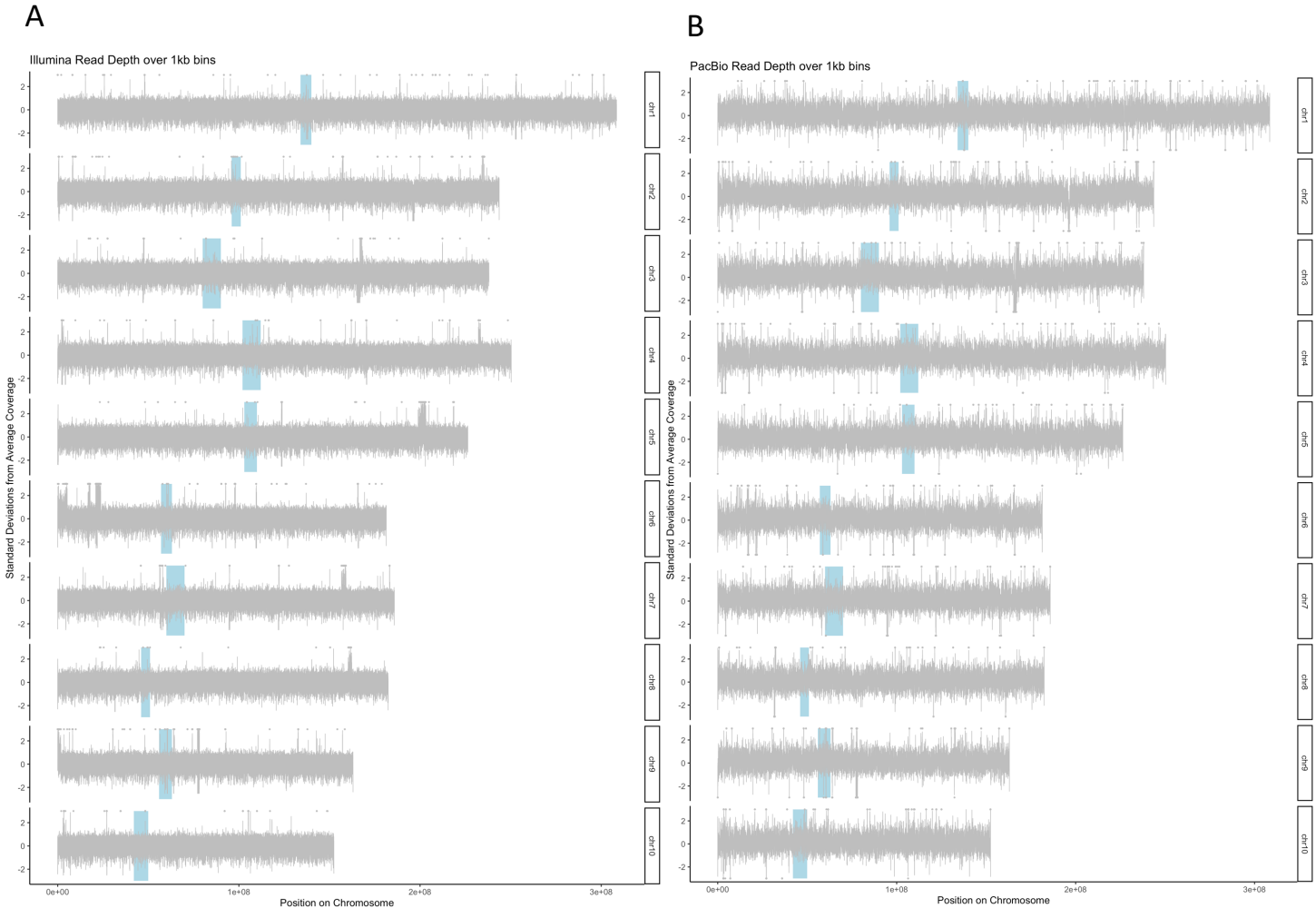
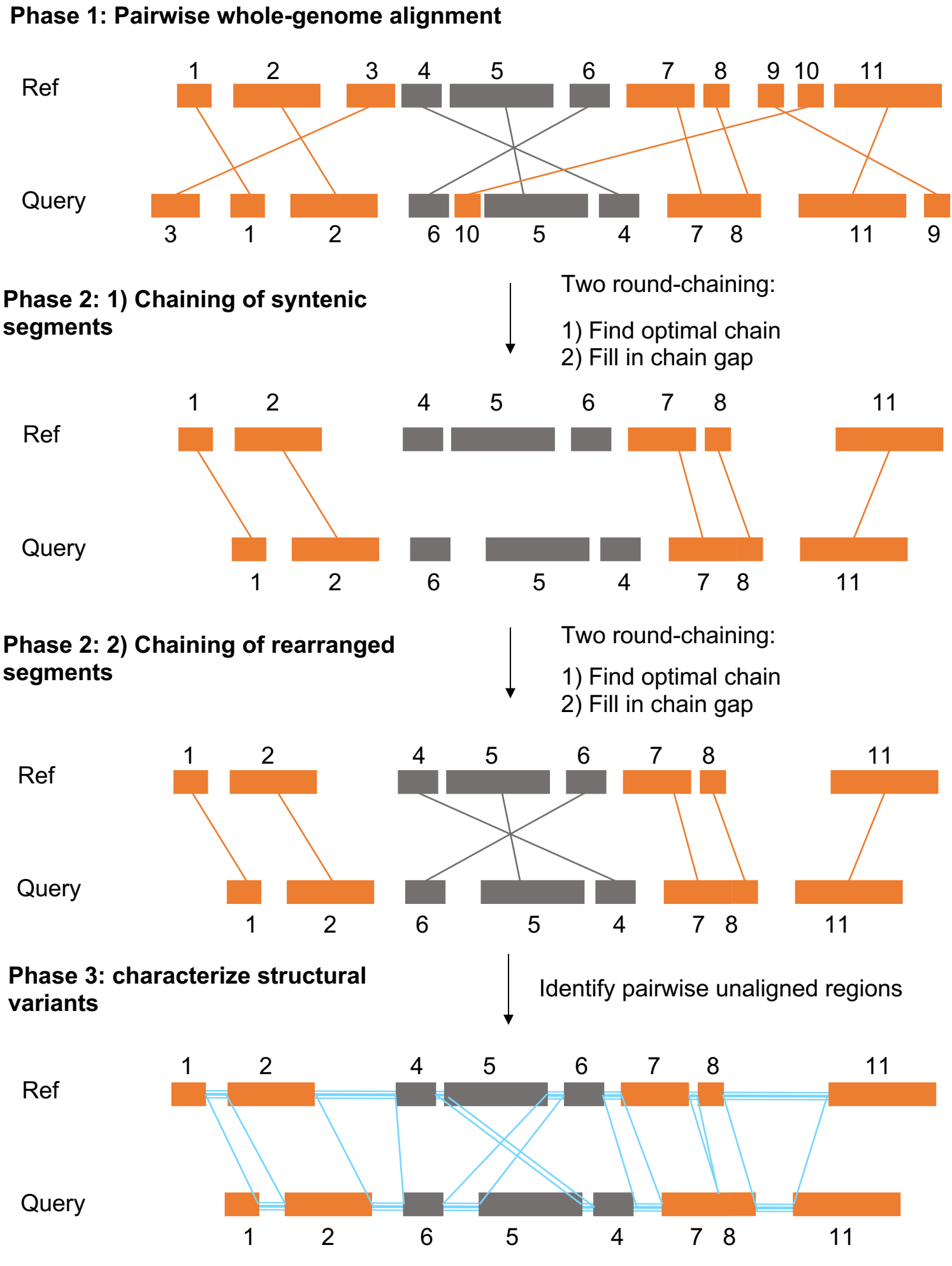
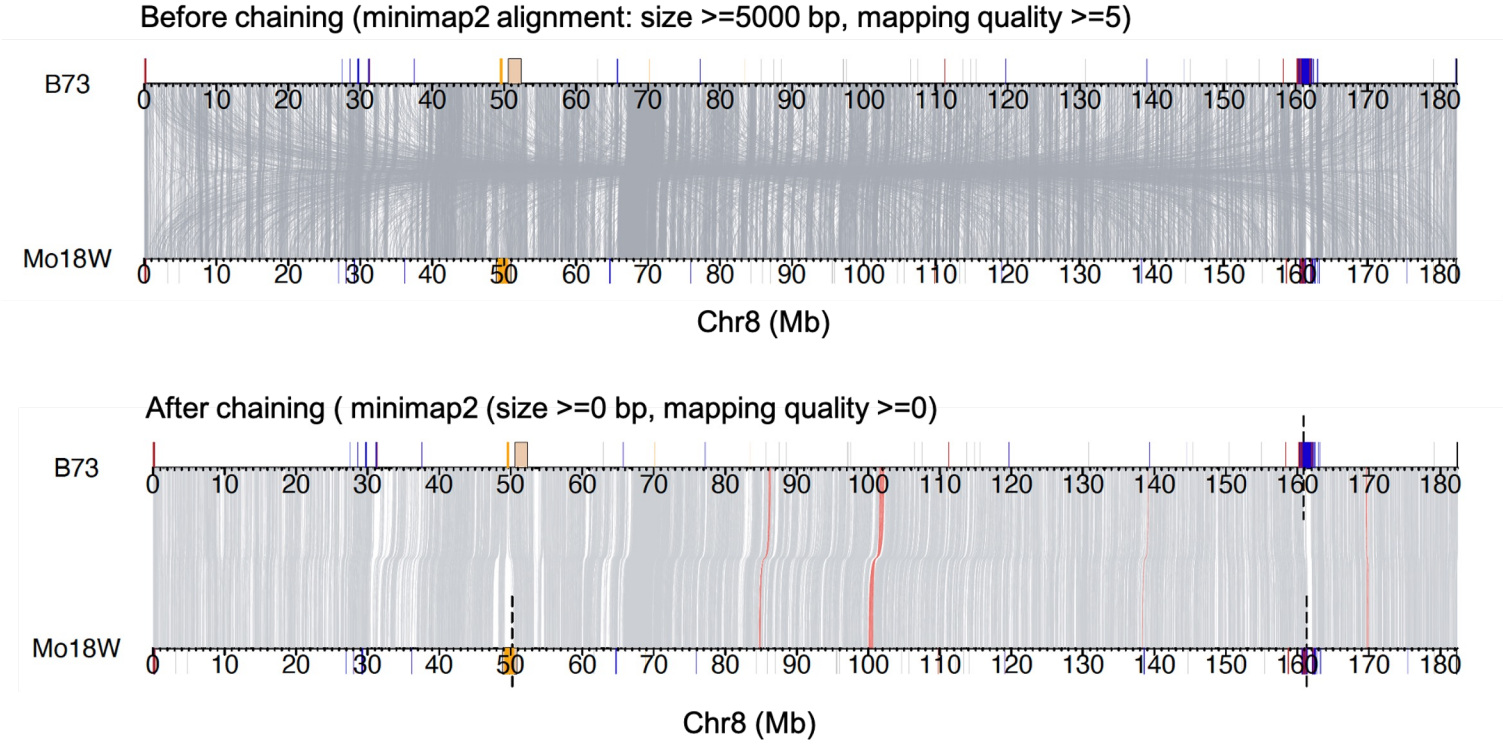


Supplemental Figure S1. Read depth distribution across the Zm-B73-REFERENCE-NAM-5.0 genome assembly. **A**) Depth of Illumina reads across all chromosomes. **B**) Depth of PacBio reads across all chromosomes. Values are averages over 1 kb bins, plotted as their standard deviations from the mean coverage of the whole genome. Bins with values greater than or less than three standard deviations are marked with dots at +/- 3 standard deviations.

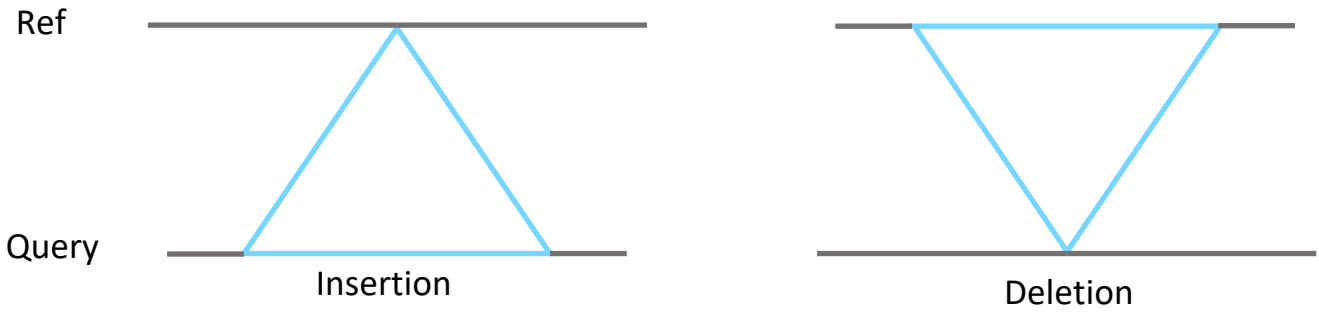


Supplemental Figure S2. Workflow for synteny detection and structural variant characterization. Syntenic and inverted genomic segments are respectively colored in orange and grey, and variant regions are noted with cyan lines.

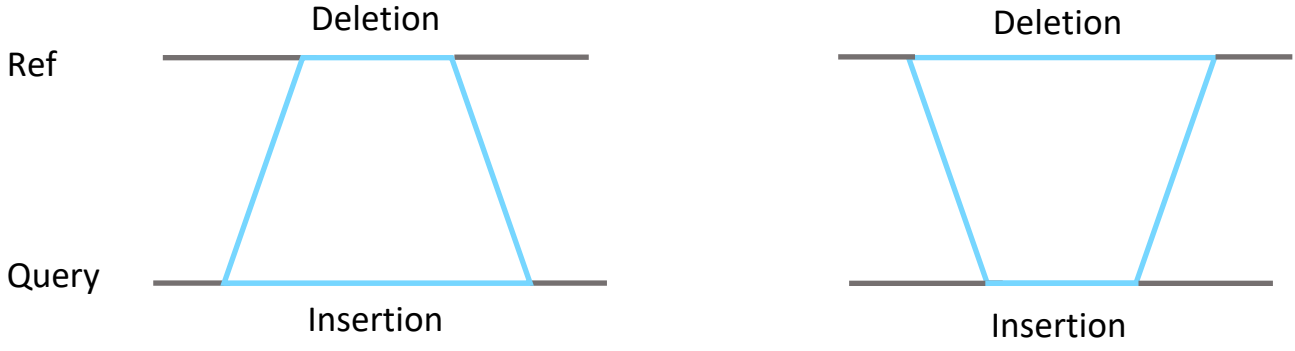


Supplemental Figure S3. Spurious alignment removal through chaining. The upper panel shows the alignment between B73 and Mo18W chromosome 8 before chaining, where alignment blocks with a size above 5 kb and a mapping quality higher than 5 are shown as links. The lower panel depicts alignments after chaining, where no size or mapping quality was applied. Inverted regions are highlighted in red. CENH3 Chip-seq regions are marked by yellow/tan boxes, and CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars above and below the alignments. Dashed lines indicate 100N gaps, which were introduced through pseudomolecule construction prior to chaining.

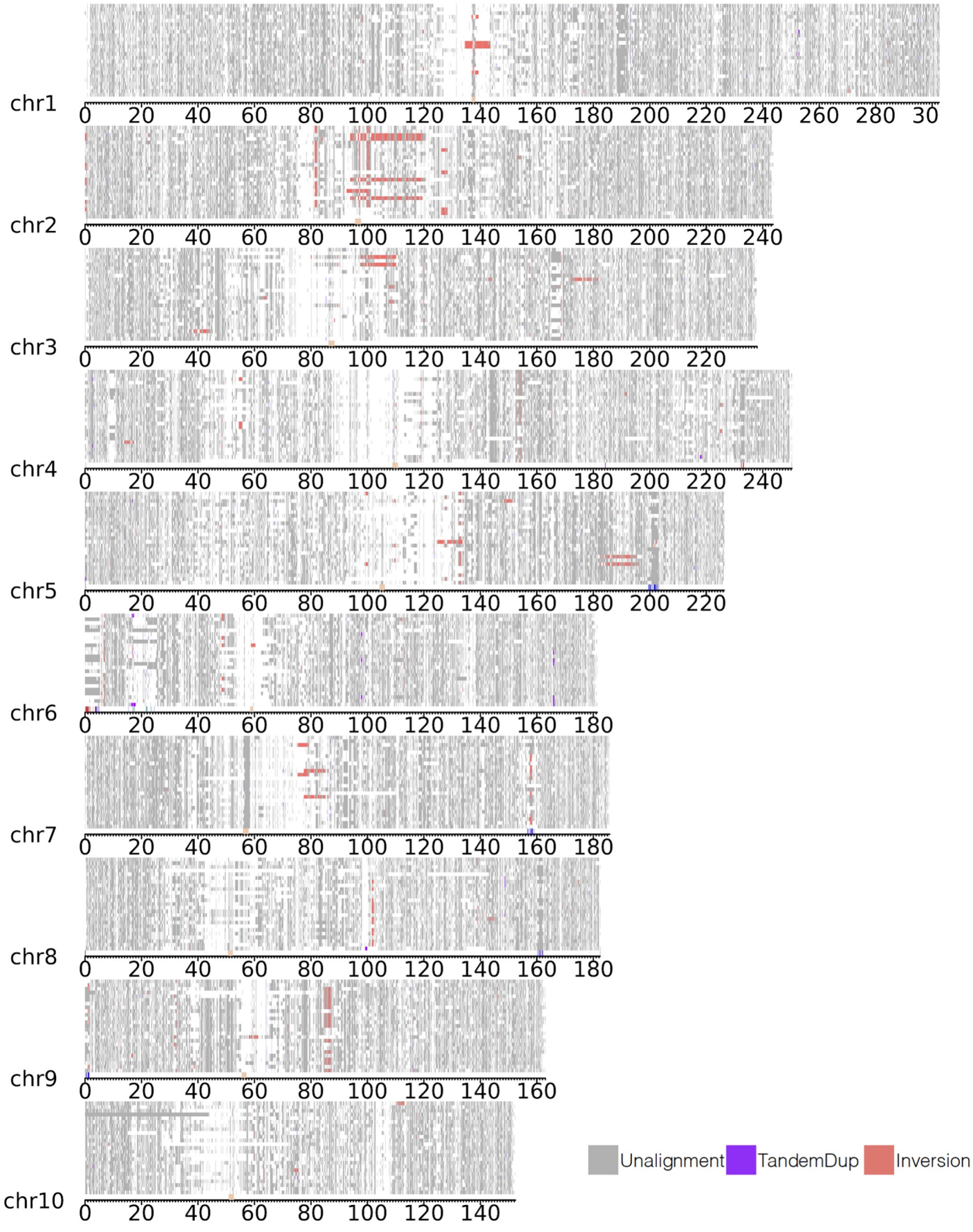
A. Indels with single-junctions



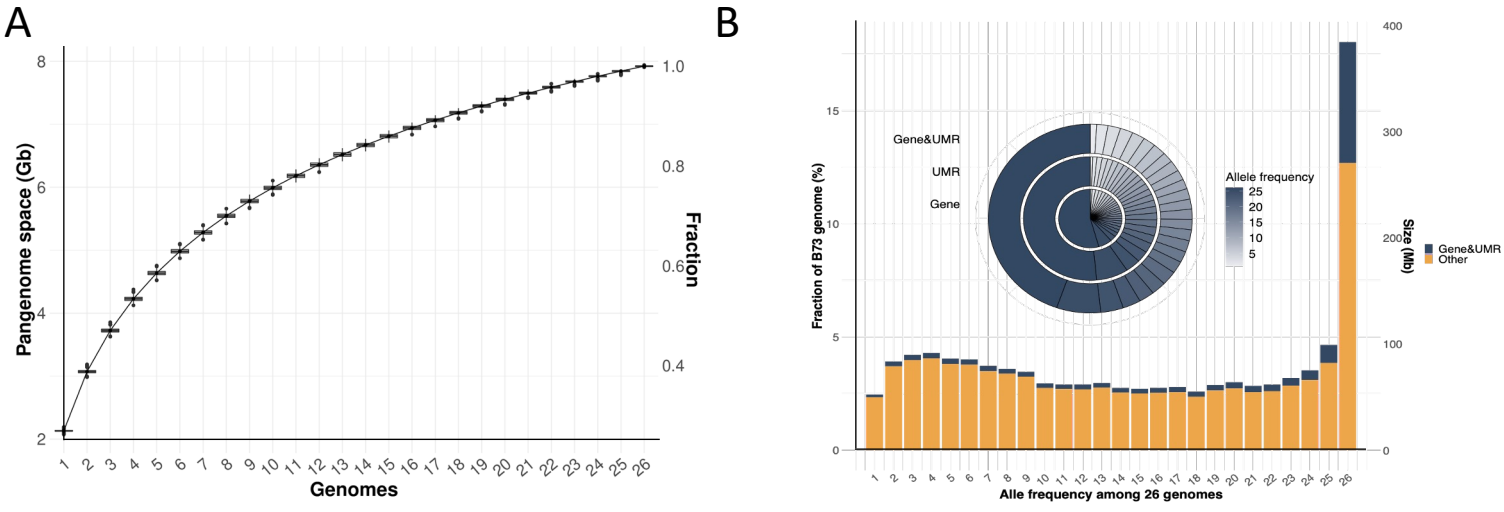
B. Pairwise un-alignment patterns



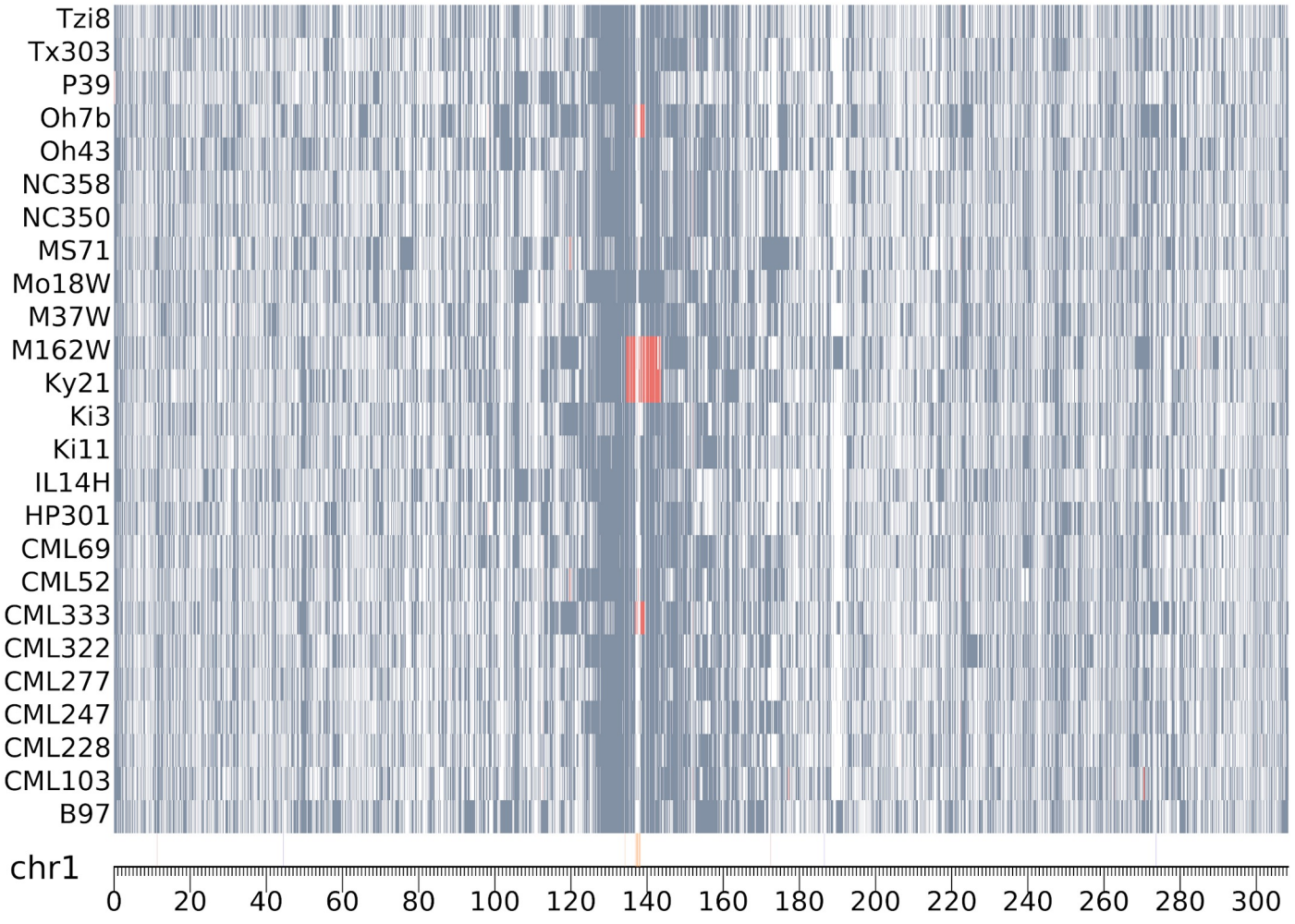
Supplemental Figure S4. SV calling. **A)** Insertions and deletions with a single-junction feature. **B)** Insertions and deletions with an unalignment pattern. Real-world examples of pairwise unalignment can be seen in Figure 1C.

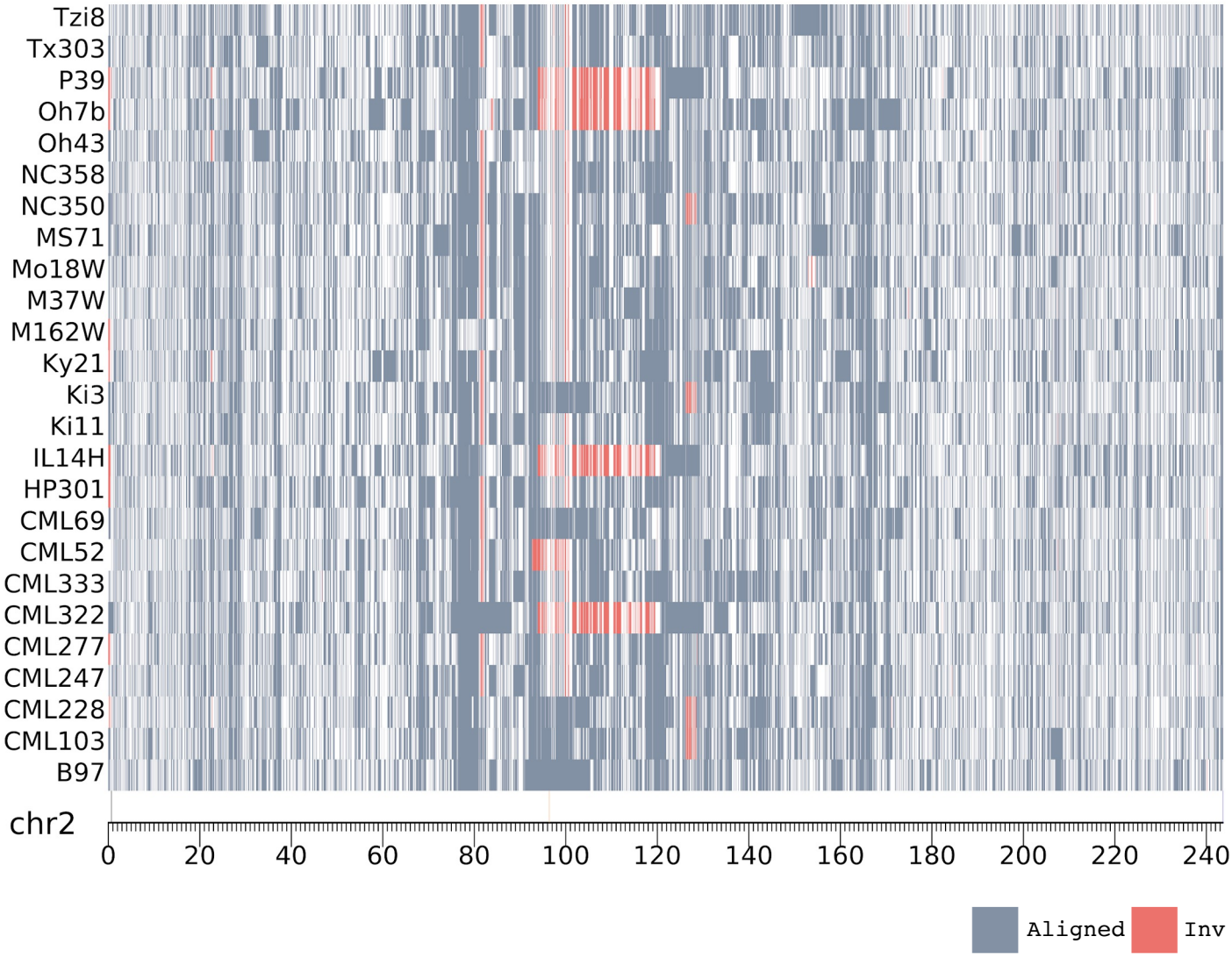


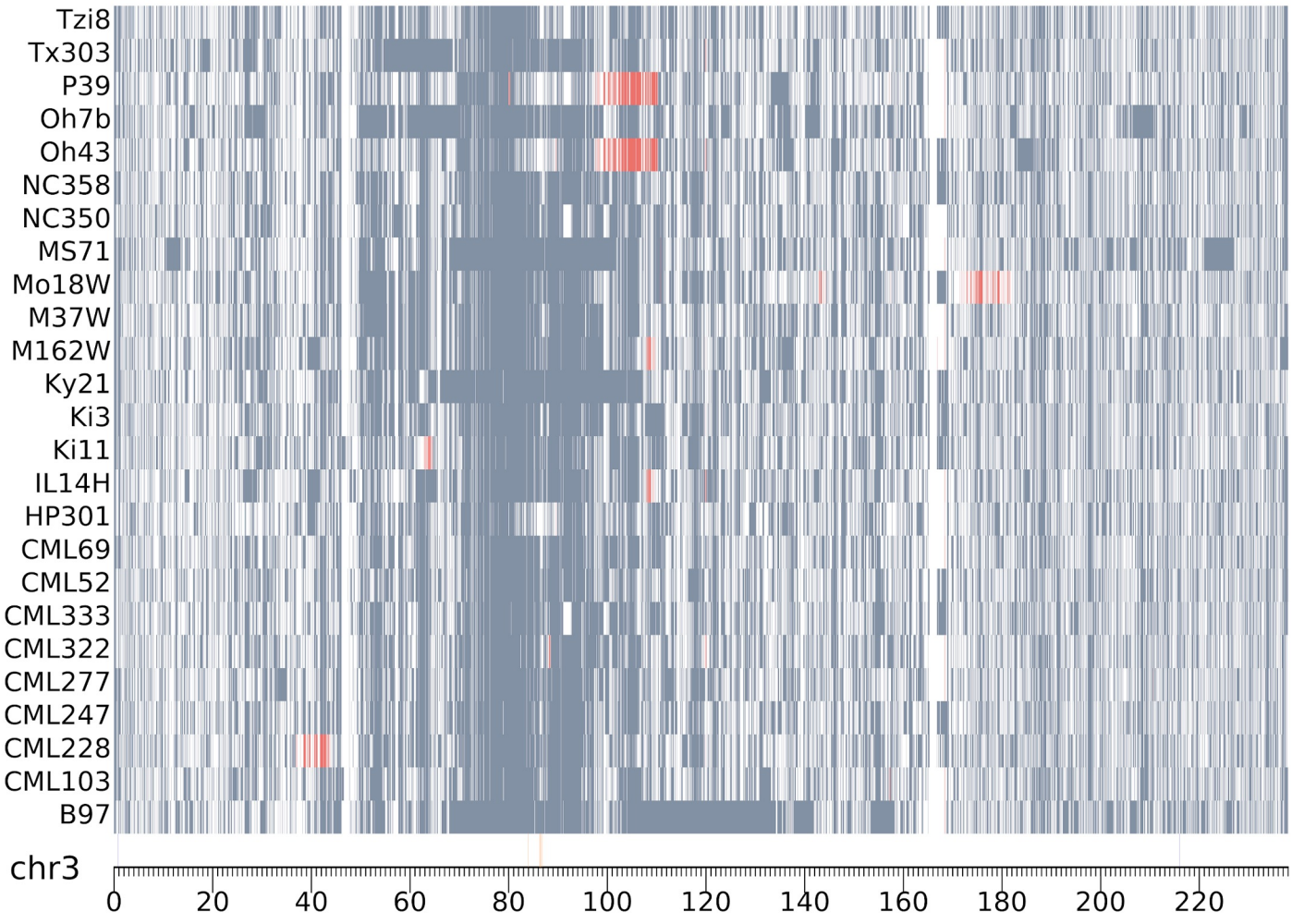
Supplemental Figure S5. Structural variants detected between NAM lines and B73 reference. The bottom to top tracks represent lines: B97, CML103, CML228, CML247, CML277, CML322, CML333, CML52, CML69, HP301, IL14H, Ki11, Ki3, Ky21, M162W, M37W, Mo18W, MS71, NC350, NC358, Oh43, Oh7b, P39, Tx303, and Tzi8. CENH3 Chip-seq regions are marked by yellow/tan boxes, and CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars above and below the alignments.

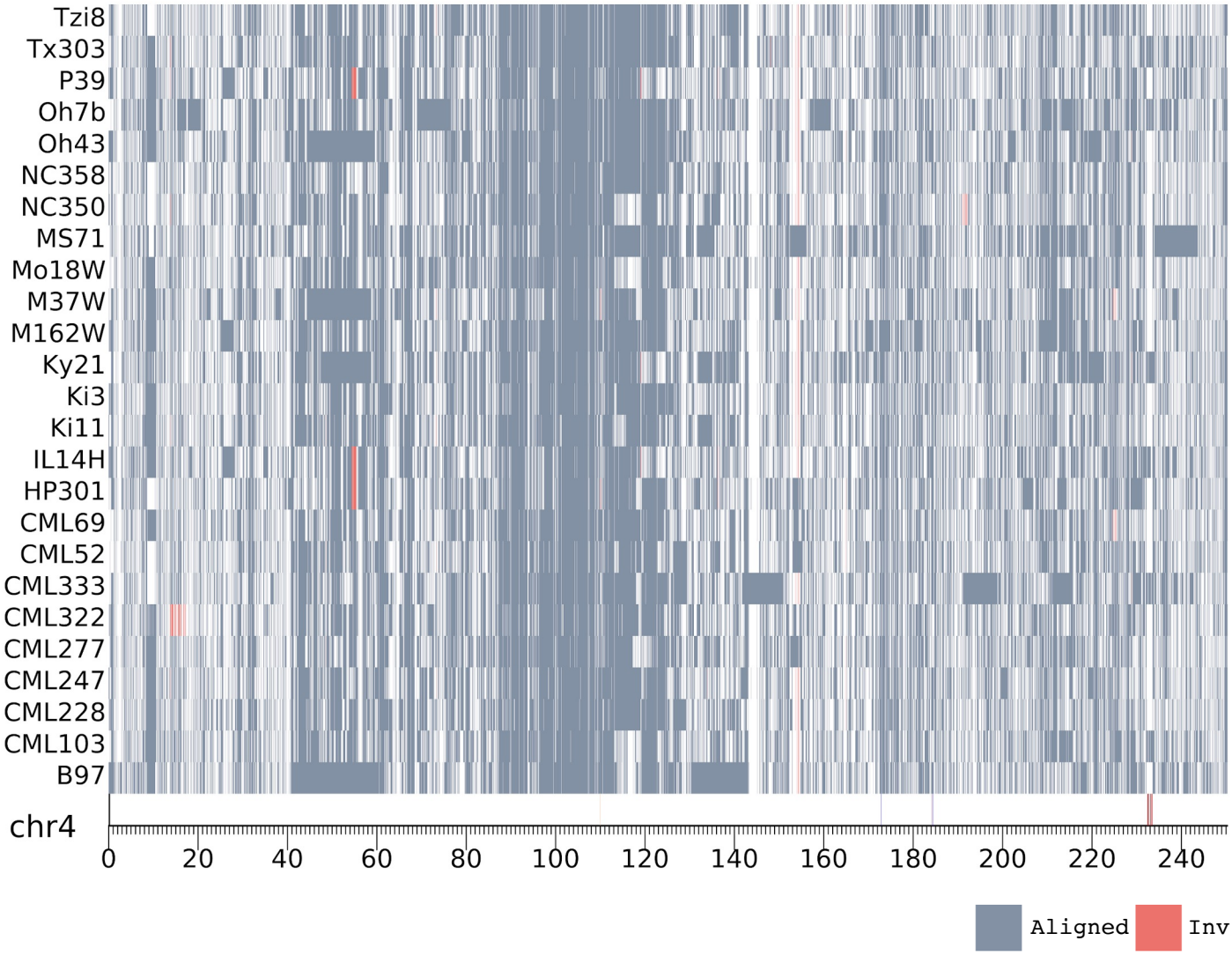


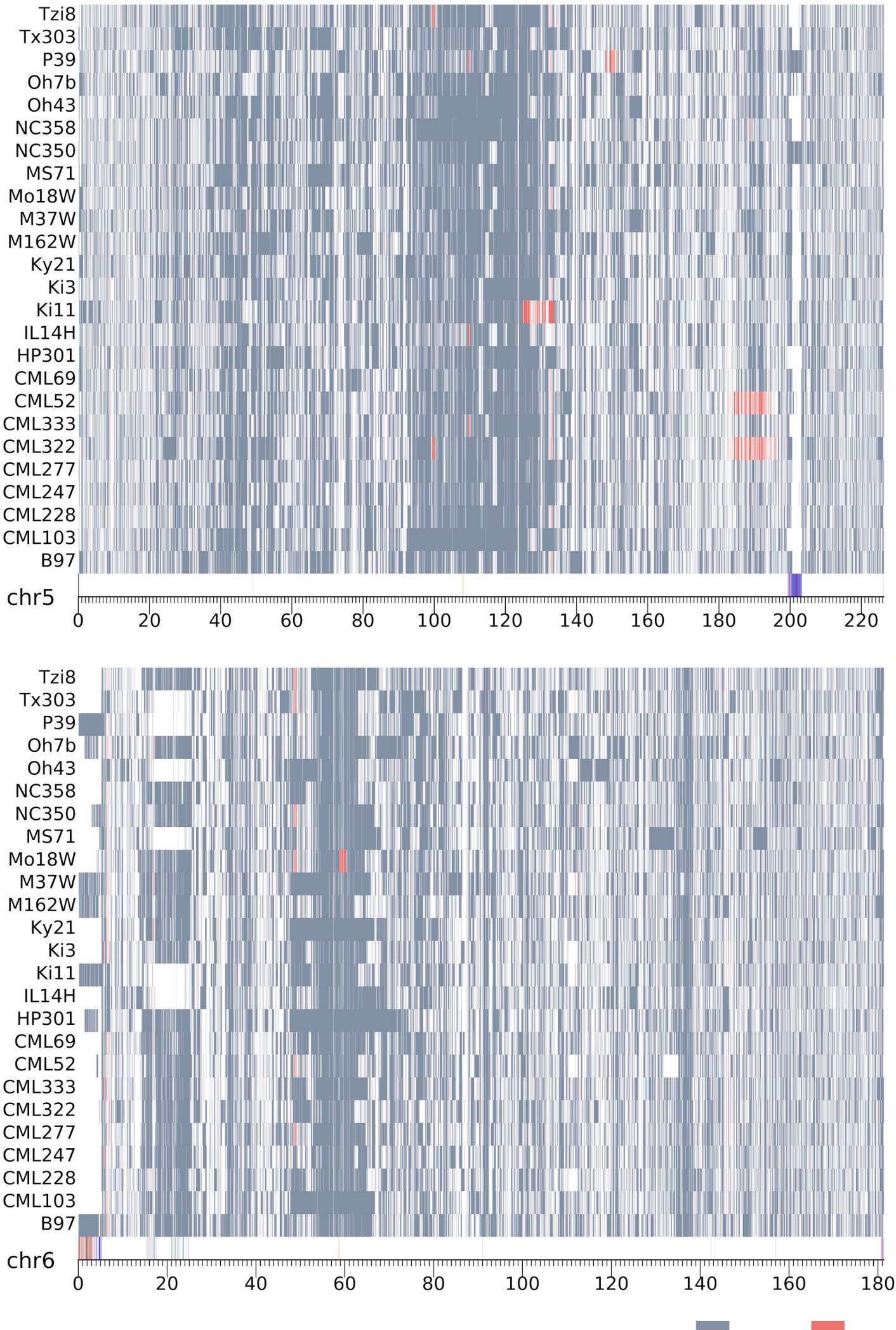
Supplemental Figure S6. Pangenome size and allele frequency of B73 segments across 26 lines. **A)** Pan-genome space with each additional line. The error bar is derived from 1000 times shuffling of the order for input genomes. **B)** Allele frequency of B73 segments among 26 lines. Genes and unmethylated regions (UMRs) highlighted in the bar plot (blue) were summarized in the pie chart, where tracks from inner to outer circle represent genes, UMRs, and total space of gene and UMRs.

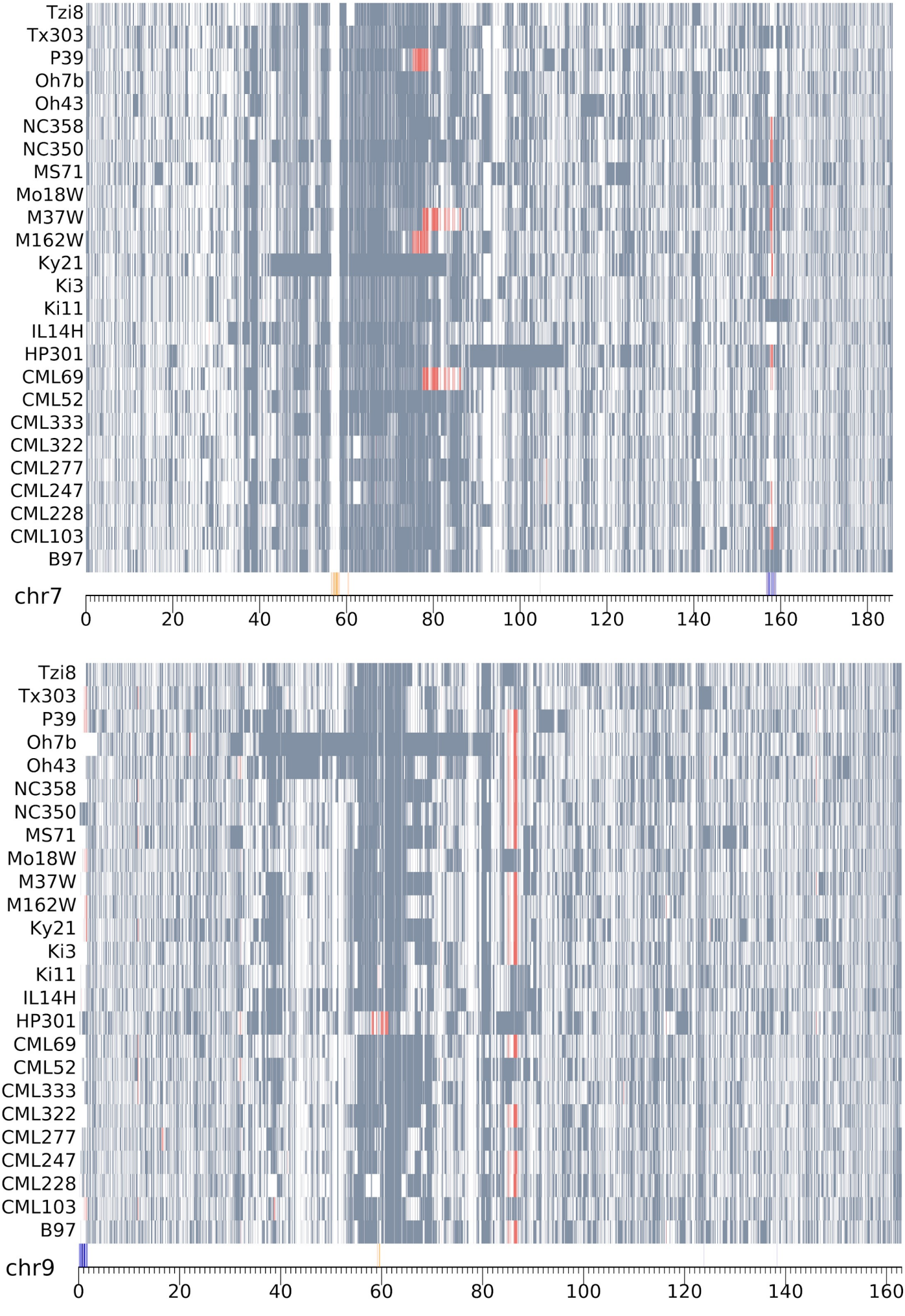

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr1
0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300

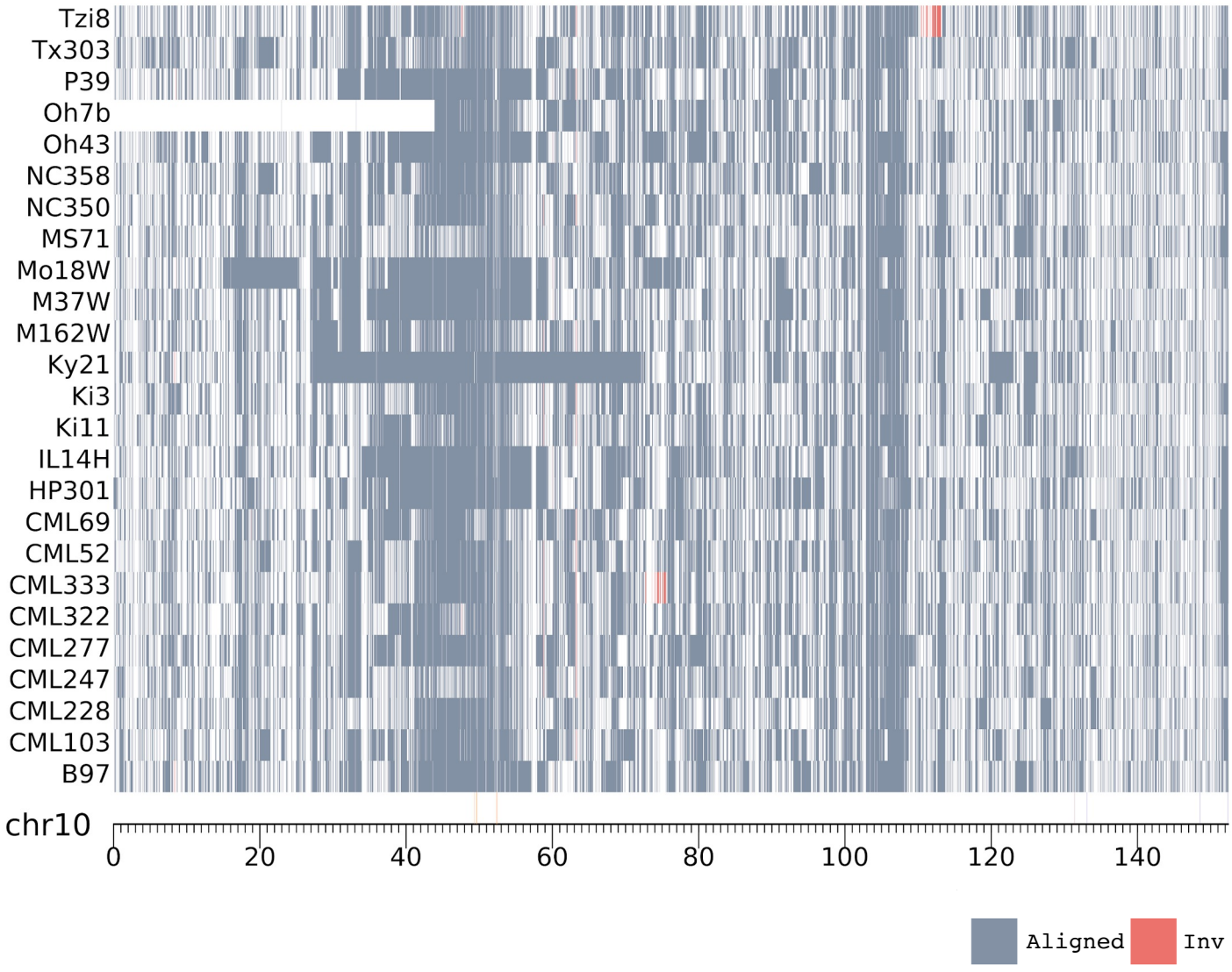

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr2
0 20 40 60 80 100 120 140 160 180 200 220 240
Aligned
Inv


Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr3
0
20
40
60
80
100
120
140
160
180
200
220

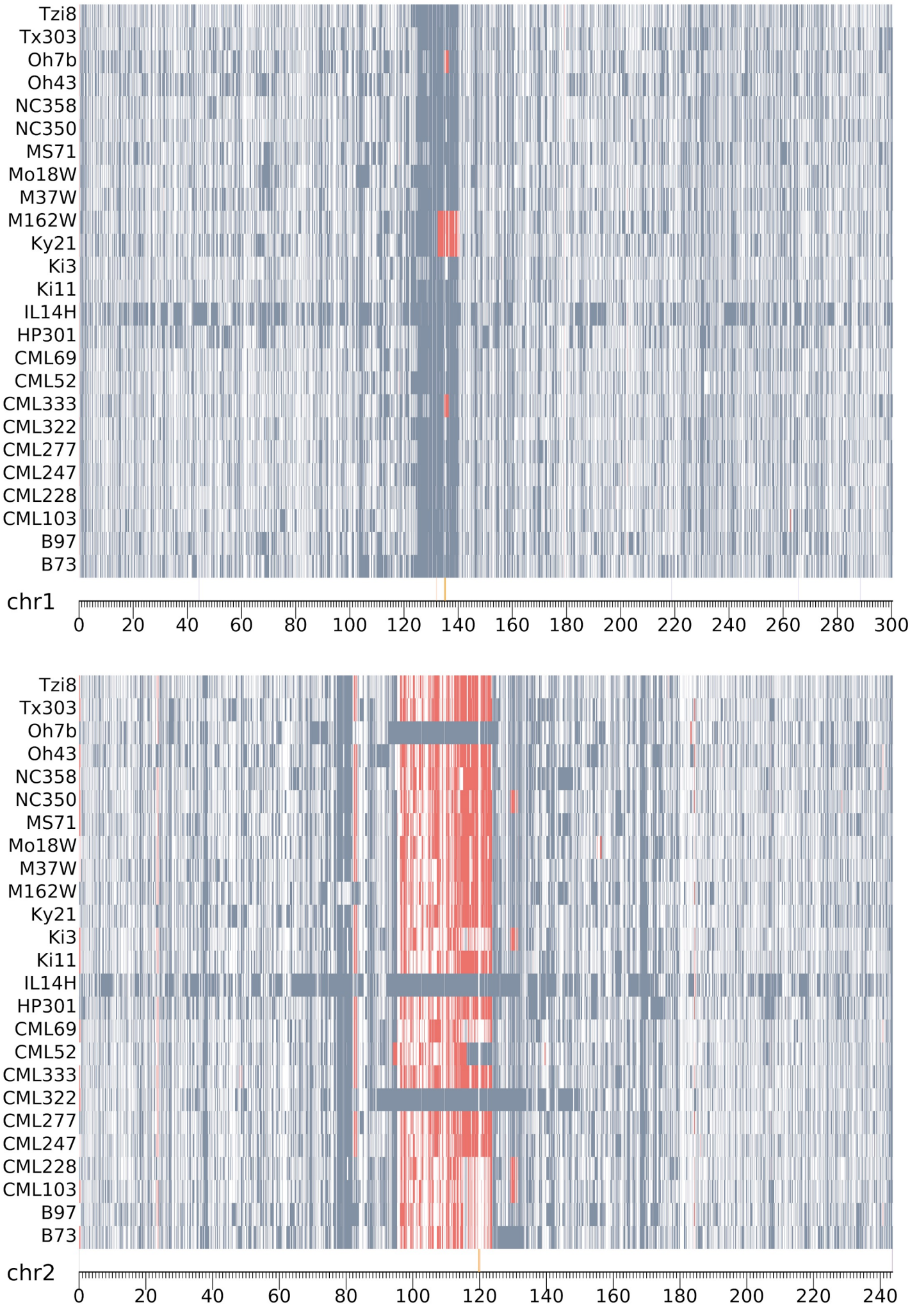

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr4
0
20
40
60
80
100
120
140
160
180
200
220
240
Aligned
Inv

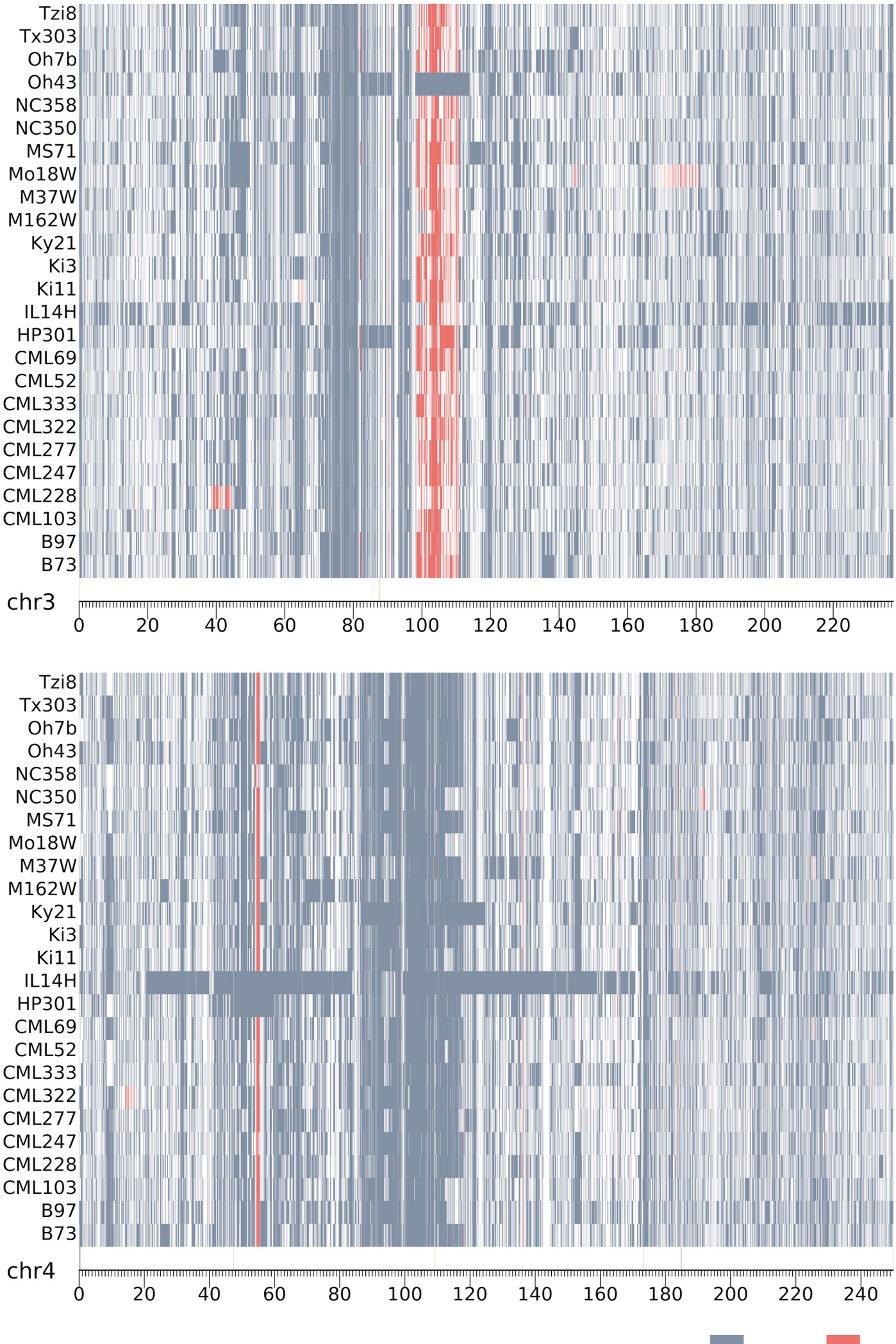

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr5
0
20
40
60
80
100
120
140
160
180
200
220
Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr6
0
20
40
60
80
100
120
140
160
180
Aligned
Inv

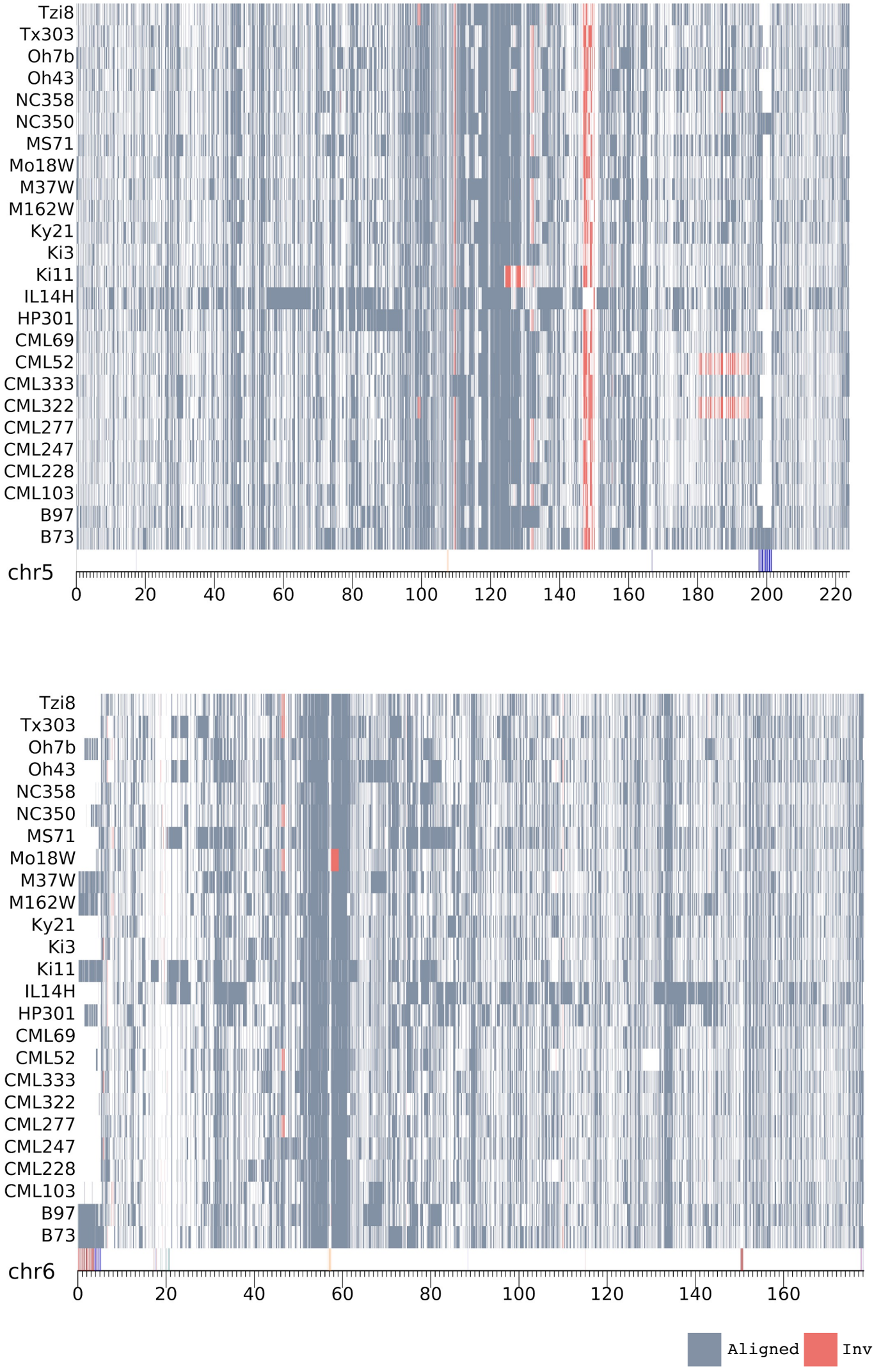

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr7
0
20
40
60
80
100
120
140
160
180
Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr9
0
20
40
60
80
100
120
140
160
Aligned
Inv

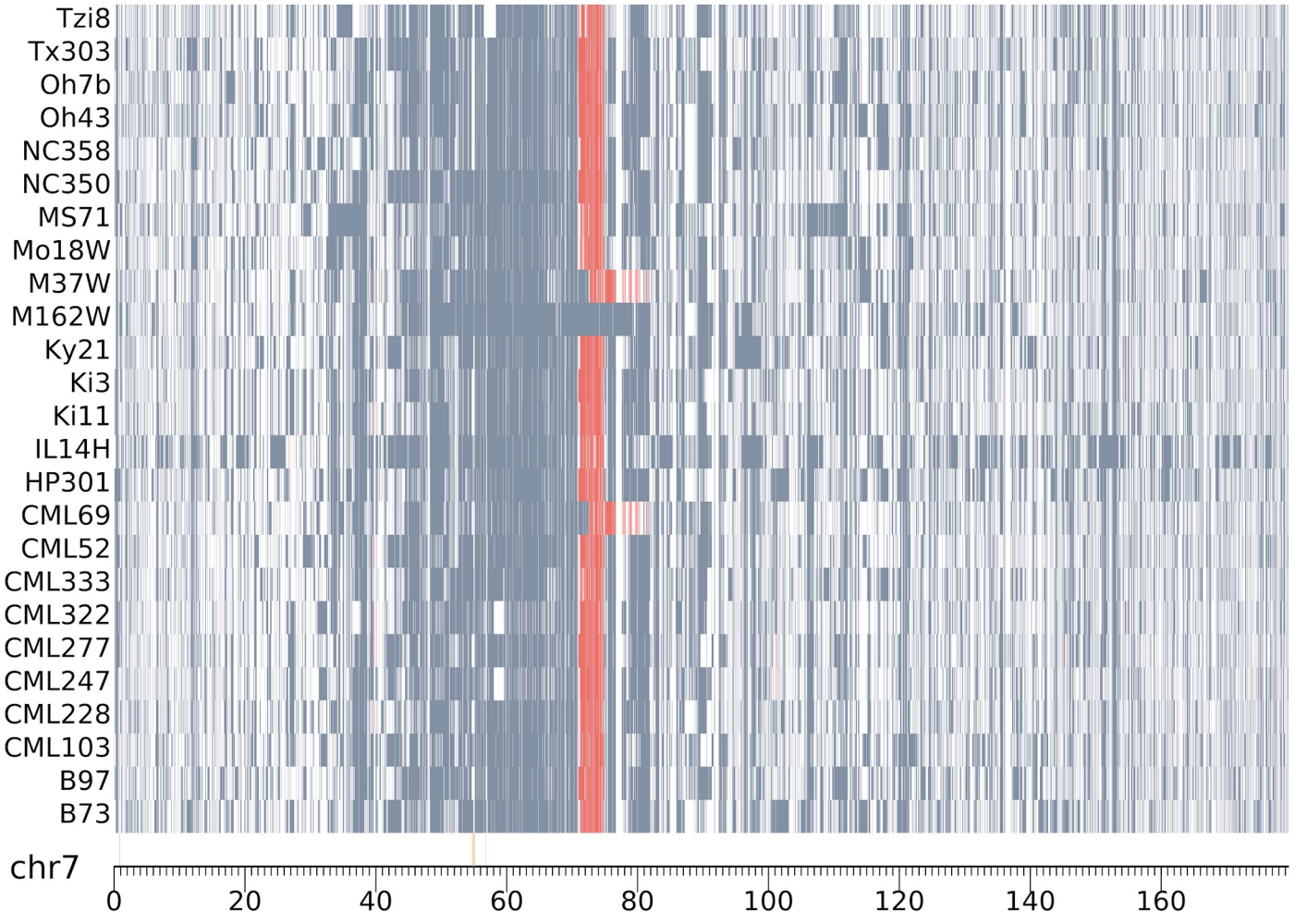


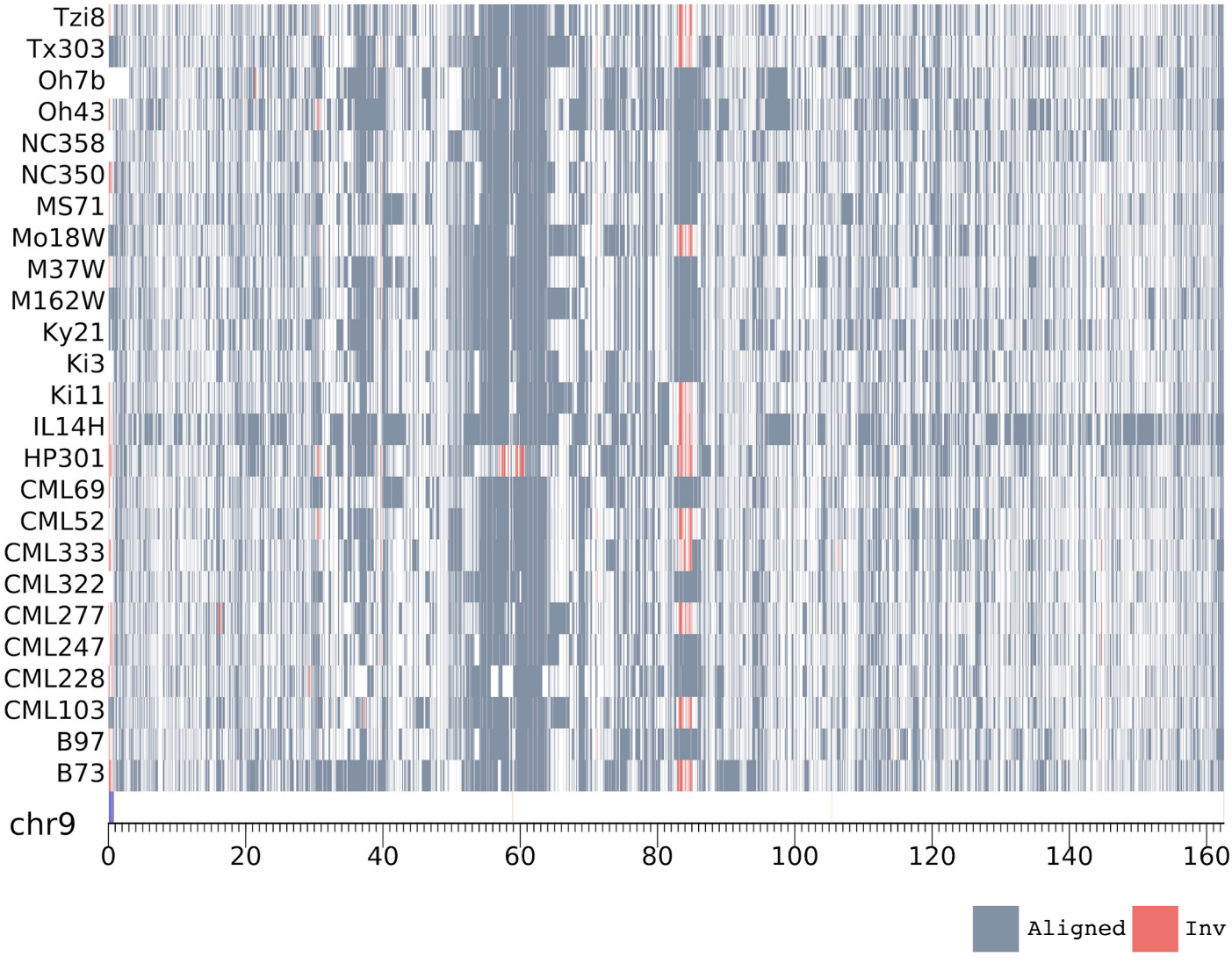
Supplemental Figure S7. Whole genome alignments using B73 as a reference. Syntenic aligned regions are colored grey and inverted segments are highlighted in red. CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars below the alignments.

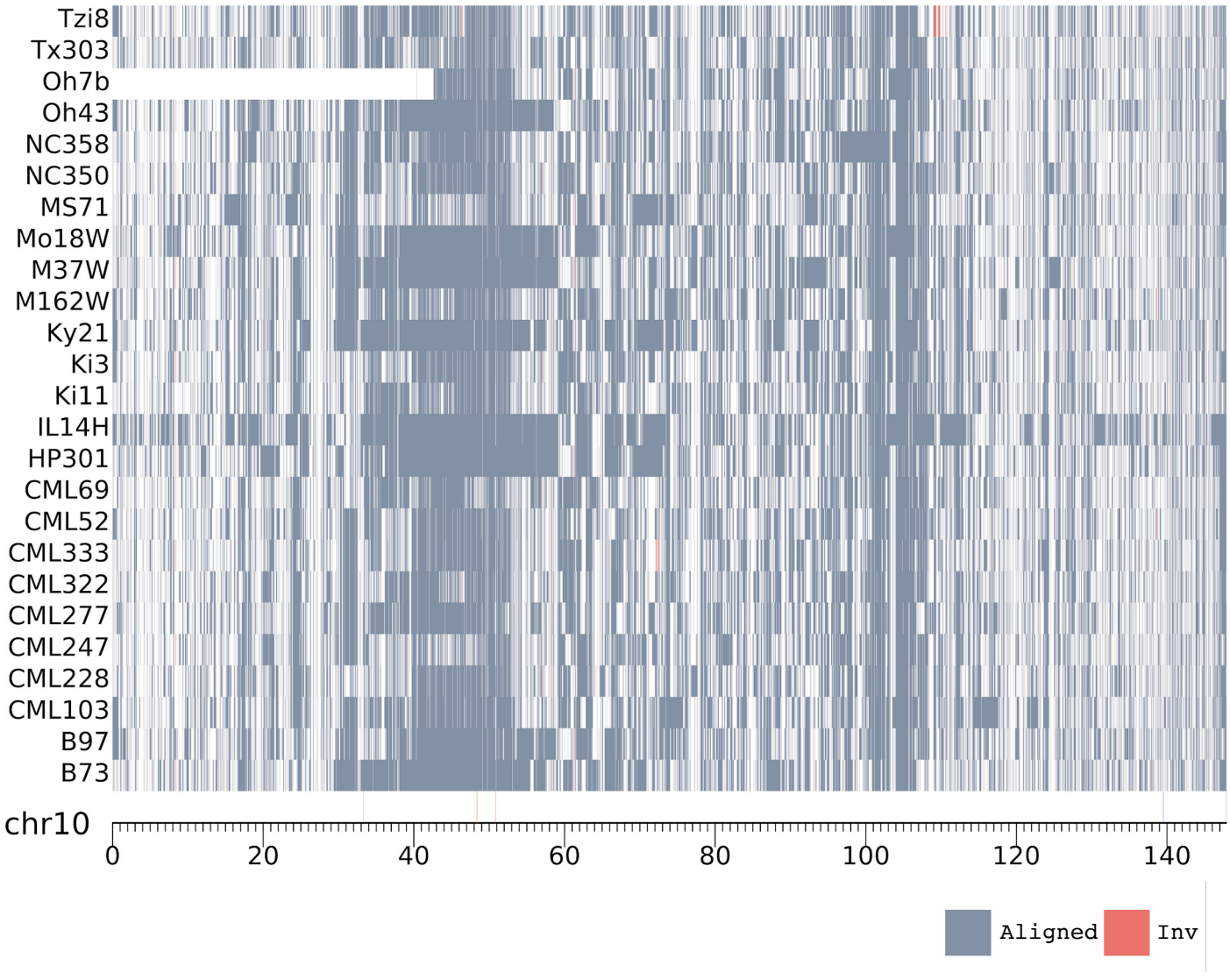

Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr1
0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300
chr2
0 20 40 60 80 100 120 140 160 180 200 220 240
Aligned
Inv


Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr3
0 20 40 60 80 100 120 140 160 180 200 220
Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr4
0 20 40 60 80 100 120 140 160 180 200 220 240
Aligned
Inv

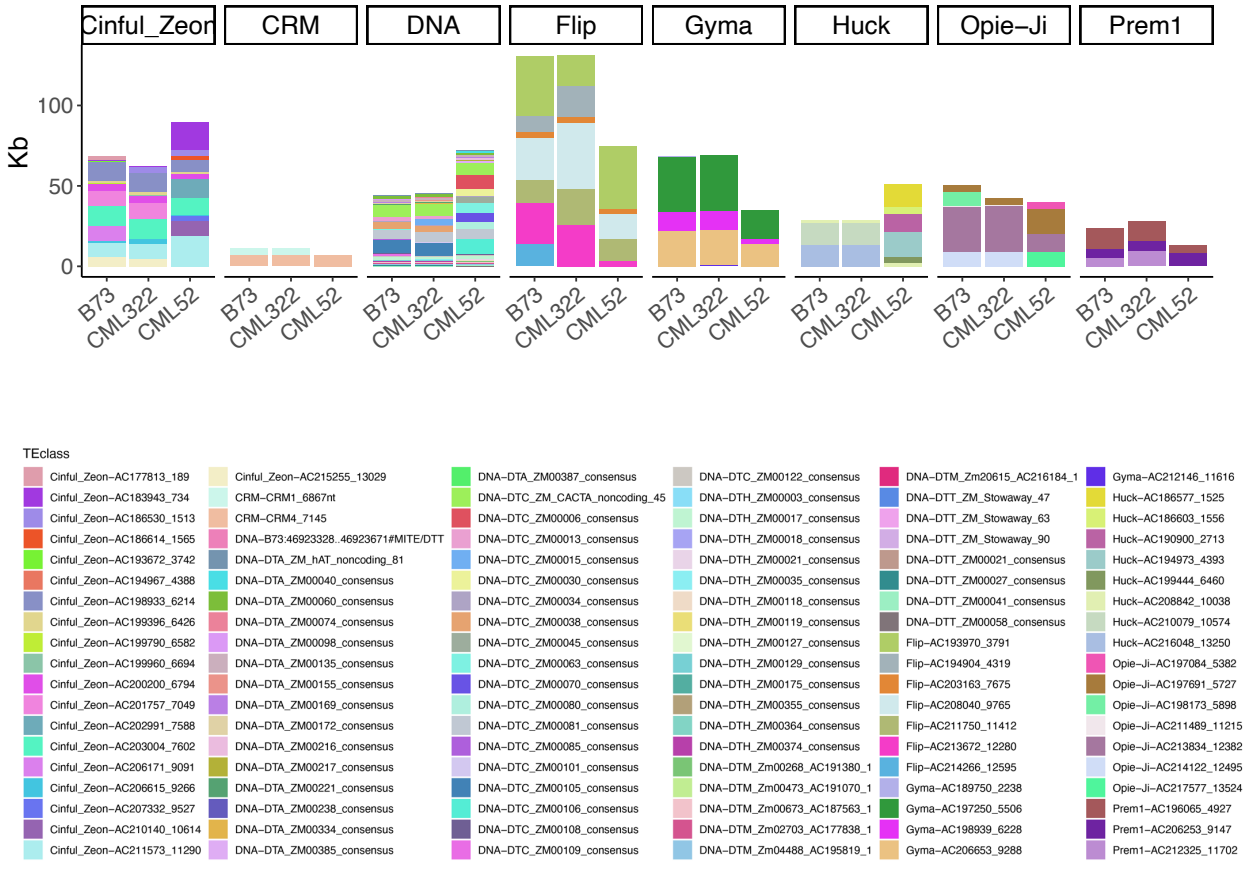

Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr5
0
20
40
60
80
100
120
140
160
180
200
220
Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr6
0
20
40
60
80
100
120
140
160
Aligned
Inv


Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr7
0
20
40
60
80
100
120
140
160


Tzi8
Tx303
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
B73
chr9
0
20
40
60
80
100
120
140
160
Aligned
Inv

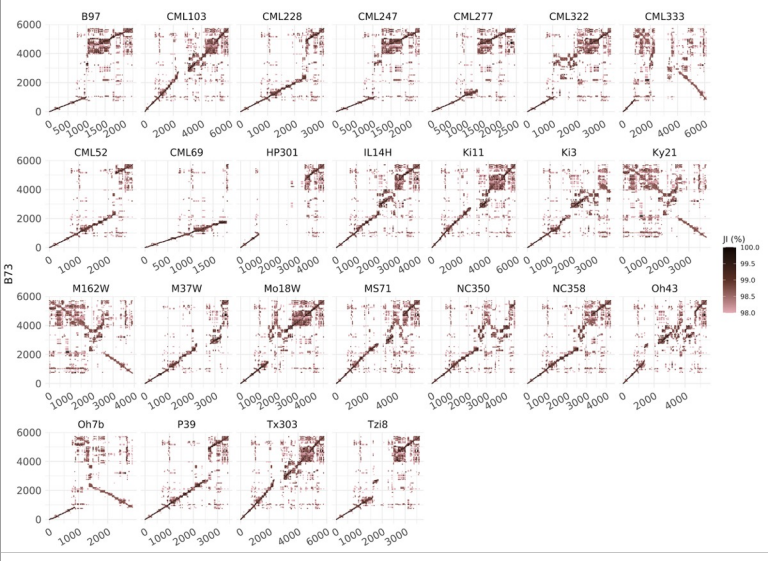


Supplemental Figure S8. Whole genome alignments using P39 as a reference. Syntenic aligned regions are colored grey and inverted segments are highlighted in red. CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars below the alignments.

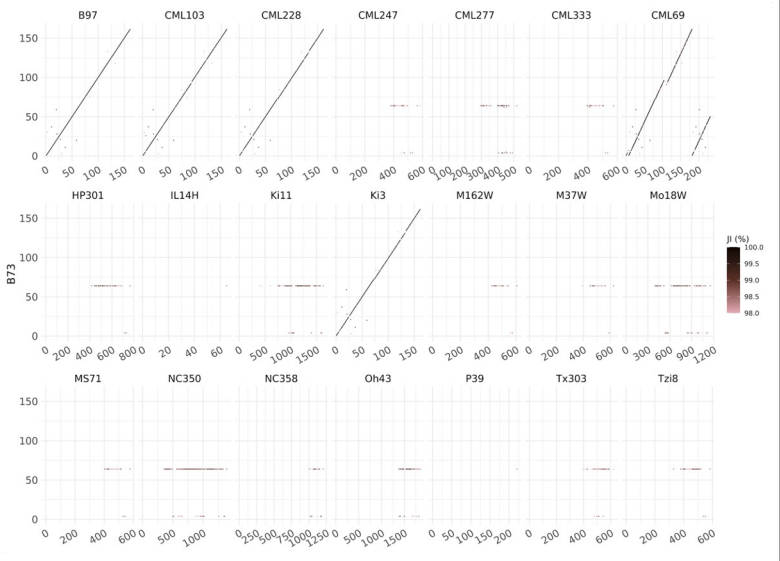


Supplemental Figure S9. Detailed view of TE subfamilies in the ~1 Mb regions of chromosome 8 shown in Figure 1C. Cinful-Zeon, CRM, Flip, Gyma, Huck, and Prem1 are Gypsy-like retroelements, while Opie-Ji is a Copia-like retroelement. DNA transposons are divided into DTA (hAT), DTC (CACTA), DTH (Pif/Harbinger) and DTT (Tc1/Mariner). Each of the major TE families also have subfamily designations (e.g. ZM00129 etc) described in Hufford et al (2021).

Chr1

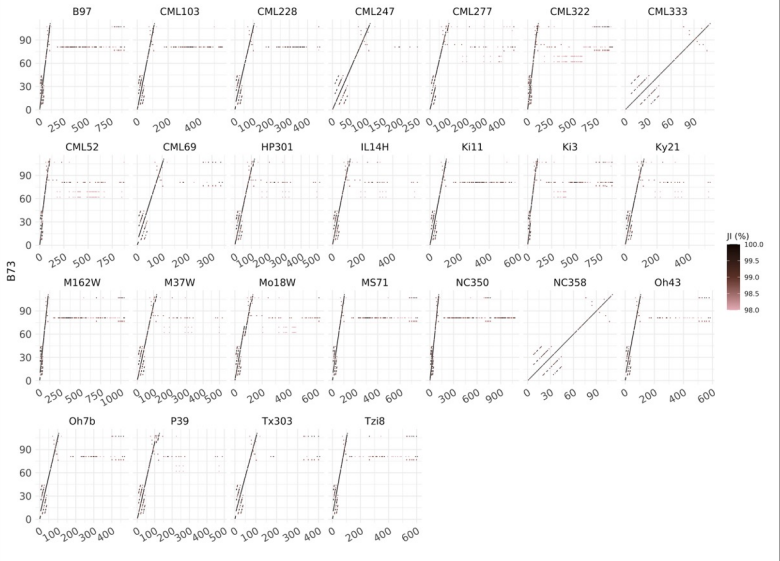


Chr2

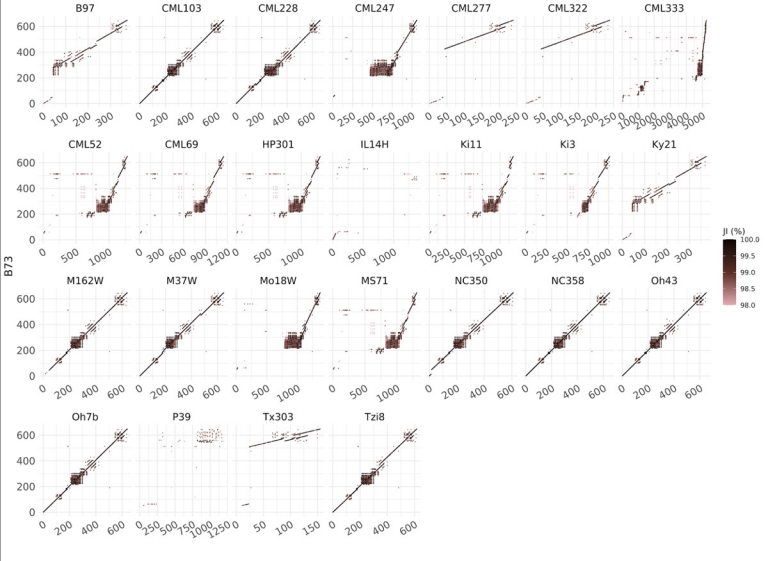


Chr3

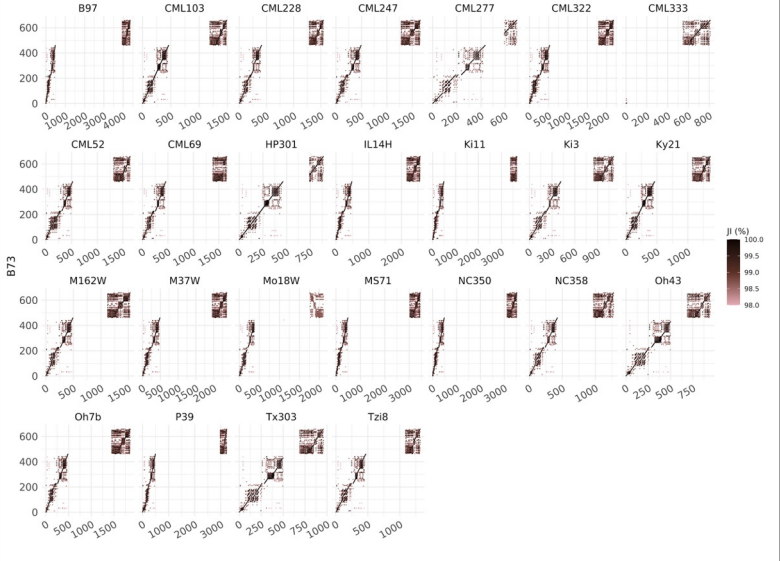
Chr4

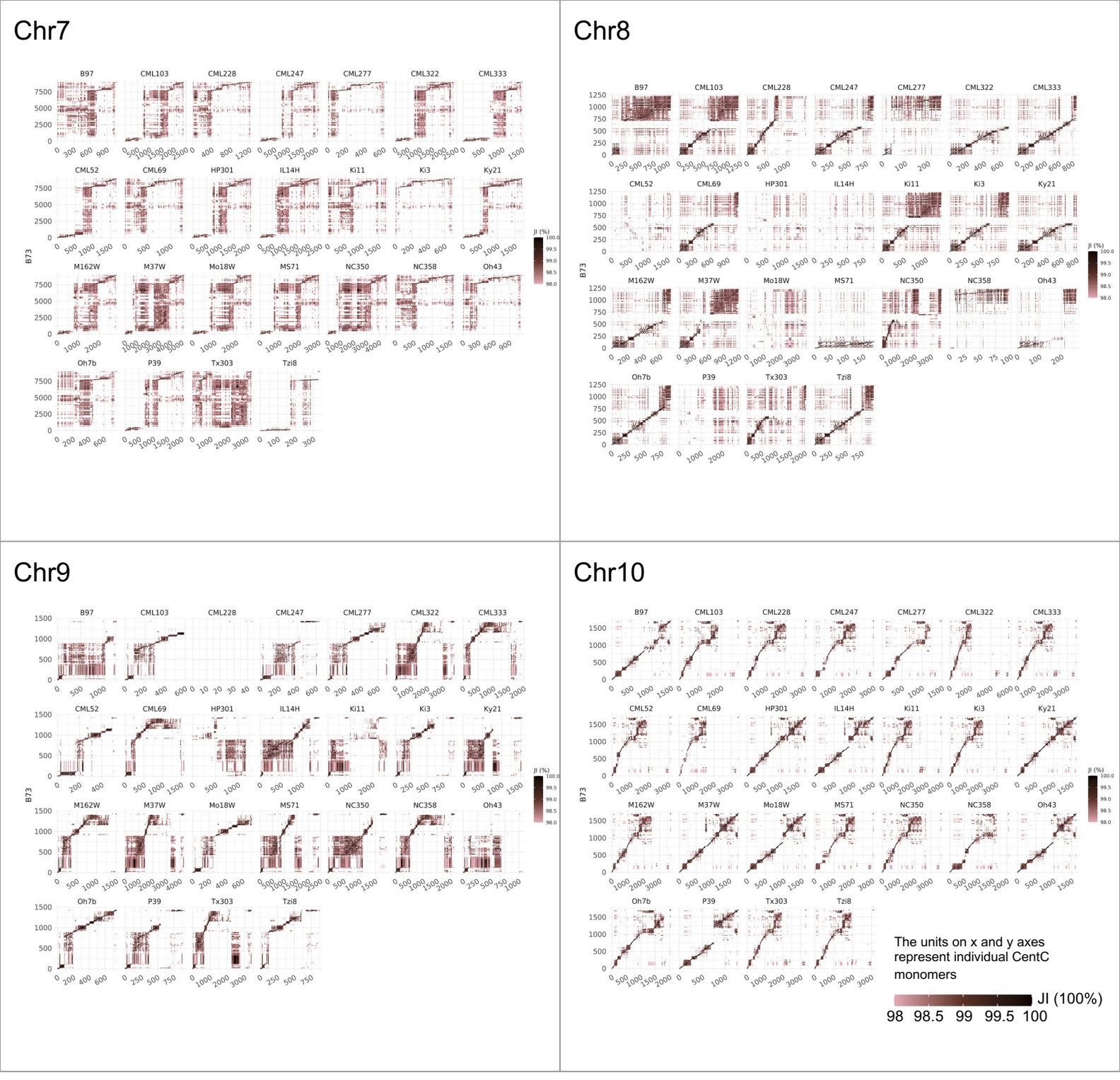


Chr5



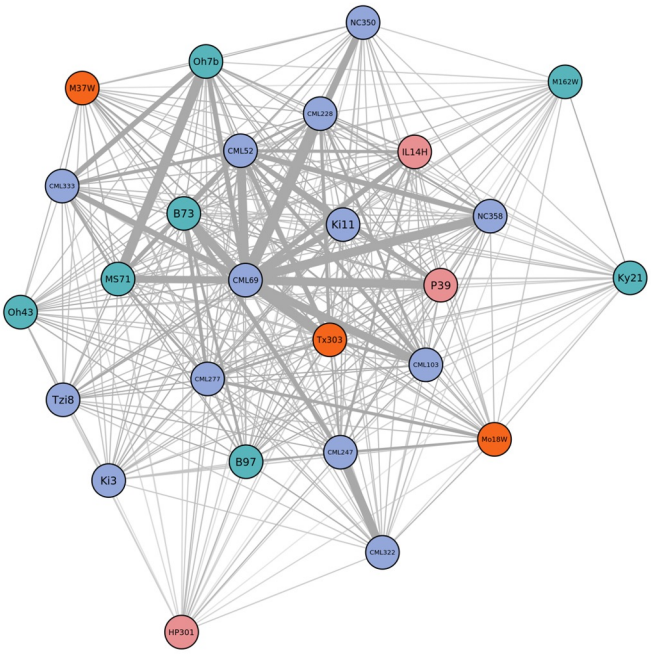
Chr6



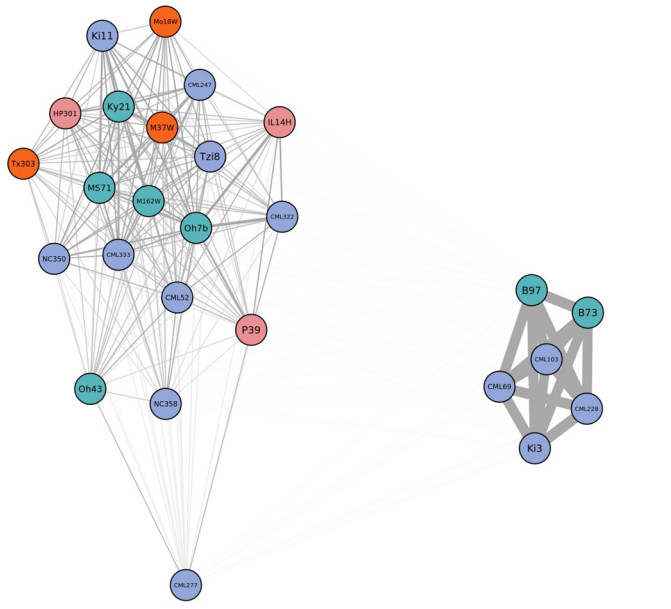


Supplemental Figure S10. Pairwise alignments between NAM lines and B73 over CentC arrays on each chromosome. X and y axes show CentC monomers, and color intensity reflects the Jaccard index (JI) between each monomer pair.

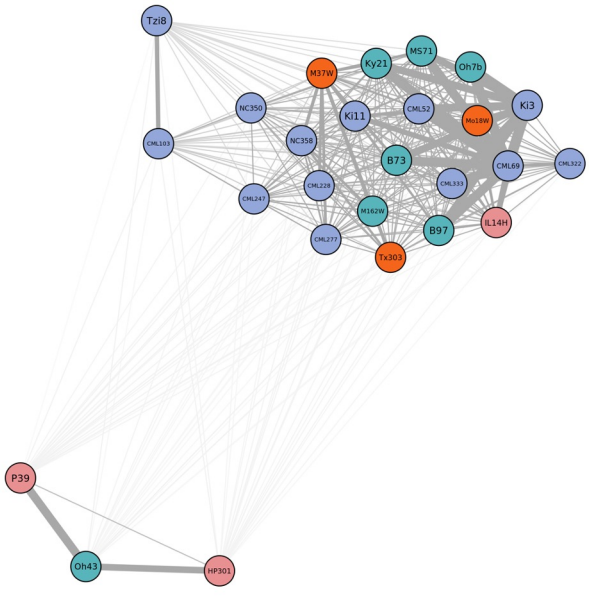
Chr1



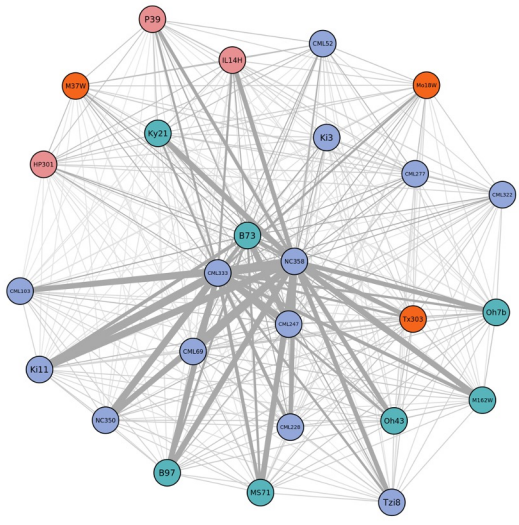
Chr2



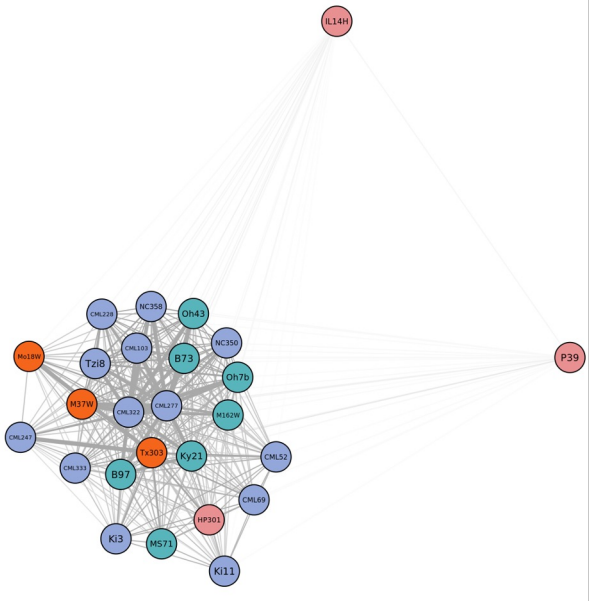
Chr3



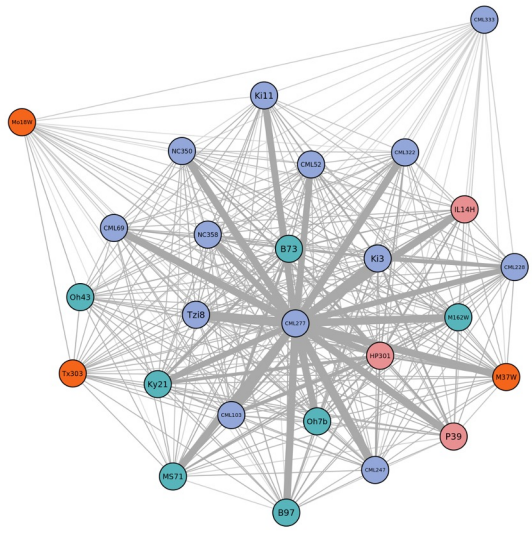
Chr4

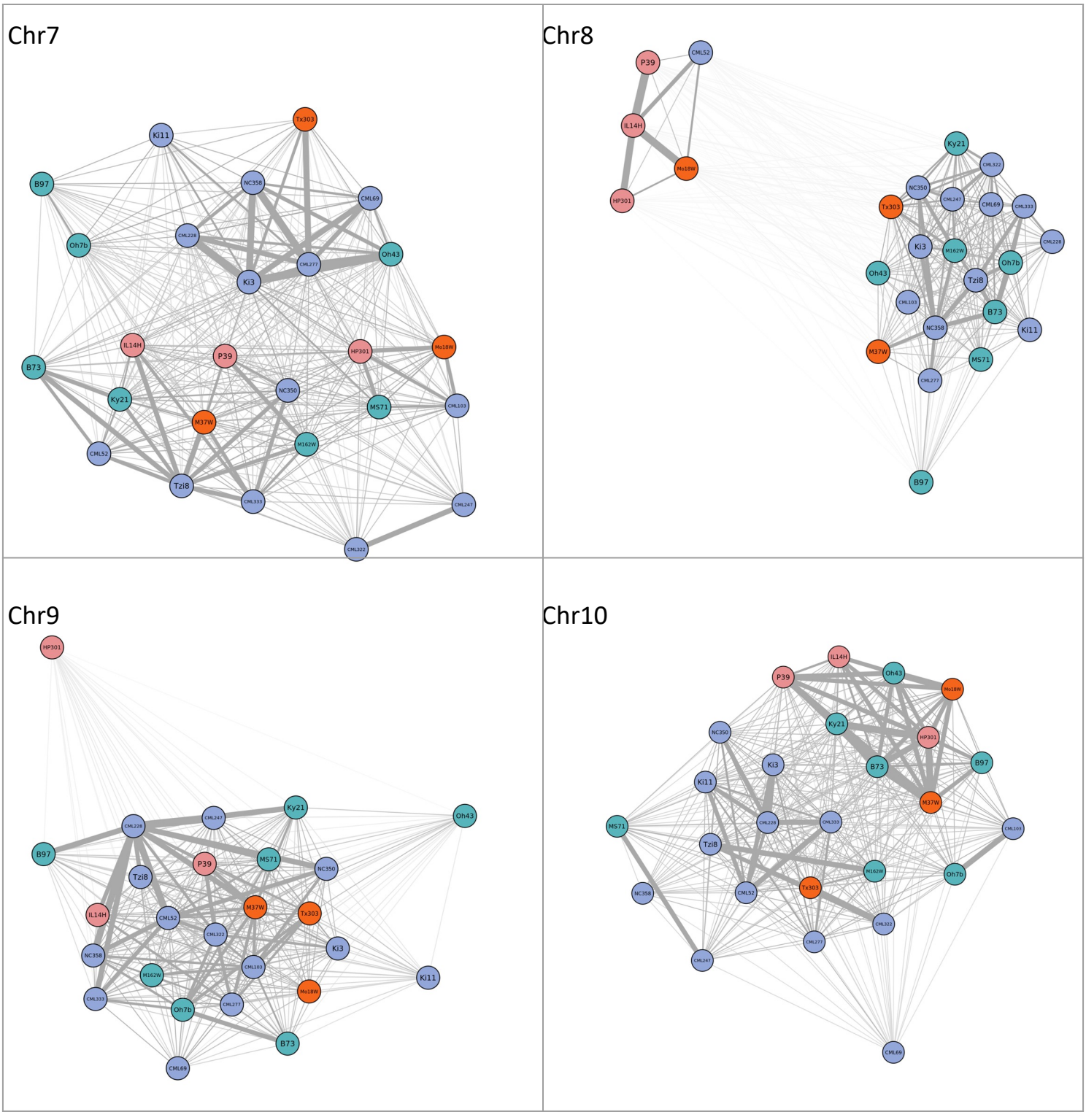


Chr5

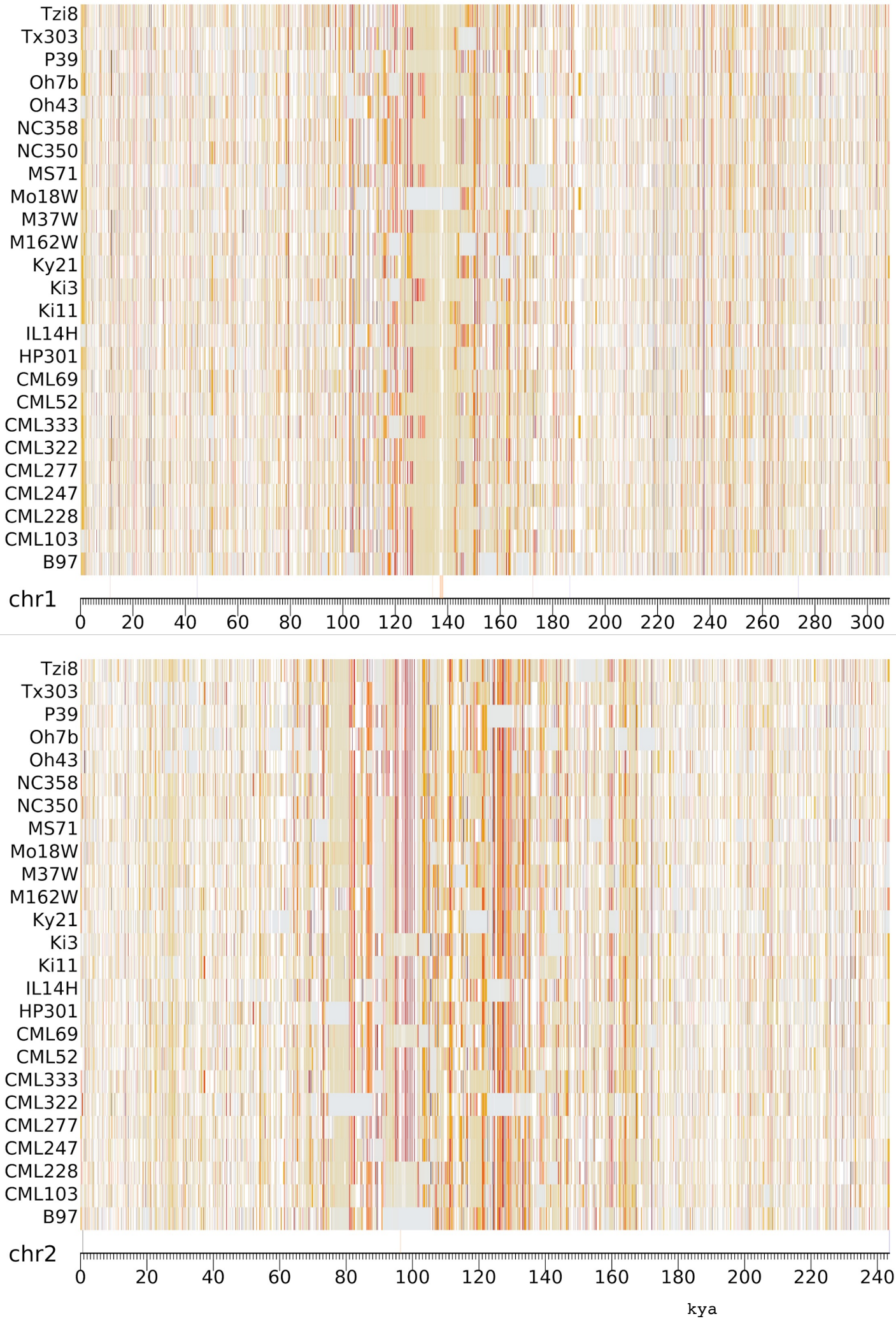


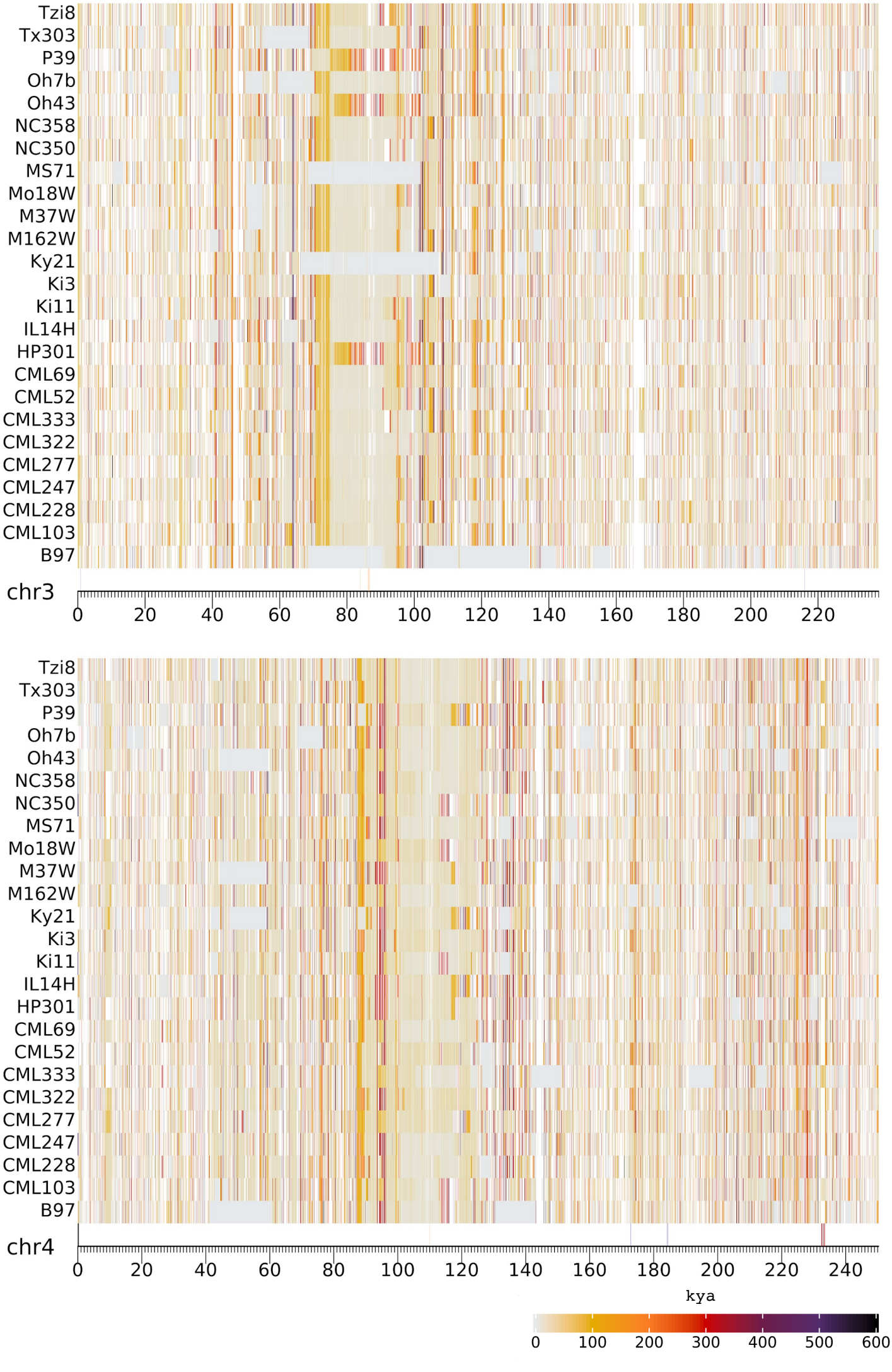
Chr6

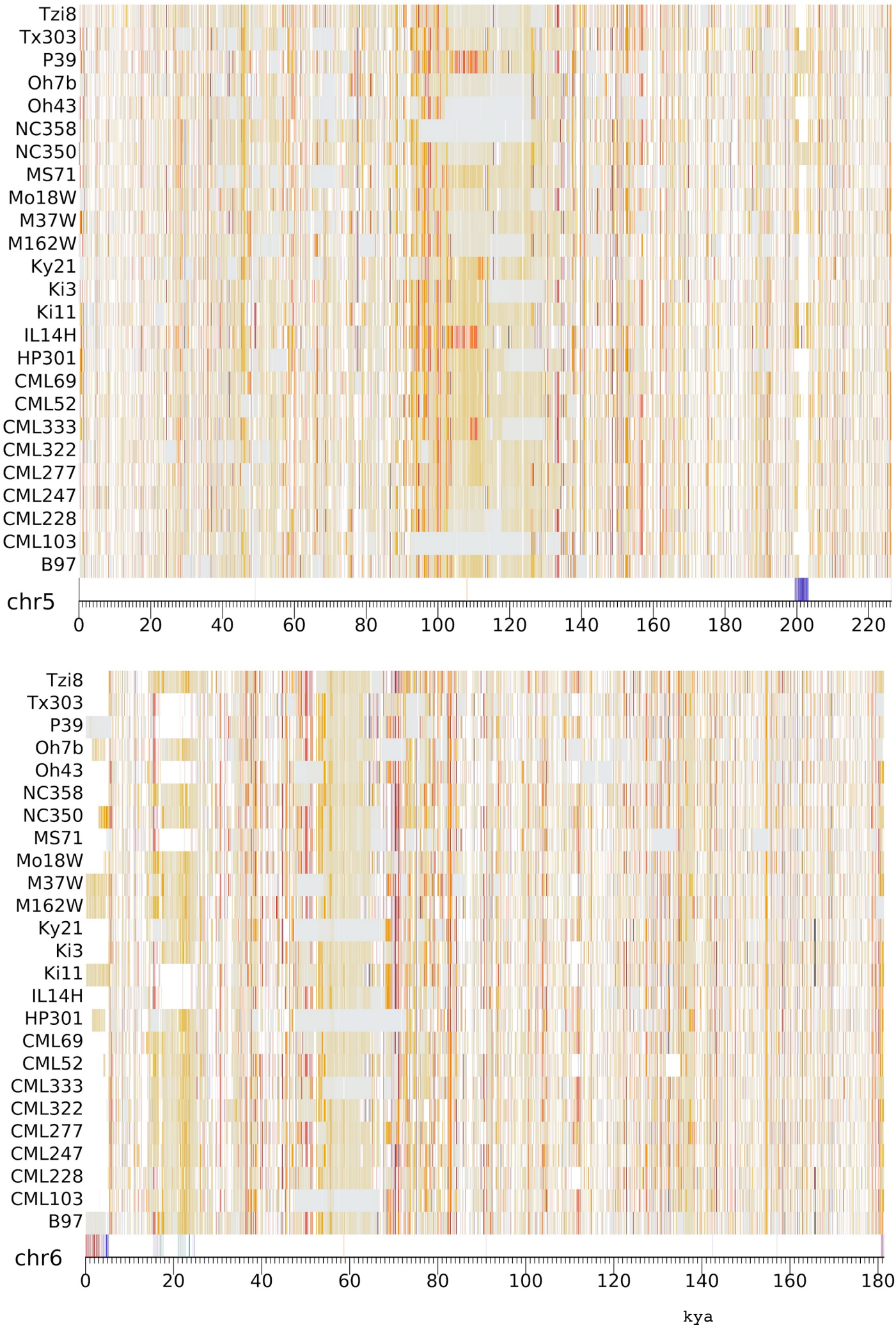


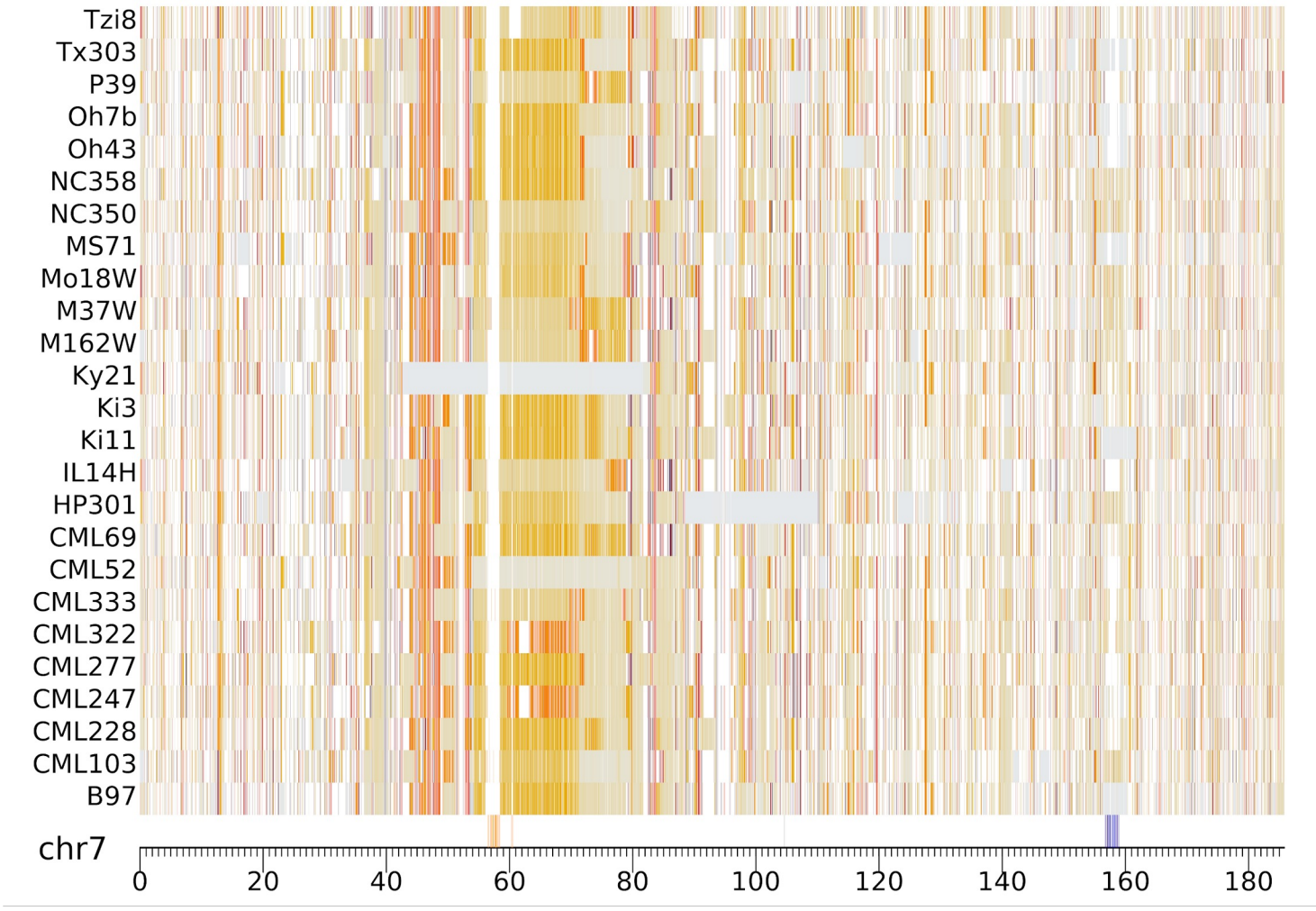


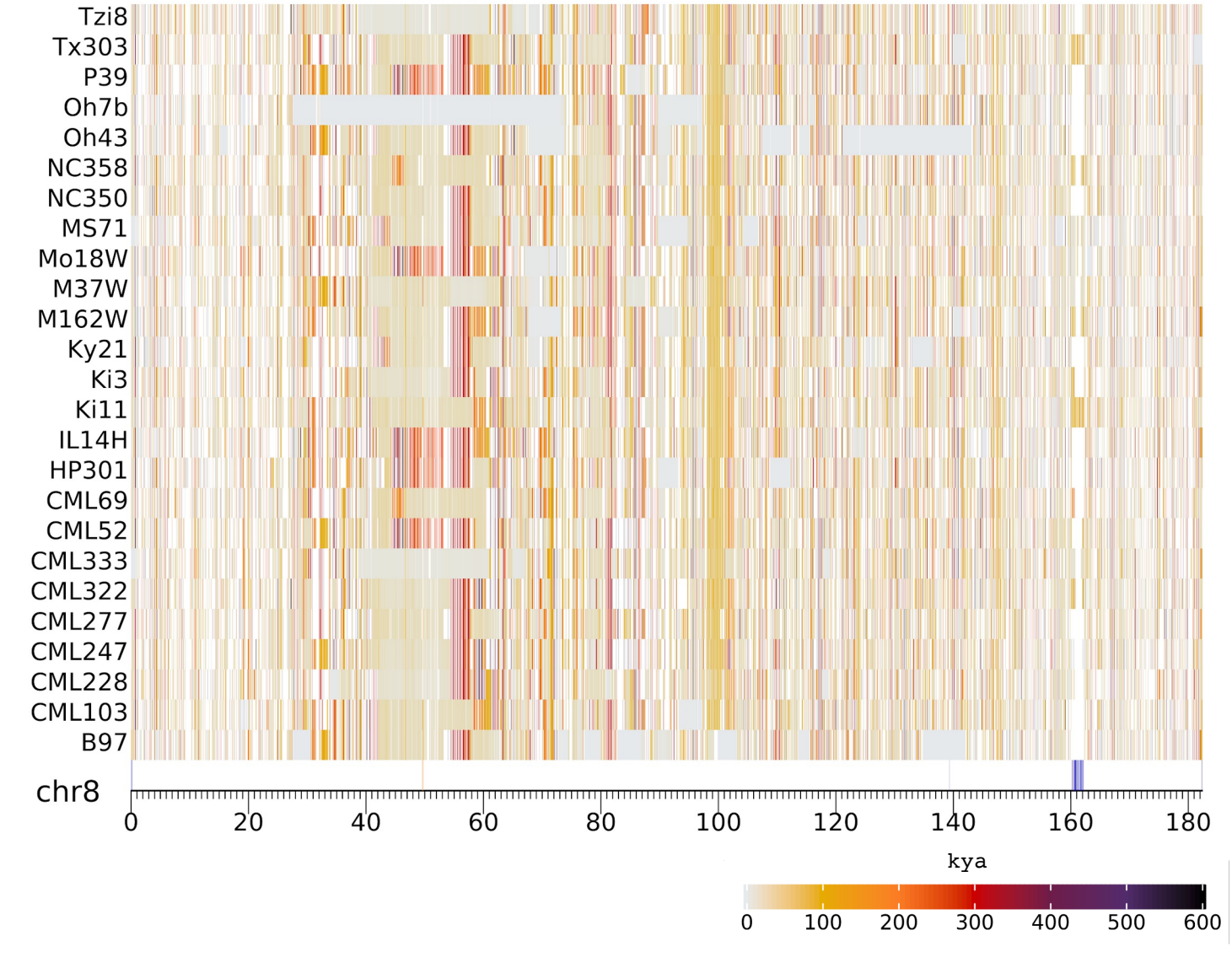
Supplemental Figure S11. Clustering of CentC arrays on all ten chromosomes. Colors over inbred names indicate varieties of corn: northern flint (pink), temperate (blue), mixed (red), and tropical (green) maize varieties.

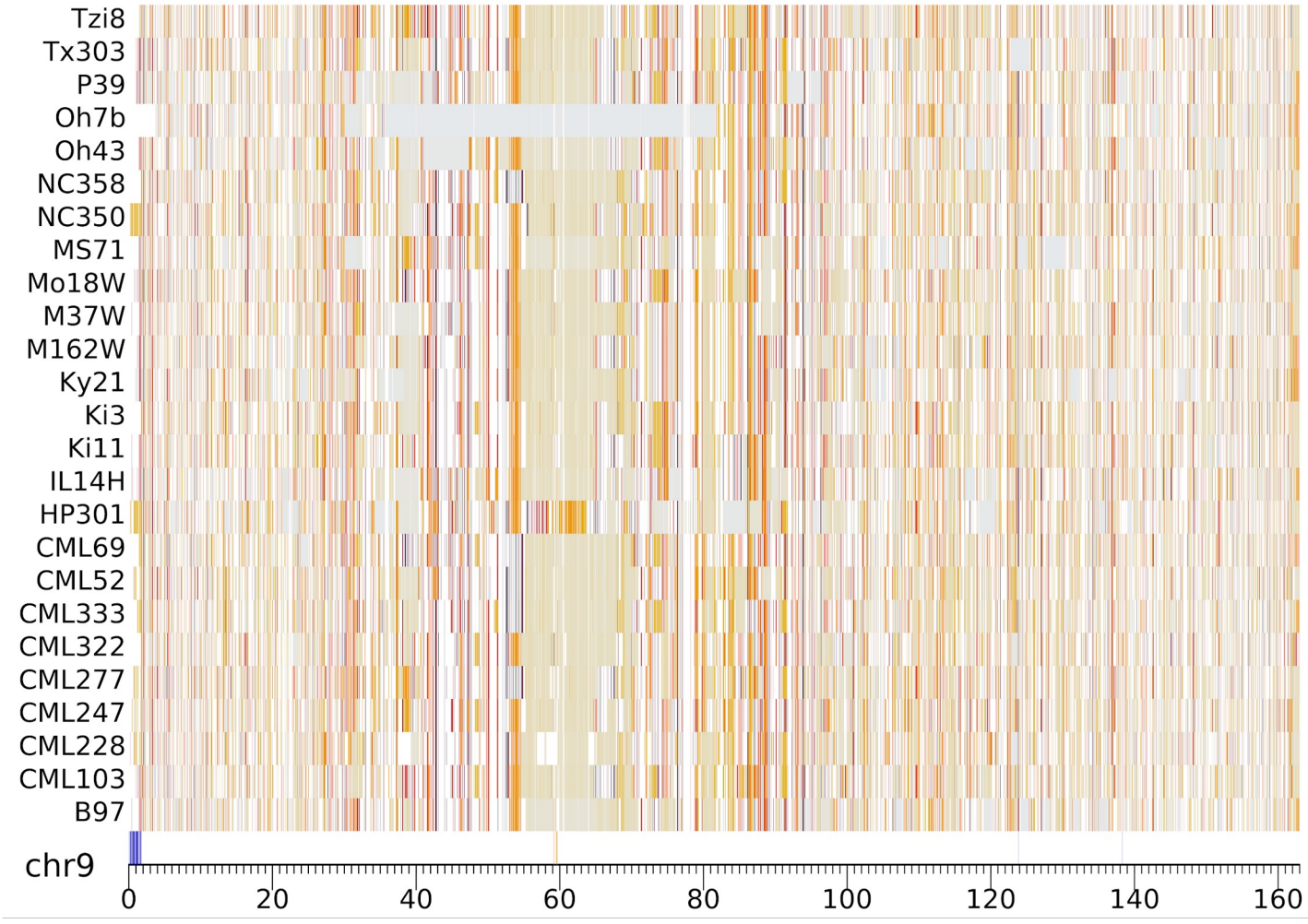

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr1
0 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300
Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr2
0 20 40 60 80 100 120 140 160 180 200 220 240
kya
0 100 200 300 400 500 600

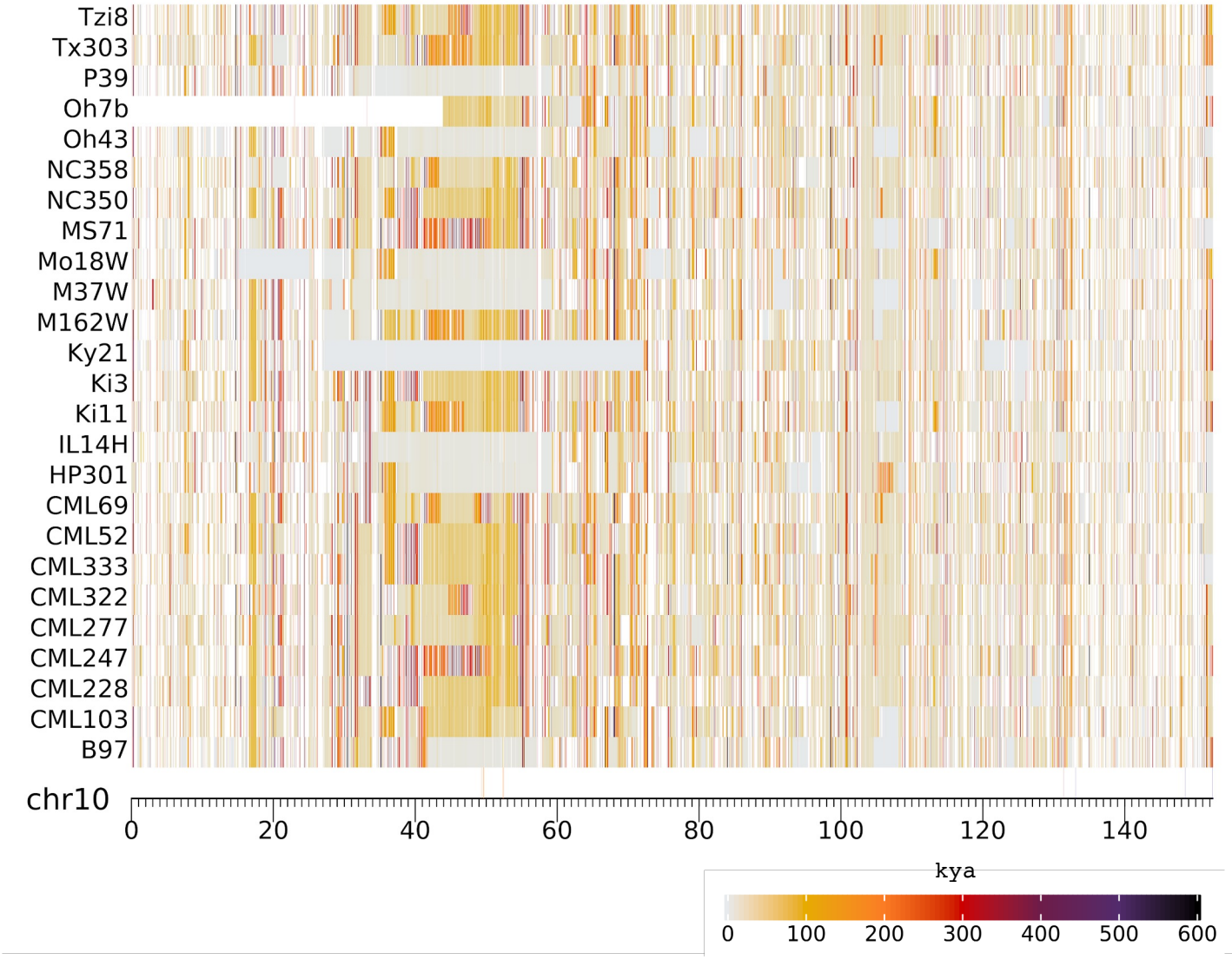

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr3
0 20 40 60 80 100 120 140 160 180 200 220
Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr4
0 20 40 60 80 100 120 140 160 180 200 220 240
kya
0 100 200 300 400 500 600


Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr5
0
20
40
60
80
100
120
140
160
180
200
220
Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr6
0
20
40
60
80
100
120
140
160
180
kya
0
100
200
300
400
500
600

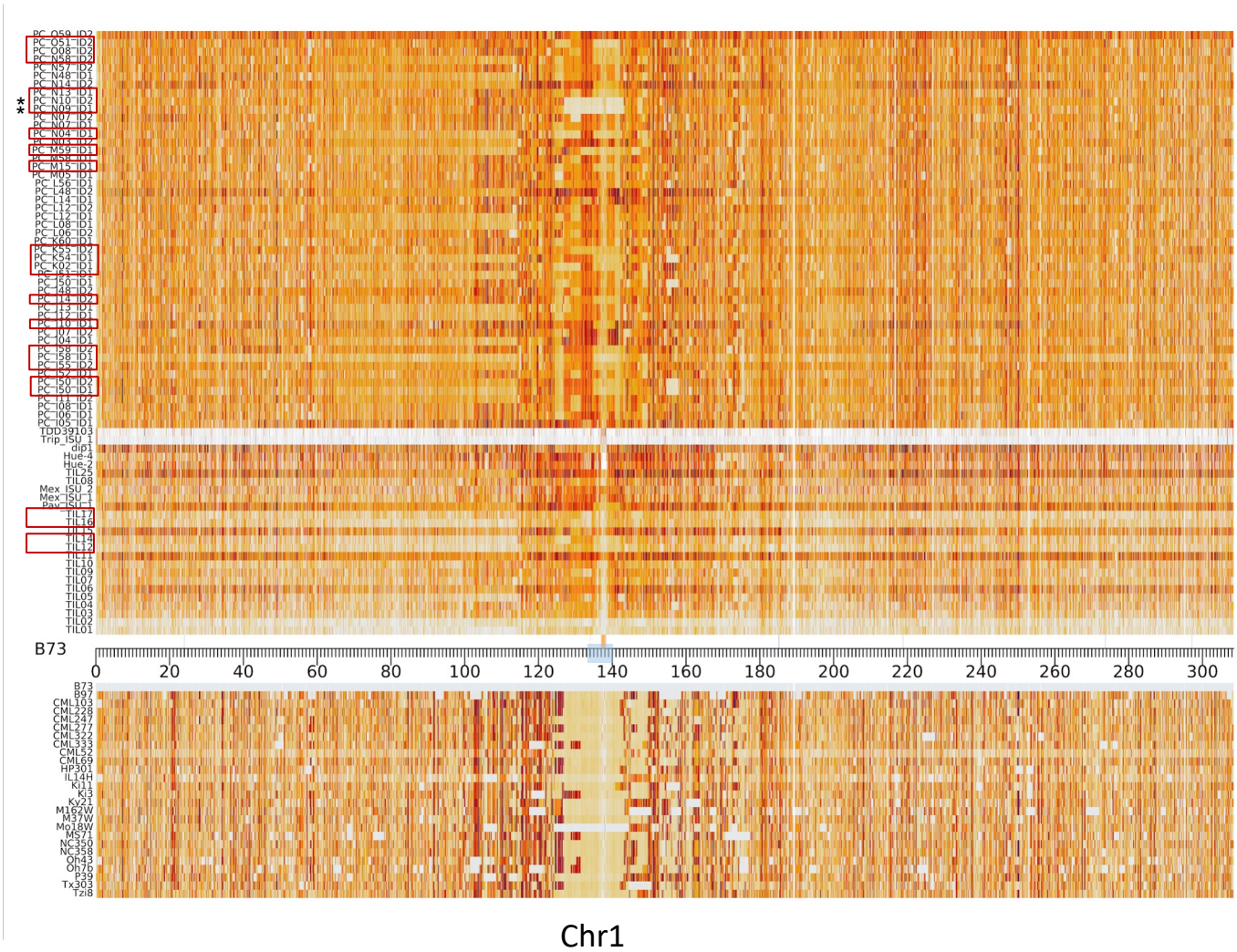

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr7
0
20
40
60
80
100
120
140
160
180


Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr8
0
20
40
60
80
100
120
140
160
180
kya
0
100
200
300
400
500
600

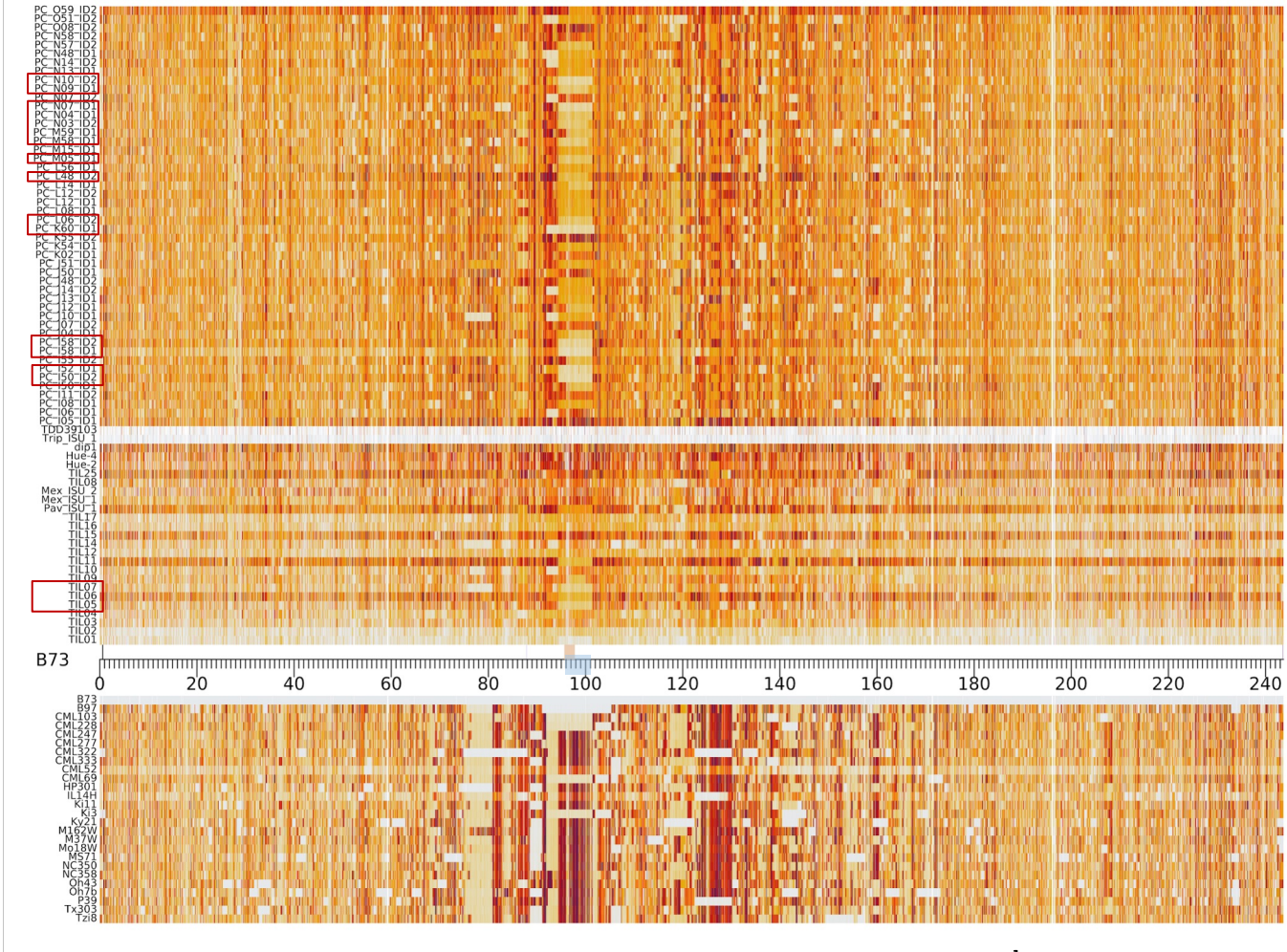

Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr9
0
20
40
60
80
100
120
140
160


Tzi8
Tx303
P39
Oh7b
Oh43
NC358
NC350
MS71
Mo18W
M37W
M162W
Ky21
Ki3
Ki11
IL14H
HP301
CML69
CML52
CML333
CML322
CML277
CML247
CML228
CML103
B97
chr10
0
20
40
60
80
100
120
140
kya
0
100
200
300
400
500
600

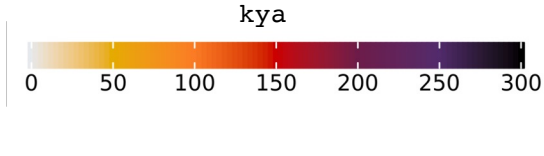
Supplemental Figure S12. Divergence times estimated with syntenic SNPs using B73 as a reference. White spaces depict unaligned regions between the query genome and B73. CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars below the alignments.

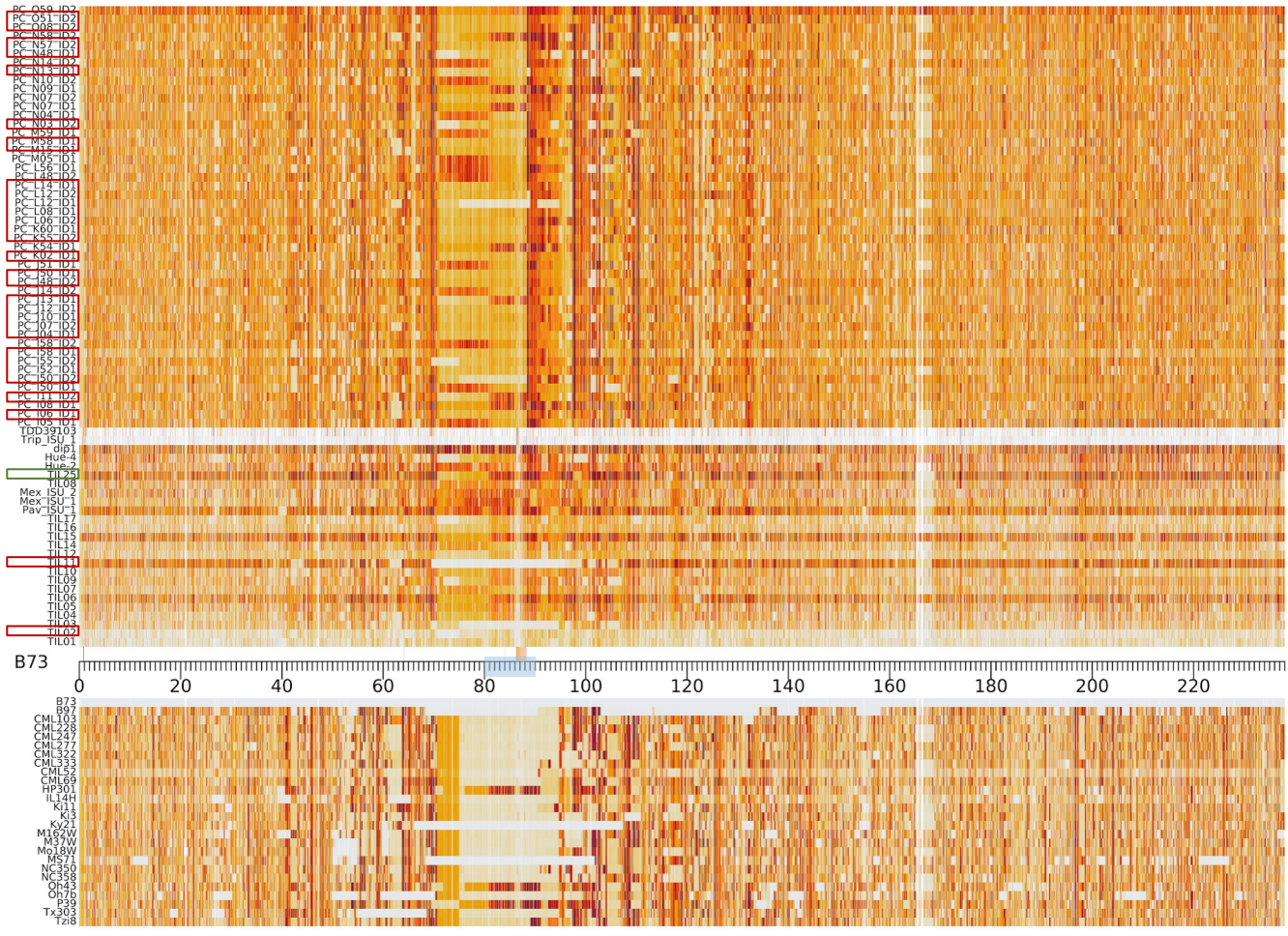

B73
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300

Chr1

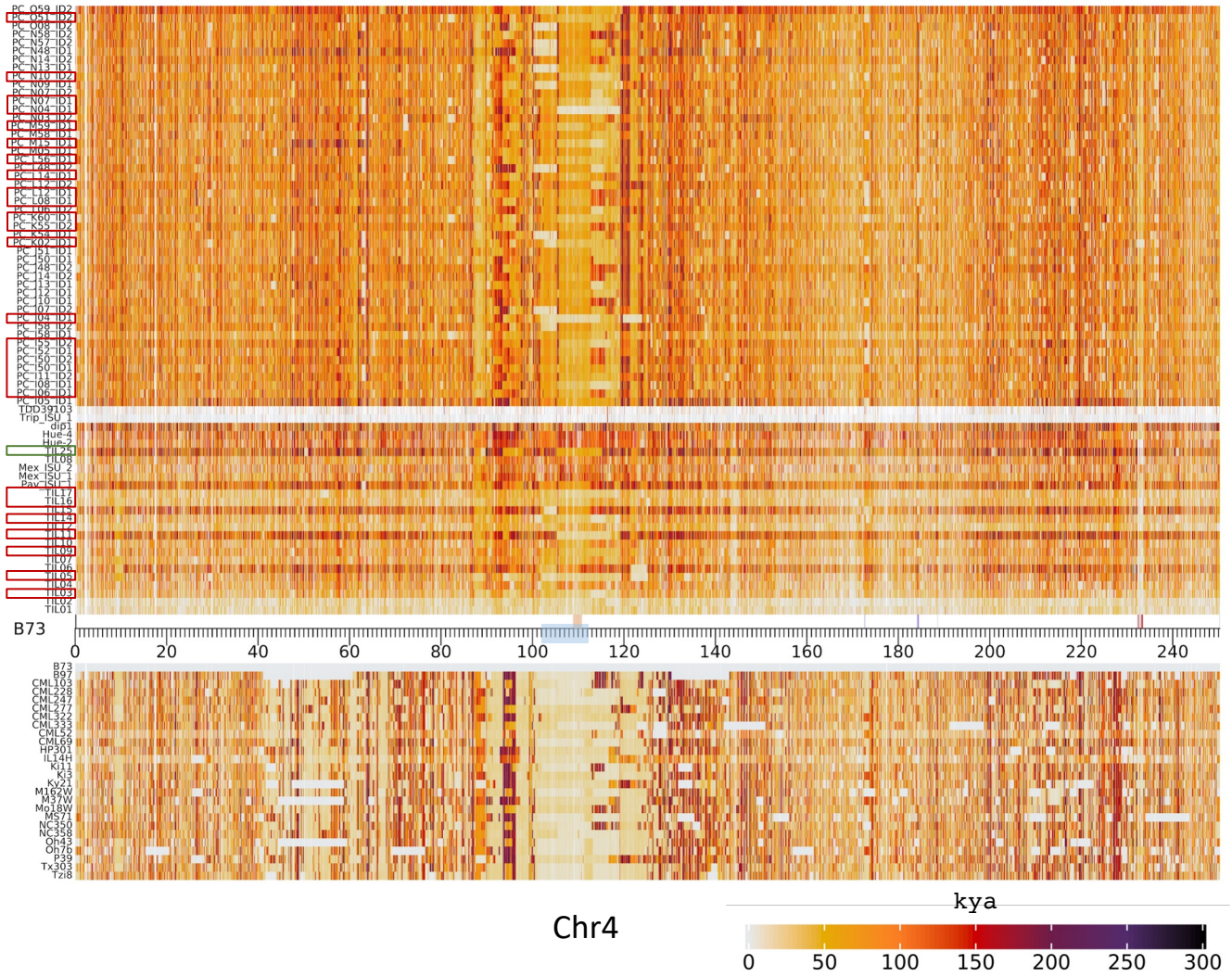

B73
0
20
40
60
80
100
120
140
160
180
200
220
240

Chr2

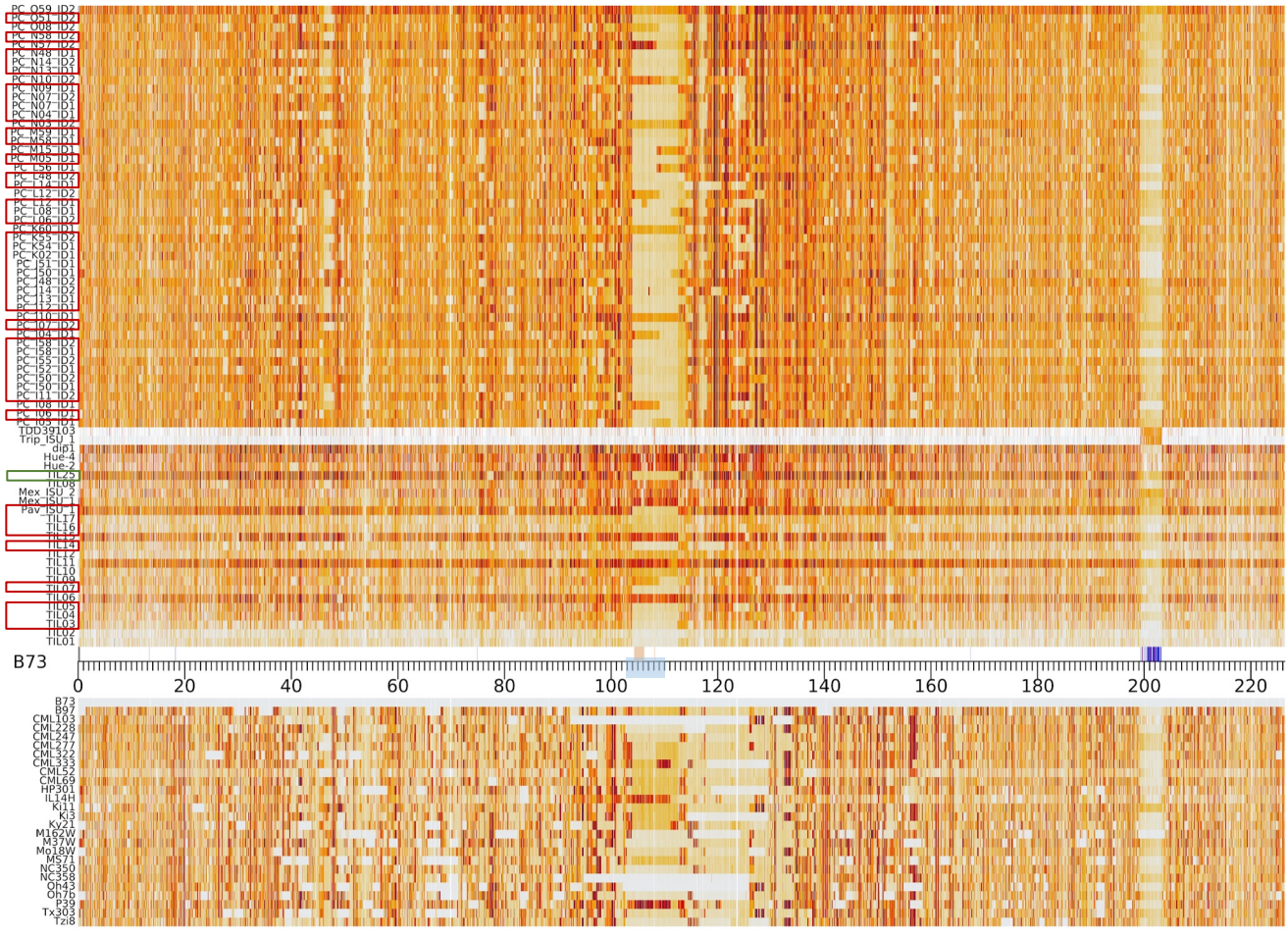

kya
0
50
100
150
200
250
300


B73
0
20
40
60
80
100
120
140
160
180
200
220

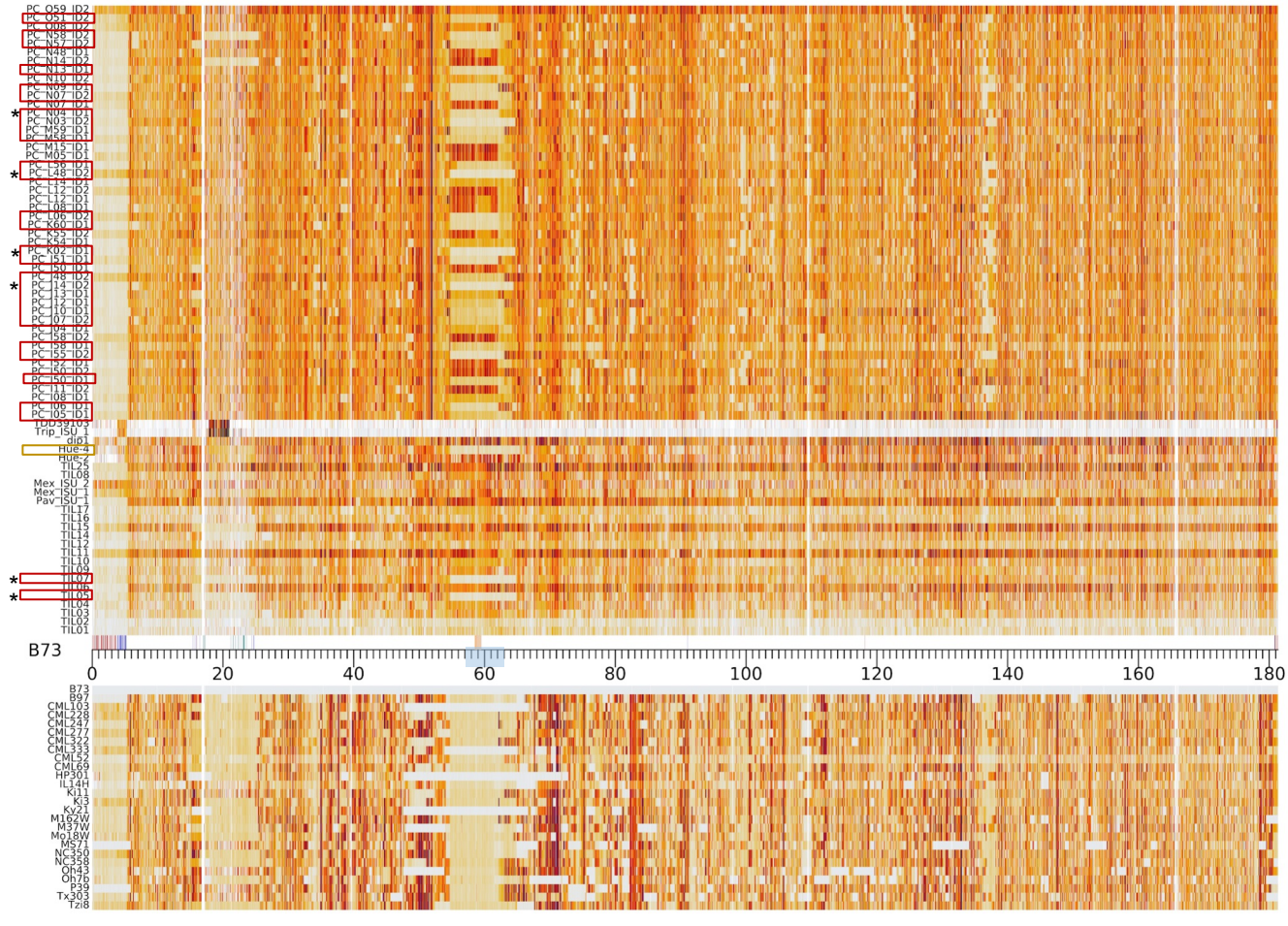
Chr3


B73
0
20
40
60
80
100
120
140
160
180
200
220
240
kya
0
50
100
150
200
250
300

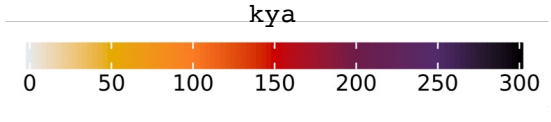
Chr4

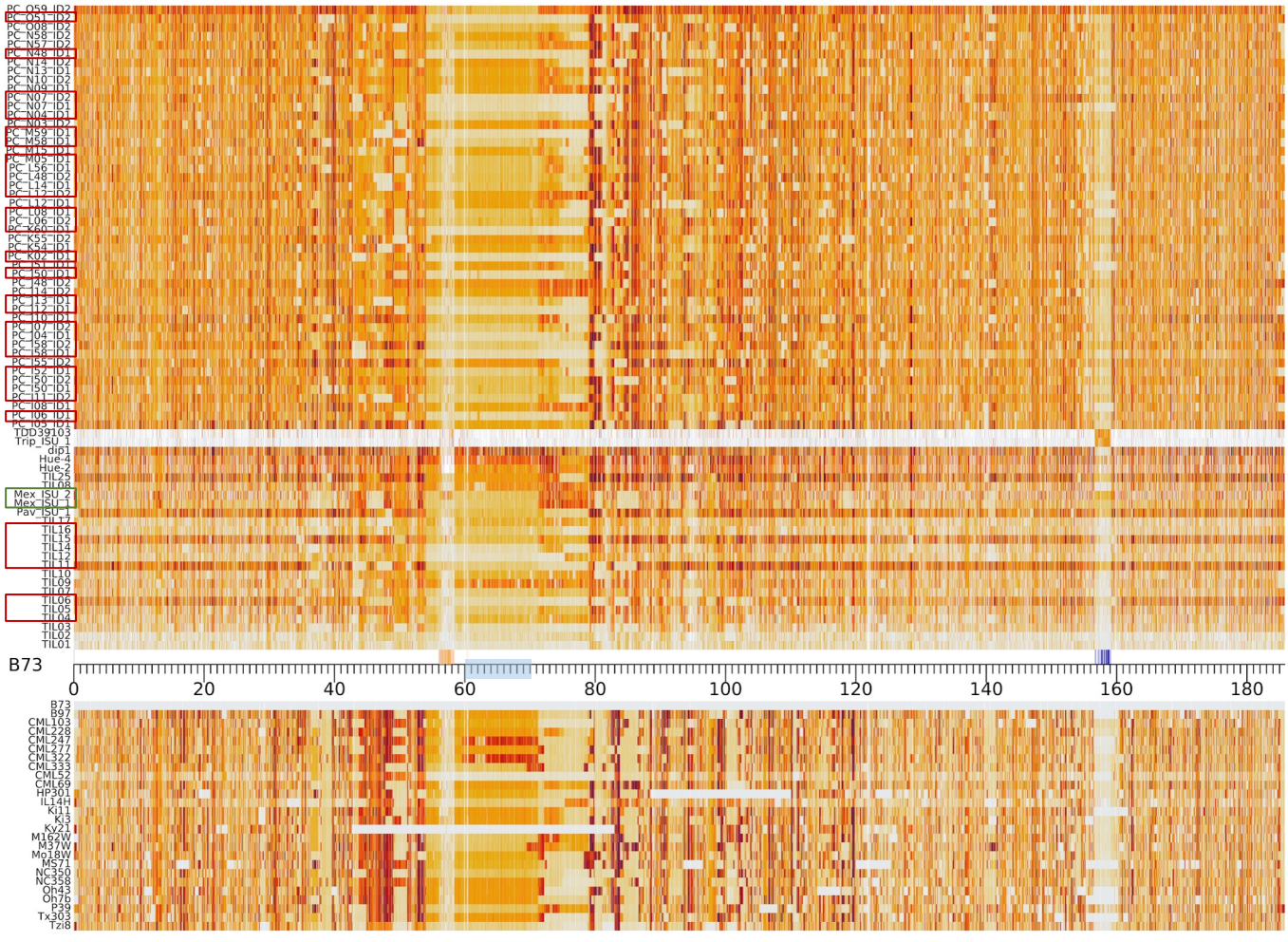

B73
0
20
40
60
80
100
120
140
160
180
200
220

Chr5

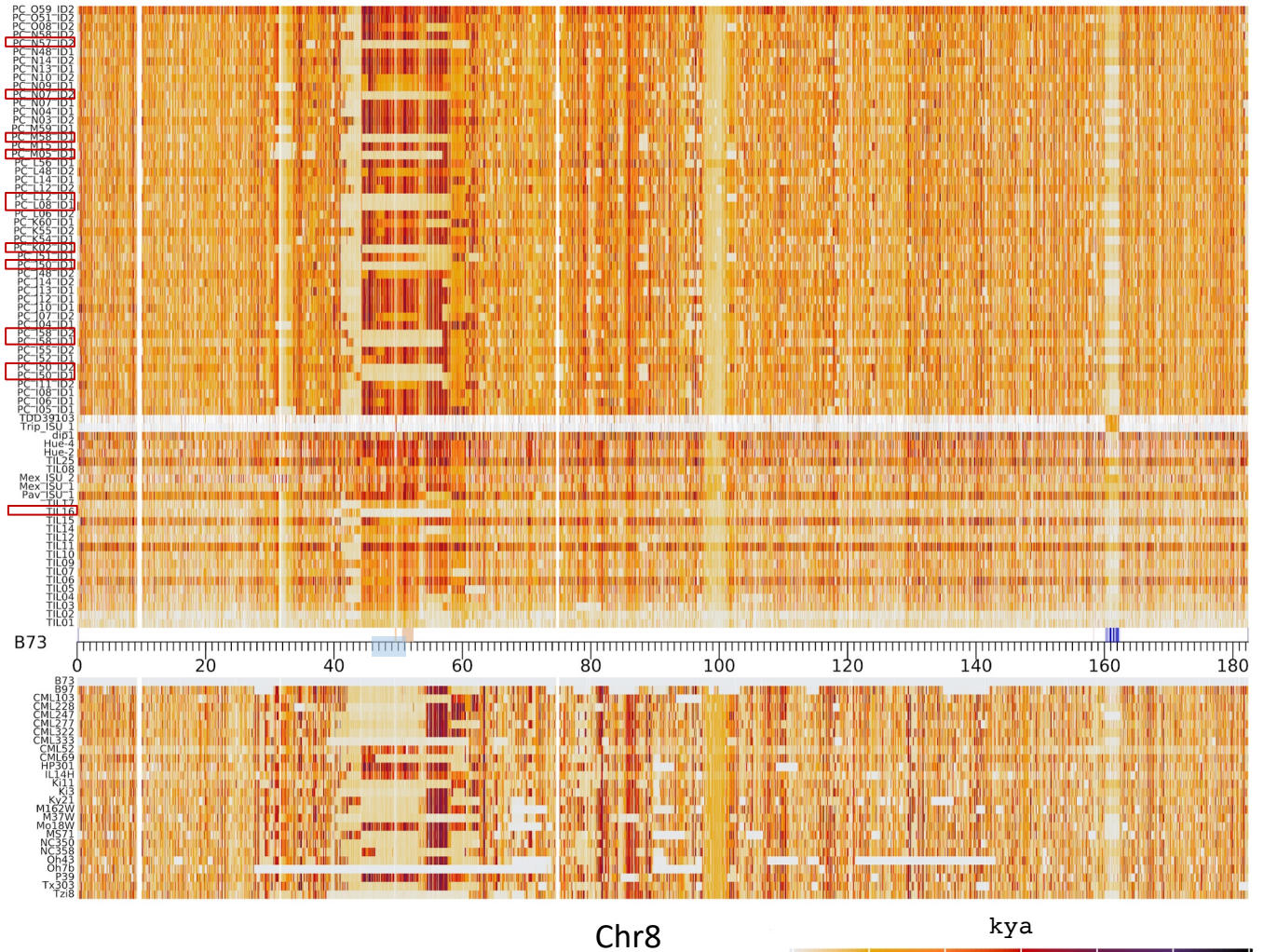

B73
0
20
40
60
80
100
120
140
160
180

Chr6

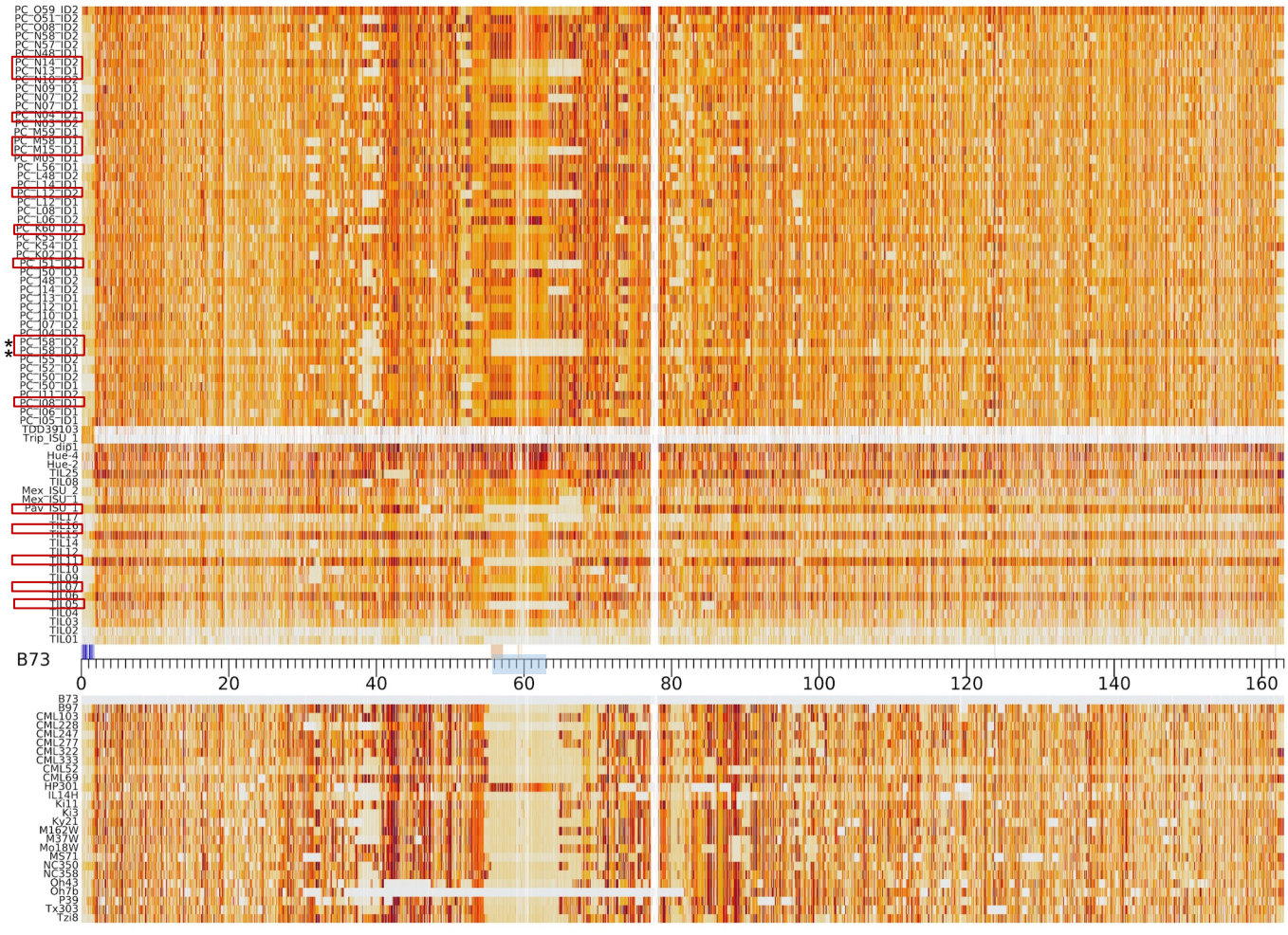

kya
0
50
100
150
200
250
300


B73
0
20
40
60
80
100
120
140
160
180

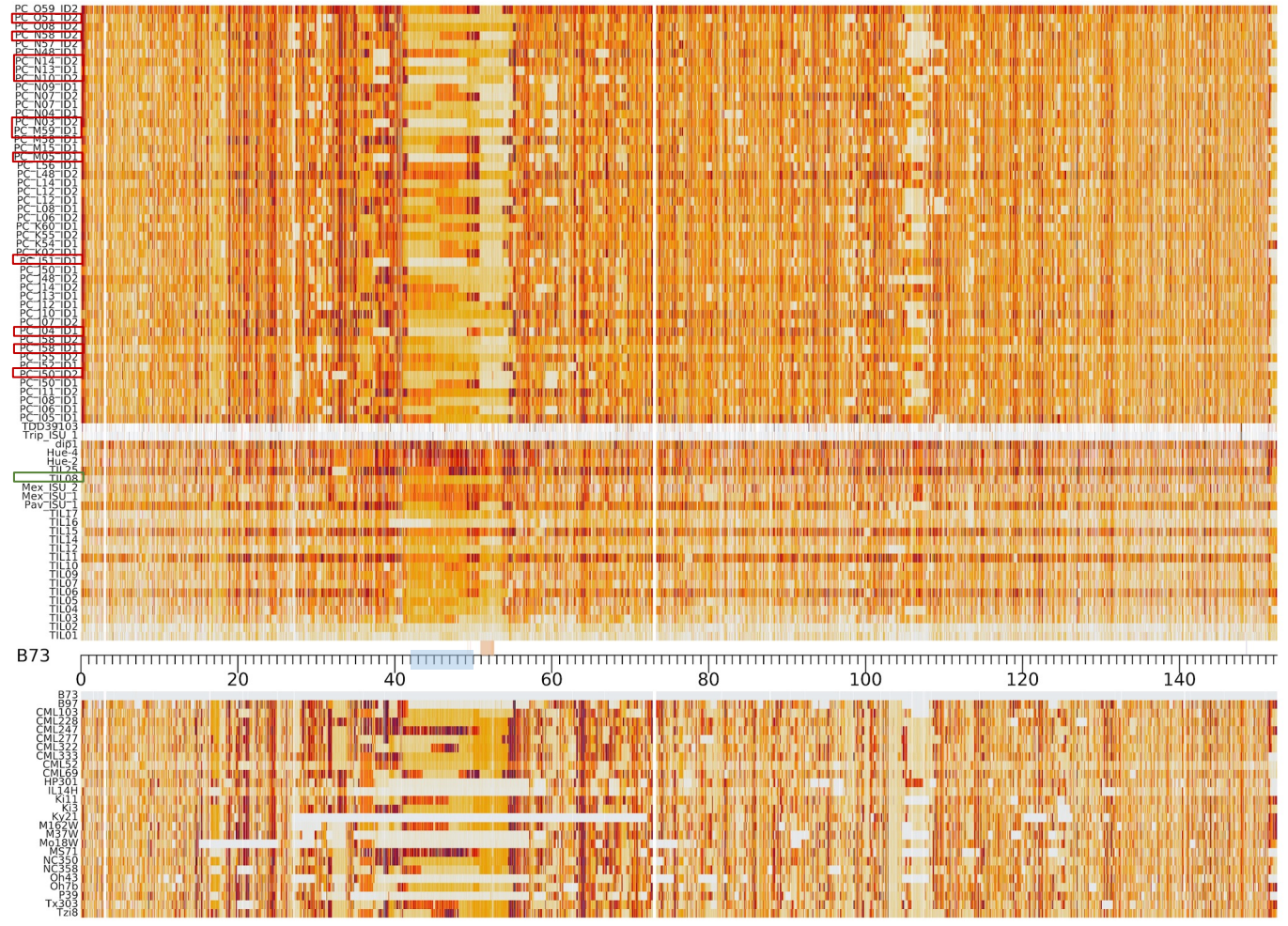
Chr7


B73
0
20
40
60
80
100
120
140
160
180
kya
0
50
100
150
200
250
300

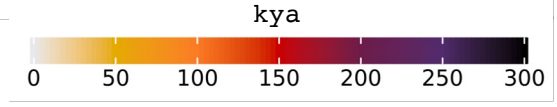
Chr8


B73
0
20
40
60
80
100
120
140
160

Chr9


B73
0
20
40
60
80
100
120
140

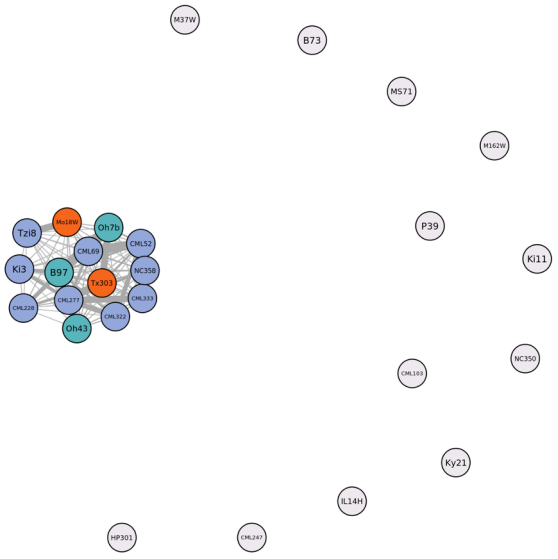
Chr10


kya
0
50
100
150
200
250
300

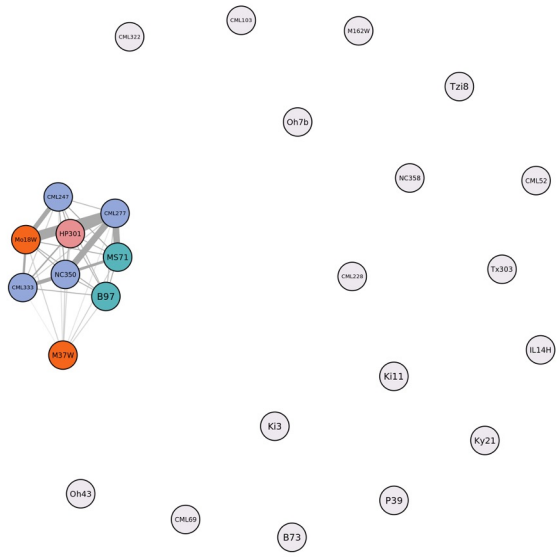
Supplemental Figure S13. Divergence times estimated with short-read mapping using B73 as a reference. Alignments using Illumina reads from the NAM lines are in the lower panel while reads from teosinte are in the upper panel. For teosinte lines, the order of samples from bottom to top is: 15 *parviglumis* lines from the HapMap II project (TIL01, TIL02, TIL03, TIL04, TIL05, TIL06, TIL07, TIL09, TIL10, TIL11, TIL14, TIL15, Pav_1), 4 *mexicana* lines (Mex1, Mex2, ,TIL08, TIL25), 2 *huehuetenangensis* samples (Hue2, Hue4*) , 1 diploperennis* (dip1), 2 *Tripsacum dactyloides* lines (Trip1, TDD39103), and 48 *parviglumis* lines from near Palmar Chico in the Balsas river valley of Mexico. Lines with the B73 haplotypes are highlighted with red (*parviglumis)*, green (*mexicana)* and yellow (*huehuetanangensis).* There is virtually no alignment between Tripsacum and maize in non-genic areas; the only area with obvious alignment is within the NOR on chromosome 6S (Huang et al, 2021). CentC (orange), knob180 (blue), TR-1 (red), and subtelomere (black) are annotated as bars below the alignments. Cenhap domains are highlighted in blue over the x-axes.

A confounding factor when analyzing teosinte accessions is heterozygosity. Reads from a heterozygous line when aligned to a single reference will reveal SNPs for both haplotypes and the age estimate will be an average. Nevertheless, we would expect occasional cenhap homozygosity. Note that in several cases, teosinte cenhaps appear to be (homozygous and) identical to the B73 cenhap, suggesting recent introgression. These are noted with asterisks *; see chromosome 1, chromosome 6, and chromosome 9.

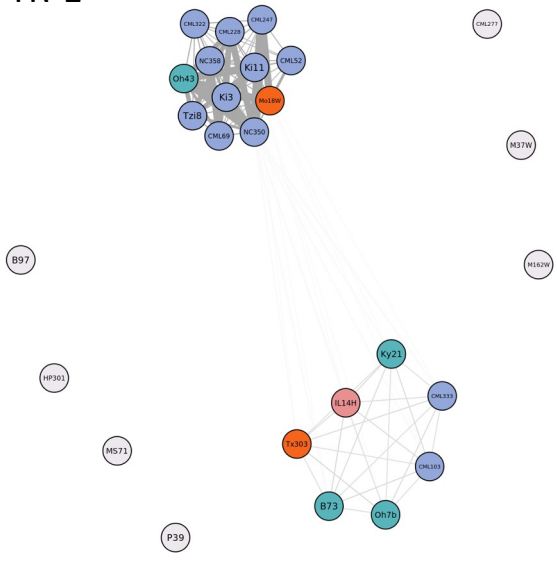
2L knob180



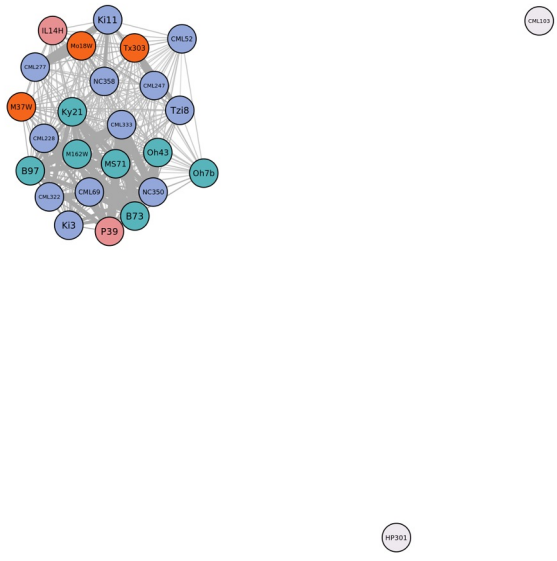
3L knob180



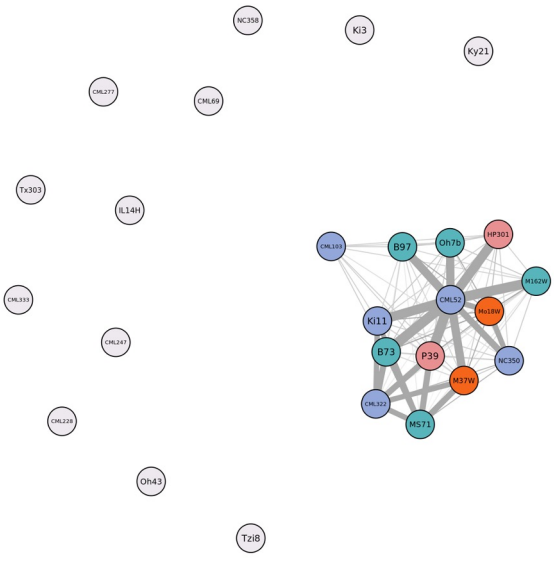
4L TR-1



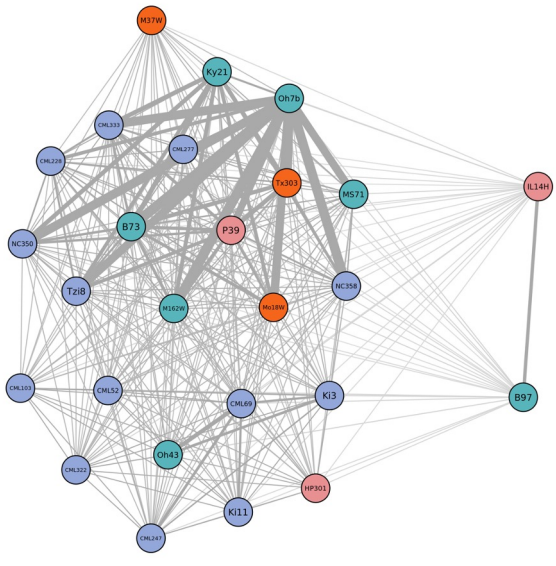
5L knob180

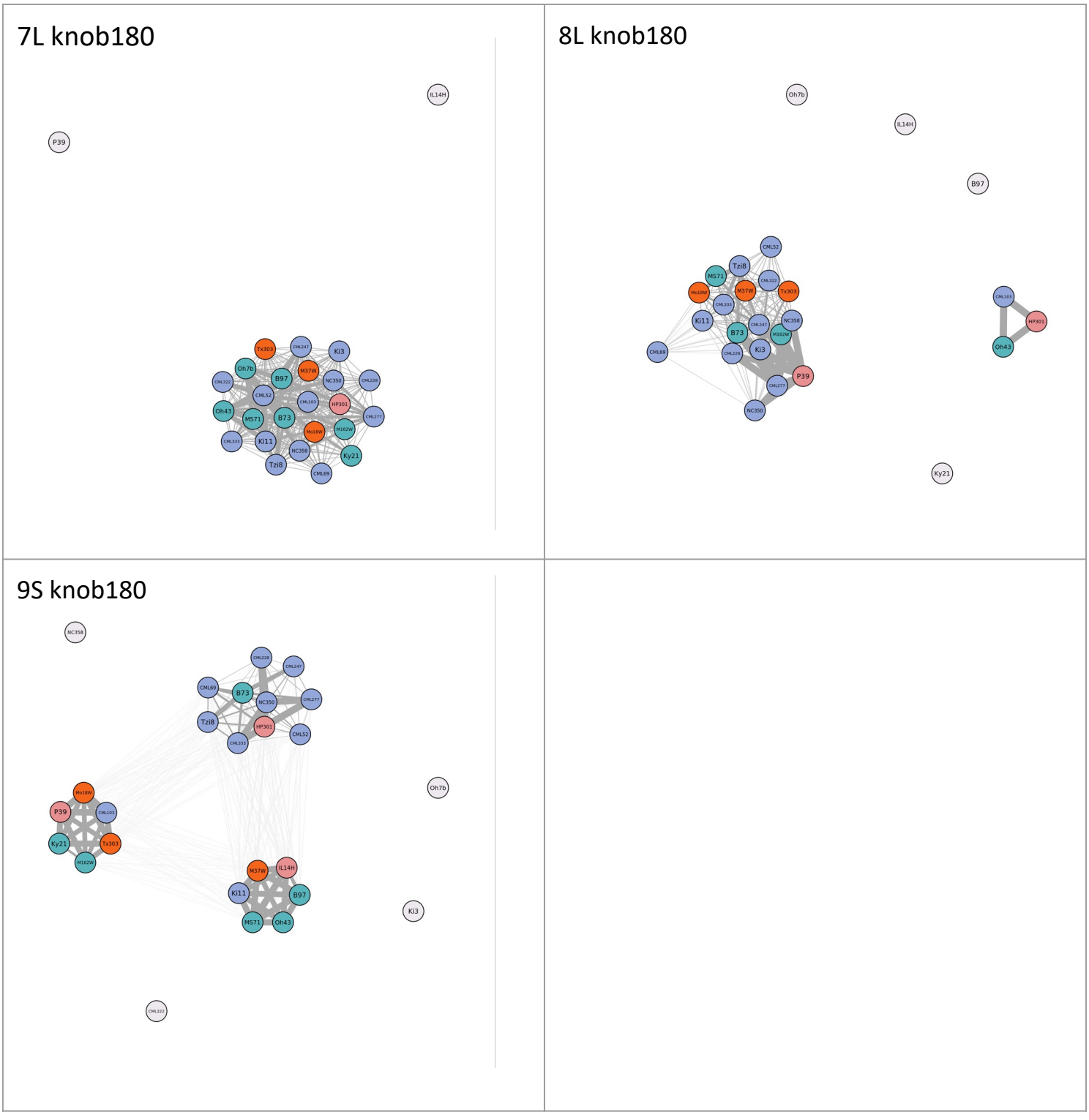


6S1 TR-1

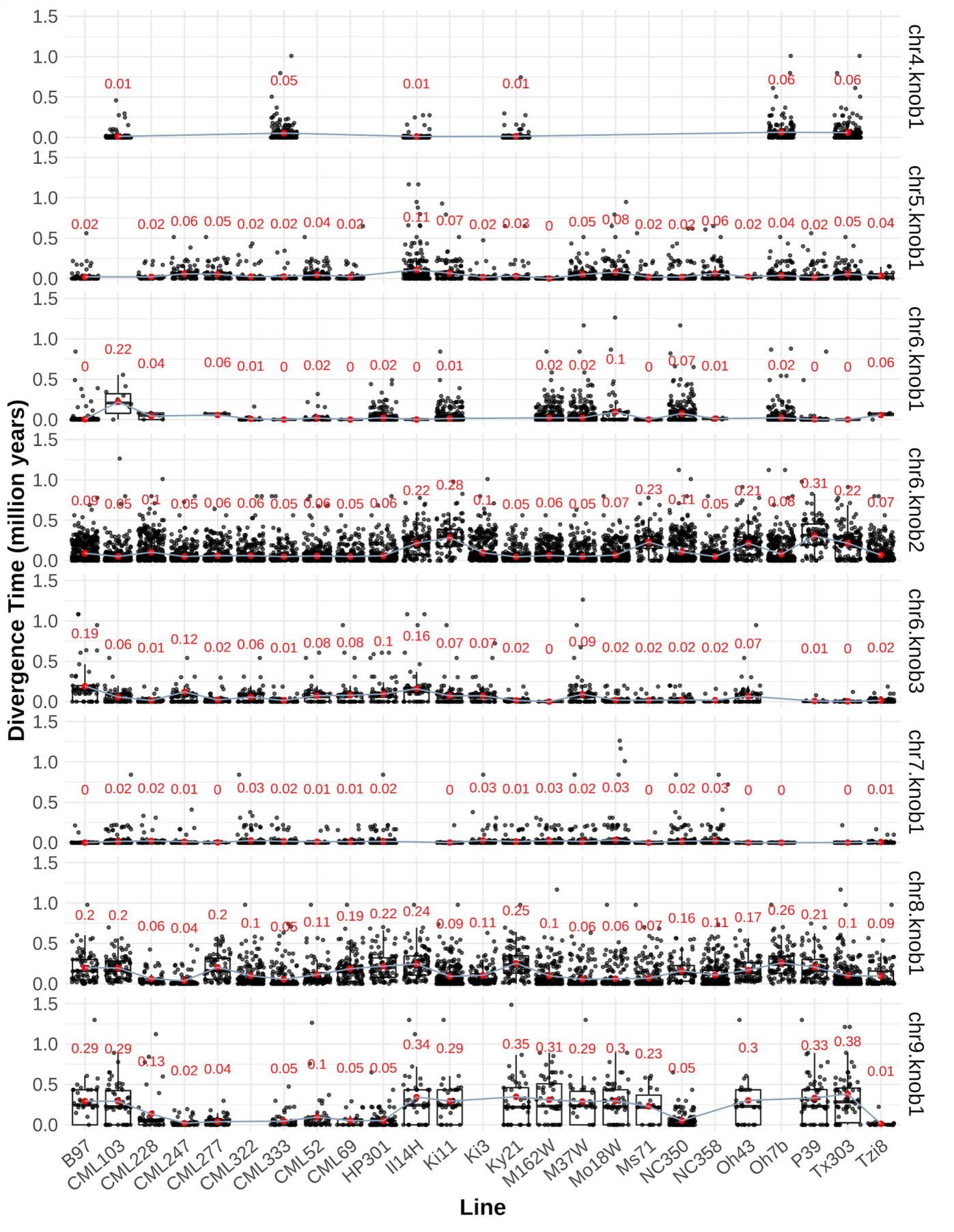


6L knob180

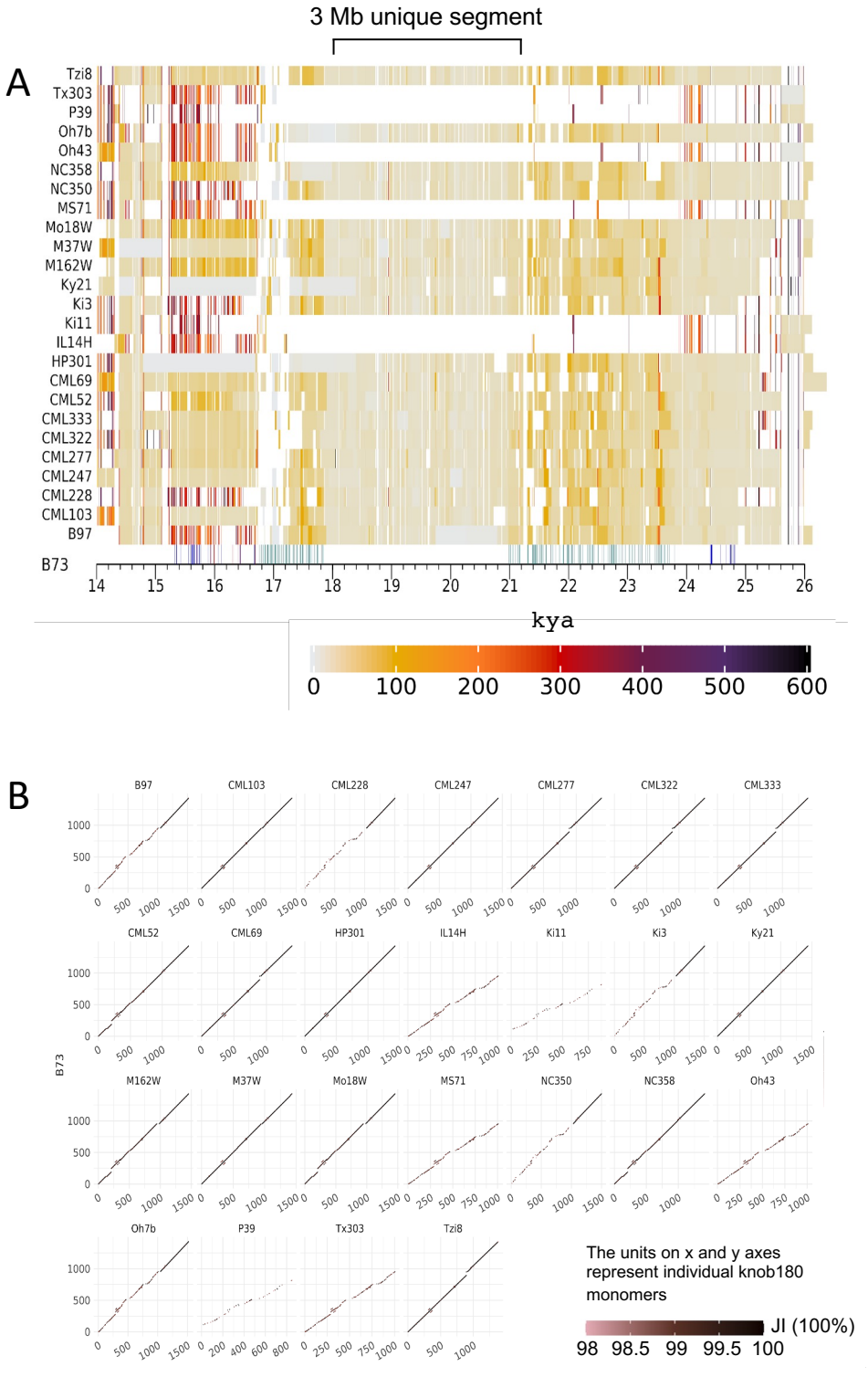




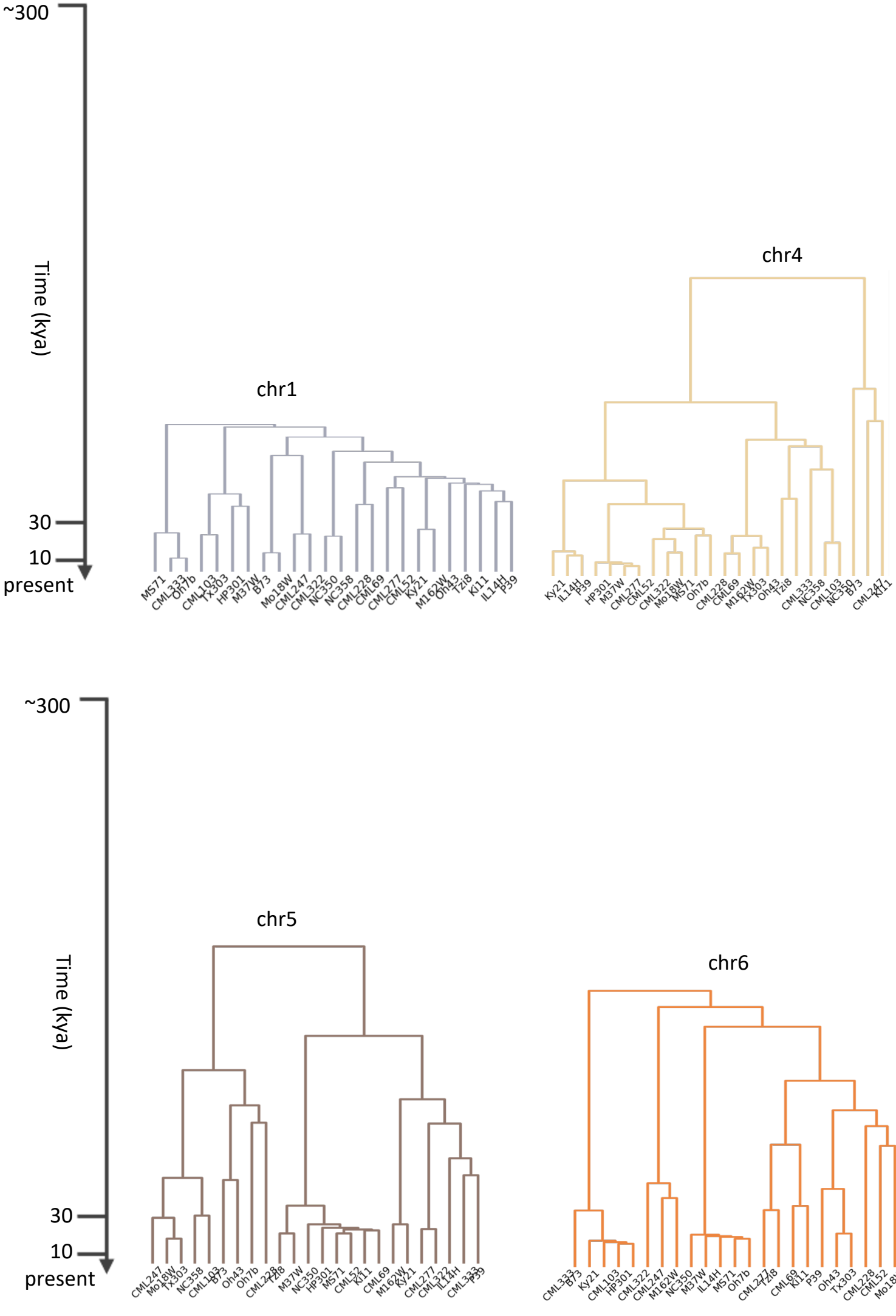
Supplemental Figure S14. Clustering of knob arrays based on all-by-all alignment. Knobs are labeled by the chromosome number, short (S) or long (L) arm, and the dominant repeat class (knob180 or TR-1). Colors over inbred names indicate varieties of corn: northern flint (pink), temperate (blue), mixed (red), and tropical (green) maize varieties. White circles indicate the lack of a syntenic knob for that line.

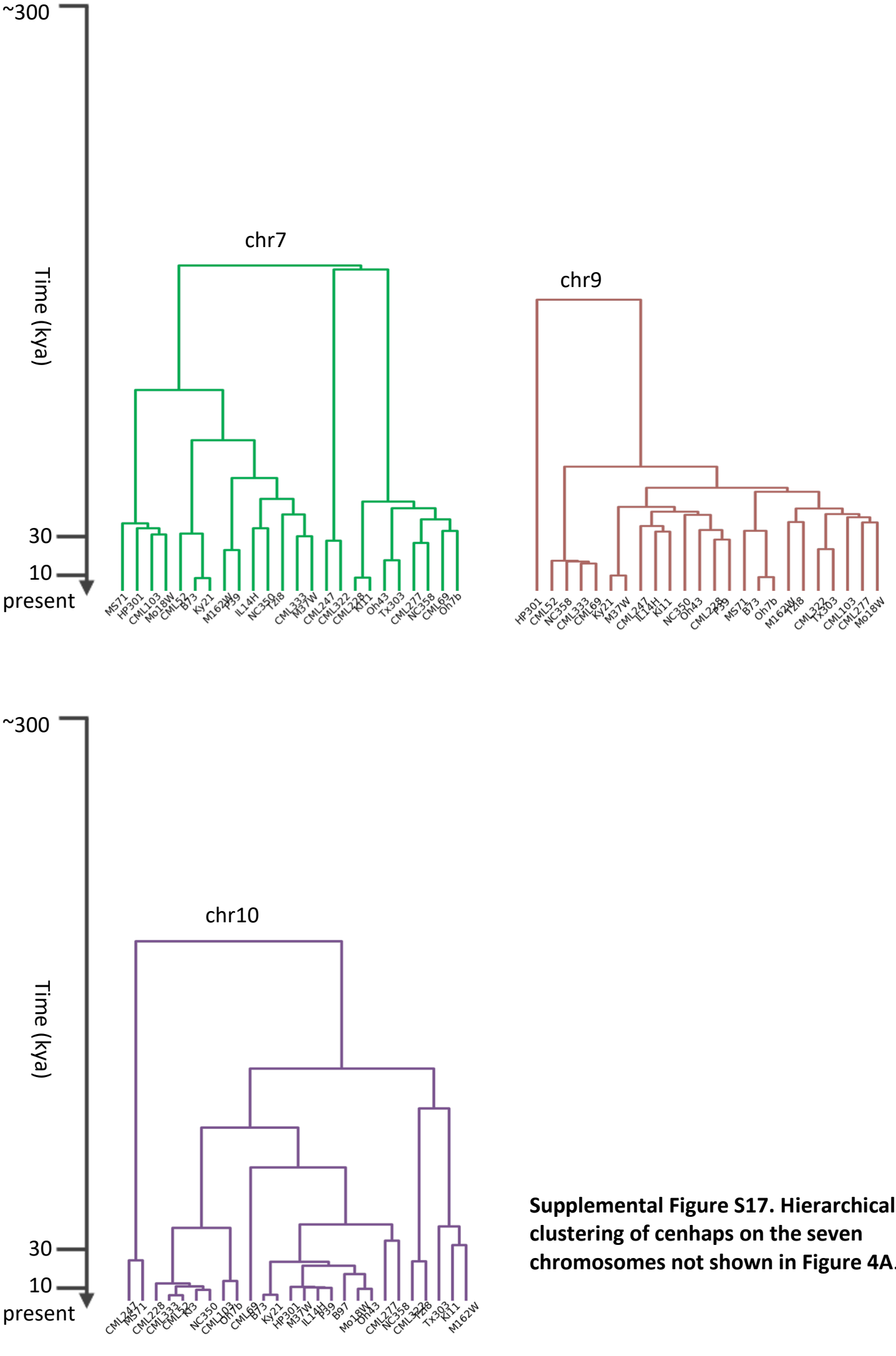


Supplemental Figure S15. Estimated divergence times of eight classical knobs relative to B73. Each dot represents the divergence value inferred from an individual syntenic aligned transposon fragment. Mean divergence times are labelled, highlighted in red and connected by a blue line across lines. Empty values indicate an absence of syntenic knobs for the corresponding lines. Note that the Chr6.knob2 corresponds to the 6L knob180 in Supplemental Fig. S14.

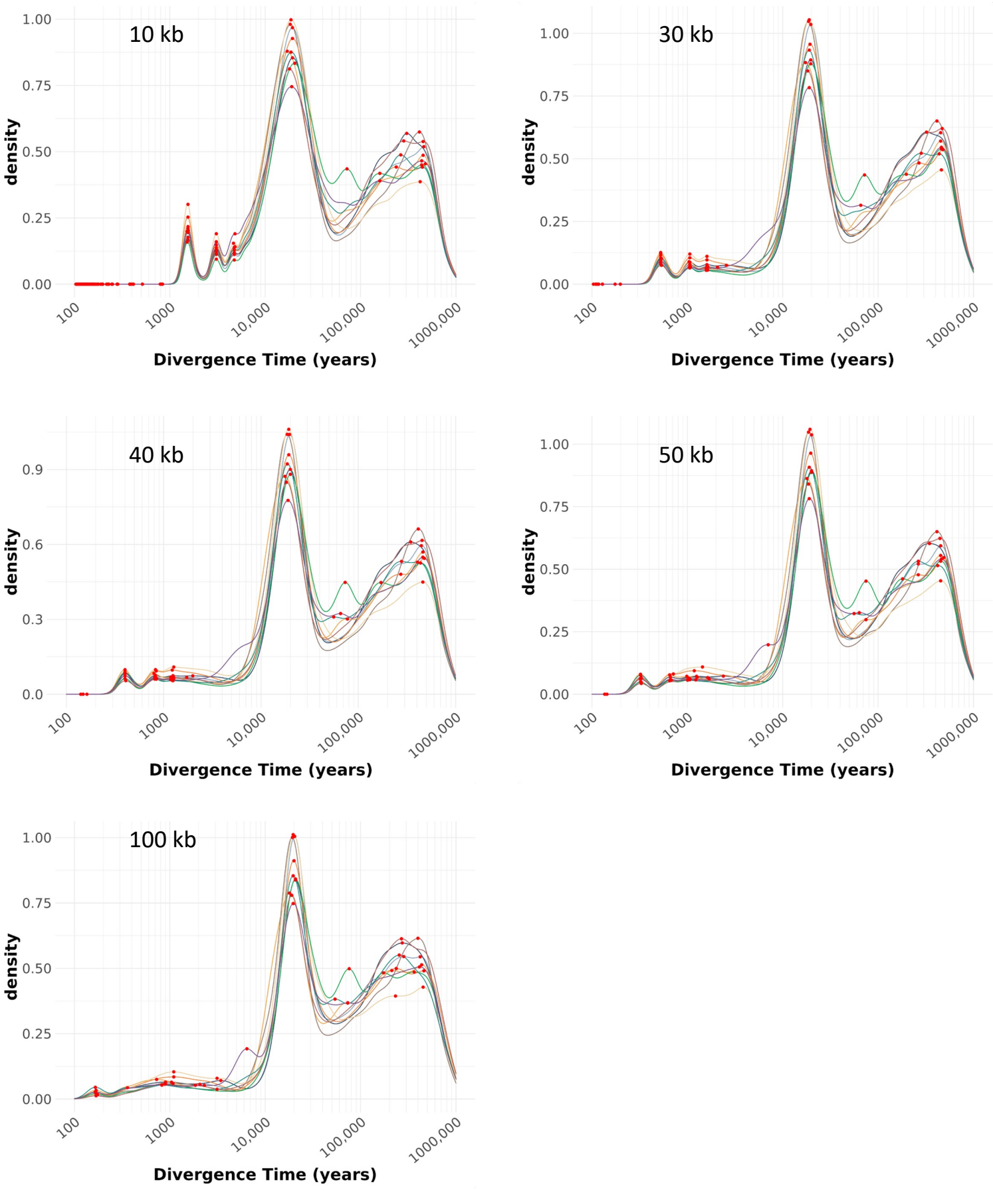


Supplemental Figure S16. Haplotypes of the NOR and adjacent 6S knob. **A)** Alignment and divergence time estimation over the NOR region using B73 as a reference. White spaces depict unaligned regions between query genome and B73. The 3 Mb unique segment is highlighted. Knob180 (blue), TR-1 (red), and rDNA (cyan) are annotated as bars below the alignments. **B)** Pairwise alignment between 25 NAM lines and B73 over the 6S knob180 array, which is adjacent to NOR (chr6:18-21 Mb).


~300
Time (kya)
30
10
present
chr1
MS71
CML333
Oh7b
CML103
TX303
HP301
M37W
B73
Mo18W
CML247
CML322
NC350
NC358
CML228
CML69
CML277
CML52
KY21
M162W
Oh43
Tzi8
Ki11
IL14H
P39
chr4
Ky21
IL14H
P39
HP301
M37W
CML277
CML52
CML322
Mo18W
MS71
Oh7b
CML228
CML69
M162W
TX303
Oh43
Tzi8
CML333
NC358
CML103
NC350
B73
CML247
Ki11
~300
Time (kya)
30
10
present
chr5
CML247
Mo18W
TX303
NC358
CML103
B73
Oh43
Oh7b
CML228
Tzi8
M37W
NC350
HP301
MS71
CML52
Ki11
CML69
M162W
KY21
CML277
CML322
IL14H
CML333
P39
chr6
CML333
B73
Ky21
CML103
HP301
CML322
CML247
M162W
NC350
M37W
IL14H
MS71
Oh7b
CML277
Tzi8
CML69
Ki11
P39
Oh43
TX303
CML228
CML52
Mo18W



Supplemental Figure S17. Hierarchical clustering of cenhaps on the seven chromosomes not shown in Figure 4A.



Supplemental Figure S18. Whole genome age density plots computed using windows of 10 kb, 30 kb, 40 kb, 50 kb and 100 kb.