACTCTATAT
TTTTTTATGCCTCGGTAATGATTTTCATTTTTTTTTTCCCCTAGCGGATGACTCTTTTTTTTCTTAGCGATTGGCATTATCACATAATGAATTATACATTATATAAAGTAATGT
GATTTCTTCGAAGAATATACTAAAAAATGAGCAGGCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGAAACCAAGATTCAGA
TTGCGATCTCTTTAAAGGGTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACACAATCGCAAGTGATTAACGTC
CACACAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGTGACTTACACATAGACGACCATCACACCA
CTGAAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAAAGAGGCCCTACTGGCGCGTGGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTTCCAGA
GCGGTGGTAGATCTTTCGAACAGGCCGTACGCAGTTGTCGAACTTGGTTTGCAAAGGGAGAAAGTAGGAGATCTCTCTTGCGAGATGATCCCGCATTTTCTTGAAAGCTTT
GCAGAGGCTAGCAGAATTACCCTCCACGTTGATTGTCTGCGAGGCAAGAATGATCATCACCGTAGTGAGAGTGCGTTCAAGGCTCTTGCGGTTGCCATAAGAGAAGCCAC
CTCGCCCAATGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATGTAGTGACACCGATTATTTAAAGCTGCAGCATACGATATATATACATGTGTATATATGTATACCT
ATGAATGTCAGTAAGTATGTATACGAACAGTATGATACTGAAGATGACAAGGTAATGCATCATTCTATACGTGTCATTCTGAACGAGGCGCGCTTTCCTTTTTTCTTTTTGCT
TTTTCTTTTTTTTTCTCTTGAACTCGACGGATCTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTTAAA
ATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCA
GTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTT
TTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAA
GCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGC
CATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCC
AGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACAATGACG
ATTGAAGGAGACGCTAGAGCAATCGTAAAAGACAATATAGCTCTAAAAAGAAAAATTGGGAGTTATAGCGGTATGAGACATACATTGCATCTGCCAGTTAGGGGCCAGCA
TACAAGAAACAACGCGAAAACTGCAAGGAAGCTTAATAAAATCGATAGACGTGGAATTCACACATTTTCTCAAGCAAAAGTACAACATAATCCGTCCCTTTGGTCTTGTATC
TTTGGCAAGTGATTATACTCTCCACATATAAACAAATTAAAGCCACGTTCATATTCTTAATGGCTAAGCTTTATGTCGTAATACATATATCCATCAAGCATCTTCAATGCACGA

GTATATTTTGTGTTGTCACTCCCTGTATATAATGTATTTGTTAATTATATAATGAATGAATATCAAAGCTAGATCATATTTCGTAATCCTTCGAAAATGAGCTCCAATCACGTA
ATTCCATTAAATGTCTCATTGTTGGACGTCCATGGGGACAATTCCAAGGCTTATCAAGTTCACTGAGATTATGAACGACTCTGGTCATAGTCTTCTTATTCAGCGGTTTACCT
ATCATTATACTACTTCTACATGCTCTCATAGCAAACATAGAACGAATCTTGGAACATCTGATGTTGTCTCTTCGTAACCCACCGTCTTCTTTGATCAAATGTATCAGTTCATTAA
AATCACCCAAATCAAAAAGGGTTTGTTTGGAAGTAGGTAAGCTTAACAGCTTAACTCTGGAGCCAAACTCTTCTTCTTCGTCAATTTTCAATTTGAATCCATTCTTTTCAAAAA
CTGGTAAATTGTCCAGCACCACAAGTTCATCAATGACACTTAATTCTACCGGCTGAGGTATTATCAATTTCTGCGATTTAAAAACTGTCACGGCCTGCAGTGTTTCGAAATTA
TACTTTTCATCACTTGCATGCTGATCGACAATAAACAGATCATATTTGTTATCAACTTTTCTGGTCACTATAATAAATCCTAAATTAAATTGTCCAACAACCTCCATTTTTTTAA
AATCATTTTTTGAAACTGTTAACGTCAGATACTTTTCTCCTTGTTCGAAATCCTCTAGTTCGTCCTTATTTTTGATAATGTTTTCTTGAGCTTCTGATTTTCTTTTACTTATGATGC
TACTCATTTGTTTGCCATTTTTACTTATTTCTTTAAAATTTTTTGTGCTTAAATTGTACTCCAATATCTCATCTTCGAATTTGCGATGCGTCAATCCATCACTCAATGACCTATAA
TTATCTTTGGAAATCGATTTACGTGAGTTCTTTAGAGGAGTTCTCACATTGATTTCAACTGGCTCAATCTCTGCGTATATAGAGTCAGCTTCATCATCTTGTTCTGTGTCAGTA
CTTCCTCTTCTTTCTTGATGACAGCAGTCATTTGTATGCTCATGACATTCATTATCTACAAAAACTAAACCATCTGCTTGTGATAGAACCGCTTTTTCTTGGAACTTTTCACCAT
CAATATCAAAATAAACTGGTTCTTCAACAACTTTTTCCAAGCTACGTGCTTTATCTGGACTAGTTATATTTTGAAATTCAGGATTGGAAAAATTATTCAGGTTCAGATCACTTA
GGTTTTGAGAATCTGTTTTGATACTTGGCAGCTTCTTAGTCGGTGTACTTGAGTTGCCCTCATCCAATACAACTGATGAATCAACCGAAACTTCACACTCCGAACCGATGACA
TCAGTGACATTAGTATAATTACCGTCCATCACAGAAGTTAGTTCTGTACCATTACTTTTGTCTATTACGCCAGTAGTACTATTAAAATGTGCATGATTGGATTGATTACTTTCG
CTTCTGGCGGTATGATAGTTTTCATTATCGGATTCGTGTGTTGTGCTTCTATCATCAAAGACTTCCGTTTTAAGTCTTTTCTGAGCTTGTTGTTCTGATTGTGAGCACATTCTTT
TGGGAAGAGCTAATTCTTGTCTATTGTAATAATCACTGAGAGTTGTTTTGAAAATATCAATGACGGCCCGTTCGTTGTGCAACAGTATCACTCTTTTGTCTGGAGTGACGTTT
ACATCAATTAAGCTCATGGGTAATTCAAGATTGAGAAATACAGCCGGAAATTGAACATTATTGAACGTTTTGTAAACCTCATTGCAGCATTTTAAAAGCGTAGAGTATTCCA
CTGGCCTTTTATTCACATAAATGAATTGCCTATCCTTACTATTTCTACCACAACCAAACGAATTCTGTGATATATAGCCCTTCACTCGAATTTTATAATCGAGGTCTAAGAAAT
CAGGATCATCGGTGTATTTCCCAAGCATTCTATTCTTGAAGGGATTTAAATCAAGCACTAAATCAACCTCTTCAAGCCCACGCATACCACCTGCACCGAATACTGAGGATATA
TTCTTTCTCATACTTGAATTTCTCATTGTTGATAGAATTAGATTTTTTTTACCCTTTGGTGTGATATTCCAAACTGAAAATTTAATGGCAGCATTTATGATTGCATAACCTTGTA
TAACTGTGAGACATTTGGTAAACTGGCGCTTAAAAGTTTTGCTAAATTCTTTCTGTCTAACGGGCAAGTTATGAAATAACTGCGAAACTAACACGGTCGTCCCTTTATTTCTT
GATGTCGTTGTCTTTGAAGTAATGTGACCGACCATATCATACTCCAACTTATCCGCTTTTGGTGGTGAAGTTGTTGTTATCACACTTAGTTTAGCTATGCCACATAAAGAAGA
TAGGGCCTCCCCTCTAAACCCTAACGTCTGTACTTTAGCAACATCTTGAAATTTCGCAATTTTTGATGTGTAATGTTTTAAAGCTAAGAACTCATAATTTGAAGGGTCTATGCC
ATCACCATTATCAGAACATTCGATAGATTCAAGACCATAATCTTTAAAGATTATTTCGATCTGGTTTGCGTTCGCATCTATACTATTATCAACGAGTTCTTTCACTGCAGTTGT
TAAGTCGGTGATAACTTGTCCAGATGTAATTCGATGAACATCTATATCGTTTATCTGATGAATTTGTGTCATGCTAAATAAATATTTCACAGGGATATACCTCTGTTCCTTTTG
CTTACACCTTTTTTCTGTCTCAATTAATAAGTTTTCGATGTGGTGAAACATTTTATCGCATAATGATTATTAAGAGAGACGCGTCTTTTCTTTTCGCGTTCTGTTATAGACGATT
TTTATAATATTTTATCAAATATCAATTAACTATTGAGGTCATTAAAATCTGTATTAGAATCGGTATTAATCTGTGCTTTCTAGAGTTCTAGCTATGTACATGTTGAAAAGAATA
TATATTATAACGAAAGCTTGGCTGACAGTACGCCCACAAACCGCAGCAGCGGTTTATAATTAAATAGAATTCACTAGTTGATACATATGCATATGTTATTATCACTACTCCAG
ATTCATGAATTTTTTCAGATTTTCGTAACATAAATAGCTGATGGAAACTGCTGGACATACTTTTGCTAAAGTTGGTACTAGACCTTTAAATAGACCTTGATAACCTTCCCTTTC
TAGCGTCTTTAACAAAACATCTTTAAAGCCGTTATACACATAAGGATGTGCATATGTTCCTTGGGCTTGTAGTCTTGTTCTTAAAAGATTGATTGGATAAACAACAGAAGCTC
CTACAGTTCCACTGAATGCACCCATTGGAAGTACAACCAGGTTGCTTAGAGTGACCTGATCTTGTGGTAGGTTCAGGGTCTTTGCCTGTTTGGCAATATACCATTTTTTTAAG
GCAGAAAAAGTCCCCAAATCTAATGCAGCATAGGGAAATATGCCCACTATACCGACTGTGACACCTCTGTAAAATAATCTGACCCCCCCCCTCTCGAAACATATCTTTTGCA
GTCTGAAAGAGCAAATTATTCCCCTTCAACTTTGTGTCCAAAGGAGCGCACTGCACCCTGAATTTCAAAGTATCTATAGGATAAACGGAAAACTGAGCTGCCATACCAGCTA
AACCACCAGCAATGTAAGTAGAGAACTTTGAAAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGAC
ATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGC
ATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGACGGTATCGATAAGCTTGAT
ATCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATC
ATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGTAACT
CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG
CTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAA
CGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG
GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA
ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC
AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATA
GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA
GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT
GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC
TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATA
CGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
TAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATA
GGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGGGTCCTTTTCATCACGTGCTATAAAAATAATTATAATTTAAATTTTTTAATATAAATATATAAATTAAAAATAGAA
AGTAAAAAAAGAAATTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAAGACTCTAGGGGGATCGCCAACAAATACTACCTTTTATCTTGCTCTTCCTGCTCTCAGGT
ATTAATGCCGAATTGTTTCATCTTGTCTGTGTAGAAGACCACACACGAAAATCCTGTGATTTTACATTTTACTTATCGTTAATCGAATGTATATCTATTTAATCTGCTTTTCTTG
TCTAATAAATATATATGTAAAGTACGCTTTTTGTTGAAATTTTTTAAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACTCTTCTACCTTCTTTATTTACTTTCTAAAATCCAAAT
ACAAAACATAAAAATAAATAAACACAGAGTAAATTCCCAAATTATTCCATCATTAAAAGATACGAGGCGCGTGTAAGTTACAGGCAAGCGATCCGTCCTAAGAAACCATTA
TTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC3’

pEAI254

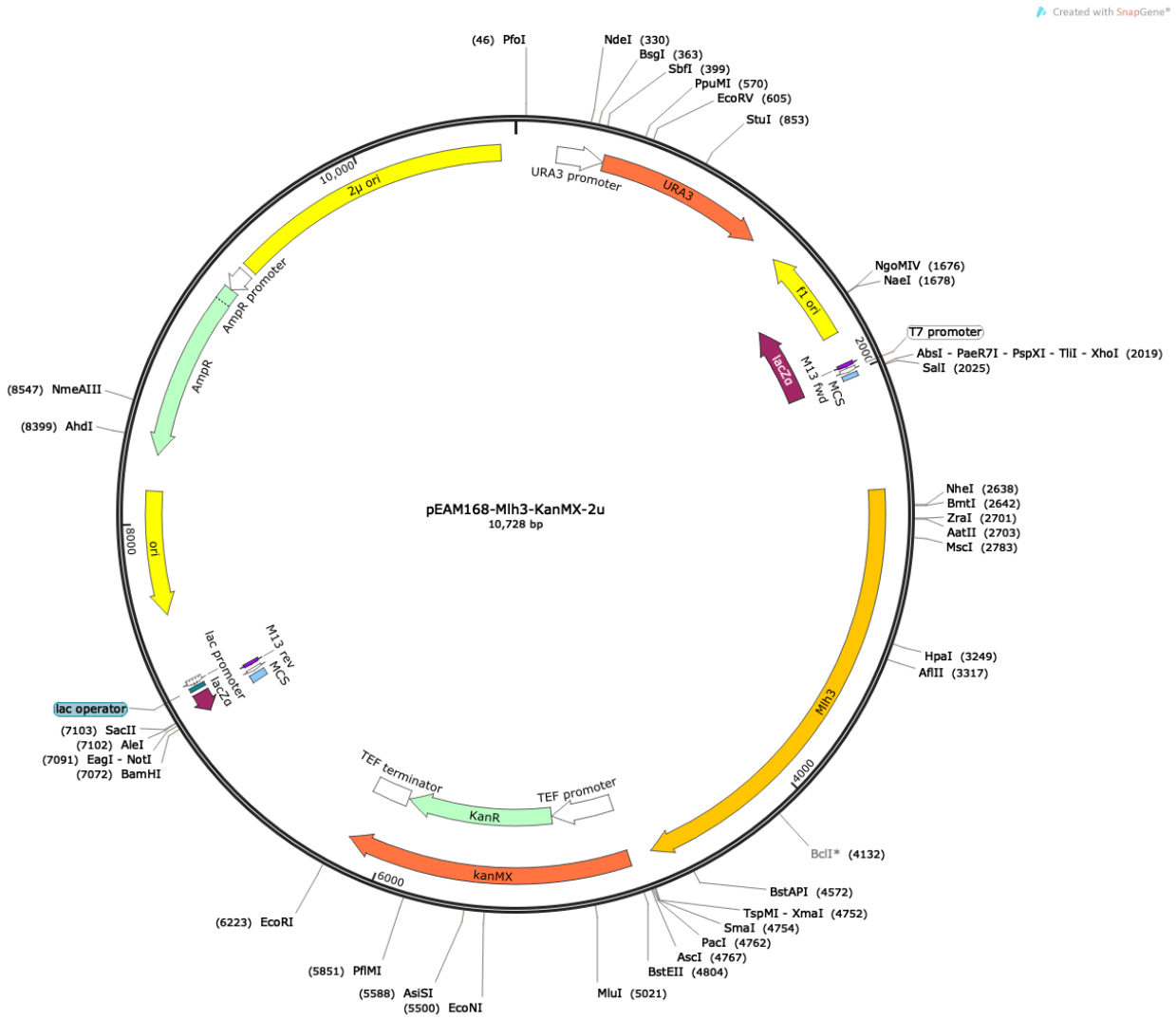


DNA sequence of pEAI254

5’TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCG
CGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAG
GAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTGCATGCCTGCAGGTCGACGCTGCAGTTACCGA
AGTGGTTACTACTACGGCGGAGCCTGTTGTTAAATACGTTACTATAACTGCCGATGCTAGTAATGTTACAGGTTCTGCTAACAACGGTACCCACATTTAATGCGTGACGTTG
AATCGAGAAAAAAAGCTACTTTAACGAAACCTTTACTAGTTATCCTATATGGGATCACTAGTATTTTTTGATTTACGATTCAATAAATAGACTAGAGACAACTTTCATATCATT
CCTTAAAAAATACATAAAGCGCAAATTCAACCCCATTGATACATATATAAGTAGTTCTATTATGACTTTCAAGAACAATAGTAGCTTTCCTAAATAATCAATAAGTAGCACAA
AATCTGTCTGTTTATACGCTTATATTTGGTTTGCGTTTATTTGCGAGCGCCACGAGAAGGGGCAGGAAAAAAAGATCAATAGTTTGCAATAAACATCGAATGATGATTTCAA
ACACCACCGATACATAAACCAGCGAGGCTTTCAAGGAAGAATGAACGTGAACTCGTCAACTCAAAAAGAAAATGAGCCAGCATATTAGGAAATTAGATTCTAATGTTTCTG
AAAGACTTAAATCTCAGGCATGCACGGTATCGCTAGCATCAGCGGTTAGAGAAATAGTTCAAAATTCTGTAGATGCACACGCTACCACTATCGACGTCATGATCGACCTCCC
TAATTTGAGCTTTGCAGTTTACGATGATGGTATTGGTTTGACTCGAAGTGACCTAAATATATTGGCCACACAAAATTATACTTCAAAAATACGAAAGATGAATGATTTAGTA
ACGATGAAAACCTACGGTTACAGAGGAGACGCCCTATATAGCATTTCTAATGTCTCTAAGCTGTTTGTTTGTTCCAAGAAAAAGGATTACAACTCTGCATGGATGAGAAAAT
TTCCATCCAAAAGCGTCATGTTGAGTGAGAATACCATACTCCCAATAGATCCTTTTTGGAAAATTTGTGCTTGGAGGCGAACAAAGTCTGGTACTGTTGTTATTGTTGAAGA
TATGCTGTATAATTTACCTGTCCGGCGCAGAATACTAAAGGAAGAACCCCCTTTCAAGACTTTCAACACAATAAAGGCAGATATGCTACAGATACTGGTAATGCACCCGATG
ATTTCACTCAACGTACAATATACAGACAAACTCAGAATCAATACAGAAGTTCTTTTCCGCTCAAAGAACATAACTGAGGGGTTAACAAAGCACCAGCAAATGTCTCAAGTAC
TGAGAAATGTGTTTGGGGCAATAATACCACCGGATATGCTTAAGAAAGTTTCCTTGAAGTTTAATGAATATCAGATTGAAGGGATCATATCGAAGATGCCCGTTCGATTGA
AGGATCTTCAATTTATATATATCAATGGCAGACGTTATGCGGACAGCGCCTTTCAAGGATACGTCGATTCTTTATTCCAAGCACAAGATTTCGGGGAGAAAGGCATGAGTTT
ACTCAAAACTAAAAGCGTGGGAAAACCGTATCGATCACATCCCGTCTTCATCCTTGATGTTCGGTGCCCCCAAACTATCGATGATCTGTTACAAGATCCAGCAAAGAAAATT
GTTAAACCTTCCCATATACGTACCATTGAACCATTGATCGTTAAGACTATTCGCTCCTTTCTGACATTTCAAGGATACTTAACGCCTGATAAATCAGATTCATCTTTCGAAATA
TTCAATTGTTCACAGAAGACGGCAACATTGCCTGATAGTCGAATCCAAATCAGTAAACGTAATCAAGTATTAAATTCAAAAATGAAGATTGCACGCATTAATTCATACATAG
GAAAACCCGTTGTCAATGGATGTAGAATAAACAATAGTACAATAAACTACGAAAAAATCAAGAATATAAGAATCGACGGTCAGAAATCAAGCTTACAGAATAAACTCTCTT

CCAGGCCCTACGATAGTGGTTTTACAGAAGATTATGACAGTATAGGAAAAACCATAACAGATTTTTCAATATCAAGATCTGTCCTGGCAAAATATGAAGTGATTAATCAAGT
AGATAAAAAATTTATTCTCATACGTTGTTTAGACCAATCTATTCATAATCGCCCACTATTAGTTCTAGTTGATCAACATGCTTGCGATGAGAGGATAAGACTAGAAGAGCTAT
TTTATAGCCTGTTGACAGAAGTTGTAACCGGAACTTTTGTTGCACGAGACCTCAAGGATTGTTGCATTGAAGTTGATCGTACTGAGGCTGACTTATTTAAGCATTATCAAAG
TGAATTTAAAAAGTGGGGAATTGGCTATGAAACAATTGAGGGAACAATGGAAACCTCACTTTTGGAAATTAAAACACTGCCTGAGATGTTGACTTCCAAATATAACGGAGA
TAAAGACTACTTAAAAATGGTCTTGCTGCAACATGCACATGACCTGAAGGATTTCAAAAAACTTCCAATGGATTTATCCCATTTCGAAAACCACACTTCGGTTGATAAATTAT
ATTGGTGGAAGTACTCAAGCTGCGTTCCAACGGTCTTTCATGAAATTCTTAACAGCAAGGCATGCAGATCTGCTGTTATGTTTGGAGATGAATTAACCCGGCAAGAGTGCAT
AATTTTAATTAGCAAGCTCTCGCGATGCCATAACCCATTCGAATGCGCCCACGGGAGACCTTCTATGGTACCCATTGCAGAATTGAAGTAAAGAGAACCTTAATCTTGGAAC
AAGGTTTGTCGCCTGCCCCGGGTTAATTAAGGCGCGCCAGATCTGTTTAGCTTGCCTCGTCCCCGCCGGGTCACCCGGCCAGCGACATGGAGGCCCAGAATACCCTCCTTG
ACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGC
GCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAG
GTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTC
ACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGA
AGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGA
CCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATT
GTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAA
CGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACT
CATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTAT
GGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTT
TCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTT
TTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGATACTAACG
CCGCCATCCAGTGTCGAAAACGAGCTCGAATTCGCGCGGATATGCAAAACAGAAGATTACCTCATCACGTTGTCAAAACCTGGAAAAATTTAAAAAACAAAAACAAAAAAG
AAACGTGCGATACAATACTGCTCCACTGTTGCATTACCGTTTTATGTATCTCCGATATTATAAATGTAGCCCTAGATTTATATGAATATATATATATACCATTTATATATGCAA
ACACAAAATAACCTAAAATTTAAAAATGACTATTGATGGAGACGTACATGAAATTTCCCCATTTTTCCAGGTTAGGCAAACTAAATGGGGCGGTAGGGCTTGTTTCAGTAAC
GGTAATATTCCTAAAGGAACAACAGTCTTGCAAGTCAGTAATTTTACAGGAACATCTATTTCCTATGAGTTCAGGAAGGAAGTTTGTCACAATTGTTTCGCATACGCTAATG
CGAAAACTATGAAGTACAAACTAAACTACGATTGCCTGAGAGATTTGGTTTGTAATGCACACTACCAAATTAACCCTAAGAAGTTTCTGGGCGCTGGGCTGTGGTTCTGTTC
AGAACATTGCAGAACTTCTTATCTTCAAATACCAAATATCATTGAATTAATTGAATGTTATGAGATATTACTGCATCATTTTCCAAGCATGCTGAAAAGATATAATTCTACCTC
TGAACAGGAGGAAAAATTAAACTCTATCTTAATTTCGGAAAATGTCATACAGTCTTCATGGGATGAGATTGAAAATAAATGGATACCGCGCATCAATAACATGAAATCTGC
GAAGAGAATAAATCAATTACCACCCATTTGCGAAGATGAATATTGCTGCATTAGGTTTGTGTGCGAGAGTTTGTTTAATCTTAAATATATGGATCCCCGGGTACCGAGCTCG
AATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAG
TGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGC
GTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG
CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG
CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT
GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT
CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT
TTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTC
GTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC
AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC
GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT
GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGG
CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCA
GTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA
GGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTC
GTC3’

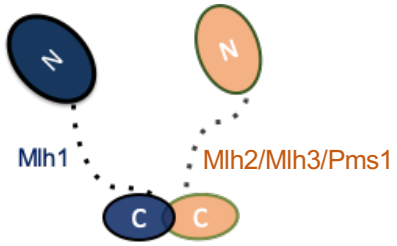
pEAM168



5’TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCG
CGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATACCACAGCTTTTCAATTCAATTCATCATTTTTTTT
TTATTCTTTTTTTTGATTTCGGTTTCTTTGAAATTTTTTTGATTCGGTAATCTCCGAACAGAAGGAAGAACGAAGGAAGGAGCACAGACTTAGATTGGTATATATACGCATAT
GTAGTGTTGAAGAAACATGAAATTGCCCAGTATTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAAACGAAGATAAATCATGTCGAAAGCTACATATAAGGAACGT
GCTGCTACTCATCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGAAAAGCAAACAAACTTGTGTGCTTCATTGGATGTTCGTACCACCAAGGAATTACTGGAGTT
AGTTGAAGCATTAGGTCCCAAAATTTGTTTACTAAAAACACATGTGGATATCTTGACTGATTTTTCCATGGAGGGCACAGTTAAGCCGCTAAAGGCATTATCCGCCAAGTAC
AATTTTTTACTCTTCGAAGACAGAAAATTTGCTGACATTGGTAATACAGTCAAATTGCAGTACTCTGCGGGTGTATACAGAATAGCAGAATGGGCAGACATTACGAATGCAC
ACGGTGTGGTGGGCCCAGGTATTGTTAGCGGTTTGAAGCAGGCGGCAGAAGAAGTAACAAAGGAACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCATGCAAGGGCTC
CCTATCTACTGGAGAATATACTAAGGGTACTGTTGACATTGCGAAGAGCGACAAAGATTTTGTTATCGGCTTTATTGCTCAAAGAGACATGGGTGGAAGAGATGAAGGTTA
CGATTGGTTGATTATGACACCCGGTGTGGGTTTAGATGACAAGGGAGACGCATTGGGTCAACAGTATAGAACCGTGGATGATGTGGTCTCTACAGGATCTGACATTATTAT
TGTTGGAAGAGGACTATTTGCAAAGGGAAGGGATGCTAAGGTAGAGGGTGAACGTTACAGAAAAGCAGGCTGGGAAGCATATTTGAGAAGATGCGGCCAGCAAAACTA
AAAAACTGTATTATAAGTAAATGCATGTATACTAAACTCACAAATTAGAGCTTCAATTTAATTATATCAGTTATTACCCTATGCGGTGTGAAATACCGCACAGATGCGTAAG
GAGAAAATACCGCATCAGGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCC
TTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCT
ATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTT
GACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA
CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG
GCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATA
GGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGCTGCAGTTACCGAAGTGGTTACTACTACGGCGGAGCCTGTTGTTAAATACGTTACTATAACTGCCGATGCTAG
TAATGTTACAGGTTCTGCTAACAACGGTACCCACATTTAATGCGTGACGTTGAATCGAGAAAAAAAGCTACTTTAACGAAACCTTTACTAGTTATCCTATATGGGATCACTA
GTATTTTTTGATTTACGATTCAATAAATAGACTAGAGACAACTTTCATATCATTCCTTAAAAAATACATAAAGCGCAAATTCAACCCCATTGATACATATATAAGTAGTTCTAT
TATGACTTTCAAGAACAATAGTAGCTTTCCTAAATAATCAATAAGTAGCACAAAATCTGTCTGTTTATACGCTTATATTTGGTTTGCGTTTATTTGCGAGCGCCACGAGAAGG
GGCAGGAAAAAAAGATCAATAGTTTGCAATAAACATCGAATGATGATTTCAAACACCACCGATACATAAACCAGCGAGGCTTTCAAGGAAGAATGAACGTGAACTCGTCA
ACTCAAAAAGAAAATGAGCCAGCATATTAGGAAATTAGATTCTAATGTTTCTGAAAGACTTAAATCTCAGGCATGCACGGTATCGCTAGCATCAGCGGTTAGAGAAATAGT
TCAAAATTCTGTAGATGCACACGCTACCACTATCGACGTCATGATCGACCTCCCTAATTTGAGCTTTGCAGTTTACGATGATGGTATTGGTTTGACTCGAAGTGACCTAAATA
TATTGGCCACACAAAATTATACTTCAAAAATACGAAAGATGAATGATTTAGTAACGATGAAAACCTACGGTTACAGAGGAGACGCCCTATATAGCATTTCTAATGTCTCTAA
GCTGTTTGTTTGTTCCAAGAAAAAGGATTACAACTCTGCATGGATGAGAAAATTTCCATCCAAAAGCGTCATGTTGAGTGAGAATACCATACTCCCAATAGATCCTTTTTGG
AAAATTTGTGCTTGGAGGCGAACAAAGTCTGGTACTGTTGTTATTGTTGAAGATATGCTGTATAATTTACCTGTCCGGCGCAGAATACTAAAGGAAGAACCCCCTTTCAAGA

CTTTCAACACAATAAAGGCAGATATGCTACAGATACTGGTAATGCACCCGATGATTTCACTCAACGTACAATATACAGACAAACTCAGAATCAATACAGAAGTTCTTTTCCG
CTCAAAGAACATAACTGAGGGGTTAACAAAGCACCAGCAAATGTCTCAAGTACTGAGAAATGTGTTTGGGGCAATAATACCACCGGATATGCTTAAGAAAGTTTCCTTGAA
GTTTAATGAATATCAGATTGAAGGGATCATATCGAAGATGCCCGTTCGATTGAAGGATCTTCAATTTATATATATCAATGGCAGACGTTATGCGGACAGCGCCTTTCAAGGA
TACGTCGATTCTTTATTCCAAGCACAAGATTTCGGGGAGAAAGGCATGAGTTTACTCAAAACTAAAAGCGTGGGAAAACCGTATCGATCACATCCCGTCTTCATCCTTGATG
TTCGGTGCCCCCAAACTATCGATGATCTGTTACAAGATCCAGCAAAGAAAATTGTTAAACCTTCCCATATACGTACCATTGAACCATTGATCGTTAAGACTATTCGCTCCTTTC
TGACATTTCAAGGATACTTAACGCCTGATAAATCAGATTCATCTTTCGAAATATTCAATTGTTCACAGAAGACGGCAACATTGCCTGATAGTCGAATCCAAATCAGTAAACGT
AATCAAGTATTAAATTCAAAAATGAAGATTGCACGCATTAATTCATACATAGGAAAACCCGTTGTCAATGGATGTAGAATAAACAATAGTACAATAAACTACGAAAAAATCA
AGAATATAAGAATCGACGGTCAGAAATCAAGCTTACAGAATAAACTCTCTTCCAGGCCCTACGATAGTGGTTTTACAGAAGATTATGACAGTATAGGAAAAACCATAACAG
ATTTTTCAATATCAAGATCTGTCCTGGCAAAATATGAAGTGATTAATCAAGTAGATAAAAAATTTATTCTCATACGTTGTTTAGACCAATCTATTCATAATCGCCCACTATTAG
TTCTAGTTGATCAACATGCTTGCGATGAGAGGATAAGACTAGAAGAGCTATTTTATAGCCTGTTGACAGAAGTTGTAACCGGAACTTTTGTTGCACGAGACCTCAAGGATT
GTTGCATTGAAGTTGATCGTACTGAGGCTGACTTATTTAAGCATTATCAAAGTGAATTTAAAAAGTGGGGAATTGGCTATGAAACAATTGAGGGAACAATGGAAACCTCAC
TTTTGGAAATTAAAACACTGCCTGAGATGTTGACTTCCAAATATAACGGAGATAAAGACTACTTAAAAATGGTCTTGCTGCAACATGCACATGACCTGAAGGATTTCAAAAA
ACTTCCAATGGATTTATCCCATTTCGAAAACCACACTTCGGTTGATAAATTATATTGGTGGAAGTACTCAAGCTGCGTTCCAACGGTCTTTCATGAAATTCTTAACAGCAAGG
CATGCAGATCTGCTGTTATGTTTGGAGATGAATTAACCCGGCAAGAGTGCATAATTTTAATTAGCAAGCTCTCGCGATGCCATAACCCATTCGAATGCGCCCACGGGAGACC
TTCTATGGTACCCATTGCAGAATTGAAGTAAAGAGAACCTTAATCTTGGAACAAGGTTTGTCGCCTGCCCCGGGTTAATTAAGGCGCGCCAGATCTGTTTAGCTTGCCTCGT
CCCCGCCGGGTCACCCGGCCAGCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCA
TACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCAC
AGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACA
GTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTAC
AGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGC
AAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTT
TAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGT
CTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTG
ATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTAT
TGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAGTTT
TTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGC
GTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGATACTAACGCCGCCATCCAGTGTCGAAAACGAGCTCGAATTCGCGCGGATATGCAAAACAGAAGATT
ACCTCATCACGTTGTCAAAACCTGGAAAAATTTAAAAAACAAAAACAAAAAAGAAACGTGCGATACAATACTGCTCCACTGTTGCATTACCGTTTTATGTATCTCCGATATTA
TAAATGTAGCCCTAGATTTATATGAATATATATATATACCATTTATATATGCAAACACAAAATAACCTAAAATTTAAAAATGACTATTGATGGAGACGTACATGAAATTTCCC
CATTTTTCCAGGTTAGGCAAACTAAATGGGGCGGTAGGGCTTGTTTCAGTAACGGTAATATTCCTAAAGGAACAACAGTCTTGCAAGTCAGTAATTTTACAGGAACATCTAT
TTCCTATGAGTTCAGGAAGGAAGTTTGTCACAATTGTTTCGCATACGCTAATGCGAAAACTATGAAGTACAAACTAAACTACGATTGCCTGAGAGATTTGGTTTGTAATGCA
CACTACCAAATTAACCCTAAGAAGTTTCTGGGCGCTGGGCTGTGGTTCTGTTCAGAACATTGCAGAACTTCTTATCTTCAAATACCAAATATCATTGAATTAATTGAATGTTA
TGAGATATTACTGCATCATTTTCCAAGCATGCTGAAAAGATATAATTCTACCTCTGAACAGGAGGAAAAATTAAACTCTATCTTAATTTCGGAAAATGTCATACAGTCTTCAT
GGGATGAGATTGAAAATAAATGGATACCGCGCATCAATAACATGAAATCTGCGAAGAGAATAAATCAATTACCACCCATTTGCGAAGATGAATATTGCTGCATTAGGTTTG
TGTGCGAGAGTTTGTTTAATCTTAAATATATGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT
GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT
GAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCG
TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA
GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGC
ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT
GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA
GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCT
CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTT
TAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG
ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTA
GTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG
AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTG
CATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGC
GTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG
TTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG
GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATA
AACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAACA
AAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTACCAACGAAGAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGC
TAATTTTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGAATCTATACTTCTTTTTTGTTCTACAAAAATGCA
TCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGA
AGAAGGCTACTTTGGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATC
CCCGATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTCATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTC
TATATACTACGTATAGGAAATGTTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTAAAGAGTAATACTAGAGATAAACATAAAA
AATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATG
TTTGTGGAAGCGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCGTTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGGTTTTCAAAAGCGCTCTGAAGT
TCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTCA
CGTCGCACCTATATCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGCGTGTTTATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAA
GGTAGTCTAGTACCTCCTGTGATATTATCCCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAGTCT
CATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC3’

A



	linker size (amino acids)	Location
Mlh1	164	335-499
Mlh2	245	372-617
Mlh3	113	376-488
Pms1	277	362-638

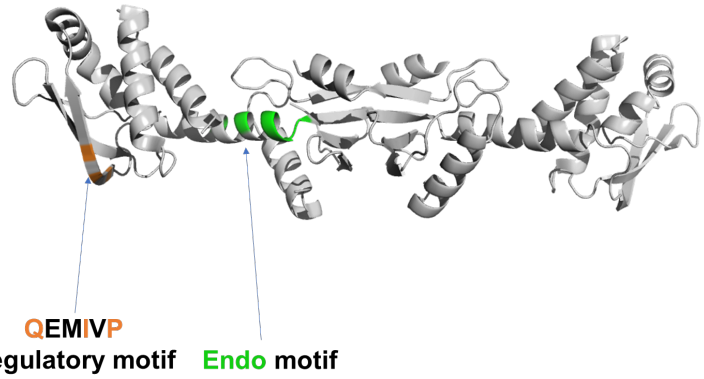
B

Boundaries for Mlh3 and Pms1 linker sequences

SK1 Mlh3 linker sequence, amino acids 376 to 488, with flanking Mlh3 N- and C-terminal sequences in purple font:
IRSFLTFQGYLTPDKSDSSFEIFNCSQKTATLPDSRIQISKRNQVLNSKMKIARINSYIGKPV
VNGCRINNSTINYEKIKNIRIDGQKSSLQNKLSSRPYDSGFTEDYDSIGKTITDFSISRSVLA
KYEVINQ

SK1 Pms1 linker sequence, amino acids 362 to 638, with flanking Pms1 N- and C-terminal sequences in purple font:
IRSFLTFQGYALPKRMCSQSEQQAQKRLKTEVVDDRSTTHESDNENYHSARSESNQSN
HAHFNSTCEGTTGVIDKSNGTELTSVMDGNYTNVTDVIGSECEVSVDSSVVLDEGNSST
PTKKLPSIKTDSQNLSDLNLNNFSNPEFQNITSPDKARSLEKVVEEPVFFDIDGEKFQEKA
VLSQADGLVFVDNECHEHTNDCCHQERRGSTDTEQDDEADSIYAEIEPVEINVRTPLKNS
RKSISKDNYRSLSDGLTHRKFEDEILEYNLSTKNFKEISKNGKQMSSIIVLAKYEVINQ

C



D

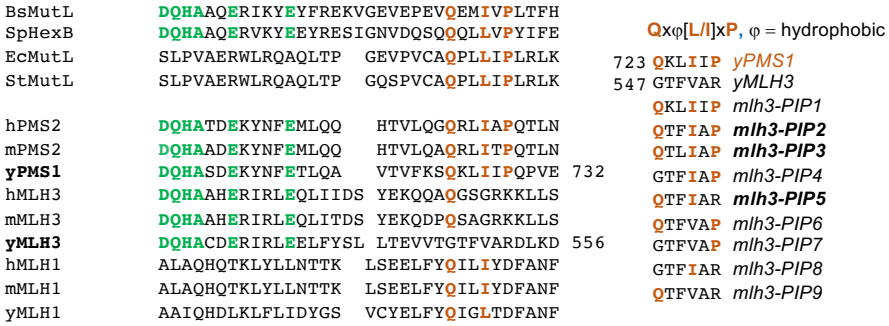


Figure S3

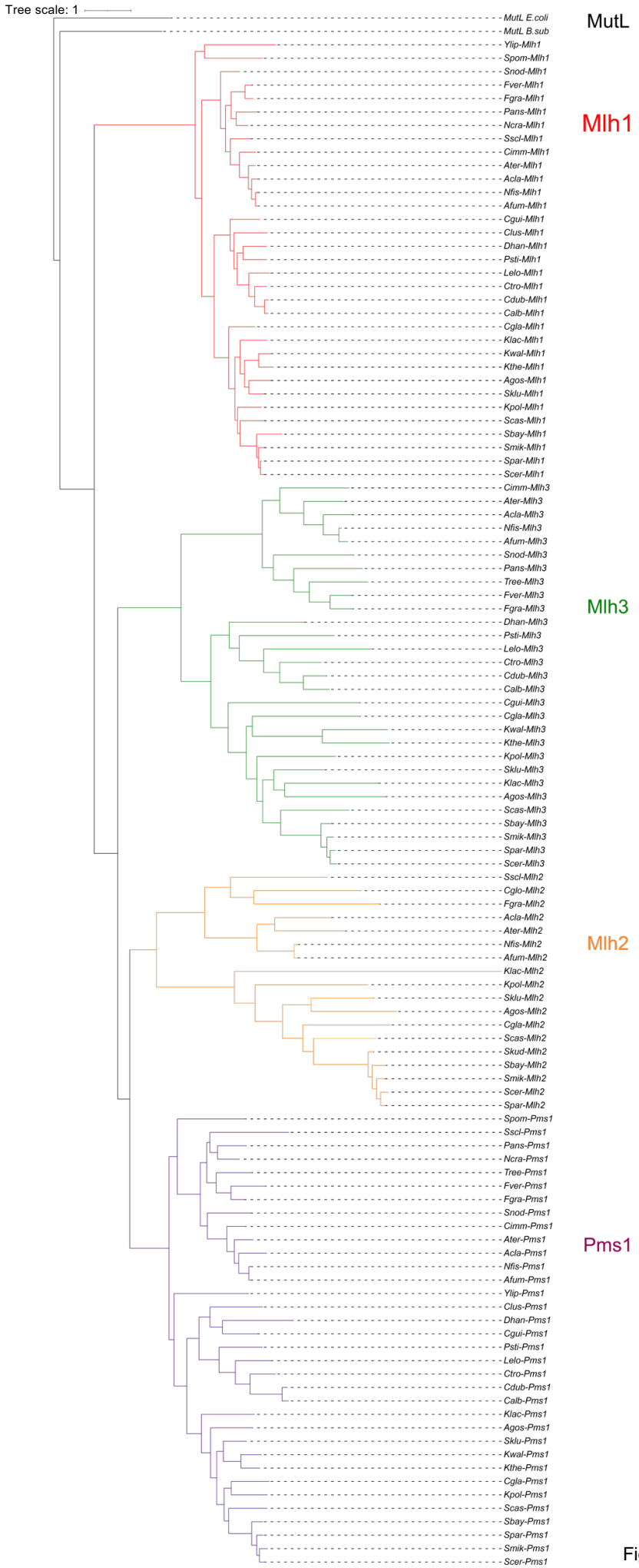

Tree scale: 1
MutL E.coli
MutL B.sub
Ylip-Mlh1
Spom-Mlh1
Snod-Mlh1
Fver-Mlh1
Fgra-Mlh1
Pans-Mlh1
Ncra-Mlh1
Sscl-Mlh1
Cimm-Mlh1
Ater-Mlh1
Acla-Mlh1
Nfis-Mlh1
Afum-Mlh1
Cgui-Mlh1
Clus-Mlh1
Dhan-Mlh1
Psti-Mlh1
Lelo-Mlh1
Ctro-Mlh1
Cdub-Mlh1
Calb-Mlh1
Cgla-Mlh1
Klac-Mlh1
Kwal-Mlh1
Kthe-Mlh1
Agos-Mlh1
Sklu-Mlh1
Kpol-Mlh1
Scas-Mlh1
Sbay-Mlh1
Smik-Mlh1
Spar-Mlh1
Scer-Mlh1
Cimm-Mlh3
Ater-Mlh3
Acla-Mlh3
Nfis-Mlh3
Afum-Mlh3
Snod-Mlh3
Pans-Mlh3
Tree-Mlh3
Fver-Mlh3
Fgra-Mlh3
Dhan-Mlh3
Psti-Mlh3
Lelo-Mlh3
Ctro-Mlh3
Cdub-Mlh3
Calb-Mlh3
Cgui-Mlh3
Cgla-Mlh3
Kwal-Mlh3
Kthe-Mlh3
Kpol-Mlh3
Sklu-Mlh3
Klac-Mlh3
Agos-Mlh3
Scas-Mlh3
Sbay-Mlh3
Smik-Mlh3
Spar-Mlh3
Scer-Mlh3
Sscl-Mlh2
Cglo-Mlh2
Fgra-Mlh2
Acla-Mlh2
Ater-Mlh2
Nfis-Mlh2
Afum-Mlh2
Klac-Mlh2
Kpol-Mlh2
Sklu-Mlh2
Agos-Mlh2
Cgla-Mlh2
Scas-Mlh2
Skud-Mlh2
Sbay-Mlh2
Smik-Mlh2
Scer-Mlh2
Spar-Mlh2
Spom-Pms1
Sscl-Pms1
Pans-Pms1
Ncra-Pms1
Tree-Pms1
Fver-Pms1
Fgra-Pms1
Snod-Pms1
Cimm-Pms1
Ater-Pms1
Acla-Pms1
Nfis-Pms1
Afum-Pms1
Ylip-Pms1
Clus-Pms1
Dhan-Pms1
Cgui-Pms1
Psti-Pms1
Lelo-Pms1
Ctro-Pms1
Cdub-Pms1
Calb-Pms1
Klac-Pms1
Agos-Pms1
Sklu-Pms1
Kwal-Pms1
Kthe-Pms1
Cgla-Pms1
Kpol-Pms1
Scas-Pms1
Sbay-Pms1
Spar-Pms1
Smik-Pms1
Scer-Pms1
MutL
Mlh1
Mlh3
Mlh2
Pms1

Fig S4A

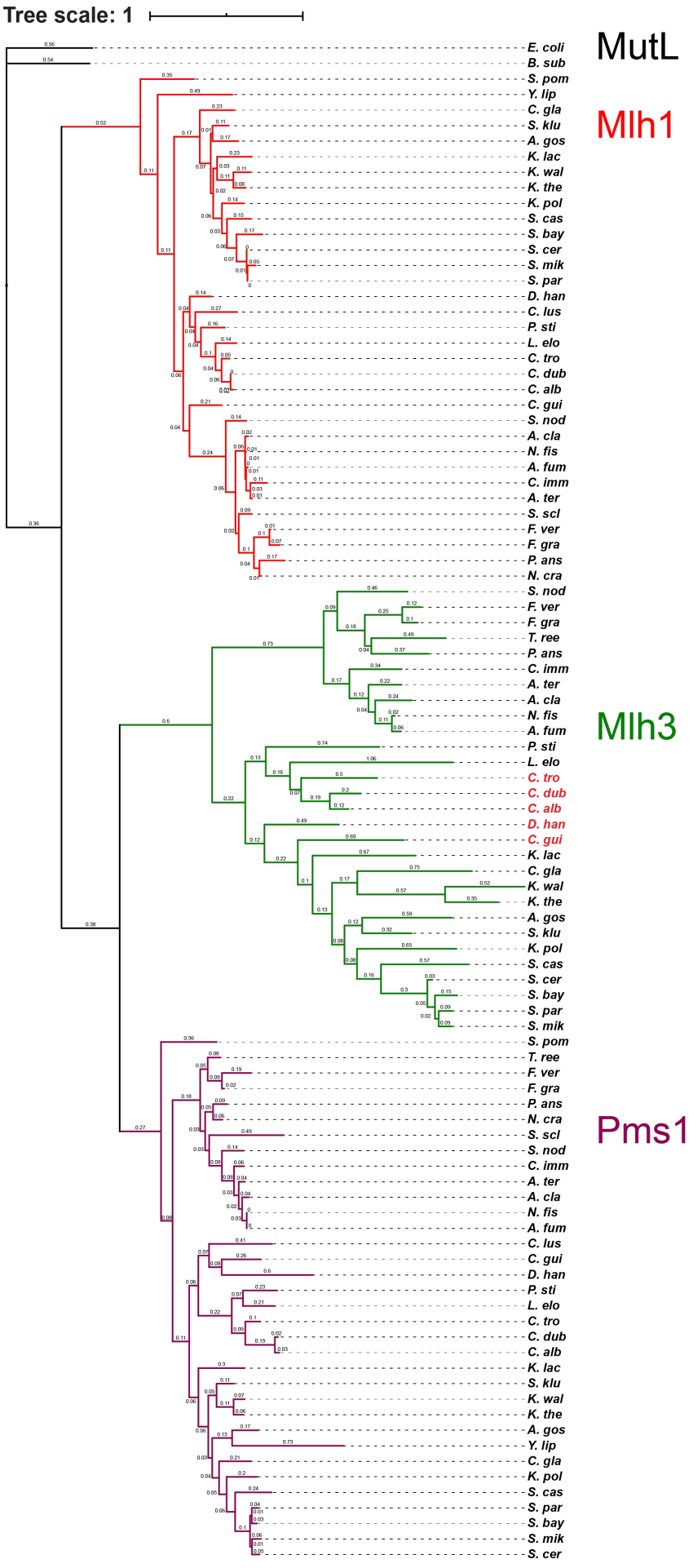

Tree scale: 1
MutL
Mlh1
Mlh3
Pms1
E. coli
B. sub
S. pom
Y. lip
C. gla
S. klu
A. gos
K. lac
K. wal
K. the
K. pol
S. cas
S. bay
S. cer
S. mik
S. par
D. han
C. lus
P. sti
L. elo
C. tro
C. dub
C. alb
C. gui
S. nod
A. cla
N. fis
A. fum
C. imm
A. ter
S. scl
F. ver
F. gra
P. ans
N. cra
S. nod
F. ver
F. gra
T. ree
P. ans
C. imm
A. ter
A. cla
N. fis
A. fum
P. sti
L. elo
C. tro
C. dub
C. alb
D. han
C. gui
K. lac
C. gla
K. wal
K. the
A. gos
S. klu
K. pol
S. cas
S. cer
S. bay
S. par
S. mik
S. pom
T. ree
F. ver
F. gra
P. ans
N. cra
S. scl
S. nod
C. imm
A. ter
A. cla
N. fis
A. fum
C. lus
C. gui
D. han
P. sti
L. elo
C. tro
C. dub
C. alb
K. lac
S. klu
K. wal
K. the
A. gos
Y. lip
C. gla
K. pol
S. cas
S. par
S. bay
S. mik
S. cer

Fig S4B

Mlh3

```
Scer 11 DVSERLKSQACTVSLASAVREIVQNSVDAHATTIDVM 47
Agos 15 SVFLRLRSQLSIVSVGAAVRELMQNSVDACCGRLEVS 51
Calb 33 SVITQIRSHTIFHSLDSVVRELLQNSLDAGANEITIK 69
Cdub 10 SVVTQIRSHTILHSLDLVVRELLQNSVDAGADKITIK 46
Cgla  8 ETWQVLRADTQITGIDSVLRELLQNAVDAGASTVDVS 44
Cgui 10 DVIQQLVAPLSVSGYASAVRELLHNSIDAQATFISVK 46
Ctro 10 SVVSQIRSHVILNSFEQVVGELLQNSLDAGANTISIR 46
Dhan 17 SVLSCLQSQITISSFQAAIRELLQNSLDANANEVIIK 53
Klac  9 TVSSRLKSHESVVSTTAVVRELVQNSVDADATKIKVD 45
Kpol 11 EVLRLLVPQSKTISLSSTVREVIQNSVDAKSTTIKIF 47
Kthe 11 DVAKSLKSSASIDCITTALRLLLQNSIDAGASHITVF 47
Kwal 11 DVVNSISSGASLSCITSAIQMALENSVDSGASEIIIW 47
Sbay 11 LVLKRLKSQVYAVSLASAVREIVQNSIDAHASNVDVL 47
Scas 11 TLSERLNSQTHVVSITSALKEVVQNSIDAKATTITIQ 47
Sklu 11 NVLRVLKSQVFIVSVTSAVRELLQNSVDAGASEIEAV 47
Smik 11 TVSEMLKSQVCAVSLASAVREIVQNSVDAHATTIDVV 47
Spar 11 NVSEKLKSQVSSVSLASAVREIVQNSVDAHATTIDVV 47
Lelo 10 DVINQIKSHLRFNTIDSIIEHLLKDSLQRSPKNIIVK 46
Psti 13 QVLSELRSQTIFNSLASVVQELLRNSLDAQAKVVEIR 49
Afum 14 EVVAKIKSSTSIINLNRVVVELVKNALDANARTVSVT 50
Nfis 14 EVVAKIKSSTSITDLNRVVVELVKNALDANAHTVSVT 50
Acla 14 DAVAKIKSSTSITHLNGVIVELVKNALDANARTVFVT 50
Ater 14 DVIAKLKSSTSITQLNGVVVELVKNALDANAHTIYVT 50
Cimm 14 KVAAQIESSTSITSLNSAILGLVKNSLDADAQTVAIT 50
Pans  9 DVVAQIKSSTIITSLNTAVCGLLRNSLDARSTKINIS 45
Fgra  9 DVVDKIKSSVNITSLNGVVCGLIANSLDAGASKVNIS 45
Fver  9 DVVDKIKSSVNITSLNGVVCGLITNSLDAGASKINIS 45
Tree  9 DVVGKIRSSSTITSLNGVVCGLVKNSLDAGATKVNIV 45
Snod 20 DVVAQIKSSTAIVSLTDVVLELLKNSLDARATKVEAS 56
```

Top z-score in Multi-harmony:
Block 1, N-terminal ATP binding domain
mlh3-K17T,A20Q,S24D,R30K, Q34D

Pms1

```
Scer 10 IDVHRITSGQVITDLTTAVKELVDNSIDANANQIEII 46
Agos 11 ADVHRITSGQVIIDLVAAVKEVVENSLDAHADKLEIT 47
Calb 14 FDISKITSSQVIIDLKAIIKELIENSIDAHADKIDII 50
Cdub 14 FDISKITSSQVIIDLKAIIKELIENSIDAHANKIDII 50
Cgla  9 QDVHRITSGQVIIDLASAVKELVENSLDAQATQIDIT 45
Cgui  9 TEVQRITSGQVIVDLVSVVKELVENAIDSGSTKIDVT 45
Clus  9 SKVAKITSGQVIIELLSVVKELVENSIDAESDKIDVV 45
Ctro 12 ADISKLTSSQVIIDLKSILKELIENSIDANADKIDIT 48
Dhan  9 GDIQRITSGQVIIDLVSIVKELVENSIDASSSKIEVL 45
Klac  9 ADIHKITSGQVIVDLRSAIKELLENSLDAKADKIEII 45
Kpol 10 SDVHRITSGQVIIDLTTAVKELIDNSIDASAKQIDIT 46
Kthe 11 ADVHRITSGQVIIDLASAVKELLDNSIDSGADQVVCT 47
Kwal  9 ADVHRITSGQVIVDLTSAVKELLDNSIDSGADHVECT 45
Sbay 14 IDVHRITSGQVITDLTTAVKELVDNSIDANASQIEIT 50
Scas 10 KDIHRLTSGQVIVDLVSAVKELVDNSIDAHSSQIELT 46
Sklu 10 TDVHRITSGQVIVDLATAVKELVENSIDAGADKIDVV 46
Smik 10 IDVHQITSGQVITDLTTAVKELVDNSIDANANQIEIT 46
Spar 10 IDVHRITSGQVITDLTTAVKELIDNSIDANASQIEIT 46
Lelo 14 TDISKITSGQVIIDLRSIVKELVENAIDASAKKILVN 50
Psti  9 KDVSKITSGQVIIDLKSIVKELVENSIDANSTKIEIN 45
Ylip  9 ASIRQITSAQVVTDLNSAVKEVVENSLDANAKNVEIK 45
Afum 10 RSVHQIQSGQVIVDLCSVVKELVENSLDAGATSIEVR 46
Nfis 10 RSVHQIQSGQVIVDLCSVVKELVENSLDAGATSIEVR 46
Acla 10 RSVHQIQSGQVIVDLCSVVKELVENSLDAGATSIEVR 46
Ater 10 RSVHQIQSGQVIVDLCSVAKELVENSLDAGATSIEVR 46
Cimm 14 RSIHQIQSGQVIVDLCSVVKELVENSLDAGATSIDIR 50
Ncra 13 STIHQIQSGQVIVDLCSVAKELVENALDAGATTIDVR 49
Pans 13 SAVHQIQAGQVIVDLCSVAKELVENSVDAGATTIEVR 49
Fgra 10 RTVHQIQSGQVIVDLCSVVKELVENSVDSGATSIDVR 46
Fver 10 RTVHQIQSGQVIVDLCSVVKELVENSVDAGATSIDVR 46
Tree 10 RT   LTSGQ                       DVR 19
Sscl 10 RTIHQIQSGQVIVDLCSVVKELVENSLDAGATSIDVR 46
Snod 10 RSVHQIQSGQVIVDLQSVCKELVENSIDAGATTVEVR 46
Spom 10 NTVHKICSGQVITDVASAVKELVENSLDSGATTIEIR 46
```

Fig S5A

Top z-score in Multi-harmony:
Block 2, Mlh1 interaction
mlh3-Y493M,N497G,V499F,D500N,K502G

mlh3-D500N, mlh3-K502G single substitutions

Mlh3

```
Scer 492 KYEVINQVDKKFILI 506
Agos 493 DCIVINQIGNKFILL 507
Calb 438 KFRVINQVDRKFILL 452
Cdub 400 RFKVINQIDRKFILL 414
Cgla 503 EVDVVNQVDQKFILL 517
Cgui 398 SFRVLKQLDRKYILF 412
Ctro 435 NYKIIKQLDKKFILV 449
Dhan 511 NYRIISQVDRKFILL 525
Klac 477 NFEVIQQVDLKFILV 491
Kpol 501 DCKLINQLDRKFILL 515
Kthe 457 RFRVVNQVDKKFVLL 471
Kwal 468 SYRAVNQVDNSFILL 482
Sbay 492 NCEVINQVDKKFILI 506
Scas 507 DCTIINQVDDKFILL 521
Sklu 525 NCHVVNQVDNKFILL 539
Smik 492 KYEVINQVDKKFVLI 506
Spar 492 KYEVINQVDKKFVLI 506
Lelo 476 NYRIIRQIDNKFILL 489
Psti 424 TYRIVRQLDSKFILV 437
Afum 657 TANVIAQVDRKFILV 671
Nfis 655 TANIIAQVDRKFILA 669
Acla 649 TAEIIAQVDRNFILA 663
Ater 619 TAEVIAQIDQKFILA 633
Cimm 704 NAQIVAQVDNKFILL 718
Pans 633 KAEVVNQVDEKFVLV 647
Fgra 602 SATVIAQVDRKFILV 616
Fver 600 RATVIEQVDRKFILV 614
Tree 625 AAEVVAQVDKKFILL 639
Snod 656 SAEVLAQVDKKFILV 670
```

Pms1

```
Scer 673 KMEVVGQFNLGFIIV 687
Agos 703 EMSIIGQFNLGFIIV 717
Calb 699 KMKLIGQFNLGFIIV 713
Cdub 693 KMKLIGQFNLGFIIV 707
Cgla 702 DMKVVGQFNLGFILV 716
Cgui 659 KMQVVGQFNLGFILV 673
Clus 675 RMSVVGQFNLGFVVV 689
Ctro 595 DMKLIGQFNLGFILV 609
Dhan 535 RDKLLS  DLGVVEV 547
Klac 723 NMTIVGQFNLGFIIV 737
Kpol 757 RMEIVGQFNLGFIIA 771
Kthe 706 KMEIVGQFNLGFILV 720
Kwal 698 NMNVVGQFNLGFIIV 712
Sbay 696 KMEIVGQFNLGFIIV 710
Scas 655 KMVIIGQFNLGFIVV 669
Sklu 678 KMEIVGQFNLGFIIV 692
Smik 674 KMEVVGQFNLGFIIV 688
Spar 680 KMEVVGQFNLGFIIV 694
Lelo 748 DMKLIGQFNLGFILV 762
Psti 614 KMKLIGQFNLGFILV 628
Ylip 699 NFNIIGQFNEAFIIV 713
Afum 785 RMRIIGQFNLGFILA 799
Nfis 786 RMRIIGQFNLGFILA 800
Acla 801 RMRIIGQFNLGFILA 815
Ater 794 QMRIIGQFNLGFILA 808
Cimm 759 KMRVVGQFNLGFILA 773
Ncra 843 KMKIVGQFNLGFIIA 857
Pans 753 KMKIIGQFNLGFILA 767
Fgra 776 RMRIAGQFNMGFIIA 790
Fver 783 KMRIAGQFNMGFIIA 797
Tree 688 RMRIIGQFNLGFIIA 702
Sscl 855 KMKIIGQFNLGFILA 869
Snod 831 EMRIIGQFNLGFIIA 845
Spom 610 RMRVVGQFNRGFIVV 624
```

Fig S5B

Mlh3

Block 3, endonuclease motif
mlh3-R530K, mlh3-R532N, mlh3-R530K,R532N

Block 4, PIP motif
mlh3-PIP1-9

```
Scer  519 LVLVDQHACDERIRLEELFYSLLTEVVTGTFVARD 553
Agos  520 LLILDQHAADERVKLEAYTRDYLFTLLTAQPSFYT 554
Calb  458 LVVLDQHASDERIRVEQYLQEFVSQPNPGLRLHSP 492
Cdub  420 IVVLDQHASDERIRVEQYLQEFVEQRHPGLRLQNP 454
Cgla  528 LYIIDQHACDERIRLESFLKQYICDIMANALAVQK 562
Cgui  422 LFIVDQHACDERIRVESYLKDYIQRTKVNQHQNKA 456
Ctro  458 IVVLDQHATDERIKVEEYLQEFVQLLQKNPGLRLK 492
Dhan  541 ILVVDQHACDERIRVEALFKDFILMLLDKTLGIEL 575
Klac  497 LLMLDQHACHERILVENMLKETIIKCMNKCFNYVK 531
Kpol  528 LLLIDQHACDERIKLEEYLKEYLTSVRSKTVTTRK 562
Kthe  483 LVLLDQHAADERIKYESLLNGFIWGMLTSPHLHIR 517
Kwal  494 LLLVDQHAADERMRFESLMSDFLYEFITAPYLQTL 528
Sbay  519 LILVDQHACDERIRLEDLLHSLLSEVLTETFVTQD 553
Scas  533 LLIVDQHACDERIKLESYLKDFFQDVINGTISSQP 567
Sklu  552 LLIVDQHACDERIRLESYIKEFVFNALSFRGLNTV 586
Smik  519 LILVDQHACDERIRLEGLLHNLLTEVISGTFVAQD 553
Spar  519 LVLVDQHACDERIRLEDLFHNLLTEIITGTFVAQD 553
Lelo  517 LVVLDQHASDERVKIEKLIKEFVDEMSANPGLRLC 451
Psti  448 LLVIDQHACDERIKVEALFKDFIFLVLDAHTNLLL 482
Afum  683 LVLIDQHAADERCRIESLFGAMFADGHRQVQSIRI 717
Nfis  681 LVLIDQHAADERCRIESLFGDMFADGHRQVQTIRI 715
Acla  675 LILVDQHAADERCRVEHLFGGLFADNDNSLCPIHT 709
Ater  645 LVLIDQHAADERCRVERLFGEFFADDRSGRVQTIT 679
Cimm  734 LVLVDQHAADERIRVERLFDELCGSSPSHTVDTTP 768
Pans  667 LVIIDQHAADERYRVEALLKEYFVPNPDDGRSLVA 701
Fgra  637 LVMLDQHAVDERCQLEELMLEYFTTDPLTNQVLPQ 671
Fver  635 LVMLDQHAADERCQLEDLMASYFAHDPSTNQTSSV 669
Tree  658 LVMLDQHAADERCRLEDLMAGYFTHDSSIGAIRAV 692
Snod  689 LVLIDQHAADERIRVEALLRELCSPMDTHCSGYQS 723
```

Pms1

```
Scer  697 LFIVDQHASDEKYNFETL QAVTVFKSQKLIIPQP 730
Agos  727 LFIVDQHASDEKYNFENL QKSTVFNSQHLIKPLT 760
Calb  721 LFIIDQHASDEKYNFEKL MASFKINYQLLIKPIK 754
Cdub  715 LFIIDQHASDEKYNFEKL MTNFRINYQSLIKPIK 748
Cgla  726 LFIVDQHASDEKFNFENL QQTTRFKSQKLISPET 759
Cgui  681 LFIVDQHASDEKYNFERLANSTTMFHSQLLVVPRN 715
Clus  695 LFIVDQHALDEIFNYERL MQSLVLRAQPLVIPRL 728
Ctro  615 LFIIDQHASDEKFNFEKL MSNFQIKHQPLMMPIN 648
Dhan  555 RRDRDNNIEDEKY                       567
Klac  747 LFIVDQHASDEKYNFEML QKNTVFKSQSLLSLKT 780
Kpol  781 LFIIDQHASDEKYNFETL QKSTVFKSQKLIVPQP 814
Kthe  730 LFIIDQHASDEKYNFEKL QKNTVFKSQKLLAPQI 763
Kwal  722 LFIVDQHASDEKYNFETL QRTTVFRSQRLIAPQV 755
Sbay  720 LFIVDQHASDEKYNFETL QAVTVFKSQKLIVPQP 753
Scas  679 LFIVDQHASDEKYNFEML QKETVFNSQRLIAPQP 712
Sklu  702 MFIVDQHASDEKYNFETL QKTTVFKSQRLLSPQI 735
Smik  698 LFIVDQHASDEKYNFETM QAVTVFKSQKLIIPQP 731
Spar  704 LFIVDQHASDEKYNFETL QAVTVFKSQKLITPQP 737
Lelo  768 LFIIDQHASDEKFNFERL LETFAVNYQPLITPLF 801
Psti  635 LFIIDQHASDEKYNFERL NQELSIKIQRLIVPQT 668
Ylip  719 LFIIDQHASDEKYNFERL QRDTKITPQPFVNPLT 752
Afum  809 LFIIDQHASDEKFNFERL QAETVVQNQRLVQPKR 842
Nfis  810 LFIIDQHASDEKFNFERL QAETVVQNQRLVQPKR 843
Acla  825 LFIIDQHASDEKFNFERL QAETVVQNQRLVQSKR 858
Ater  818 LFIIDQHASDEKFNFERL QAETVVQNQRLVQPKR 851
Cimm  783 LFIIDQHASDEKYNFERL QAETVVQNQRLVRPKT 816
Ncra  867 LFIIDQHASDEKYNFERL QSTTTVQSQRLVQPKP 900
Pans  777 LFIIDQHASDEKYNFERL QSTTTVQSQRLVQPKP 810
Fgra  800 LFIIDQHATDEKYNFERL QEVQTVQSQRLVHPKR 833
Fver  807 LFIIDQHATDEKFNFERL QEIQTVQSQRLVHPKR 840
Tree  712 LFIIDQHASDEKYNYERL QRTTEIQSQRLVHPKR 745
Sscl  879 VFIIDQHSSDEKYNFERL QATTIVQSQRLVYPQN 912
Snod  855 LFIIDQHASDEKYNFERL SATTTLVSQRLVHPHP 888
Spom  630 LFIIDQHASDEKFNYEHL KSNLVINSQDLVLPKR 663
```

Fig S5C

Block 5, Helix 2 *mlh3-V660K, N666A, F676I, D678K* *mlh3-D678K* single substitution

Block 6 , Helix1 to linker. *mlh3-C695L, F699W, A702P, S707T, V709R, P710H*

Mlh3

Top z-scores in Multi-harmony

```
Scer 658 PTVFHEILNSKACRSAVMFGDELTRQECIILISKLSRCHNPFECAHGRPSMVPIA 712
Agos 660 PTVFLEILNSKACRSAIMFGDKLNHDECLFLVRQLSTCNMPLRCAHGRPSVIPIA 714
Calb 579 PRIITELINSKACRSAIMFGDILTKDEMQDLVNKLSRCKLPFQCAHGRPSIVPIA 633
Cdub 538 PRIITELINSKACRSAIMFGDILTKDEMYQIVTKLSQCKLPFQCAHGRPSIVPIA 592
Cgla 664 PSFIRRFFDSKACRSAIMFGDTLNLQECRDLVRRLNGCIQPNFCAHGRPSVVELI 718
Cgui 546 PQIILDSINMRACRSAIMFGIPLTLAEMNYMLQCLFRCQHPFHCAHGRPSVVEVR 600
Ctro 582 PRIIIELINSKACRSAIMFGDELNHDDMERLVGKLRHCKLPFQCAHGRPSIVPLA 636
Dhan 669 PQVIIDCINSKACRSAIMFGDKLTVEEMMYLIKSLSECNQPFQCAHGRPSIVPLA 723
Klac 632 PRVYTEIINSKSCRSAIMFGTSLSRTECDVMISDLSKCQQPFHCAHGRPSVVPIV 686
Kpol 668 PIILLELFNSRACRSAIMFGDKLDKEECRILIRQLSDCHFPFQCAHGRPSTIPLA 722
Kthe 621 PTMITDSLKSKACRSAIMFGDYLSPQETTLLVEMLGKCRNPFYCAHGRPSLVPIF 675
Kwal 632 PNVLIDFFKSKACRSAIMFGDVLSMQECNLLVEMLGKCRNPFFCAHGRPSVAPLA 686
Sbay 658 PTVFHEILNSKACRSAIMFGDELSRQECIILVNKLSQCHNPFQCAHGRPSMVPIA 712
Scas 673 PKVFQEIFNSKACRSAIMFGDKLTQQECTILIKKLSECKVPFQCAHGRPSVIPLT 727
Sklu 693 PTLFMEIFNSKACRSAIMFGDALTKAECQLLIRELCTCQLPFQCAHGRPSVIPIV 747
Smik 658 PTVFHEILNSKACRSAIMFGDELTRQECMILISKLSLCHNPFQCAHGRPSMVPIA 712
Spar 658 PTVFHEILNSKACRSAVMFGDELTRQECIILISKLSRCHNPFQCAHGRPSMVPIA 712
Lelo 643 PRAIIDLLNSKACRSAIMFGDPLTFTEMSSLIQELSRCKLPFQCAHGRPSVVPLA 697
Psti 577 PTFLIDIINSKACHSSVVFGEVLEYSEMEKMVRQLLHCRLPFQCAHGRPSIVPLV 631
Afum 831 PQGIIDLLNSRACRTAIMFNDMLTAEECKSL           CAHGRPSMVPIL 873
Nfis 829 PQGIIDLLNSRACRTAIMFNDMLTAEECKSLIGRLARCVLPFQCAHGRPSMVPIL 883
Acla 818 PRGIIDLLNSRACRTAIMFNDVLTVDECQSLVSRLRQCVLPFQCAHGRPSMVPIL 872
Ater 780 PKGIVDLLNSRACRTAIMFNDALAVDECQRLVMQLARCLFPFQCAHGRPSMIPIL 834
Cimm 884 PRGIIDLLNSRACRSAIMFNDKLSKKECKELISTLAKCVFPFQCAHGRPSMVPTM 938
Pans 812 PEGIIELIHSRACRSSIMFNDVLTKEQCFQLVQNLATCAFPFQCAHGRPSMVPLV 866
Fgra 774 P            RGAIMFNDILTTQQCEELIARLSRCAFPFQCAHGRPSMAPLV 816
Fver 771 PRGILELLHSRACRSAIMFNDVLSVNQLR          AWSTQYG        806
Tree 801 PRGILELLHSRACRKRADDGD             LCRCHHVQRCSFG      G 835
Snod 848 PEGLVDIINSRACRSAIMFNDELDMHQSRGLVQKLATCAFPFVCAHGRPSMVPLG 902
```

Pms1

```
Scer 805 CSKIRSMFAMRACRSSIMIGKPLNKKTMTRVVHNLSELDKPWNCPHGRPTMRHLM 859
Agos 835 CSKIRSMFAMRACRMSIMIGKPLTRRTMTEVVRKLSELDKPWNCPHGRPTMRHLM 889
Calb 828 CSKIKKILAMKACRSSIMIGTFLSKSKMREIISNLSTLDKPWNCPHGRPTMRHLI 882
Cdub 822 CSKIKQILAMKACRSSIMIGTFLSKSKMKEIISNLSTLDKPWNCPHGRPTMRHLI 876
Cgla 834 CSKIRAMFAMRACRSSIMVGKPLNMRTMTRVVQNLSTLDKPWNCPHGRPTMRHLM 888
Cgui 790 CSKIRTILALRSCRSSIMIGQPLSTSTMKKVVHNLSHLDKPWNCPHGRPTMRHLT 844
Clus 803 CSKVDKMIALRACRRSIMIGQSLSKNTMAKVVRHLSRLEKPWNCPHGRPTMRHLA 857
Ctro 721 ISKIRKILAMKACRSSIMIGSSLKKSKMNEIVKNLSTLDKPWNCPHGRPTMRHLI 775
Dhan ---
Klac 855 CSKIRSMFAMRACRSSIMIGKPLSMRTMKKVVNNLSDLEKPWNCPHGRPTLRHLM 909
Kpol 889 CTKIRSMFAMRACRTSIMIGKPLTKKTMSKVVKHLSELHKPWNCPHGRPTMRHLM 943
Kthe 838 CSKIRAMHAMRACRSSIMVGRPLVKKSMLRVVRNLSELDKPWNCPHGRPTMRHLM 892
Kwal 830 CSKIRAMHAMRACRSSIMVGKPLIKKAMLRVVRNLSELDKPWNCPHGRPTMRHLM 884
Sbay 828 CSKIRSMFAMRACRSSIMIGKPLNRKTMSRVVHNLSGLDKPWNCPHGRPTMRHLM 882
Scas 787 CSKIRSMFAMRACRSSIMIGKPLTTKIMTRVVHHLGDLDKPWNCPHGRPTMRHLA 841
Sklu 810 CSKIRAMFAMRACRSSIMIGKPLTKKTMVGVVRHLSELDKPWNCPHGRPTMRHLM 864
Smik 806 CSKIRSMFAMRACRSSIMIGKPLNKKTMTRVVHNLSALDKPWNCPHGRPTMRHLM 860
Spar 812 CSKIRSMFAMRACRSSIMIGKPLNKKTMTRVVHNLSELDKPWNCPHGRPTMRHLM 866
Lelo 874 CSKIRKIVAMKACRSSIMIGSFLSKQRMQKVVANLSKLDKPWNCPHGRPTMRHLI 928
Psti 741 CSKIRNLLAMRACRSSIMIGQPLTRGRMTKVVQNLSQLDKPWNCPHGRPTMRHLV 795
Ylip 825 PKKVRDVFASRACRGSVMVGTALKEKEMDRIVRNLAGLDKPWNCPHGRPTMRHLM 879
Afum 919 PSKVRKMFAMRACRSSIMIGKSLTQTQMVRVVRNMGTIDKPWNCPHGRPTMRHLM 973
Nfis 920 PSKVRKMFAMRACRSSIMIGKSLTQTQMVRVVRNMGTIDKPWNCPHGRPTMRHLM 974
Acla 935 PSKVRKMFAMRACRSSIMIGKNLTQKQMVRVVRNMGTIDKPWNCPHGRPTMRHLM 989
Ater 928 PSKVRKMFAMRACRSSIMIGKTLTTKQMERVVRNMGTIDKPWNCPHGRPTMRHLM 982
Cimm 893 PGKVRKMFAMRACRSSIMIGKSLTVKQMERVVRHMGMIDKPWNCPHGRPTMRHLM 947
Ncra 977 PSKVRKMFAMRACRSSIMIGRALSRPQMEKVVRHMGEMEKPWNCPHGRPTMRHLC1031
Pans 887 PSKVRKMFAMRACRSSIMIGRALSGRQMERVVRNMGGMEKPWNCPHGRPTMRHLC 941
Fgra 910 PSKVRKMFASRACRSSVMIGKALTHGQMETLVRHMAELDKPWNCPHGRPTMRHLC 964
Fver 909 YGRVRQAMELPAVPTGLVGCEGVEGRQLARSIGILASVLE WQISSGR        955
Tree 822 PSKVRKMLAMRACRSSIMIGKAMTRSQMYTLVNHMGELDKPWNCPHGRPTLRHLS 876
Sscl 922             NRS                                        924
Snod 965 PSKVRKLLASRACRSSVMIGKTLKTARMREIVRHMGSMDKPWSCPHGRPTMRHLF1019
Spom 735 SSRLERMLASKACRSSVMIGRALTISEMNTIVRHLAELSKPWNCPHGRPTMRHLL 789
```

Fig S5D

N-terminal domain
Mlh3 (1-375)
Pms1 (1-361)

```
Mlh3   1 MSQHIRKLDSDVSERLKSQACTVSLASAVREIVQNSVDAHATTIDVMI-D    49
         |:|..:..|.|| .|:.|......|.:||:|:|.||:||:|..|:::. |
Pms1   1 MTQIHQINDIDV-HRITSGQVITDLTTAVKELVDNSIDANANQIEIIFKD    49

      50 LPNLSFAVYDDGIGLTRSDLNILATQNYTSKIRKMNDLVTMKTYGYRGDA    99
         ....|....|:|.|:..|:...||.::|||||.|..|:..::|.|:||:|
      50 YGLESIECSDNGDGIDPSNYEFLALKHYTSKIAKFQDVAKVQTLGFRGEA    99

     100 LYSISNVSNLFVCSKKKDYNSAWMRKFPSKSVMLSENTILPIDPFWKICP    149
         |.|:..::.|.|.:...          |.|:..|..:.:..|.
     100 LSSLCGIAKLSVITTTS----------PPKADKLEYDMVGHIT-------    132

     150 WSRT----KSGTVVIVEDMLYNLPVRRRILKEEPPFKTFNTIKADMLQIL    195
          |:|    ..||.|:|..:.:|||||::...     |||.......|.::
     133 -SKTTTSRNKGTTVLVSQLFHNLPVRQKEFS-----KTFKRQFTKCLTVI    176

     196 VMHPMISLNVQY-----TDKLRINTEVLFRSKNITEGLTKHQQMSQVLRN    240
         ..:.:|:..:::     |.|         ..||:.....:.:..|.:.:.:
     177 QGYAIINAAIKFSVWNITPK---------GKKNLILSTMRNSSMRKNISS    217

     241 VFGA--------------IIP------------PDMLKKVSLKFNEYQIE    264
         ||||              :.|            ||.|   .|.: :.:::
     218 VFGAGGMRGLEEVDLVLDLNPFKNRMLGKYTDDPDFL---DLDY-KIRVK    263

     265 GIISKMPVGL----KDLQFIYINGRRYADSAFQGYVDSLFQAQDFGEKGM    310
         |.||:...|.    ||.||||:|.|....|
     264 GYISQNSFGCGRNSKDRQFIYVNKRPVEYS---------------------    293

     311 SLLK-TKSVGKPYRS--HPVFILDVRCPQTIDDLLQDPAKKIVKPSHIRT    357
         :||| ...|.|.:.:  .|...|::..|.::.|:...|.|:::...:.|.
     294 TLLKCCNEVYKTFNNVQFPAVFLNLELPMSLIDVNVTPDKRVILLHNERA    343

     358 IEPLIVKTIRSFLTFQGY    375
         :..:...|:..:...|..
     344 VIDIFKTTLSDYYNRQEL    361
```

24.6% identity over 418 amino acid alignment with gaps

Linker
Mlh3 (376-488)
Pms1 (362-638)

```
Mlh3   1 --------------------------------------------------      0
Pms1   1 ALPKRMCSQSEQQAQKRLKTEVFDDRSTTHESDNENYHTARSESNQSNHA     50

         1 --------------------------------------------------      0

      51 HFNSTTGVIDKSNGTELTSVMDGNYTNVTDVIGSECEVSVDSSVVLDEGN    100

         1 --------------------------------------------------      0

     101 SSTPTKKLPSIKTDSQNLSDLNLNNFSNPEFQNITSPDKARSLEKVVEEP    150

         1 ----------------------------------------------LTP      3
                                                          ...
     151 VYFDIDGEKFQEKAVLSQADGLVFVDNECHEHTNDCCHQERRGSTDTEQD    200

         4 DKSDSSF---EIVNCSQKTATLPDSRIQISKRN-QVLNSKMKIARINSYI     49
           |::||.:   |.|.:.:.:| .|.:||..|||.| :.|:..:...:....|
     201 DEADSIYAEIEPVEINVRT-PLKNSRKSISKDNYRSLSDGLTHRKFEDEI    249

      50 GKPAVNGCRINNSTINYEKI-KNIRIDGQKSRLRNKLSSRPYDSGFTEDY     98
                ...|.||.|:::| ||.:   |.|.:.:|..
     250 -------LEYNLSTKNFKEISKNGK---QMSSIISKRK------------    277

      99 DSIGKTITDFSISRS    113

     278 ---------------    277
```

7.9% identity over 315 amino acid alignment with gaps

C-terminal domain
Mlh3 (489-715)
Pms1 (639-873)

```
Mlh3   1 --------------------------------VLAKYEVINQVDKKFILIR     19
                                         ...|.||:.|.:..||::
Pms1   1 SEAQENIIKNKDELEDFEQGEKYLTLTVSKNDFKKMEVVGQFNLGFIIV-     49

      20 CLDQSIHNCPLLVLVDQHACDERIRLEEL-----FYSLLTEVVTGTFVAR     64
           .:.:.|...|.:|||||.||:...|.|     |.|  .:::....|..
      50 --TRKVDNKYDLFIVDQHASDEKYNFETLQAVTVFKS--QKLIIPQPVEL     95

      65 DLKDCCIEVDRT---EADLFKHYQSEFKKWGIGYETIEGTMETSLLEIKT     111
         .:.|..:.:|..   |.:.||....|.:::|            |.::::.:
      96 SVIDELVVLDNLPVFEKNGFKLKIDEEEEFG-----------SRVKLLS     133

     112 LPEMLTSKYNGDKDYLKMVLLQHAHDLKDFKKLPMDLSHFENYTSVDKLY     161
         ||   |||             |...||.||.:| :.|...:.....|.:.
     134 LP---TSK-------------QTLFDLGDFNEL-IHLIKEDGGLRRDNIR     166

     162 WWKYSSCVPTVFHEILNSKACRSAVMFGDELTRQECIILISKLSRCHNPF     211
         ..|..|        :...:||||::|.|..|.::....::..||....|:
     167 CSKIRS--------MFAMRACRSSIMIGKPLNKKTMTRVVHNLSELDKPW     208

     212 ECAHGRPSMVPIAELK-----------    227
         .|.||||:|..:.||:
     209 NCPHGRPTMRHLMELRDWSSFSKDYEI    235
```

22.0% identity over 277 amino acid alignment with gaps

Figure S6

Table S1. Plasmids used in this study.

Plasmid	Promoter	Relevant Genotype	Source
pRS415		*ARS-CEN, LEU2*	Christianson et al. (1992)
pRS425		*2μ, LEU2*	Christianson et al. (1992)
pJH481	*PMS1*	*PMS1$_{S288c}$, ARS-CEN, LEU2*	James Haber
pEAA238	*PMS1*	*PMS1$_{S288c}$, ARS-CEN, HIS3*	
pEAM50	*PMS1*	*PMS1$_{S288c}$, 2μ, LEU2*	
pEAM65	*MLH3*	*MLH3, 2μ, LEU2*	
pEAM168	*MLH3*	*MLH3::KanMX, 2μ, URA3*	
pEAI254	*MLH3*	*MLH3::KanMX*	
pEAA566	*MLH3*	*MLH3, ARS-CEN, LEU2*	
pEAA636	*MLH3*	*MLH3, ARS-CEN, HIS3*	
Block mutations in *MLH3:* Integration vectors			
pEAM313	*MLH3*	Block 1, ATP binding-*mlh3-K17T,A20Q,S24D,R30K,Q34D ::KanMX*	
pEAM314	*MLH3*	Block 2, Mlh1 interaction-*mlh3-Y493M,N497G,V499F, D500N,K502G::KanMX*	
pEAM320	*MLH3*	Block 2-*mlh3-D500N*::*KanMX*	
pEAM321	*MLH3*	Block 2-*mlh3-K502G::KanMX*	
pEAM324	*MLH3*	Block 3, Endonuclease motif-*mlh3-R530K::KanMX*	
pEAM322	*MLH3*	Block 3-*mlh3-R532N::KanMX*	
pEAM325	*MLH3*	Block 3-*mlh3-R530K,R532N::KanMX*	
pEAI439	*MLH3*	Block 4, PCNA interaction motif-*mlh3-PIP1::KanMX*	
pEAI440	*MLH3*	Block 4-*mlh3-PIP2::KanMX*	
pEAM315	*MLH3*	Block 5, Helix 2-*mlh3-V660K,N666A,F676I,D678K:: KanMX*	
pEAM323	*MLH3*	Block 5-*mlh3-D678K::KanMX*	
pEAI479	*MLH3*	Block 6, Helix 1-*mlh3-C695L,F699W,A702P,S707T, V709R,P710H::KanMX*	
pEAM316	*MLH3*	Block 1 (*ATP binding*), Block 4 *(PIP2*)*::KanMX* combination	
pEAM326	*MLH3*	Block 5 *(Helix 2),* Block 6 (*Helix 1*)*::KanMX* combination	
Block mutations in *MLH3: 2μ, LEU2* expression vectors			
pEAM303	*MLH3*	Block 1-*mlh3-K17T,A20Q,S24D,R30K,Q34D*	
pEAM304	*MLH3*	Block 2-*mlh3-Y493M,N497G,V499F,D500N,K502G*	
pEAM307	*MLH3*	Block 2-*mlh3-D500N*	
pEAM308	*MLH3*	Block 2-*mlh3-K502G*	
pEAM311	*MLH3*	Block 3-*mlh3-R530K*	
pEAM309	*MLH3*	Block 3-*mlh3-R532N*	
pEAM312	*MLH3*	Block 3-*mlh3-R530K,R532N*	
pEAM278	*MLH3*	Block 4-*mlh3-PIP1::KanMX*	
pEAM279	*MLH3*	Block 4-*mlh3-PIP2::KanMX*	
pEAM291	*MLH3*	Block 4-*mlh3-PIP3::KanMX*	
pEAM285	*MLH3*	Block 4-*mlh3-PIP4::KanMX*	

pEAM287	*MLH3*	Block 4-*mlh3-PIP5::KanMX*
pEAM286	*MLH3*	Block 4-*mlh3-PIP6::KanMX*
pEAM288	*MLH3*	Block 4-*mlh3-PIP7::KanMX*
pEAM289	*MLH3*	Block 4-*mlh3-PIP8::KanMX*
pEAM290	*MLH3*	Block 4-*mlh3-PIP9::KanMX*
pEAM305	*MLH3*	Block 5-*mlh3-V660K,N666A,F676I,D678K*
pEAM310	*MLH3*	Block 5-*mlh3-D678K*
pEAM306	*MLH3*	Block 6-*mlh3-C695L,F699W,A702P,S707T,V709R,P710H*
pEAM316	*MLH3*	Block 1 (*ATP binding*), Block 4 *(PIP2*) combination
pEAM292	*MLH3*	Block 2 (*D500N),* Block 4 *(PIP2)* combination
pEAM293	*MLH3*	Block 2 (*K502G),* Block 4 *(PIP2)* combination
pEAM283	*MLH3*	Block 3 (*R530K*), Block 4 (*PIP2*) combination
pEAM294	*MLH3*	Block 3 (*R532N),* Block 4 *(PIP2)* combination
pEAM284	*MLH3*	Block 3 (*R530K, R532N),* Block *4 (PIP2)* combination
pEAM295	*MLH3*	Block 4 (*PIP*2), Block 5 *(D678K)* combination
pEAM319	*MLH3*	Block 5 (*Helix 2*), Block 6 *(Helix1)* combination

Block mutations in *MLH3: ARS-CEN, HIS3* expression vectors

pEAA701	*MLH3*	Block 4-*mlh3-PIP1*
pEAA702	*MLH3*	Block 4-*mlh3-PIP2*

Chimera constructs containing *MLH3(M)* and *PMS1(P)* domains

pEAI398	*MLH3*	*PMM::KanMX*
pEAI399	*MLH3*	*PPP::KanMX*
pEAI400	*MLH3*	*MMP::KanMX*
pEAI401	*MLH3*	*PPM::KanMX*
pEAI402	*MLH3*	*MPP::KanMX*
pEAI404	*MLH3*	*PMP::KanMX*
pEAI403	*MLH3*	*MPM::KanMX*
pEAM296	*MLH3*	*PMM::KanMX, 2μ, LEU2*
pEAM297	*MLH3*	*PPP::KanMX, 2μ, LEU2*
pEAM298	*MLH3*	*MMP::KanMX, 2μ, LEU2*
pEAM299	*MLH3*	*PPM::KanMX, 2μ, LEU2*
pEAM300	*MLH3*	*MPP::KanMX, 2μ, LEU2*
pEAM301	*MLH3*	*PMP::KanMX, 2μ, LEU2*
pEAM302	*MLH3*	*MPM::KanMX, 2μ, LEU2*

Unless indicated, *MLH3* and *PMS1* genes were derived from the SK1 strain background and all plasmids with the pEA plasmid designation were built in the Alani laboratory (Materials and Methods). The designation of the *M* (*MLH3*-derived) and *P* (*PMS1*-derived) abbreviations for the *MLH3-PMS1* chimeras, and the detailed description of the block mutations can be found in the Materials and Methods and Figures 2 and S3.

Table S2. Strains used in this study.

Parental S288c strains (Tables 1 and 2)	
EAY1269	*MATa, ura3-52, leu2Δ1, trp1Δ63, lys2::insE-A$_{14}$*
EAY4595	EAY1269*, mlh3Δ::KanMX*
EAY3097	*MATa, ura3-52, leu2Δ1, trp1Δ63, his3Δ200, lys::insE-A$_{14}$, pms1Δ::KanMX4*
Parental SK1 strains (Table 1)	
EAY3252	*MATα ho::hisG, ura3, leu2::hisG, trp1::hisG, ADE2, HIS4, CEN8Tomato::LEU2, MLH3, lys2::insE-A$_{14}$*
EAY3255	*MATα ho::hisG, ura3, leu2::hisG, trp1::hisG, ADE2, his4xB, CEN8Tomato::LEU2, mlh3Δ::NATMX, lys2::insE-A$_{14}$*
EAY3486	*MATa ho::LYS2, lys2, ura3, leu2::hisG, trp1::hisG, THR1::m-Cerulean-TRP1, mlh3Δ::NatMX*
Block mutation integrations at the *MLH3* locus in EAY3255 (Table 1)	
EAY4563-4564	Block 1 (ATP binding)-*mlh3-K17T,A20Q,S24D,R30K,Q34D::KanMX*
EAY4565-4567	Block 2 (Mlh1 interaction)-*mlh3-Y493M,N497G,V499F,D500N,K502G::KanMX*
EAY4575-4577	Block 2-*mlh3-D500N::KanMX*
EAY4578-4580	Block 2-*mlh3-K502G::KanMX*
EAY4581-4583	Block 3 (Endonuclease motif)-*mlh3-R532N::KanMX*
EAY4587-4589	Block 3-*mlh3-R530K::KanMX*
EAY4590-4591	Block 3-*mlh3-R530K,R532::KanMX*
EAY4121-4123	Block 4 (PCNA interaction motif)-*mlh3-PIP1::KanMX*
EAY4124-4126	Block 4 -*mlh3-PIP2::KanMX*
EAY4568-4570	Block 5 (Helix 2)-*mlh3-V660K,N666A,F676I,D678K::KanMX*
EAY4584-4586	Block 5-*mlh3-D678K::KanMX*
EAY4571-4572	Block 6 (Helix 1)-*mlh3-C695L,F699W,A702P,S707T,V709R,P710H::KanMX*
EAY4573-4574	Block 1 (*ATP binding),* Block 4 *(PIP2)::KanMX* combination
EAY4592-4594	Block 5 (*Helix 2*), Block 6 (*Helix 1*)*::KanMX* combination
Chimera integrations in the *MLH3* locus in EAY3255 (Table 1)	
EAY3937	*PMM::KanMX*
EAY3938	*PPP::KanMX*
EAY3939	*MMP::KanMX*
EAY3940	*PPM::KanMX*
EAY3941	*MPP::KanMX*
EAY3943	*PMP::KanMX*
EAY3942	*MPM::KanMX*

lys2-A$_{14}$ reversion assays were performed on the SK1 strains EAY3252 (*wild-type*), EAY3255 (*mlh3Δ*), and EAY3255 derivatives containing integrated *mlh3$_{SK1}$::KanMX* alleles (Table 1). Meiotic crossover assays were performed on the following diploids: EAY3252/EAY3486 (*wild-type*), EAY3255/EAY3486 (*mlh3Δ*), and EAY3255::*mlh3* alleles/EAY3486 (Table 1). *lys2-A$_{14}$* reversion assays were performed on the S288c strains EAY1269, EAY4595::*mlh3Δ,* and EAY3097 transformed with plasmids as indicated (Tables 1 and 2).

Table S3. Pearson's Chi-Squared Contingency Test for meiotic crossing over data presented in Table 1.

mlh3 allele	PD	TT	PD+TT	TT fraction	P-value to *MLH3*	P-value to null	Phenotype
MLH3	641	382	1023	0.37	**1**	<0.0001	Plus
mlh3Δ	1002	227	1229	0.18	<0.0001	**1**	minus
PPP (PMS1)	451	98	549	0.18	<0.0001	0.7548	minus
PMM	431	102	533	0.19	<0.0001	0.7415	minus
MMP	412	112	524	0.21	<0.0001	0.1588	minus
PPM	661	163	824	0.20	<0.0001	0.4578	minus
MPP	650	187	837	0.22	<0.0001	0.0309	minus
PMP	1020	267	1287	0.21	<0.0001	0.1509	minus
MPM	940	256	1196	0.21	<0.0001	0.0705	minus
mlh3-K17T, A20Q, S24D, R30K, Q34D	330	183	513	0.36	0.5224	<0.0001	Plus
mlh3-Y493M, N497G, V499F, D500N, K502G	545	224	769	0.29	0.0003	<0.0001	Plus/minus
mlh3-D500N	536	265	801	0.33	0.0593	<0.0001	Plus
mlh3-K502G	616	145	761	0.19	<0.0001	0.7456	minus
mlh3-R530K	576	195	771	0.25	<0.0001	0.0003	Plus/minus
mlh3-R532N	546	228	774	0.29	0.0005	<0.0001	Plus/minus
mlh3-R530K,R532N	405	104	509	0.20	<0.0001	0.3432	minus
mlh3-PIP1	513	229	742	0.31	0.0047	<0.0001	Plus/minus
mlh3-PIP2	498	294	792	0.37	0.9234	<0.0001	Plus
mlh3-V660K,N666A,F676I,D678K	622	138	760	0.18	<0.0001	0.8612	minus
mlh3-D678K	524	248	772	0.32	0.0219	<0.0001	Plus/minus
mlh3-C695L,F699W,A702P, S707T,V709R,P710H	434	103	527	0.20	<0.0001	0.7246	minus
Block 1 (*ATP binding*), Block 4 (*PIP2*)	348	161	509	0.32	0.0277	<0.0001	Plus/minus
Block 5 (*Helix 2*), Block 6 (*Helix 1*)	422	92	514	0.18	<0.0001	0.7784	minus

P-values were determined using a Pearson's Chi-Squared contingency test, with a Benjamini-Hochberg correction for 32 comparisons. A $p<0.002$ cut off was used for statistical significance.

Table S4. Uniprot, NCBI, and RefSeq identifications for the genes used to perform the phylogenetic analysis in Figure 1B.

	UniProt	Gene IC (NCBI)	RefSeq (NCBI)
E. coli mutL	P23367	948691	
B. subtilis mutL	P49850	939455	
S. cerevisiae MLH1	P38920	855203	
S. cerevisiae MLH2	Q07980	850722	
S. cerevisiae MLH3	Q12083	855939	
S. cerevisiae PMS1	P14242	855642	
A. thaliana MLH1	Q9ZRV4	826493	
A. thaliana MLH3	F4JN26	829704	
A. thaliana PMS1	Q941I6	827997	
D. discoideum MLH1	Q54KD8		XP_637285.1
D. discoideum PMS1	1Q54QA0		XP_638844.1
H. sapien MLH1	P40692	4292	
H. sapien PMS1	P54277	5378	
H. sapien PMS2	P54278	5395	
H. sapien MLH3	Q9UHC1	27030	
C. cinerea MLH1	A8PCM6		XP_001840436.2
C. cinerea MLH3	D6RMZ1		XP_002911247.1
C. cinerea PMS1	A8P227		XP_001838229.2