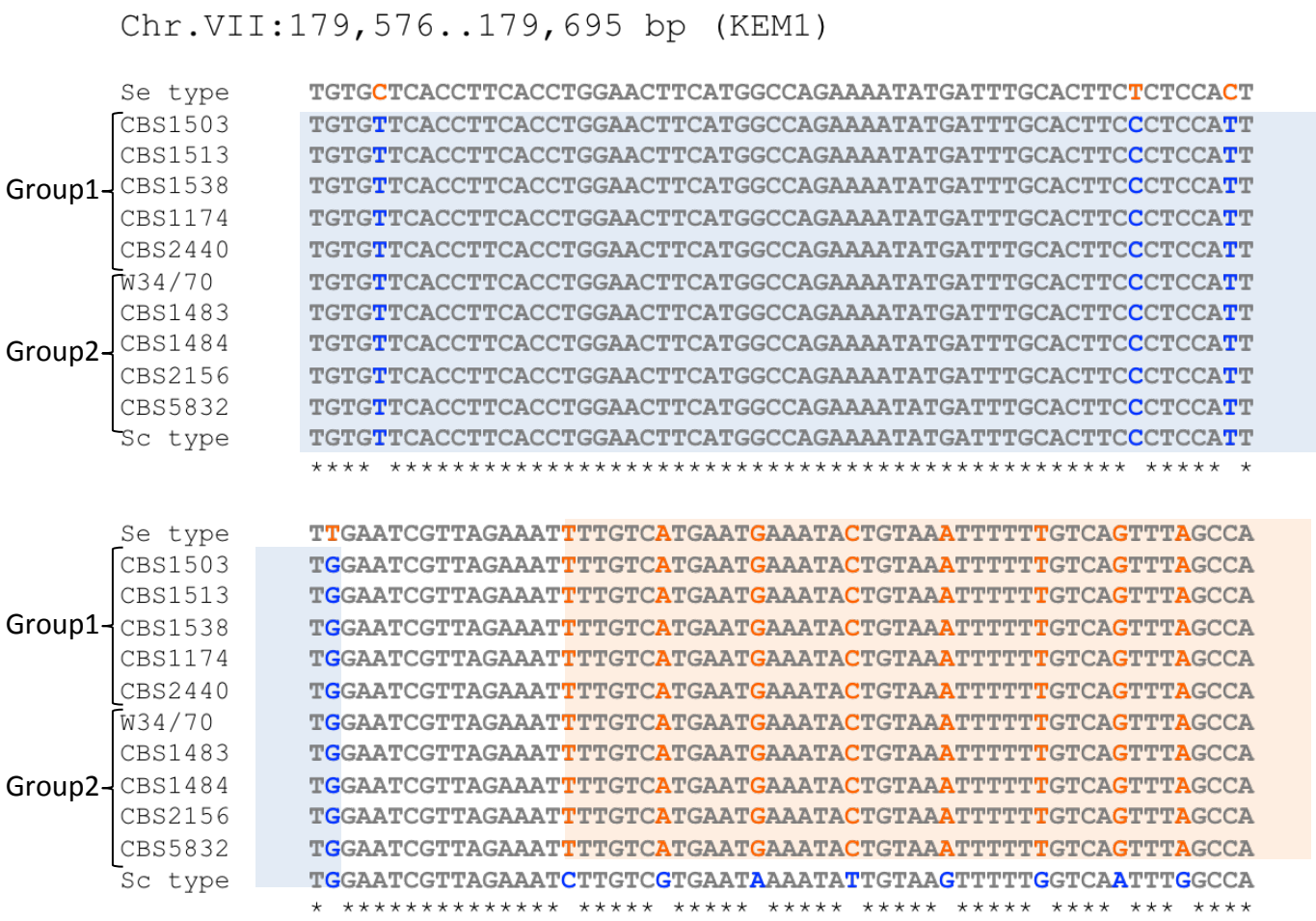
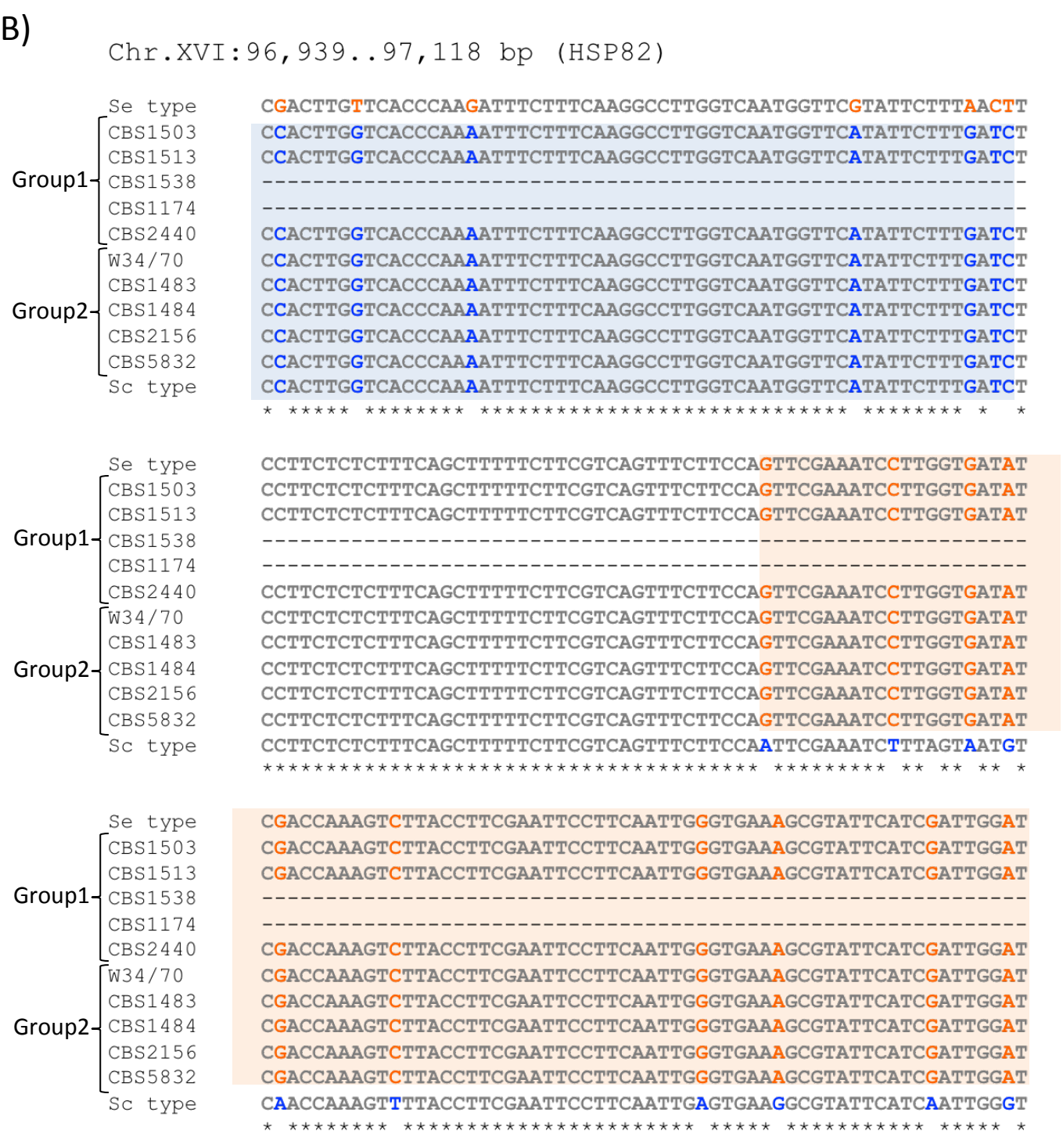
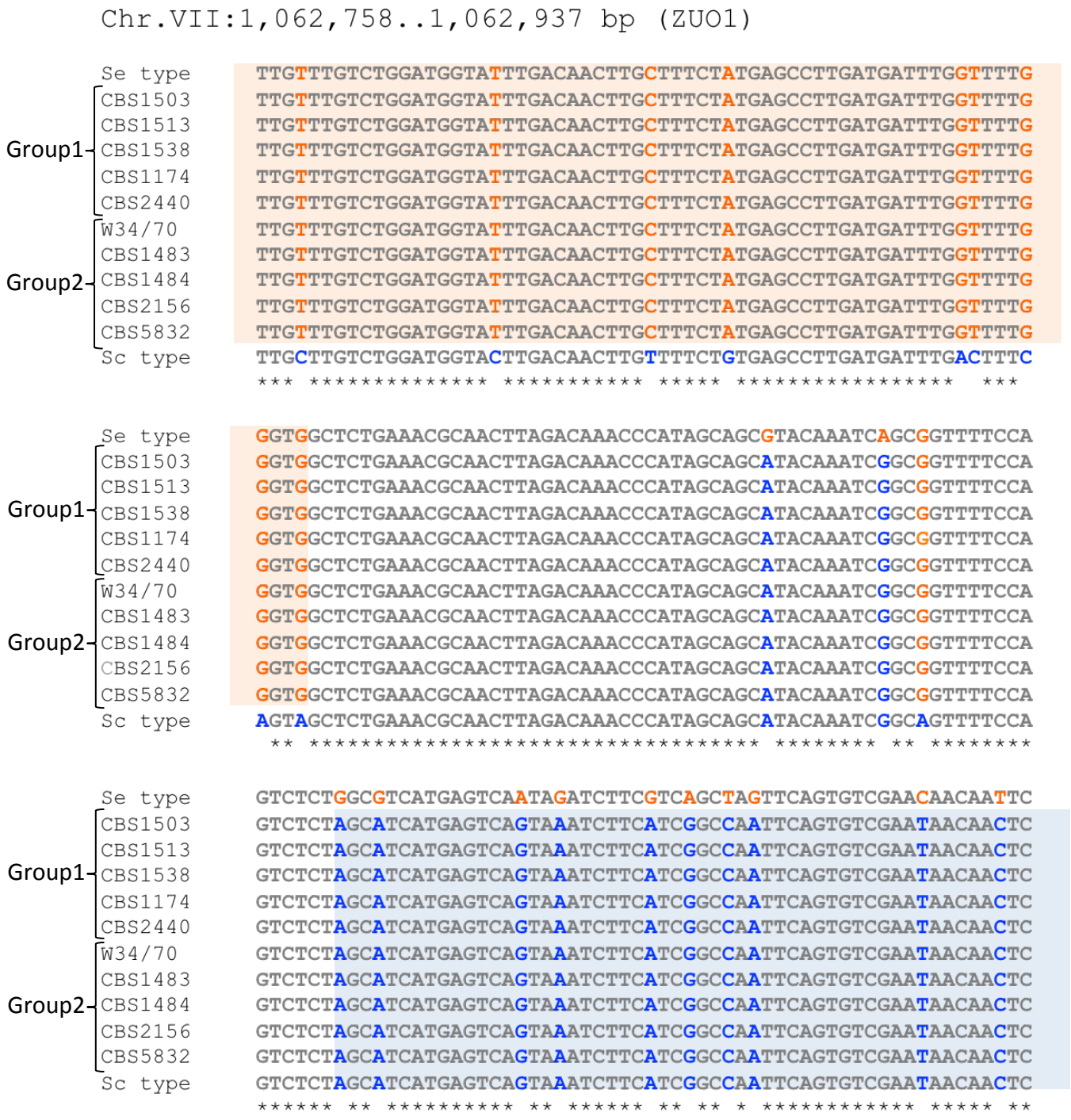
A)


Chr.VII:179,576..179,695 bp (KEM1)
Se type
CBS1503
CBS1513
Group1
CBS1538
CBS1174
CBS2440
W34/70
CBS1483
Group2
CBS1484
CBS2156
CBS5832
Sc type
TGTGCTCACCTTCACCTGGAACTTCATGGCCAGAAAATATGATTTGCACTTCTCTCCACT
TGTGTTCACCTTCACCTGGAACTTCATGGCCAGAAAATATGATTTGCACTTCCCTCCATT
TGGAATCGTTAGAAATTTTGTCATGAATGAAATACTGTAAATTTTTTGTCAGTTTAGCCA
TGGAATCGTTAGAAATCTTGTCGTGAATAAAATATTGTAAGTTTTTGGTCAATTTGGCCA



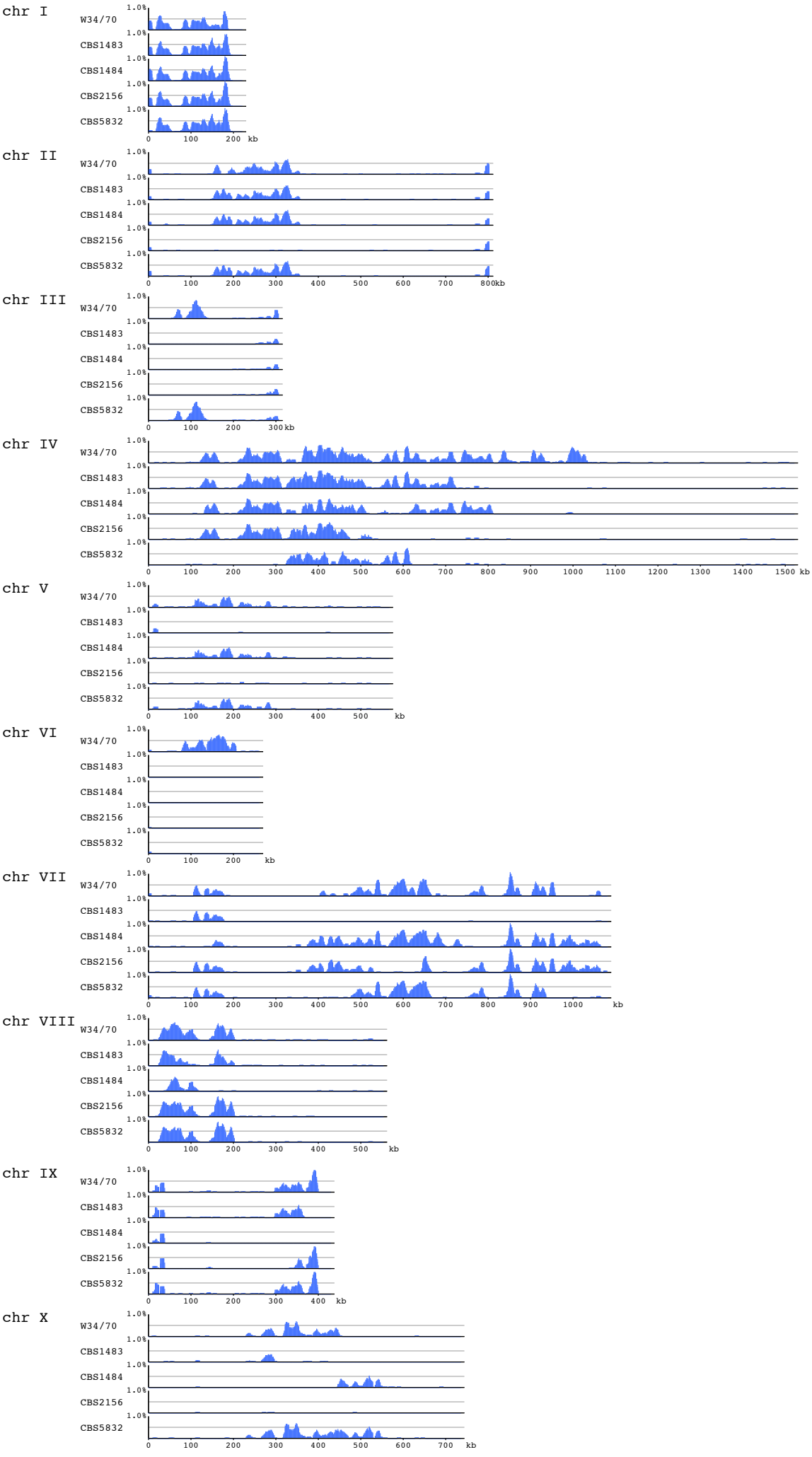
Supplementary Figure 1. Multiple sequence alignment of 10 *S. pastorianus* strains surrounding the breakpoint (A.) KEM1 locus B.) HSP82 locus). Shown are multiple sequence alignments of *S. pastorianus* strains and parental species *S. cerevisiae* and *S. eubayanus* (Sc and Se types, respectively) surrounding the breakpoint regions. Sc and Se type sequences (SNVs) are indicated by blue and orange characters, respectively. Based on SNV distribution, blue and orange background regions were presumed to be derived from Sc- and Se-type parents. The breakpoint was matched at the nucleotide level across all Group 1 and 2 strains.

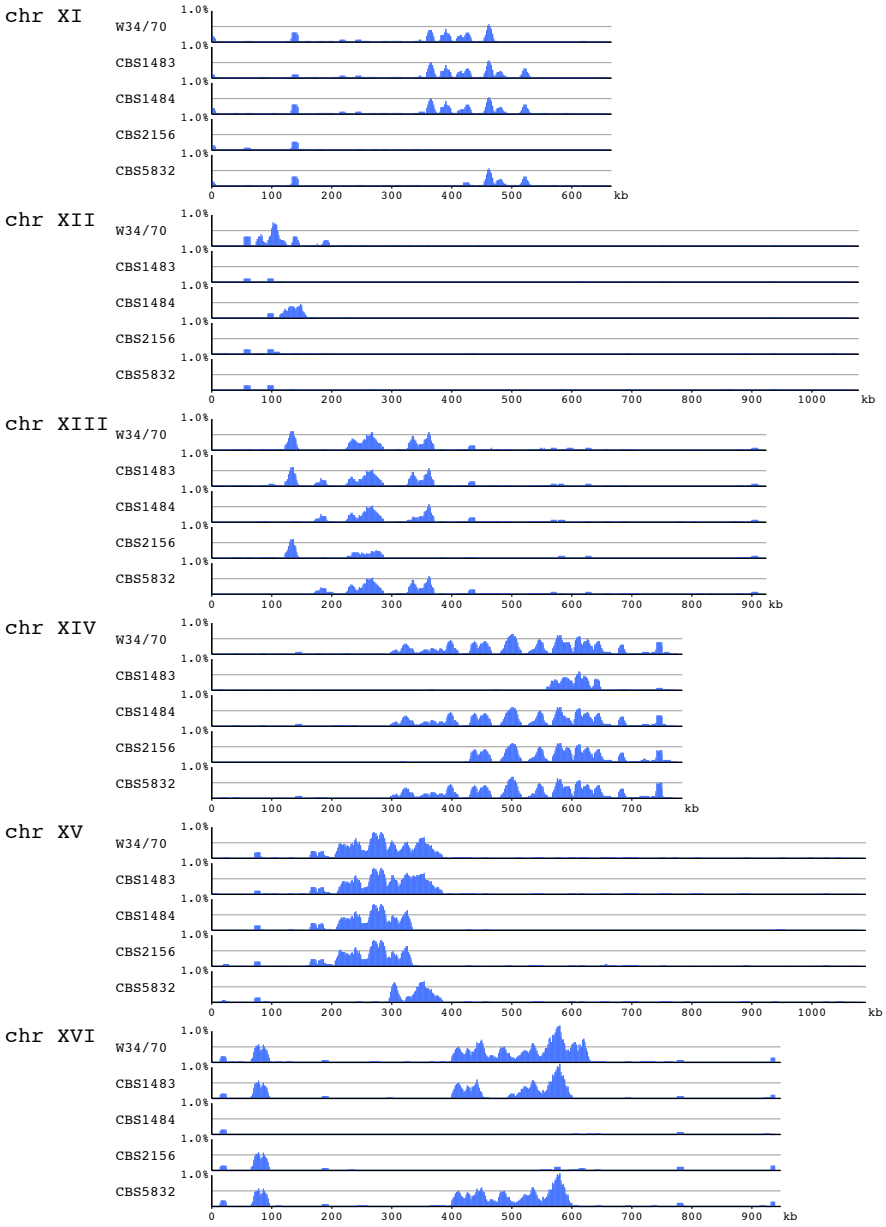


Supplemental Figure 2. Multiple sequence alignment of 10 *S. pastorianus* strains surrounding the novel discovered breakpoint (ZUO1 locus)

Multiple sequence alignments of 10 *S. pastorianus* strains and the parental species *S. cerevisiae* (Sc type) and *S. eubayanus* (Se type) surrounding the breakpoint regions are shown. Blue characters (SNPs) indicate Sc type sequence and orange characters indicate Se type. According to the distribution of SNPs, blue background region is implied to be derived from Sc type parent and orange background region is from Se type parent.

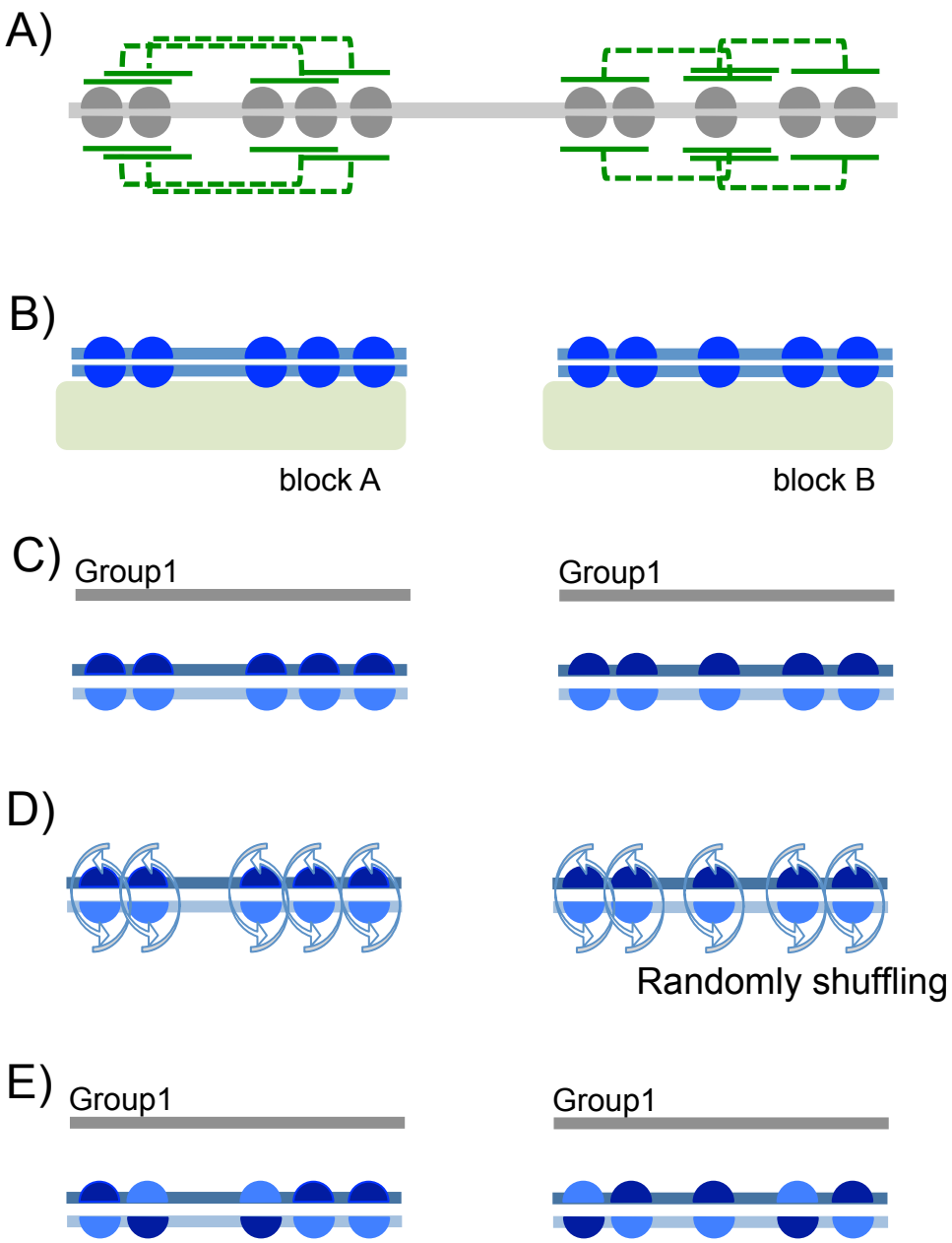
Supplemental Figure 3. Sequence alignments surrounding the novel discovered breakpoint in W34/70 (PRP8 locus)
Multiple sequence alignments of W34/70 and the parental species *S. cerevisiae* (Sc type) and *S. eubayanus* (Se type) surrounding the breakpoint regions are shown. Blue characters (SNPs) indicate Sc type sequence and orange characters indicate Se type. According to the distribution of SNPs, blue background region is implied to be derived from Sc type parent and orange background region is from Se type parent.


chr I
chr II
chr III
chr IV
chr V
chr VI
chr VII
chr VIII
chr IX
chr X
W34/70
CBS1483
CBS1484
CBS2156
CBS5832
1.0%
kb



Supplementary Figure 4. Distribution of hetero SNVs in Sc type chromosome of Group 2 strains

Distribution of heterozygosity along with Sc type chromosomes in five Group 2 strains (CBS1483, CBS1484, CBS2156, CBS5832 and W34/70, from the top in each chromosome) were shown by calculating moving average of SNVs density (window size 10 kbp and step size 1 kbp). Hetero SNVs are distributed unevenly and it might be due to LOH. In many of regions, low/none heterozygosity are observed in all 5 strains, but in some regions, the distribution patterns of low/none heterozygosity regions are different between them. So, LOH is considered the relatively occurring events and common low/none heterozygosity regions were shaped in the common ancestor or as a result of overlapping LOH events.



Supplementary Figure 5. Schematic view of making haplotype phased blocks and "random shuffling" test

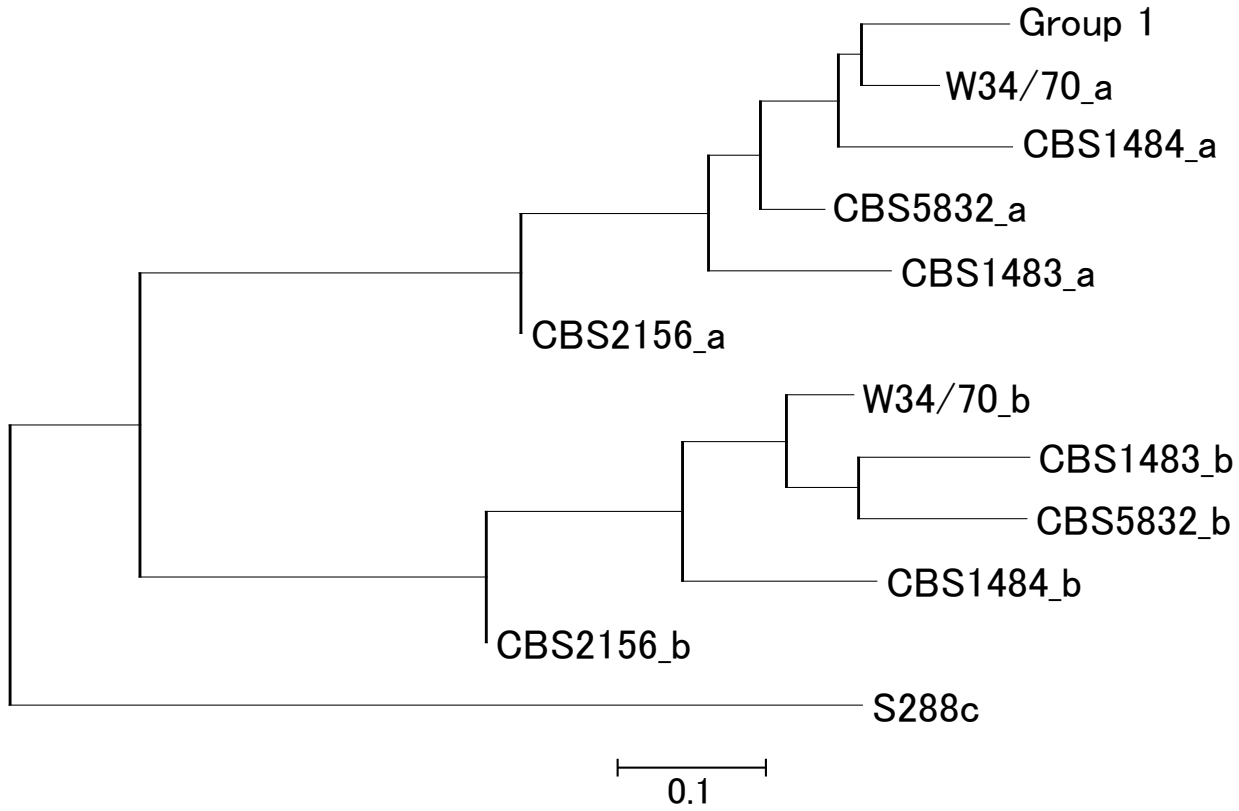
A) linkage information between adjacent hetero SNVs using paired-end or mate-pair reads

B) according to solved linkage information, haplotype phased blocks were built. In each block, all the including hetero SNVs are separated with linkage information. 2 haplotype phased chromosomes were obtained.

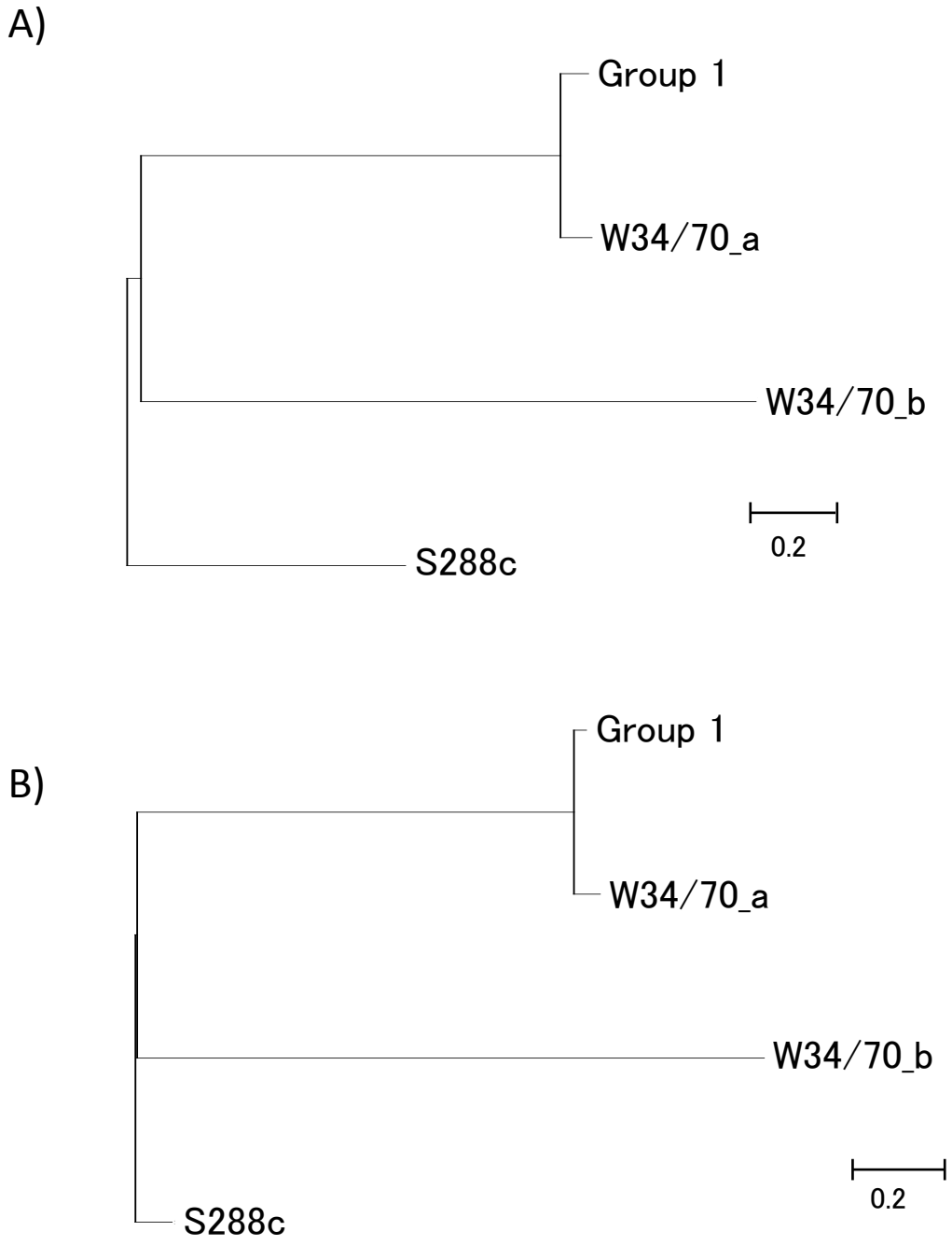
C) comparing with Group 1 sequences, haplotype phased chromosomes were divided (_a : similar with Group 1, _b: another chromosomes)

D) schematic view of random shuffling test. allele sequences are randomly shuffled dropping off the linkage information

E) as mentioned in C), 2 haplotype phased chromosomes with randomly shuffled alleles were divided according to similarity with Group 1



Supplementary Figure 6. Phylogenetic tree haplotype phased Group 2 sequences and Group 1 consensus sequences

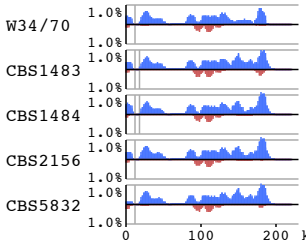


Supplementary Figure 7. Phylogenetic tree based on haplotype phased W34/70 sequences and Group 1 sequences

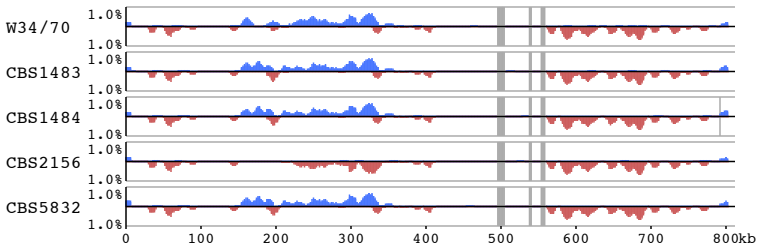
A) Threshold length =1000 bp B) Threshold length = 300 bp.
If the distance between adjacent SNV loci exceeded the threshold, the blocks were divided.

CBS1503

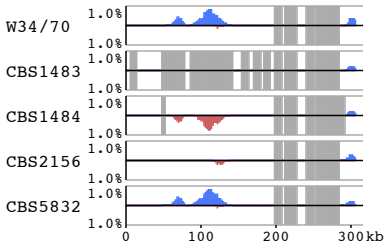
chr I



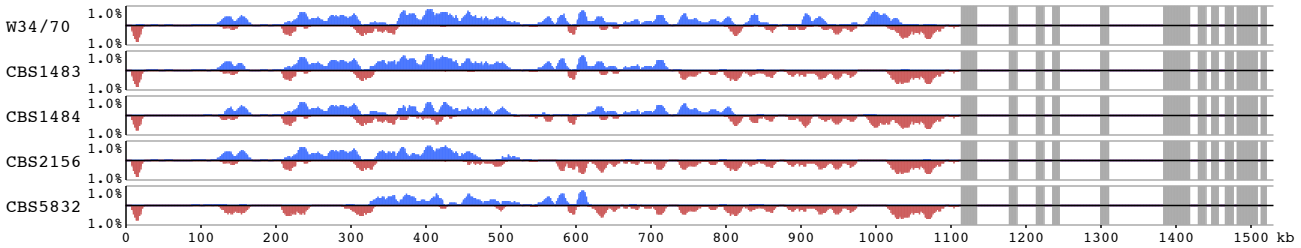
chr II



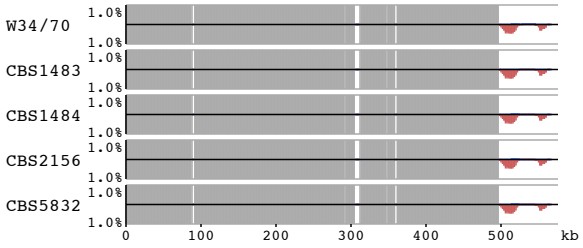
chr III



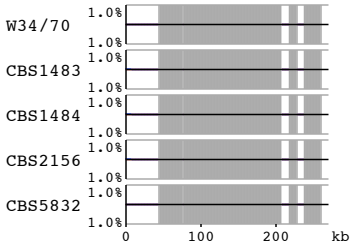
chr IV



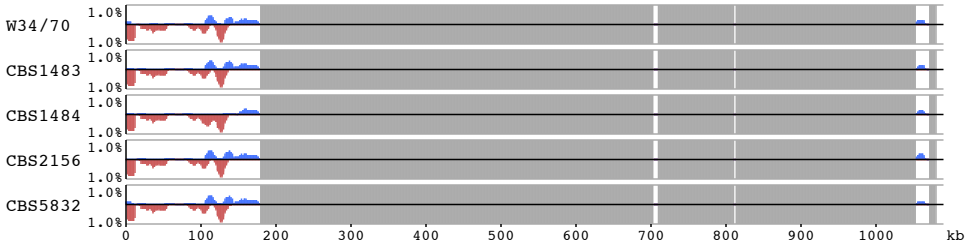
chr V



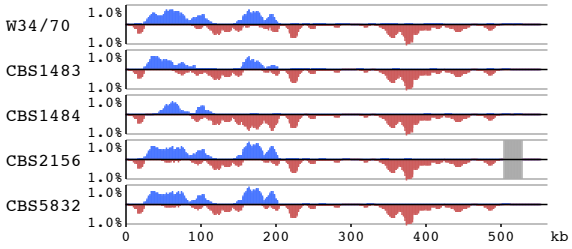
chr VI



chr VII

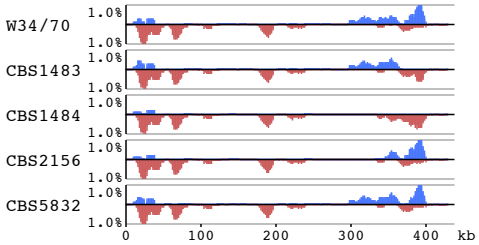


chr VIII

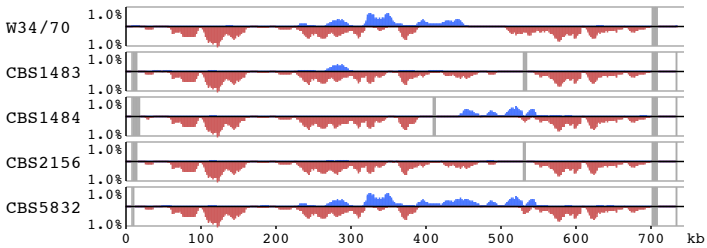


CBS1503

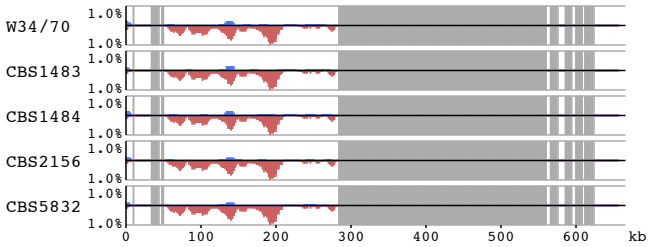
chr IX



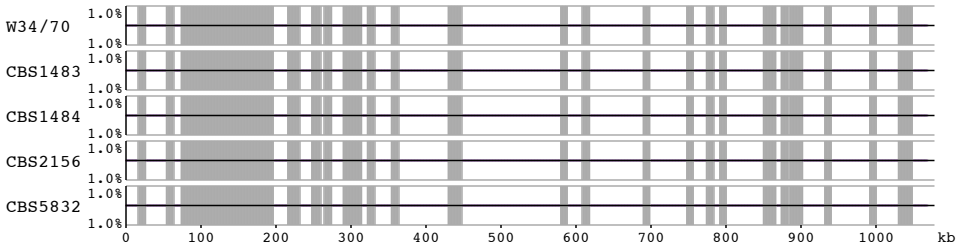
chr X



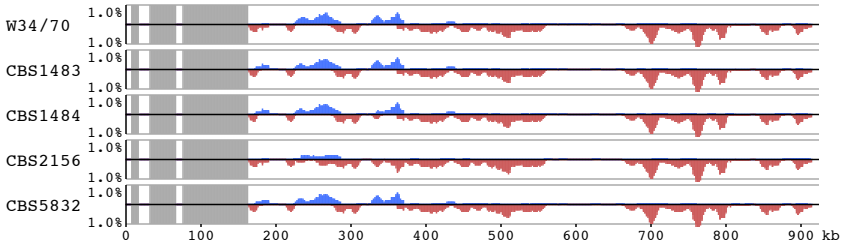
chr XI



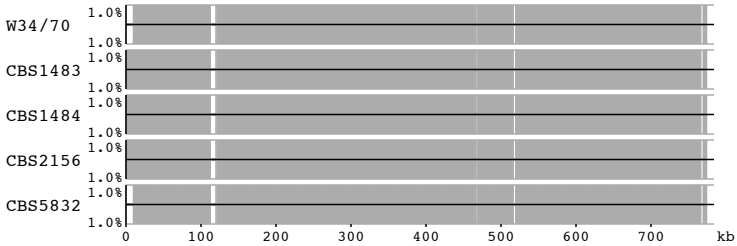
chr XII



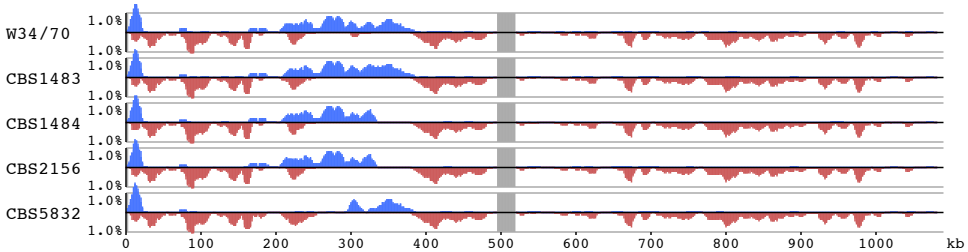
chr XIII



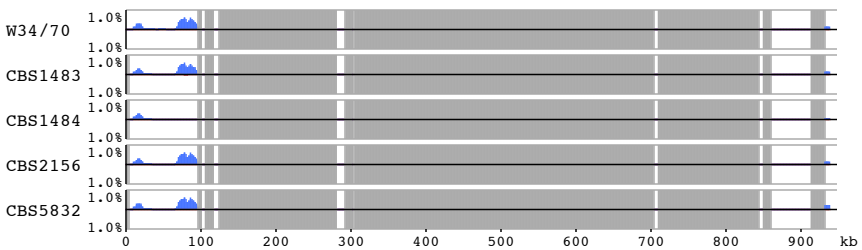
chr XIV



chr XV

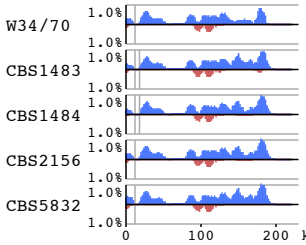


chr XVI

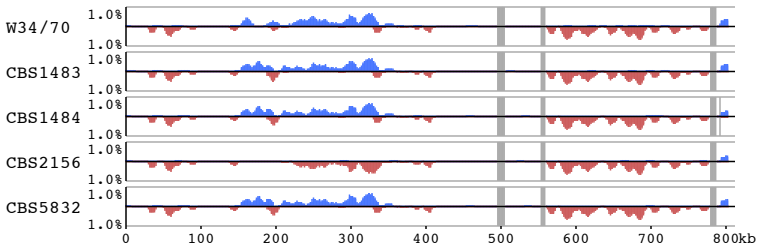


CBS1513

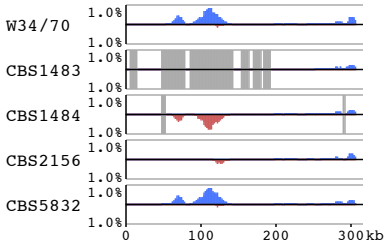
chr I



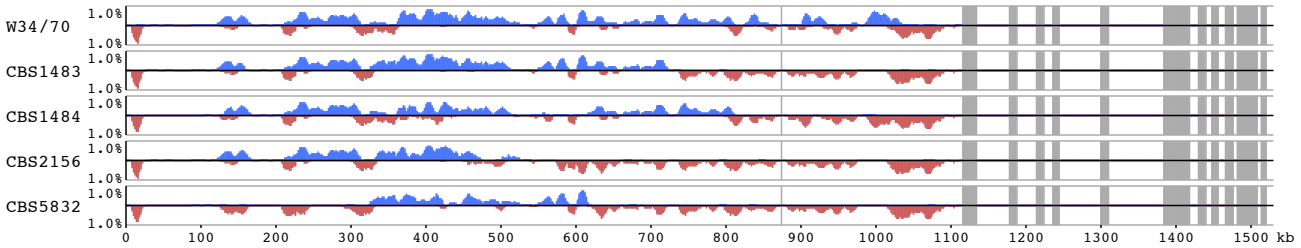
chr II



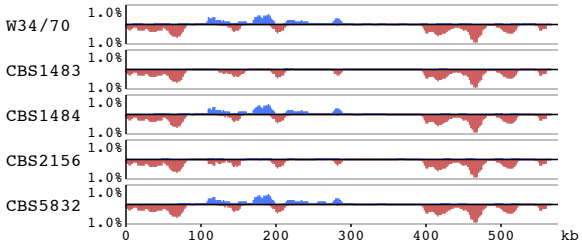
chr III



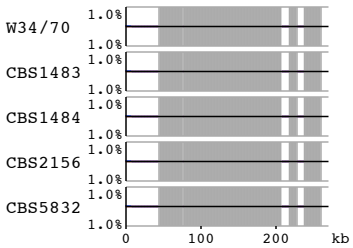
chr IV



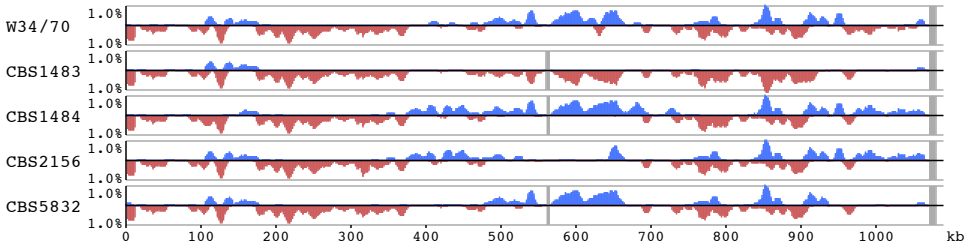
chr V



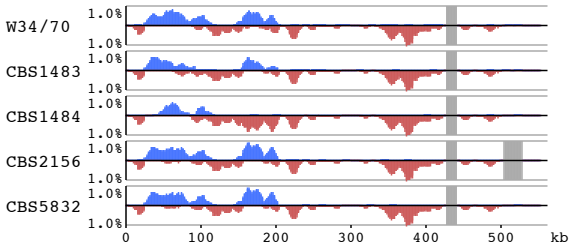
chr VI



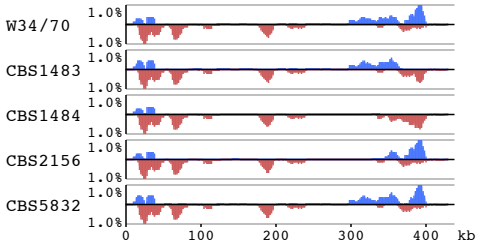
chr VII



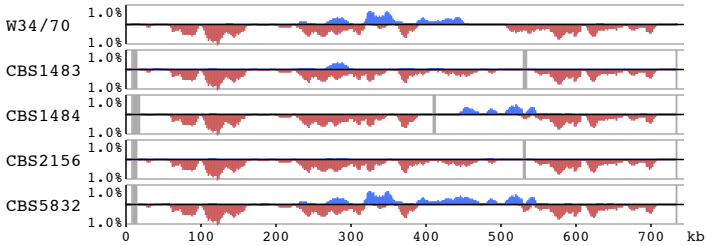
chr VIII



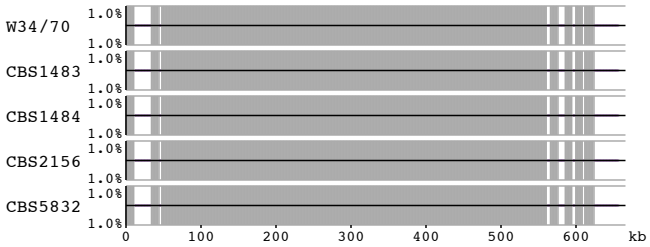
chr IX



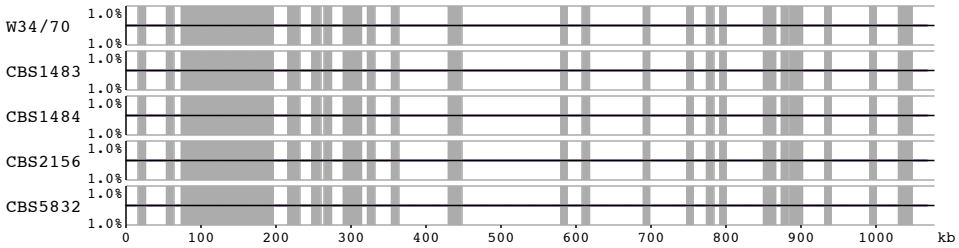
chr X



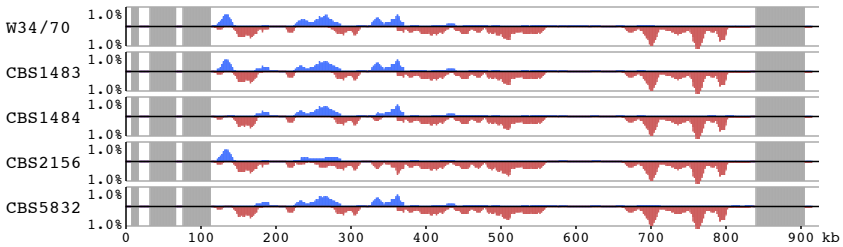
chr XI



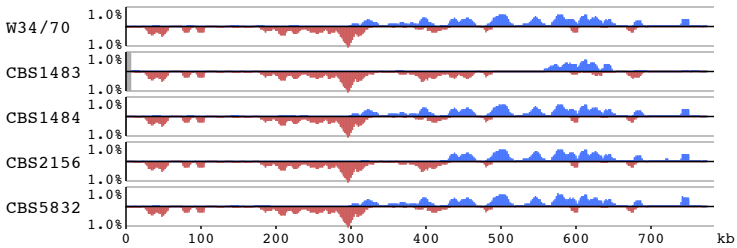
chr XII



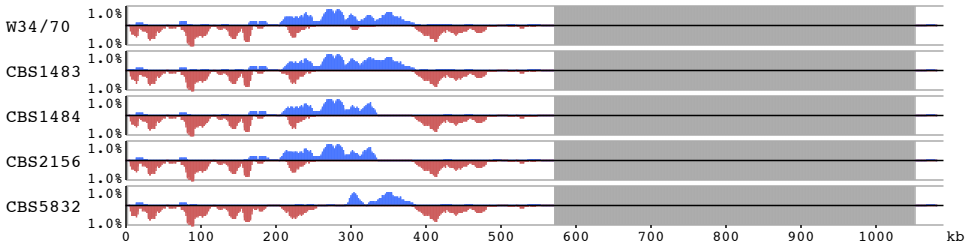
chr XIII



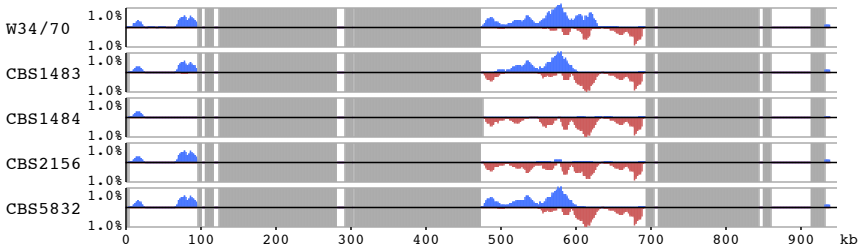
chr XIV



chr XV

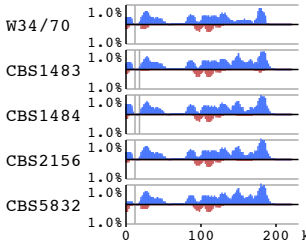


chr XVI

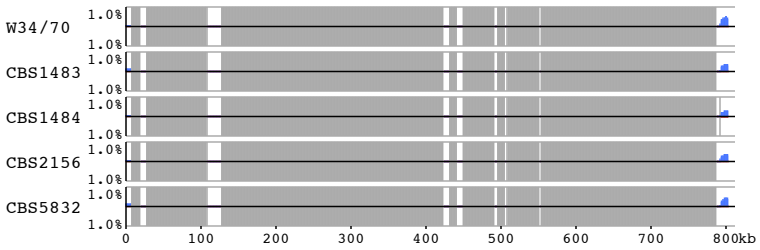


CBS1538

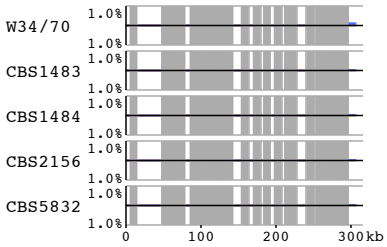
chr I



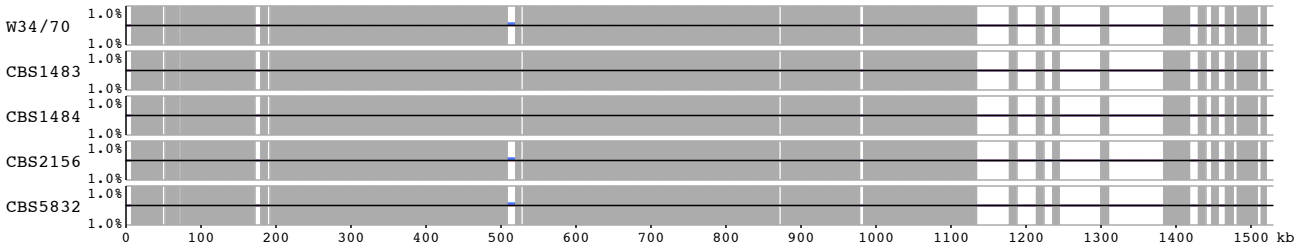
chr II



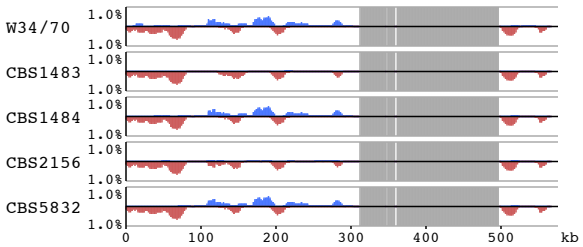
chr III



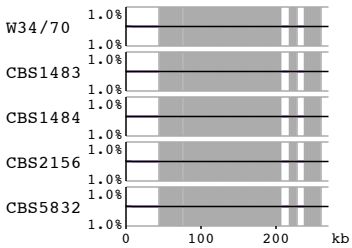
chr IV



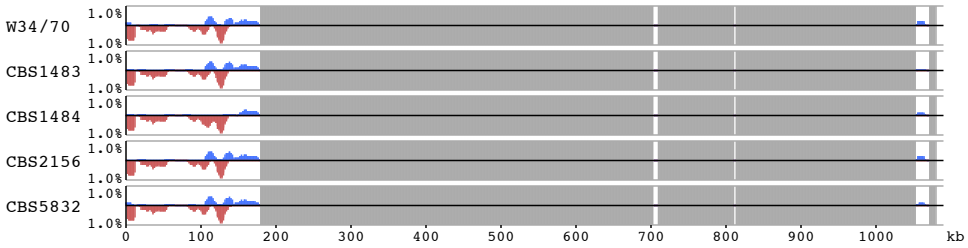
chr V



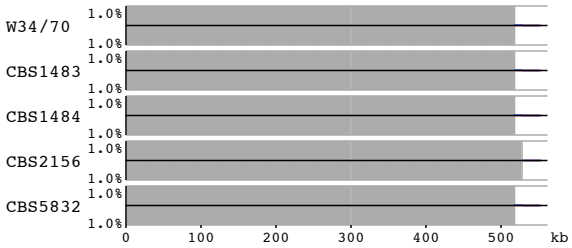
chr VI



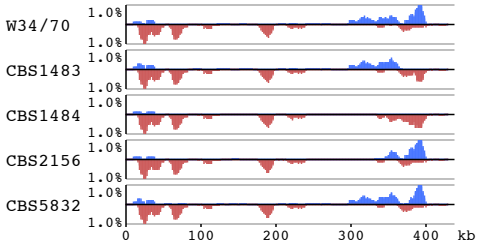
chr VII



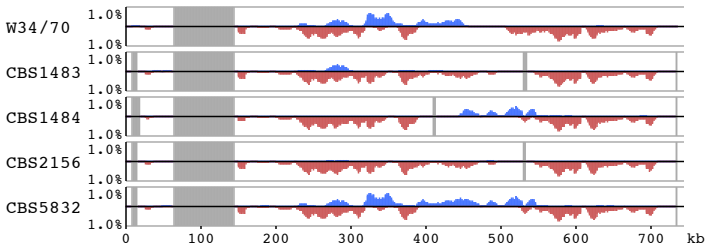
chr VIII



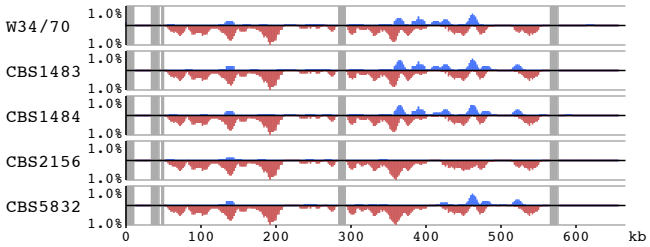
chr IX



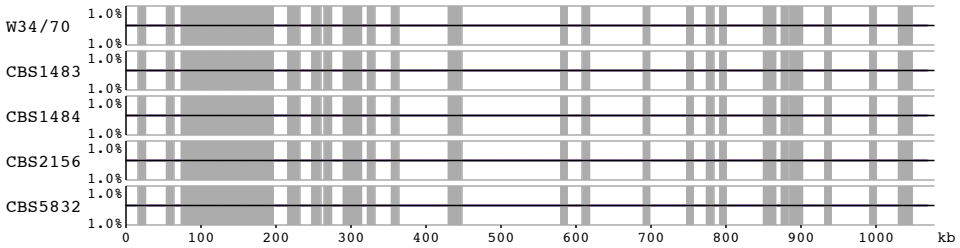
chr X



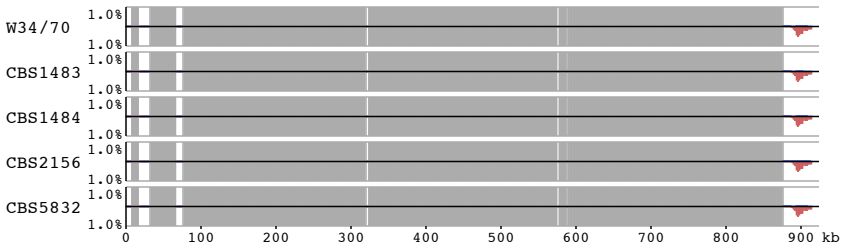
chr XI



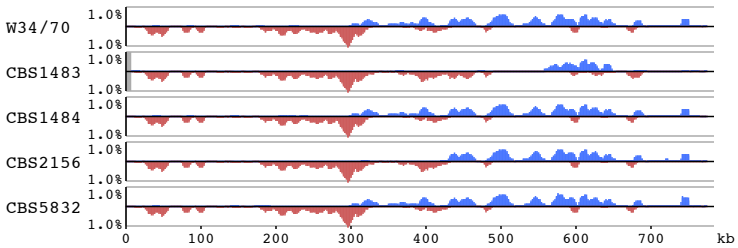
chr XII



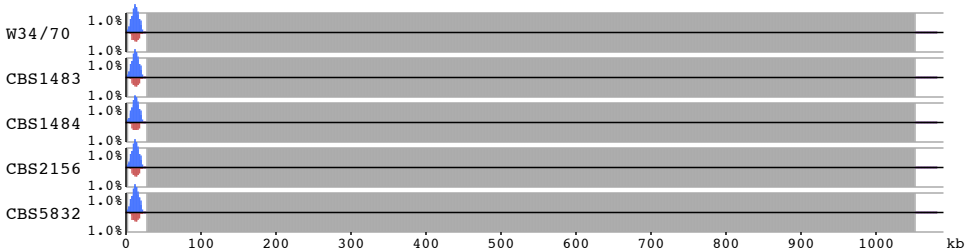
chr XIII



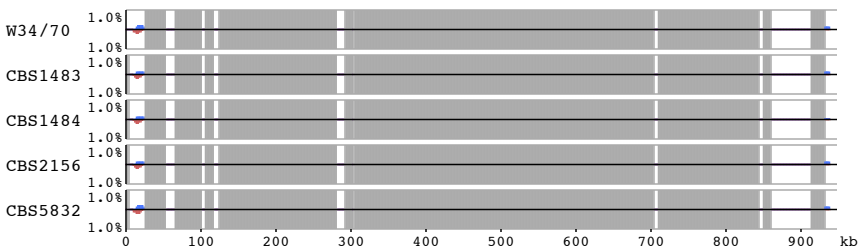
chr XIV



chr XV

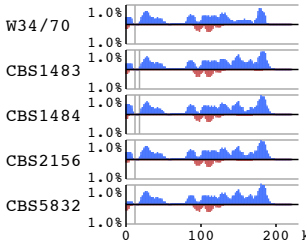


chr XVI

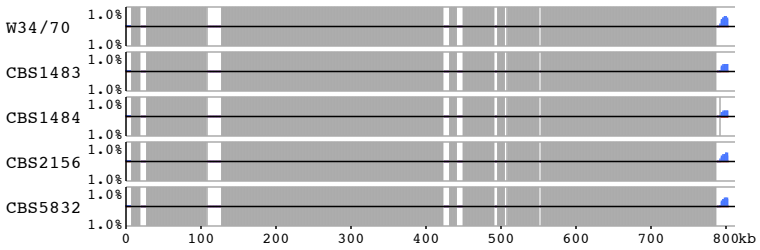


CBS1174

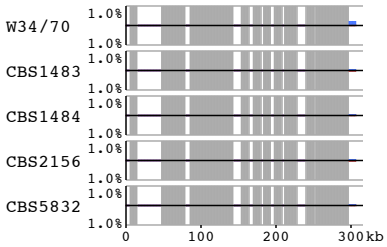
chr I



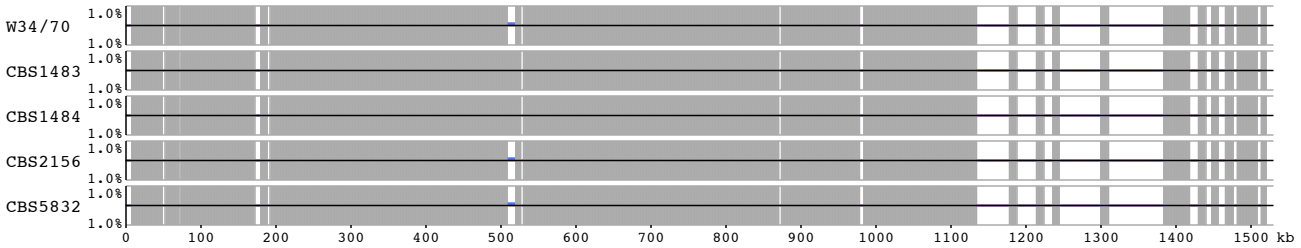
chr II



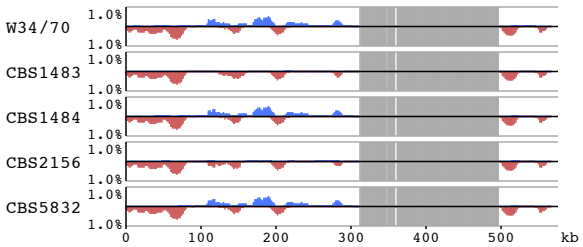
chr III



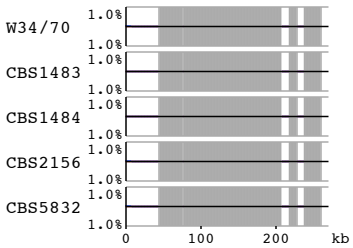
chr IV



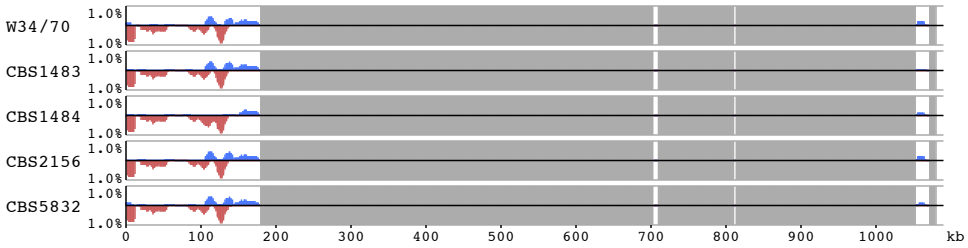
chr V



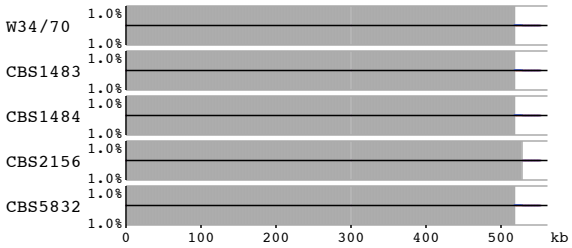
chr VI



chr VII

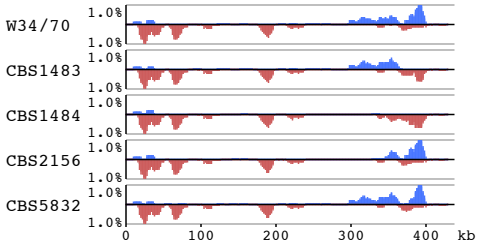


chr VIII

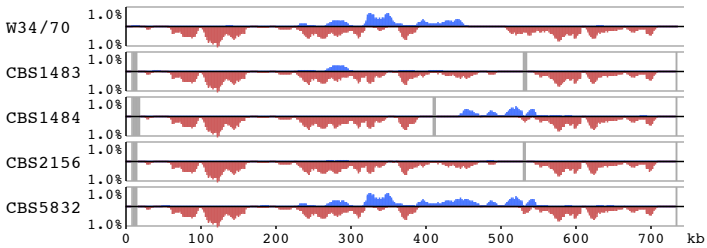


CBS1174

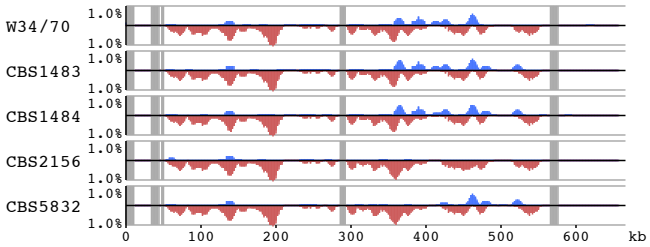
chr IX



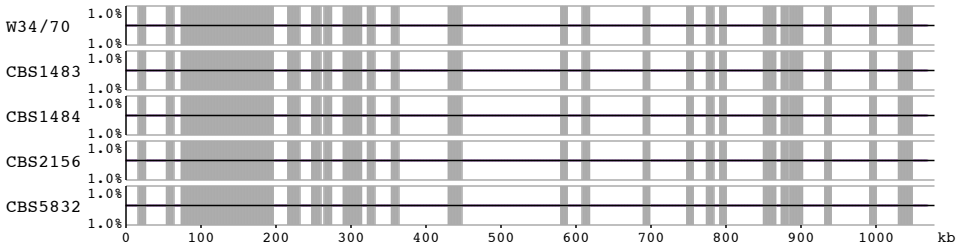
chr X



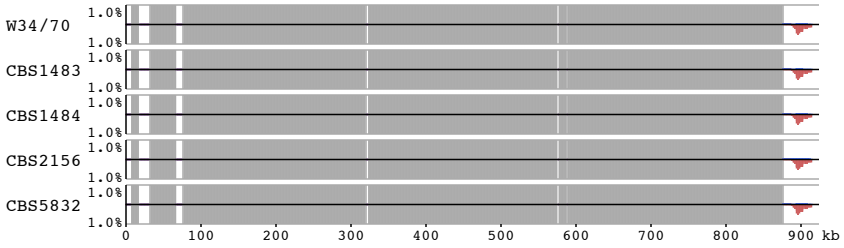
chr XI



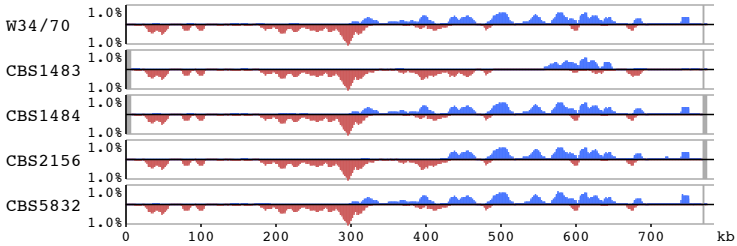
chr XII



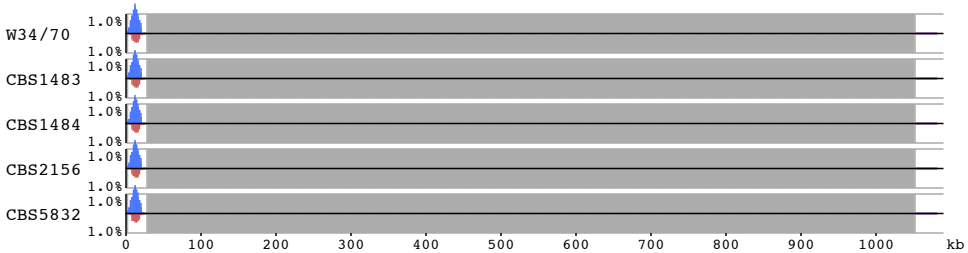
chr XIII



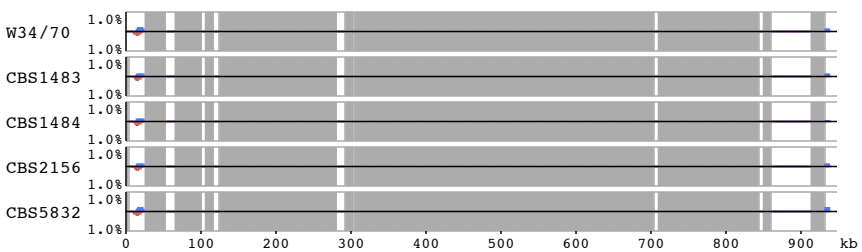
chr XIV



chr XV

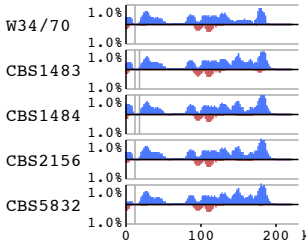


chr XVI

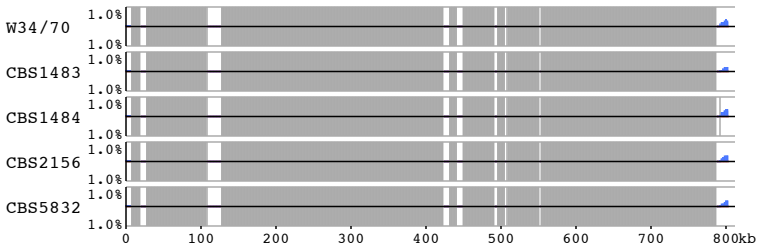


CBS2440

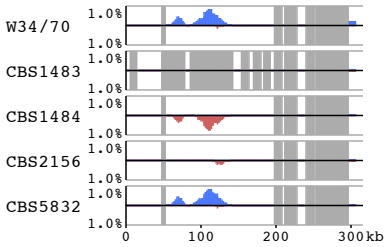
chr I



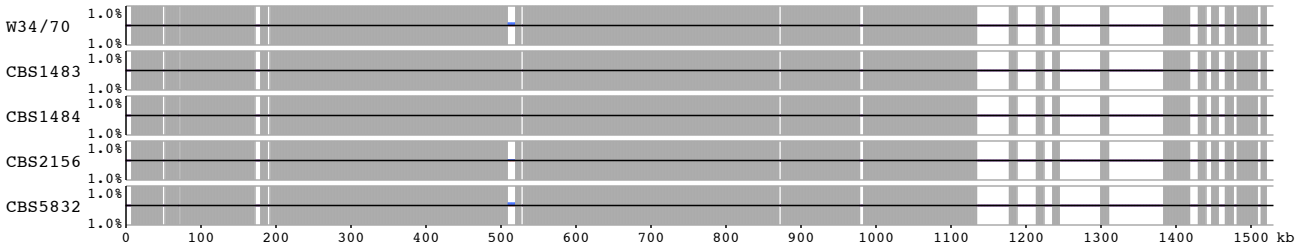
chr II



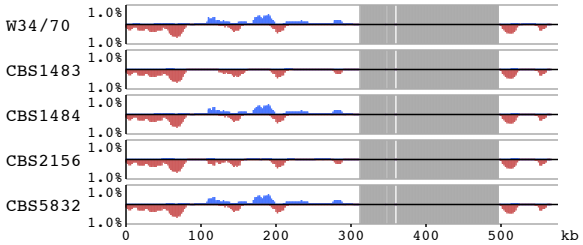
chr III



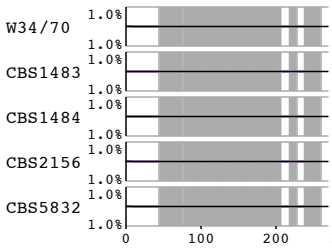
chr IV



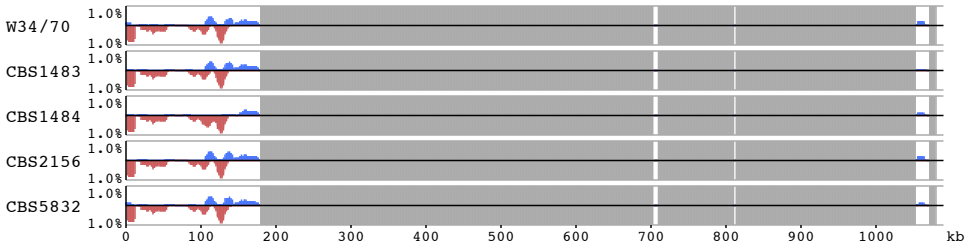
chr V



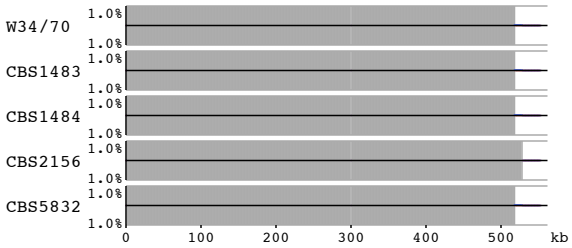
chr VI



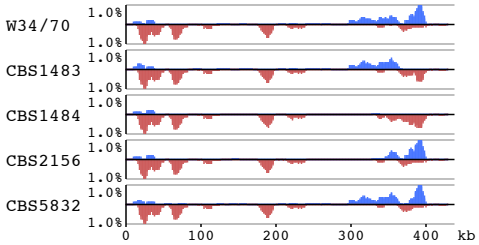
chr VII



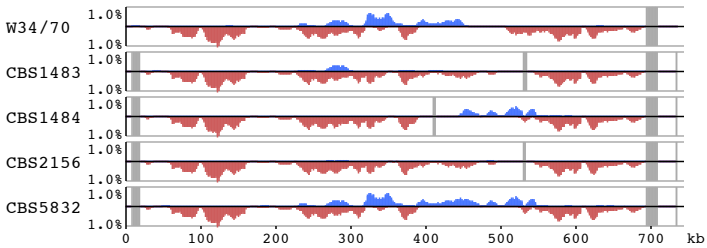
chr VIII



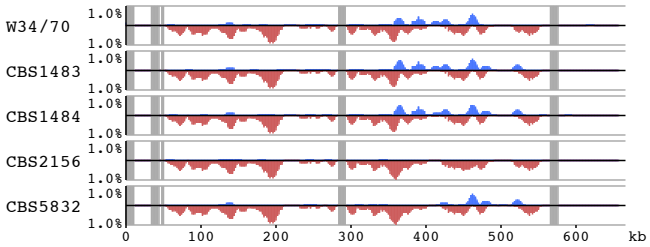
chr IX



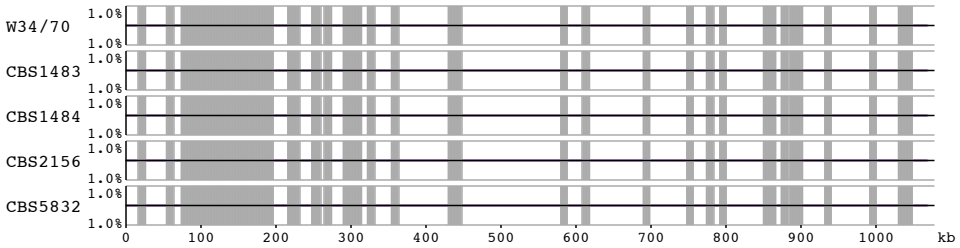
chr X



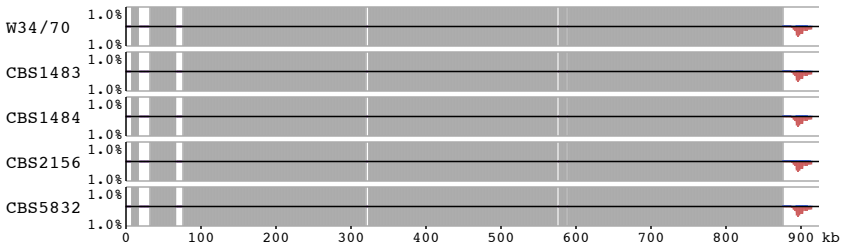
chr XI



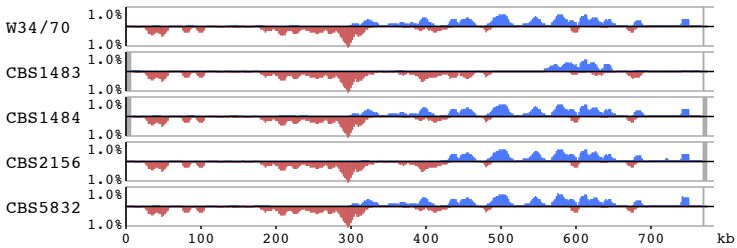
chr XII



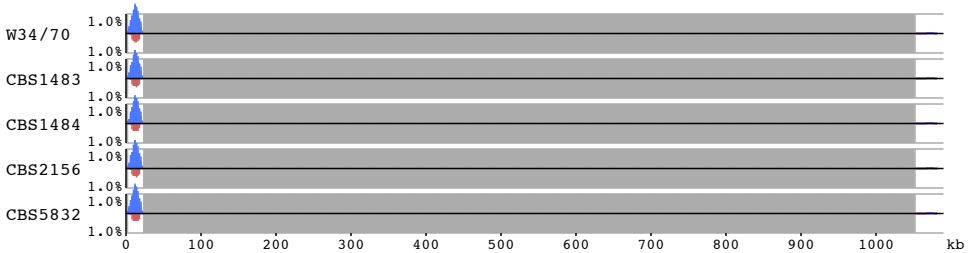
chr XIII



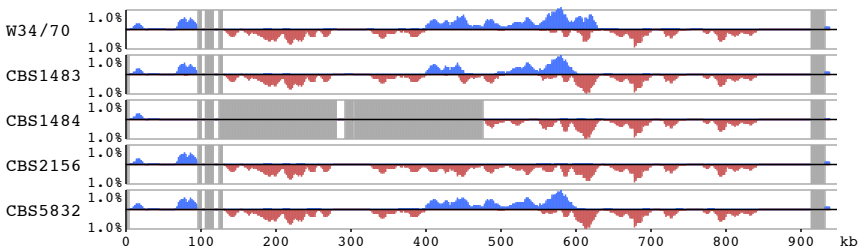
chr XIV



chr XV



chr XVI



Supplementary Figure 8. Distribution of hetero and homo SNVs between Group 1 and 2 strains.

Blue (upper) and red (lower) lines indicate the distribution of hetero and homo SNVs, respectively, between Group 1 and 2 strains. Grey regions indicate chromosome deletion in either Group 1 or 2. Hetero SNVs, homo SNV regions, and regions with no SNVs co-segregated as continuous blocks, possibly as a result of LOH in these regions.

A)

	strains	positions in S288C chromosome X 435,102	435,236	435,254	CBS strain name
Group 1	GSY509	A	A	T	CBS2440
	GSY133	A	A	T	
	GSY501	A	A	T	CBS1174
	GSY131	A	A	T	CBS1538
	GSY137	A	A	T	
	GSY129	A	A	T	CBS1513
	GSY134	A	A	T	CBS1503
Group 2	GSY132	T	T	A	
	GSY138	T	T	A	
	GSY139	T	T	A	
	GSY135	T	T	A	
	GSY136	T	T	A	
	GSY516	T	T	A	
	GSY515	T	T	A	CBS5832
	GSY503	T	T	A	CBS1483
	GSY504	A	A	T	CBS1484
	GSY508	T	T	A	CBS2156
	S288C	T	A	A	

B)

	strains	positions in S288C chromosome X 435,102	435,236	435,254
Group 1	CBS1503	A	A	T
	CBS1513	A	A	T
	CBS1538	A	A	T
	CBS1174	A	A	T
	CBS2440	A	A	T
Group 2	W34/70	T,A	A,T	A,T
	CBS1483	T	T	A
	CBS1484	T	T	A
	CBS2156	A	A	T
	CBS5832	T,A	A,T	A,T
	S288C	T	A	A

Supplementary Figure 9. Re-analysis of previously reported SNV sites.
A) Data are from previous report; B) NGS data from this study. In the *PRE3* region, W34/70 shows hetero SNVs consistent with two homo SNVs. CBS5832 shows the same hetero SNVs, whereas in previous report, CBS5832 showed only homo SNVs. The minor allele frequency at these sites was about 1/3, implying that one of three homologous chromosomes may have a different allele; these minor alleles may have been overlooked in the previous analysis.

```
CBS1503         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS1513         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS1538         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS1174         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS2440         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
W34/70          AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS1483         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS1484         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS2156         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
CBS5832         AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT
S288C   434950  AACACCGATCATTTTTGATGATTTGGCGTTTTTGTTATATTGATAGATTTCTTTTGAATT 435009
                ************************************************************

CBS1503         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS1513         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS1538         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS1174         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS2440         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
W34/70          TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS1483         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS1484         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS2156         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
CBS5832         TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG
S288C   435010  TTGTCATTTTCACTTTTCCACTCGCAACGGAATCCGGTGGCAAAAAAGGGAAAAGCATTG 435069
                ************************************************************

CBS1503         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
CBS1513         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
CBS1538         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
CBS1174         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
CBS2440         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
W34/70          AAATGCAATCTTTAACAGTATTTTAAACAAGTWGCGAGACGGTGTACAATTACGATAAGA
CBS1483         AAATGCAATCTTTAACAGTATTTTAAACAAGTTGCGAGACGGTGTACAATTACGATAAGA
CBS1484         AAATGCAATCTTTAACAGTATTTTAAACAAGTAGCGAGACGGTGTACAATTACGATAAGA
CBS2156         AAATGCAATCTTTAACAGTATTTTAAACAAGTTGCGAGACGGTGTACAATTACGATAAGA
CBS5832         AAATGCAATCTTTAACAGTATTTTAAACAAGTWGCGAGACGGTGTACAATTACGATAAGA
S288C   435070  AAATGCAATCTTTAACAGTATTTTAAACAAGTTGCGAGACGGTGTACAATTACGATAAGA 435129
                ******************************* *************************
```

```
CBS1503         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS1513         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS1538         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS1174         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS2440         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
W34/70          ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS1483         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS1484         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS2156         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
CBS5832         ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT
S288C   435130  ATTGCTACTTCAAAGTACACACAGAAAGTTAACATGAATGGAATTCAAGTGGACATCAAT 435189
                ************************************************************
```

```
CBS1503         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
CBS1513         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
CBS1538         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
CBS1174         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
CBS2440         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
W34/70          CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATWTAAGAATTTTTCC
CBS1483         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATTTAAGAATTTTTCC
CBS1484         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC
CBS2156         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATTTAAGAATTTTTCC
CBS5832         CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATWTAAGAATTTTTCC
S288C   435190  CGTTTGAAAAAGGGCGAAGTCAGTTTAGGTACCTCAATGTATGTATATAAGAATTTTTCC 435249
                ******************************************** *************
```

```
CBS1503         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS1513         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS1538         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS1174         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS2440         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
W34/70          TCCCWCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS1483         TCCCACTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS1484         TCCCTCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS2156         TCCCACTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
CBS5832         TCCCWCTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA
S288C   435250  TCCCACTTTATTGTTTCTAAAAGTTCAATGAAGTAAAGTCTCAATTGGCCTTATTACTAA 435309
                **** **************************************************
```

```
CBS1503         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS1513         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS1538         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS1174         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS2440         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
W34/70          CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS1483         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS1484         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS2156         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
CBS5832         CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG
S288C   435310 CTAATAGGTATCTTATAATCACCTAATAAAATAGTATGGCCGTGACATTTAAGGATGGTG 435369
               ************************************************************

CBS1503         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS1513         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS1538         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS1174         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS2440         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
W34/70          TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS1483         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS1484         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS2156         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
CBS5832         TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG
S288C   435370 TGATACTAGGTGCTGATTCACGTACCACCACTGGTGCGTACATAGCTAACCGTGTGACAG 435429
               ************************************************************

CBS1503         ATAAATTAACGAGAGTA
CBS1513         ATAAATTAACGAGAGTA
CBS1538         ATAAATTAACGAGAGTA
CBS1174         ATAAATTAACGAGAGTA
CBS2440         ATAAATTAACGAGAGTA
W34/70          ATAAATTAACGAGAGTA
CBS1483         ATAAATTAACGAGAGTA
CBS1484         ATAAATTAACGAGAGTA
CBS2156         ATAAATTAACGAGAGTA
CBS5832         ATAAATTAACGAGAGTA
S288C  435430  ATAAATTAACGAGAGTA 435446
               *****************
```

Supplementary Figure 10. Re-analysis of previously reported SNV sites (PRE3 locus).

Full length multiple alignment of *PRE3* PCR-region from this study are shown. SNV sites are indicated as red. W34/70 shows hetero SNVs consistent with two homo SNVs. CBS5832 shows the same hetero SNVs, whereas in previous report, CBS5832 showed only homo SNVs. The minor allele frequency at these sites was about 1/3, implying that one of three homologous chromosomes may have a different allele; these minor alleles may have been overlooked in the previous analysis.