Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/50/eabd9230/DC1

Supplementary Materials for

Rapid and ongoing evolution of repetitive sequence structures in human centromeres

Yuta Suzuki*, Eugene W. Myers, Shinichi Morishita*

*Corresponding author. Email: yuta_suzuki@edu.k.u-tokyo.ac.jp (Y.S.); moris@edu.k.u-tokyo.ac.jp (S.M.)

Published 11 December 2020, *Sci. Adv.* **6**, eabd9230 (2020)
DOI: 10.1126/sciadv.abd9230

The PDF file includes:

Materials and Methods
Figs. S1 to S20
Tables S1 to S3
Legends for data files S1 to S4

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/6/50/eabd9230/DC1)

Data files S1 to S4

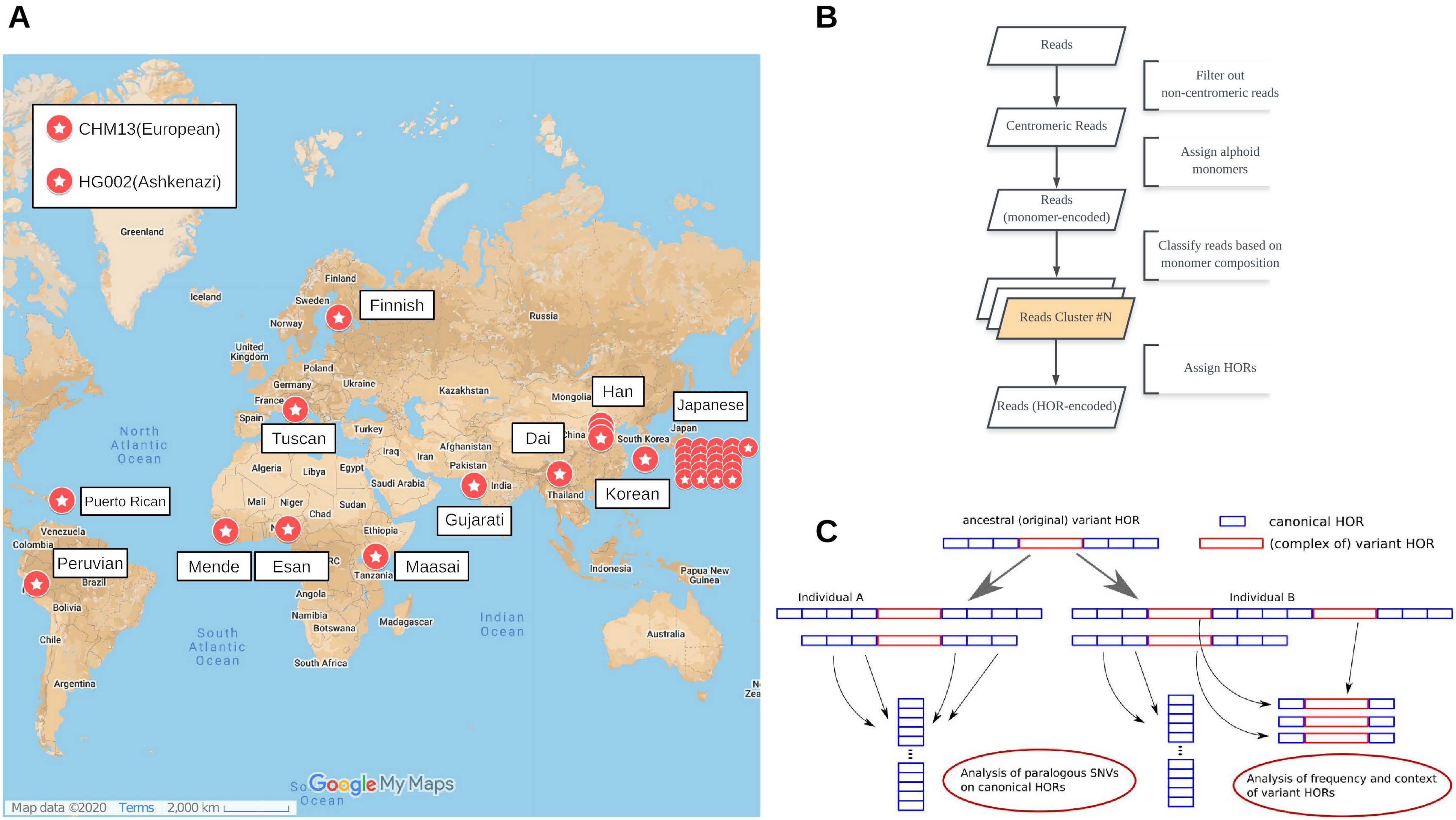
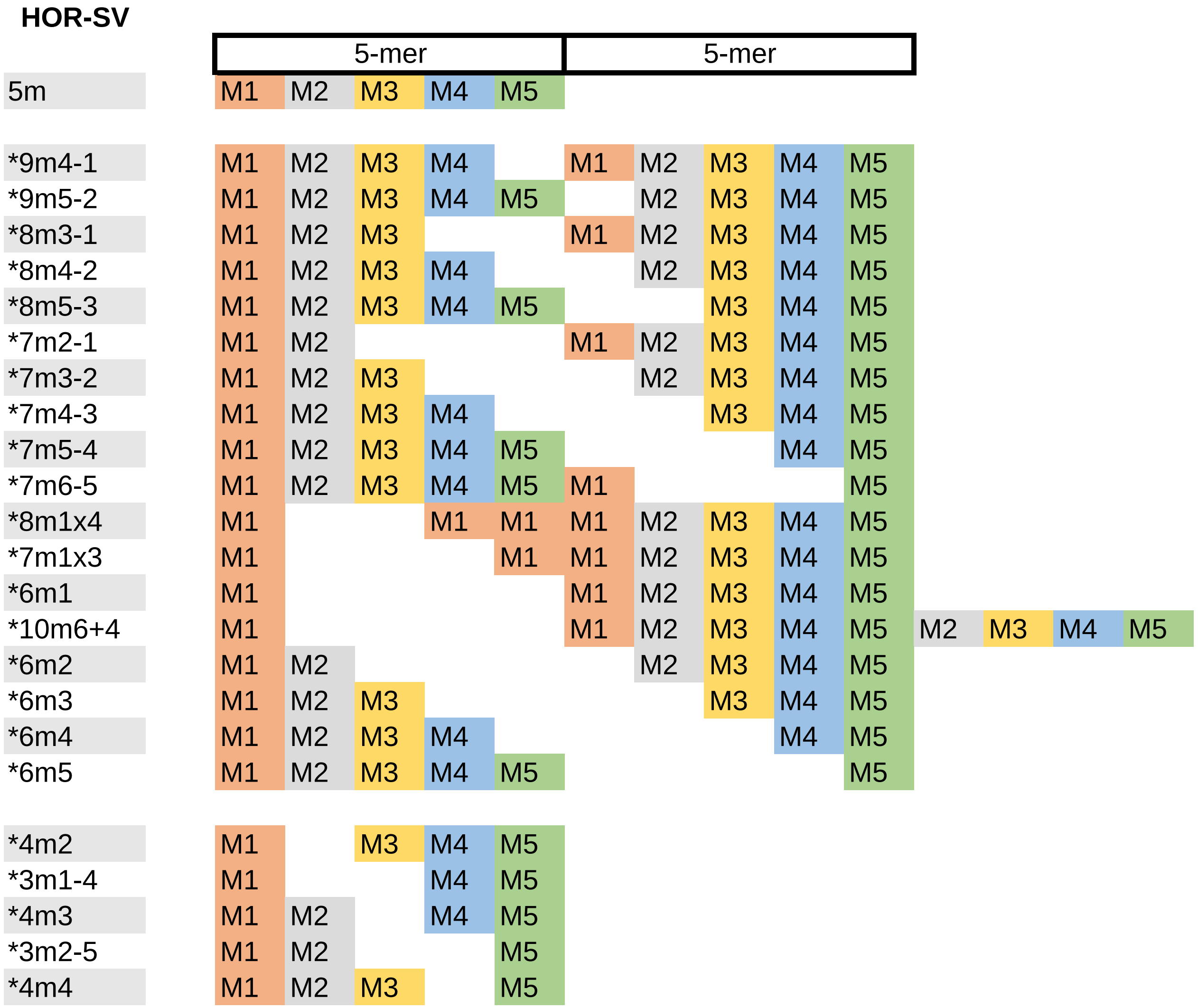
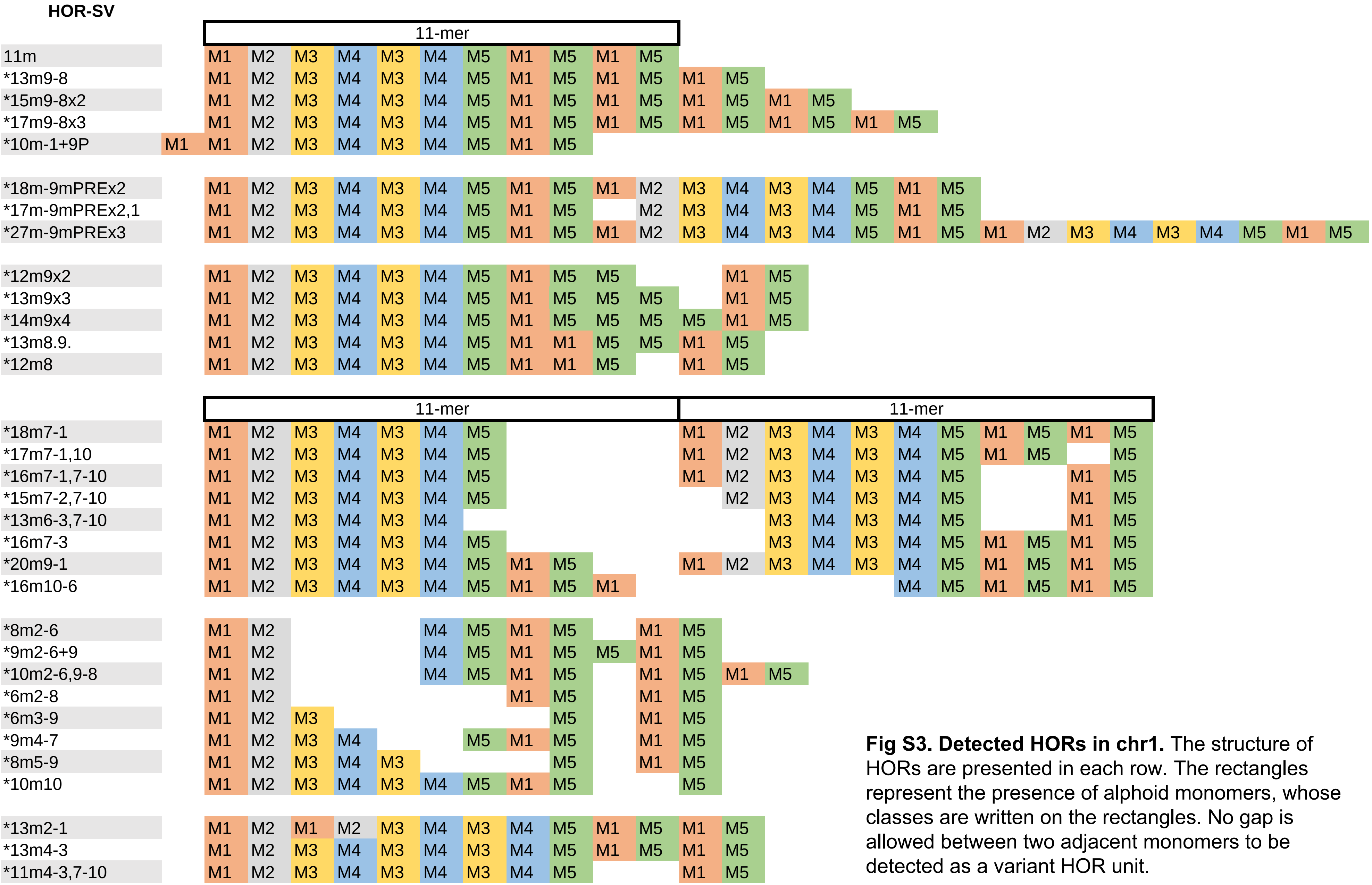


Fig S1. Summary of study design.
(A) Geographic locations of samples.
(B) Stepwise encoding strategy, specifically for the HOR analysis via long reads.
(C) In exploring tandem repeats, all homologous (orthologs/paralogs, if it applies at all) elements must be considered equally important.



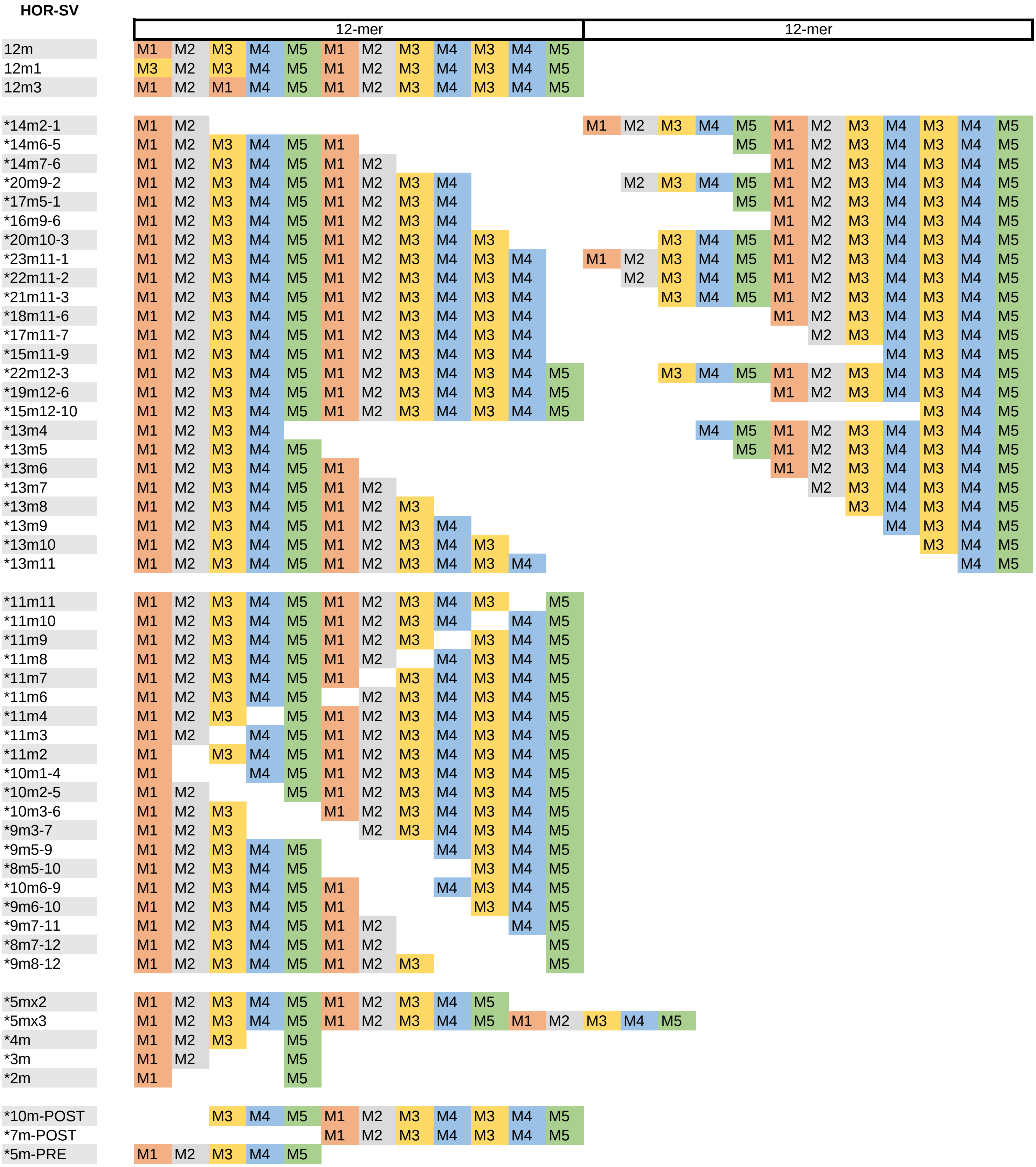
HOR-SV	5-mer					5-mer								
5m	M1	M2	M3	M4	M5									
*9m4-1	M1	M2	M3	M4		M1	M2	M3	M4	M5				
*9m5-2	M1	M2	M3	M4	M5		M2	M3	M4	M5				
*8m3-1	M1	M2	M3			M1	M2	M3	M4	M5				
*8m4-2	M1	M2	M3	M4			M2	M3	M4	M5				
*8m5-3	M1	M2	M3	M4	M5			M3	M4	M5				
*7m2-1	M1	M2				M1	M2	M3	M4	M5				
*7m3-2	M1	M2	M3				M2	M3	M4	M5				
*7m4-3	M1	M2	M3	M4				M3	M4	M5				
*7m5-4	M1	M2	M3	M4	M5				M4	M5				
*7m6-5	M1	M2	M3	M4	M5	M1				M5				
*8m1x4	M1			M1	M1	M1	M2	M3	M4	M5				
*7m1x3	M1				M1	M1	M2	M3	M4	M5				
*6m1	M1					M1	M2	M3	M4	M5				
*10m6+4	M1					M1	M2	M3	M4	M5	M2	M3	M4	M5
*6m2	M1	M2					M2	M3	M4	M5				
*6m3	M1	M2	M3					M3	M4	M5				
*6m4	M1	M2	M3	M4					M4	M5				
*6m5	M1	M2	M3	M4	M5					M5				
*4m2	M1		M3	M4	M5									
*3m1-4	M1			M4	M5									
*4m3	M1	M2		M4	M5									
*3m2-5	M1	M2			M5									
*4m4	M1	M2	M3		M5									

Fig S2. Detected HORs in chr11. The structure of HORs are presented in each row. The rectangles represent the presence of alphoid monomers, whose classes are written on the rectangles. For example, variant HOR named "4m2" is characterized by deletion of the second monomer, resulting in the 4-mer HOR variant. No gap is allowed between two adjacent monomers to be detected as a variant HOR unit.



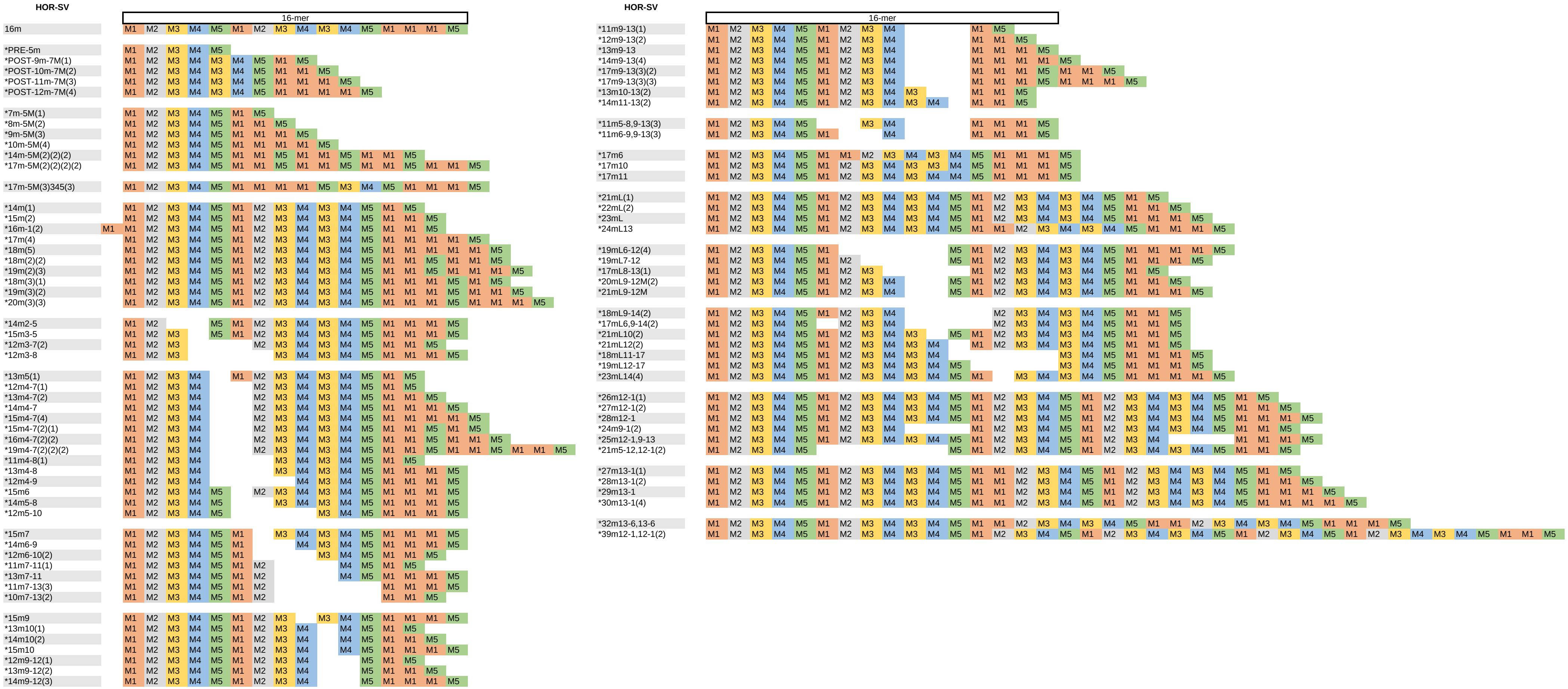
HOR-SV																												
		11-mer																										
11m		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5																
*13m9-8		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5	M1	M5														
*15m9-8x2		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5	M1	M5	M1	M5												
*17m9-8x3		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5	M1	M5	M1	M5	M1	M5										
*10m-1+9P	M1	M1	M2	M3	M4	M3	M4	M5	M1	M5																		
*18m-9mPREx2		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5									
*17m-9mPREx2,1		M1	M2	M3	M4	M3	M4	M5	M1	M5		M2	M3	M4	M3	M4	M5	M1	M5									
*27m-9mPREx3		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5
*12m9x2		M1	M2	M3	M4	M3	M4	M5	M1	M5	M5			M1	M5													
*13m9x3		M1	M2	M3	M4	M3	M4	M5	M1	M5	M5	M5		M1	M5													
*14m9x4		M1	M2	M3	M4	M3	M4	M5	M1	M5	M5	M5	M5	M1	M5													
*13m8.9.		M1	M2	M3	M4	M3	M4	M5	M1	M1	M5	M5	M1	M5														
*12m8		M1	M2	M3	M4	M3	M4	M5	M1	M1	M5		M1	M5														
		11-mer											11-mer															
*18m7-1		M1	M2	M3	M4	M3	M4	M5					M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5					
*17m7-1,10		M1	M2	M3	M4	M3	M4	M5					M1	M2	M3	M4	M3	M4	M5	M1	M5		M5					
*16m7-1,7-10		M1	M2	M3	M4	M3	M4	M5					M1	M2	M3	M4	M3	M4	M5			M1	M5					
*15m7-2,7-10		M1	M2	M3	M4	M3	M4	M5						M2	M3	M4	M3	M4	M5			M1	M5					
*13m6-3,7-10		M1	M2	M3	M4	M3	M4								M3	M4	M3	M4	M5			M1	M5					
*16m7-3		M1	M2	M3	M4	M3	M4	M5							M3	M4	M3	M4	M5	M1	M5	M1	M5					
*20m9-1		M1	M2	M3	M4	M3	M4	M5	M1	M5			M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5					
*16m10-6		M1	M2	M3	M4	M3	M4	M5	M1	M5	M1							M4	M5	M1	M5	M1	M5					
*8m2-6		M1	M2				M4	M5	M1	M5		M1	M5															
*9m2-6+9		M1	M2				M4	M5	M1	M5	M5	M1	M5															
*10m2-6,9-8		M1	M2				M4	M5	M1	M5		M1	M5	M1	M5													
*6m2-8		M1	M2						M1	M5		M1	M5															
*6m3-9		M1	M2	M3						M5		M1	M5															
*9m4-7		M1	M2	M3	M4			M5	M1	M5		M1	M5															
*8m5-9		M1	M2	M3	M4	M3				M5		M1	M5															
*10m10		M1	M2	M3	M4	M3	M4	M5	M1	M5			M5															
*13m2-1		M1	M2	M1	M2	M3	M4	M3	M4	M5	M1	M5	M1	M5														
*13m4-3		M1	M2	M3	M4	M3	M4	M3	M4	M5	M1	M5	M1	M5														
*11m4-3,7-10		M1	M2	M3	M4	M3	M4	M3	M4	M5			M1	M5														

Fig S3. Detected HORs in chr1. The structure of HORs are presented in each row. The rectangles represent the presence of alphoid monomers, whose classes are written on the rectangles. No gap is allowed between two adjacent monomers to be detected as a variant HOR unit.



HOR-SV	12-mer												12-mer											
12m	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5												
12m1	M3	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5												
12m3	M1	M2	M1	M4	M5	M1	M2	M3	M4	M3	M4	M5												
*14m2-1	M1	M2											M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*14m6-5	M1	M2	M3	M4	M5	M1											M5	M1	M2	M3	M4	M3	M4	M5
*14m7-6	M1	M2	M3	M4	M5	M1	M2											M1	M2	M3	M4	M3	M4	M5
*20m9-2	M1	M2	M3	M4	M5	M1	M2	M3	M4					M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*17m5-1	M1	M2	M3	M4	M5	M1	M2	M3	M4								M5	M1	M2	M3	M4	M3	M4	M5
*16m9-6	M1	M2	M3	M4	M5	M1	M2	M3	M4									M1	M2	M3	M4	M3	M4	M5
*20m10-3	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3					M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*23m11-1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*22m11-2	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4			M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*21m11-3	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4				M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*18m11-6	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4							M1	M2	M3	M4	M3	M4	M5
*17m11-7	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4								M2	M3	M4	M3	M4	M5
*15m11-9	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4										M4	M3	M4	M5
*22m12-3	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5			M3	M4	M5	M1	M2	M3	M4	M3	M4	M5
*19m12-6	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5						M1	M2	M3	M4	M3	M4	M5
*15m12-10	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5										M3	M4	M5
*13m4	M1	M2	M3	M4												M4	M5	M1	M2	M3	M4	M3	M4	M5
*13m5	M1	M2	M3	M4	M5												M5	M1	M2	M3	M4	M3	M4	M5
*13m6	M1	M2	M3	M4	M5	M1												M1	M2	M3	M4	M3	M4	M5
*13m7	M1	M2	M3	M4	M5	M1	M2												M2	M3	M4	M3	M4	M5
*13m8	M1	M2	M3	M4	M5	M1	M2	M3												M3	M4	M3	M4	M5
*13m9	M1	M2	M3	M4	M5	M1	M2	M3	M4												M4	M3	M4	M5
*13m10	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3												M3	M4	M5
*13m11	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4												M4	M5
*11m11	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3		M5												
*11m10	M1	M2	M3	M4	M5	M1	M2	M3	M4		M4	M5												
*11m9	M1	M2	M3	M4	M5	M1	M2	M3		M3	M4	M5												
*11m8	M1	M2	M3	M4	M5	M1	M2		M4	M3	M4	M5												
*11m7	M1	M2	M3	M4	M5	M1		M3	M4	M3	M4	M5												
*11m6	M1	M2	M3	M4	M5		M2	M3	M4	M3	M4	M5												
*11m4	M1	M2	M3		M5	M1	M2	M3	M4	M3	M4	M5												
*11m3	M1	M2		M4	M5	M1	M2	M3	M4	M3	M4	M5												
*11m2	M1		M3	M4	M5	M1	M2	M3	M4	M3	M4	M5												
*10m1-4	M1			M4	M5	M1	M2	M3	M4	M3	M4	M5												
*10m2-5	M1	M2			M5	M1	M2	M3	M4	M3	M4	M5												
*10m3-6	M1	M2	M3			M1	M2	M3	M4	M3	M4	M5												
*9m3-7	M1	M2	M3				M2	M3	M4	M3	M4	M5												
*9m5-9	M1	M2	M3	M4	M5				M4	M3	M4	M5												
*8m5-10	M1	M2	M3	M4	M5					M3	M4	M5												
*10m6-9	M1	M2	M3	M4	M5	M1			M4	M3	M4	M5												
*9m6-10	M1	M2	M3	M4	M5	M1				M3	M4	M5												
*9m7-11	M1	M2	M3	M4	M5	M1	M2				M4	M5												
*8m7-12	M1	M2	M3	M4	M5	M1	M2					M5												
*9m8-12	M1	M2	M3	M4	M5	M1	M2	M3				M5												
*5mx2	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5														
*5mx3	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5									
*4m	M1	M2	M3		M5																			
*3m	M1	M2			M5																			
*2m	M1				M5																			
*10m-POST			M3	M4	M5	M1	M2	M3	M4	M3	M4	M5												
*7m-POST						M1	M2	M3	M4	M3	M4	M5												
*5m-PRE	M1	M2	M3	M4	M5																			

Fig S4. Detected HORs in chrX. The structure of HORs are presented in each row. The rectangles represent the presence of alphoid monomers, whose classes are written on the rectangles. No gap is allowed between two adjacent monomers to be detected as a variant HOR unit.



HOR-SV		16-mer																				
16m		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5					
*PRE-5m		M1	M2	M3	M4	M5																
*POST-9m-7M(1)		M1	M2	M3	M4	M3	M4	M5	M1	M5												
*POST-10m-7M(2)		M1	M2	M3	M4	M3	M4	M5	M1	M1	M5											
*POST-11m-7M(3)		M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5										
*POST-12m-7M(4)		M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M5									
*7m-5M(1)		M1	M2	M3	M4	M5	M1	M5														
*8m-5M(2)		M1	M2	M3	M4	M5	M1	M1	M5													
*9m-5M(3)		M1	M2	M3	M4	M5	M1	M1	M1	M5												
*10m-5M(4)		M1	M2	M3	M4	M5	M1	M1	M1	M1	M5											
*14m-5M(2)(2)(2)		M1	M2	M3	M4	M5	M1	M1	M5	M1	M1	M5	M1	M1	M5							
*17m-5M(2)(2)(2)(2)		M1	M2	M3	M4	M5	M1	M1	M5	M1	M1	M5	M1	M1	M5	M1	M1	M5				
*17m-5M(3)345(3)		M1	M2	M3	M4	M5	M1	M1	M1	M1	M5	M3	M4	M5	M1	M1	M1	M5				
*14m(1)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5							
*15m(2)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5						
*16m-1(2)	M1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5						
*17m(4)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M5				
*18m(5)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M1	M5			
*18m(2)(2)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5	M1	M1	M5			
*19m(2)(3)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5	M1	M1	M1	M5		
*18m(3)(1)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5	M1	M5			
*19m(3)(2)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5	M1	M1	M5		
*20m(3)(3)		M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5	M1	M1	M1	M5	
*14m2-5		M1	M2			M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5					
*15m3-5		M1	M2	M3		M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5					
*12m3-7(2)		M1	M2	M3				M2	M3	M4	M3	M4	M5	M1	M1	M5						
*12m3-8		M1	M2	M3					M3	M4	M3	M4	M5	M1	M1	M1	M5					
*13m5(1)		M1	M2	M3	M4		M1	M2	M3	M4	M3	M4	M5	M1	M5							
*12m4-7(1)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M5							
*13m4-7(2)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M5						
*14m4-7		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M1	M5					
*15m4-7(4)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M5				
*15m4-7(2)(1)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M5	M1	M5				
*16m4-7(2)(2)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M5	M1	M1	M5			
*19m4-7(2)(2)(2)		M1	M2	M3	M4			M2	M3	M4	M3	M4	M5	M1	M1	M5	M1	M1	M5	M1	M1	M5
*11m4-8(1)		M1	M2	M3	M4				M3	M4	M3	M4	M5	M1	M5							
*13m4-8		M1	M2	M3	M4				M3	M4	M3	M4	M5	M1	M1	M1	M5					
*12m4-9		M1	M2	M3	M4					M4	M3	M4	M5	M1	M1	M1	M5					
*15m6		M1	M2	M3	M4	M5		M2	M3	M4	M3	M4	M5	M1	M1	M1	M5					
*14m5-8		M1	M2	M3	M4	M5			M3	M4	M3	M4	M5	M1	M1	M1	M5					
*12m5-10		M1	M2	M3	M4	M5					M3	M4	M5	M1	M1	M1	M5					
*15m7		M1	M2	M3	M4	M5	M1		M3	M4	M3	M4	M5	M1	M1	M1	M5					
*14m6-9		M1	M2	M3	M4	M5	M1			M4	M3	M4	M5	M1	M1	M1	M5					
*12m6-10(2)		M1	M2	M3	M4	M5	M1				M3	M4	M5	M1	M1	M5						
*11m7-11(1)		M1	M2	M3	M4	M5	M1	M2				M4	M5	M1	M5							
*13m7-11		M1	M2	M3	M4	M5	M1	M2				M4	M5	M1	M1	M1	M5					
*11m7-13(3)		M1	M2	M3	M4	M5	M1	M2						M1	M1	M1	M5					
*10m7-13(2)		M1	M2	M3	M4	M5	M1	M2						M1	M1	M5						
*15m9		M1	M2	M3	M4	M5	M1	M2	M3		M3	M4	M5	M1	M1	M1	M5					
*13m10(1)		M1	M2	M3	M4	M5	M1	M2	M3	M4		M4	M5	M1	M5							
*14m10(2)		M1	M2	M3	M4	M5	M1	M2	M3	M4		M4	M5	M1	M1	M5						
*15m10		M1	M2	M3	M4	M5	M1	M2	M3	M4		M4	M5	M1	M1	M1	M5					
*12m9-12(1)		M1	M2	M3	M4	M5	M1	M2	M3	M4			M5	M1	M5							
*13m9-12(2)		M1	M2	M3	M4	M5	M1	M2	M3	M4			M5	M1	M1	M5						
*14m9-12(3)		M1	M2	M3	M4	M5	M1	M2	M3	M4			M5	M1	M1	M1	M5					

HOR-SV	16-mer																																						
*11m9-13(1)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M5																									
*12m9-13(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M5																								
*13m9-13	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M1	M5																							
*14m9-13(4)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M1	M1	M5																						
*17m9-13(3)(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M1	M5	M1	M1	M5																				
*17m9-13(3)(3)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M1	M5	M1	M1	M1	M5																			
*13m10-13(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3			M1	M1	M5																								
*14m11-13(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4		M1	M1	M5																								
*11m5-8,9-13(3)	M1	M2	M3	M4	M5			M3	M4				M1	M1	M1	M5																							
*11m6-9,9-13(3)	M1	M2	M3	M4	M5	M1			M4				M1	M1	M1	M5																							
*17m6	M1	M2	M3	M4	M5	M1	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5																						
*17m10	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M3	M4	M5	M1	M1	M1	M5																						
*17m11	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M4	M5	M1	M1	M1	M5																						
*21mL(1)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5																		
*22mL(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*23mL	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5																
*24mL13	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5															
*19mL6-12(4)	M1	M2	M3	M4	M5	M1						M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M5															
*19mL7-12	M1	M2	M3	M4	M5	M1	M2					M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5																
*17mL8-13(1)	M1	M2	M3	M4	M5	M1	M2	M3					M1	M2	M3	M4	M3	M4	M5	M1	M5																		
*20mL9-12M(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4			M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*21mL9-12M	M1	M2	M3	M4	M5	M1	M2	M3	M4			M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5																
*18mL9-14(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4					M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*17mL6,9-14(2)	M1	M2	M3	M4	M5		M2	M3	M4					M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*21mL10(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3		M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*21mL12(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4		M1	M2	M3	M4	M3	M4	M5	M1	M1	M5																	
*18mL11-17	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4						M3	M4	M5	M1	M1	M1	M5																
*19mL12-17	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5					M3	M4	M5	M1	M1	M1	M5																
*23mL14(4)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1		M3	M4	M3	M4	M5	M1	M1	M1	M1	M5															
*26m12-1(1)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5													
*27m12-1(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5												
*28m12-1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5											
*24m9-1(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5												
*25m12-1,9-13	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4				M1	M1	M1	M5											
*21m5-12,12-1(2)	M1	M2	M3	M4	M5							M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5												
*27m13-1(1)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M5												
*28m13-1(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5											
*29m13-1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5										
*30m13-1(4)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M1	M5									
*32m13-6,13-6	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M3	M4	M5	M1	M1	M2	M3	M4	M3	M4	M5	M1	M1	M1	M5							
*39m12-1,12-1(2)	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M3	M4	M5	M1	M1	M5

Fig S5. Detected HORs in chr17. The structure of HORs are presented in each row. The rectangles represent the presence of alphoid monomers, whose classes are written on the rectangles. No gap is allowed between two adjacent monomers to be detected as a variant HOR unit.

Fig S6. SNV landscape over the 11-mer canonical HOR in chromosome 1. SNV with frequency of >3% are shown. x-axis is labelled with monomer index, but the actual coordinate represents position and base; for example, the alternative "G" at 20th base of 2nd monomer is plotted at x = 3 (G for G) + (20 * 4) + (2 * 800) = 1683. y-axis is the observed frequency in percentage. Four colors are used to distinguish four bases.

Fig S7. SNV landscape over the 5-mer canonical HOR in chromosome 11.
SNV with frequency of >3% are shown. x-axis is labelled with monomer index, but the actual coordinate represents position and base; for example, the alternative "G" at 20th base of 2nd monomer is plotted at x = 3 (G for G) + (20 * 4) + (2 * 800) = 1683. y-axis is the observed frequency in percentage. Four colors are used to distinguish four bases.

Fig S8. SNV landscape over the 12-mer canonical HOR in chromosome X. SNV with frequency of >3% are shown. x-axis is labelled with monomer index, but the actual coordinate represents position and base; for example, the alternative "G" at 20th base of 2nd monomer is plotted at x = 3 (G for G) + (20 * 4) + (2 * 800) = 1683. y-axis is the observed frequency in percentage. Four colors are used to distinguish four bases.

Fig S9. SNV landscape over the 16-mer canonical HOR in chromosome 17. SNV with frequency of >3% are shown. x-axis is labelled with monomer index, but the actual coordinate represents position and base; for example, the alternative "G" at 20th base of 2nd monomer is plotted at x = 3 (G for G) + (20 * 4) + (2 * 800) = 1683. y-axis is the observed frequency in percentage. Four colors are used to distinguish four bases.

Fig S10. Comparison of SNV frequencies on the canonical 5-mer HOR in chromosome 11. Each dot represents a single SNV (a position and an alternative base). The left panel (A) shows the SNVs with the frequencies of >3% in both samples in x-, and y-axis, and the right panel (B) shows those of the frequencies within 1-10%.

Fig S11. Comparison of SNV frequencies on the canonical 11-mer HOR in chromosome 1. Each dot represents a single SNV (a position and an alternative base). The left panel (A) shows the SNVs with the frequencies of >3% in both samples in x-, and y-axis, and the right panel (B) shows those of the frequencies within 1-10%.

Fig S12. Comparison of SNV frequencies on the canonical 12-mer HOR in chromosome X. Each dot represents a single SNV (a position and an alternative base). The left panel (A) shows the SNVs with the frequencies of >3% in both samples in x-, and y-axis, and the right panel (B) shows those of the frequencies within 1-10%.

Fig S13. Comparison of SNV frequencies on the canonical 16-mer HOR in chromosome 17. Each dot represents a single SNV (a position and an alternative base). The left panel (A) shows the SNVs with the frequencies of >3% in both samples in x-, and y-axis, and the right panel (B) shows those of the frequencies within 1-10%.

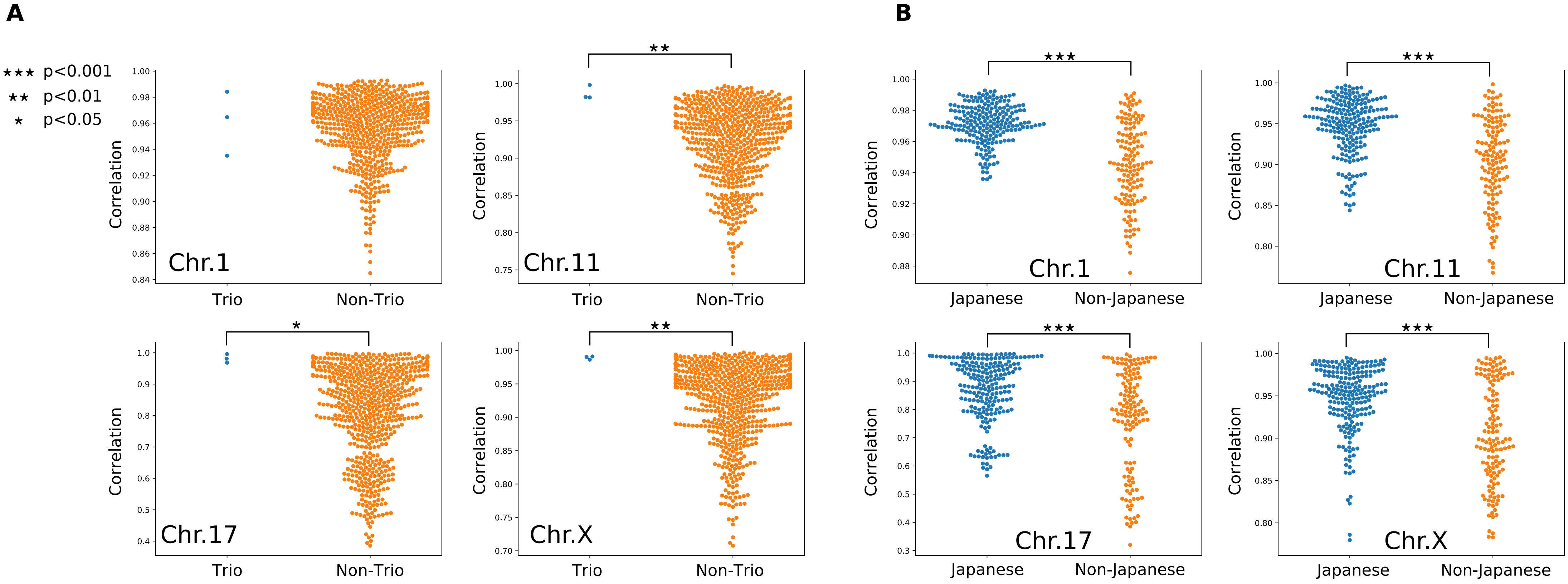


Fig S14. Pairwise (Pearson) correlation of SNV frequencies among samples. Statistically significant differences are marked with asterisks. (A) Comparison between the correlation coefficients among samples within a trio (blue; HG005-7, Han Chinese) and others (orange). (B) Coparison between the correlation among Japanese (blue), and those among others (orange).

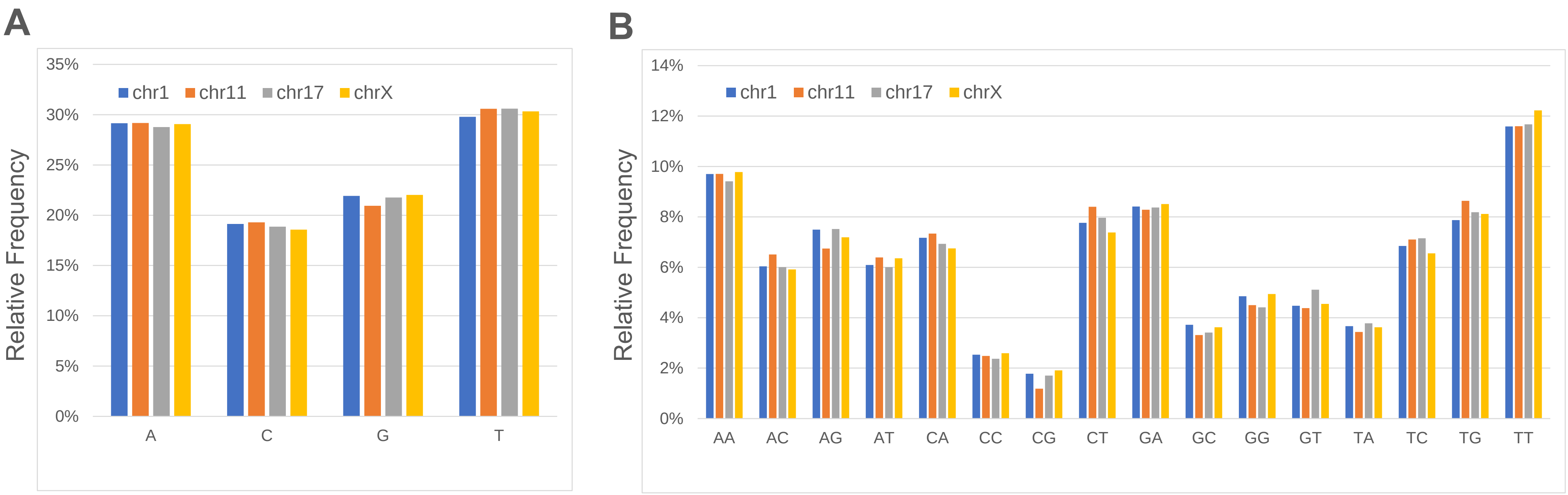


Fig S15. Sequence composition of canonical HORs of the four chromosomes.
(A) Base frequency and (B) 2-mer frequency are shown.

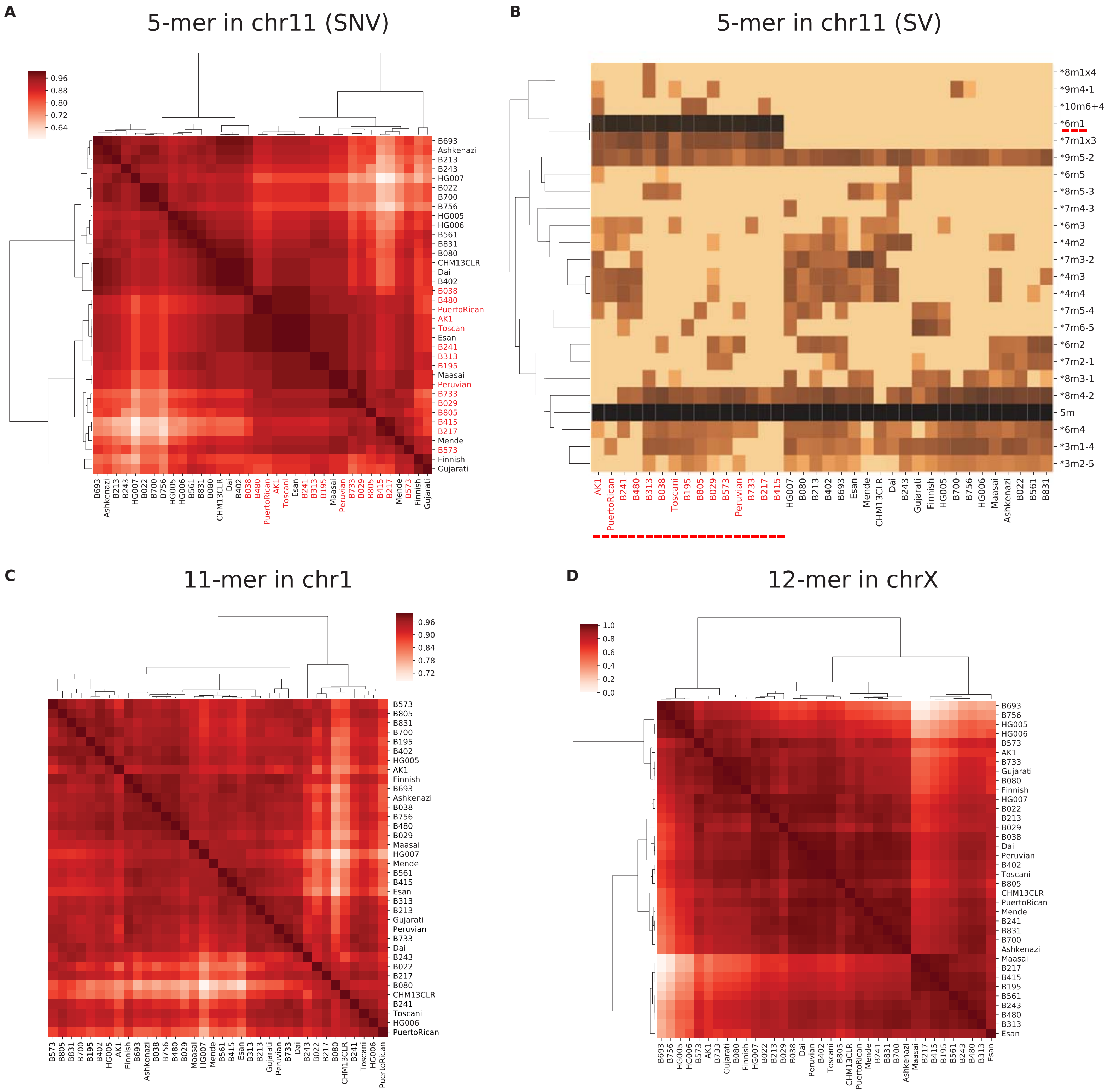
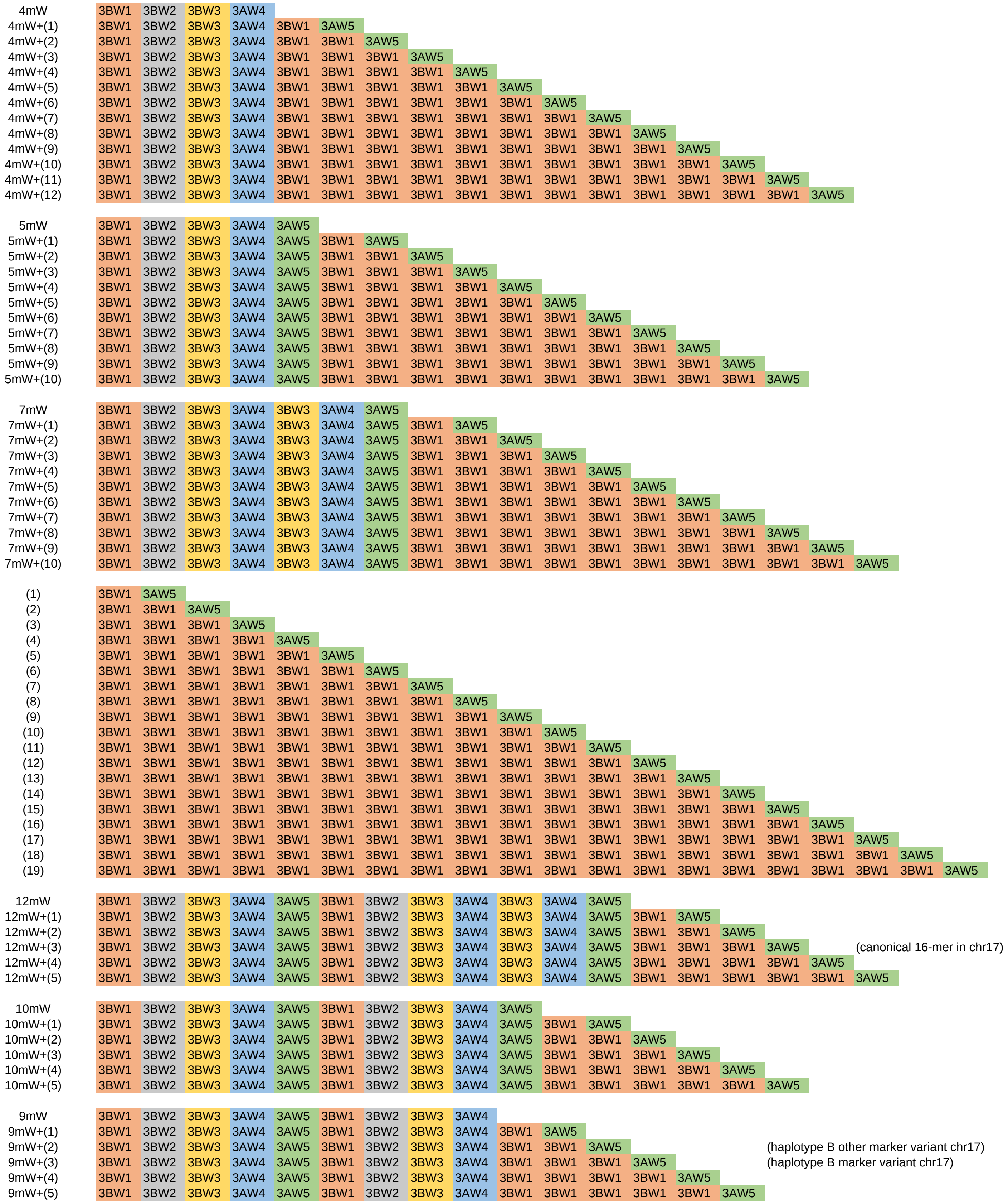


Fig S16. Analysis of SNVs detected on the canonical HORs. (A, C, D) Correlation of SNV frequencies among samples on the canonical HORs, (A) 5-mer in chromosome 11, (C) 11-mer in chromosome 1, and (D) 12-mer in chromosome X. In (A), sample labels are colored with red if the sample contains 6m1 variant HORs.

(B) Occurrence of variant HORs in each sample serves as a fingerprint of the distinct haplotypes. SVs were clustered by co-occurrence over the samples. Two marker variant HORs, 6m1 and 7m1x3, were observed only in a subset of samples defining a characteristic haplotype, which are labelled with red.



4mW	3BW1	3BW2	3BW3	3AW4													
4mW+(1)	3BW1	3BW2	3BW3	3AW4	3BW1	3AW5											
4mW+(2)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3AW5										
4mW+(3)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3AW5									
4mW+(4)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3AW5								
4mW+(5)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5							
4mW+(6)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5						
4mW+(7)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5					
4mW+(8)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5				
4mW+(9)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5			
4mW+(10)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5		
4mW+(11)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	
4mW+(12)	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5

5mW	3BW1	3BW2	3BW3	3AW4	3AW5											
5mW+(1)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3AW5									
5mW+(2)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3AW5								
5mW+(3)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3AW5							
5mW+(4)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3AW5						
5mW+(5)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5					
5mW+(6)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5				
5mW+(7)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5			
5mW+(8)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5		
5mW+(9)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	
5mW+(10)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5

7mW	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5											
7mW+(1)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3AW5									
7mW+(2)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3AW5								
7mW+(3)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3AW5							
7mW+(4)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3AW5						
7mW+(5)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5					
7mW+(6)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5				
7mW+(7)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5			
7mW+(8)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5		
7mW+(9)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	
7mW+(10)	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5

(1)	3BW1	3AW5																		
(2)	3BW1	3BW1	3AW5																	
(3)	3BW1	3BW1	3BW1	3AW5																
(4)	3BW1	3BW1	3BW1	3BW1	3AW5															
(5)	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5														
(6)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5													
(7)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5												
(8)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5											
(9)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5										
(10)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5									
(11)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5								
(12)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5							
(13)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5						
(14)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5					
(15)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5				
(16)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5			
(17)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5		
(18)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	
(19)	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5

12mW	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5							
12mW+(1)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3AW5					
12mW+(2)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3AW5				
12mW+(3)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3AW5			(canonical 16-mer in chr17)
12mW+(4)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3AW5		
12mW+(5)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	

10mW	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5						
10mW+(1)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3AW5				
10mW+(2)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3AW5			
10mW+(3)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3AW5		
10mW+(4)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3AW5	
10mW+(5)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5

9mW	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4							
9mW+(1)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW1	3AW5					
9mW+(2)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3AW5				(haplotype B other marker variant chr17)
9mW+(3)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3AW5			(haplotype B marker variant chr17)
9mW+(4)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3AW5		
9mW+(5)	3BW1	3BW2	3BW3	3AW4	3AW5	3BW1	3BW2	3BW3	3AW4	3BW1	3BW1	3BW1	3BW1	3BW1	3AW5	

Fig S17. HOR patterns defined for studying centromeric repeats in humans and chimpanzees. The structure of HORs are presented in each row. The rectangles represent the presence of alphoid monomers, whose classes are written on the rectangles. No gap is allowed between two adjacent monomers to be detected as a variant HOR unit.

AA	AB	BB

*	*	Clint	Esan	Maasai	Gujarati	Mende	Toscani	Ashkenazi	Finnish	Dai	HG005	PuertoRican	Peruvian	
1	nMons	3845573	19253041	16812971	17412128	19686565	21792096	14454006	26619252	15015579	21598986	35456165	26396834	
1	nReads	63041	201918	168447	172521	371798	280734	228847	324782	158328	393744	313677	388264	
4	4mW	25939	12359	12503	13637	20280	27632	8810	26087	10436	18066	34745	26131	
6	4mW+(1)	1641	31	54	28	28	41	31	37	22	54	84	29	
7	4mW+(2)	11792	175	187	41	1	1	0	0	0	1	59	1	
8	4mW+(3)	1867	848	1610	562	2	0	2	0	0	0	1833	0	
9	4mW+(4)	246	18	21	6	0	0	0	0	0	0	0	0	
10	4mW+(5)	214	0	0	0	0	0	0	0	0	0	0	0	
11	4mW+(6)	345	0	0	0	0	0	0	0	0	0	0	0	
12	4mW+(7)	53	0	0	0	0	0	0	0	0	0	0	0	
13	4mW+(8)	29	0	0	0	0	0	0	0	0	0	0	0	
14	4mW+(9)	25	0	0	0	0	0	0	0	0	0	0	0	
15	4mW+(10)	13	0	0	0	0	0	0	0	0	0	0	0	
16	4mW+(11)	0	0	0	0	0	0	0	0	0	0	0	0	
17	4mW+(12)	3	0	0	0	0	0	0	0	0	0	0	0	
5	5mW	4437	26953	30170	31087	47923	63167	17959	58842	27660	39699	89417	66527	
7	5mW+(1)	21	97	96	262	91	120	47	239	52	136	193	120	
8	5mW+(2)	17	226	26	44	61	83	34	116	108	82	126	86	
9	5mW+(3)	329	44	24	61	101	76	53	99	57	59	133	125	
10	5mW+(4)	288	0	40	0	0	0	0	0	0	0	0	33	
11	5mW+(5)	364	0	0	0	0	0	0	0	0	0	0	0	
12	5mW+(6)	34	0	0	0	0	0	0	0	0	0	0	0	
13	5mW+(7)	39	0	0	0	0	0	0	0	0	0	0	0	
14	5mW+(8)	15	0	0	0	0	0	0	0	0	0	0	0	
15	5mW+(9)	10	0	0	0	0	0	0	0	0	0	0	0	
16	5mW+(10)	0	0	0	0	0	0	0	0	0	0	0	0	
7	7mW	9936	19669	19681	19678	24628	22008	17751	26769	17277	19845	39494	26427	
9	7mW+(1)	1464	15104	14389	14064	14740	12683	12289	16759	10579	10502	27162	15223	
10	7mW+(2)	3664	423	303	278	390	483	323	565	308	543	922	660	
11	7mW+(3)	3270	845	1287	1212	3366	3057	2511	3306	2852	3638	3363	3647	
12	7mW+(4)	83	69	19	19	95	61	67	91	46	65	236	105	
13	7mW+(5)	77	1	8	7	8	6	3	11	16	7	27	13	
14	7mW+(6)	34	0	0	0	0	0	0	0	0	0	0	0	
15	7mW+(7)	2	0	0	0	0	0	0	0	0	0	0	0	
16	7mW+(8)	0	0	0	0	0	0	0	0	0	0	0	0	
17	7mW+(9)	0	0	0	0	0	0	0	0	0	0	0	0	
18	7mW+(10)	0	0	0	0	0	0	0	0	0	0	0	0	
2	(1)	17002	40092	40283	38132	44386	45730	34463	56340	32473	37851	80224	51640	
3	(2)	25068	8400	8744	5204	6503	7343	5296	9761	5975	7983	17572	10423	
4	(3)	10229	25458	42479	27371	37767	37862	30532	48342	41228	40193	79032	49082	
5	(4)	5667	1492	560	316	623	545	496	746	434	389	1808	683	
6	(5)	1866	18	95	111	71	83	55	104	94	61	352	160	
7	(6)	761	0	0	1	4	0	0	0	0	0	0	0	
8	(7)	228	0	0	0	0	0	0	0	0	0	0	0	
9	(8)	125	0	0	0	0	0	0	0	0	0	0	0	
10	(9)	65	0	0	0	0	0	0	0	0	0	0	0	
11	(10)	32	0	0	0	0	0	0	0	0	0	0	0	
12	(11)	18	0	0	0	0	0	0	0	0	0	0	0	
13	(12)	7	0	0	0	0	0	0	0	0	0	0	0	
14	(13)	5	0	0	0	0	0	0	0	0	0	0	0	
15	(14)	36	0	0	0	0	0	0	0	0	0	0	0	
16	(15)	24	0	0	0	0	0	0	0	0	0	0	0	
17	(16)	5	0	0	0	0	0	0	0	0	0	0	0	
18	(17)	1	0	0	0	0	0	0	0	0	0	0	0	
19	(18)	0	0	0	0	0	0	0	0	0	0	0	0	
20	(19)	1	0	0	0	0	0	0	0	0	0	0	0	
12	12mW	1973	43558	57130	50133	59740	59163	41108	75436	62653	53842	113338	78414	
14	12mW+(1)	0	6934	7786	4686	5215	6571	5388	7547	6395	5101	10542	7797	
15	12mW+(2)	69	3278	3285	2274	2680	2713	2391	4297	2889	3308	9177	4495	
16	12mW+(3)	250	9648	14100	13326	21292	20891	19318	28346	27502	21872	34745	28288	(canonical)
17	12mW+(4)	1	865	293	166	307	235	287	329	252	163	1021	294	
18	12mW+(5)	0	9	69	75	37	54	29	47	59	33	213	91	
10	10mW	6154	67673	57128	96882	69534	94046	39560	91049	55455	61340	175824	104190	
12	10mW+(1)	0	1	2	363	1	0	15	188	0	32	22	0	
13	10mW+(2)	0	0	0	12	0	13	8	0	0	0	0	0	
14	10mW+(3)	0	0	0	217	427	382	227	587	338	235	1021	527	
15	10mW+(4)	0	0	0	0	0	0	0	0	0	0	0	12	
16	10mW+(5)	0	0	0	0	0	0	0	0	0	0	0	0	
9	9mW	846	29451	45389	30175	28413	31976	13805	37534	20020	25802	74545	37673	
11	9mW+(1)	0	0	160	4	2	9	2	7	2	0	10	5	
12	9mW+(2)	0	1823	2446	588	4	4	0	5	2	2	729	3	(haplotype B marker)
13	9mW+(3)	0	9266	19420	6268	22	0	9	0	0	1	20227	0	(haplotype B marker)
14	9mW+(4)	0	249	97	44	0	0	0	0	0	0	0	0	
15	9mW+(5)	0	0	0	0	0	0	0	0	0	0	0	0	
2	2mJ	563347	1950235	1692087	1721596	2059205	2180100	1738661	2778859	1528487	2310682	4126444	2645085	
2	2mD	115735	3312078	2695348	2775544	2958758	3182308	1784263	4186888	2347457	3367703	5033049	4509999	

Fig S18. Frequencies of detected HOR patterns (definitions in Fig.S15) in centromeric repeats in humans and chimpanzees. Data for human samples are colored based on the inferred genotypes for the chromosome 17 centromere. 12mW+(3) is the same pattern as the canonical "16m" in chromosome 17, while 9mW+(3) and 9mW+(2) are markers for haplotype B. The chimpanzee (Clint) shares some haplotype-B-specific patterns such as 4mW+(2) and 4mW+(3) with humans, and contains tandemly repeating "3BW1" (numbers in red).

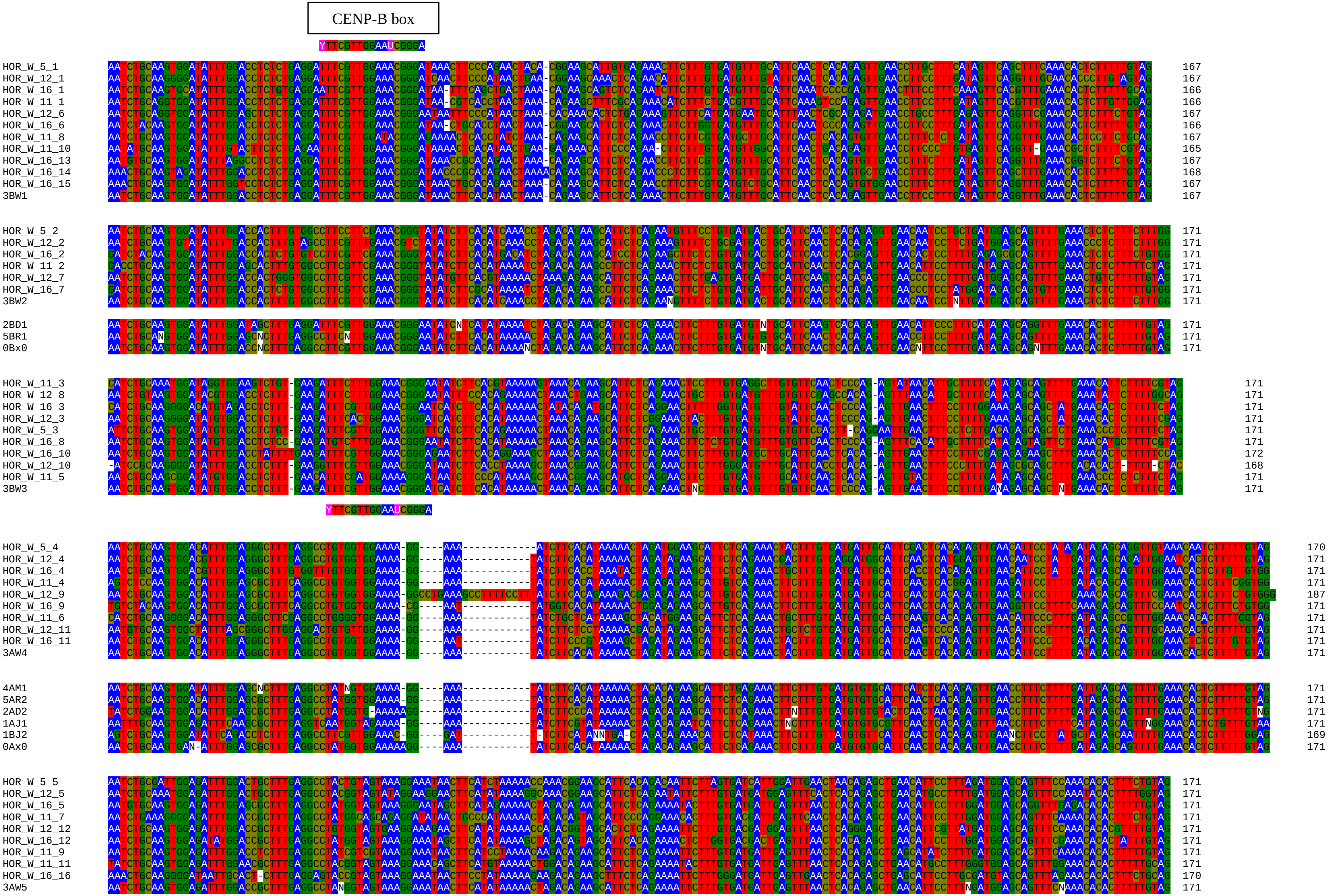


Fig S19. Structures of 58 chromosome-specific monomers.
Six clusters of multiple alignments of 58 monomers and generic monomers are shown.

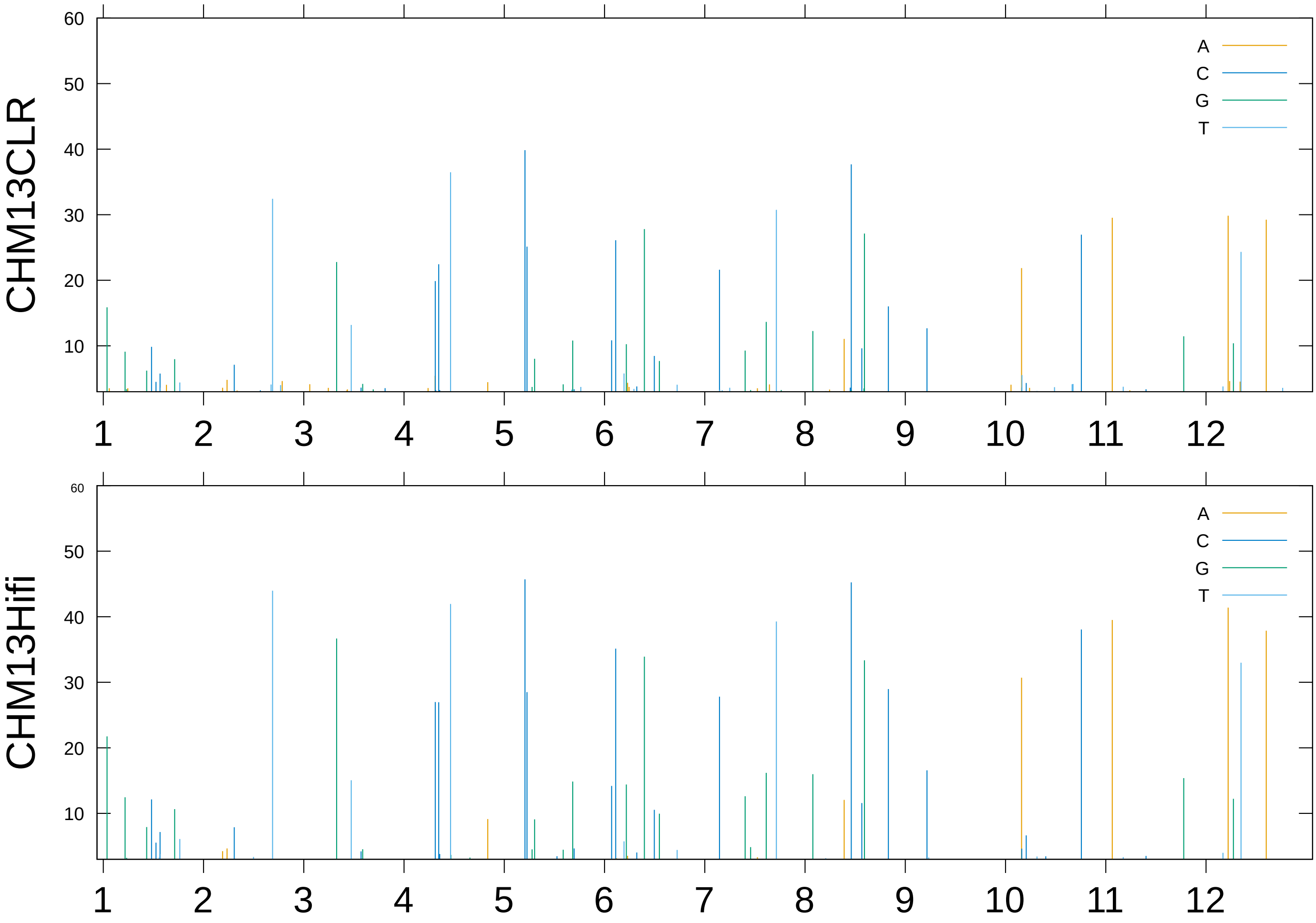


Fig S20. Comparison of SNV landscape over the 12-mer canonical HOR in chromosome X detected with CLR reads and HiFi reads for the CHM13.

SNV with frequency of >3% are shown. x-axis is labelled with monomer index, but the actual coordinate represents position and base; for example, the alternative "G" at 20th base of 2nd monomer is plotted at x = 3 (G for G) + (20 * 4) + (2 * 800) = 1683. y-axis is the observed frequency in percentage. Four colors are used to distinguish four bases.

	Mende Sierra Leone	**Esan Kenya**	**Maasai**	**Ashkenazi (HG002)**	**Toscani Italy**	**Finnish**	**Gujarati India**	**Dai Chinese**	**Han Chinese (HG005)**	**Han Chinese (HG006)**	**Han Chinese (HG007)**	**Puerto Rican**	**Peruvian**	**Chimpanzee (Clint)**	**NA12878 (HiFi)**
Project / Experiment	PRJNA438669	PRJNA530216	PRJNA530217	PRJNA558394	PRJNA530212	PRJNA480712	PRJNA530214	PRJNA547614	SRX4739017	SRX4739121	SRX4739122	PRJNA483067	PRJNA480858		PRJNA540705
Runs	SRR8035307	SRR9685168	SRR9685183	SRR9972588	SRR9677471	SRR7511979	SRR9683639	SRR9212151	SRR8955269	SRR8955270	SRR8955271	SRR7615963	SRR7515657	SRR5269206	SRR9001768
	SRR8035309	SRR9685169	SRR9685184	SRR9972589	SRR9677472	SRR7511980	SRR9683640						SRR7515658	SRR5269207	SRR9001769
	SRR8035310	SRR9685170	SRR9685185		SRR9677473	SRR7511981	SRR9683641						SRR7515659	SRR5269208	SRR9001770
	SRR8035311	SRR9685171	SRR9685186		SRR9677474	SRR7511982	SRR9683642						SRR7515660	SRR5269209	SRR9001771
	SRR8035312	SRR9685172	SRR9685187		SRR9677475	SRR7511983	SRR9683643						SRR7515661	SRR5269210	SRR9001772
	SRR8035313	SRR9685173	SRR9685188		SRR9677476	SRR7511984	SRR9683644						SRR7515662	SRR5269211	SRR9001773
	SRR8035314	SRR9685174	SRR9685189		SRR9677477	SRR7511985	SRR9683645						SRR7515663	SRR5269212	
	SRR8035315	SRR9685175	SRR9685190		SRR9677478	SRR7511986	SRR9683646						SRR7515664	SRR5269213	
	SRR8035316	SRR9685176	SRR9685191		SRR9677479	SRR7511987	SRR9683647						SRR7515665	SRR5269214	
	SRR8035317	SRR9685177	SRR9685192		SRR9677480	SRR7511988	SRR9683648						SRR7515666	SRR5269215	
	SRR8035318	SRR9685178	SRR9685193		SRR9677481	SRR7511989	SRR9683649						SRR7515667	SRR5269216	
	SRR8035319	SRR9685179	SRR9685194		SRR9677482	SRR7511990	SRR9683650						SRR7515668	SRR5269217	
	SRR8035320	SRR9685180			SRR9677483	SRR7511991	SRR9683651						SRR7515669	SRR5269218	
	SRR8035321	SRR9685181			SRR9677484	SRR7511992							SRR7515670	SRR5269219	
	SRR8035322	SRR9685182			SRR9677485	SRR7511993							SRR7515671	SRR5269439	
	SRR8035323				SRR9677486	SRR7511994							SRR7515672	SRR5269440	
	SRR8035324				SRR9677487	SRR7511995							SRR7515673	SRR5269441	
	SRR8035325				SRR9677488	SRR7511996							SRR7515674	SRR5269442	
	SRR8035326				SRR9677489	SRR7511997							SRR7515675	SRR5269443	
	SRR8035327				SRR9677490	SRR7511998							SRR7515676	SRR5269444	
	SRR8035328				SRR9677491	SRR7511999							SRR7515677	SRR5269445	
	SRR8035329				SRR9677492	SRR7512000							SRR7515678	SRR5269446	
	SRR8035330				SRR9677493	SRR7512001							SRR7515679	SRR5269447	
	SRR8035331				SRR9677494	SRR7512002							SRR7515680	SRR5269448	
	SRR8035332				SRR9677495	SRR7512003							SRR7515681	SRR5269449	
	SRR8035333					SRR7512004							SRR7515682	SRR5269450	
	SRR8035334					SRR7512005							SRR7515683	SRR5269451	
	SRR8035335					SRR7512006							SRR7515684	SRR5269452	
	SRR8035336					SRR7512007							SRR7515685	SRR5269453	
	SRR8035337					SRR7512008							SRR7515686	SRR5269454	
	SRR8035338												SRR7515687	SRR5269455	
	SRR8035339												SRR7515688	SRR5269456	
	SRR8035340												SRR7515689	SRR5269457	
	SRR8035341												SRR7515690	SRR5269458	
	SRR8035342												SRR7515691	SRR5269459	
	SRR8035343												SRR7515692	SRR5269460	
	SRR8035344												SRR7515693	SRR5269461	
	SRR8035345												SRR7515694	SRR5269462	
	SRR8035346												SRR7515695	SRR5269463	
	SRR8035347												SRR7515696	SRR5269464	
	SRR8035348												SRR7515697	SRR5269465	
	SRR8035349												SRR7515698	SRR5269466	
	SRR8035350												SRR7515699	SRR5269467	
	SRR8035351												SRR7515700	SRR5269468	
	SRR8035352												SRR7515701	SRR5269469	
	SRR8035353													SRR5269470	
	SRR8035354													SRR5269471	
	SRR8035355													SRR5269472	
	SRR8035356														
	SRR8035357														
	SRR8035358														
	SRR8035359														

Table S1. Accession IDs for publicly available sequencing data used in this study. For 15 datasets, Project / Experiment IDs and Run IDs in the Sequence Read Archive are shown. Each dataset was sequenced with Sequel I/II in CLR mode unless otherwise mentioned.

Monomer ID	Sequence
2BD1	AATCTGCAAGTGGATATTTGGATAGCTTTGAGGATTTCGTTGGAAACGGGAATATCNTCATATAAAATCTAGACAGAAGCATTCTCAGAAACTTCTTTGTGATGTNTGCATTCAAGTCACAGAGTTGAACATTCCCTTTCATAGAGCAGGTTTGAAACACTCTTTTTGTAG
2AD2	TATCTGGAAGTGGACATTTGGAGCGCTTTGAGGCCTATGGTGAAAAAGGAAATATCTTCCCATAAAAACTAGACAGAAGCATTCTCAGAAACTTNTTTGTGATGTGTGTACTCAACTAACAGAGTTGAACCTTTCTTTTGATAGAGCAGTTTTGAAACACTCTTTTTGTNG
1AJ1	AATTTGCAAGTGGAGATTTCAAGCGCTTTGAGGTCAATGGTAGAAAAGGAAATATCTTCGTATAAAAACTAGACAGAATCATTCTCAGAAACTNCTTTGTGATGTGTGCGTTCAACTCACAGAGTTTAACCTTTCTTTTCATAGAGCAGTTNGGAAACACTCTGTTTGTAA
1BJ2	AGTCTGCAAGTGGATATTCAGACCTCTTTGAGGCCTTCGTTGGAAACGGGATTTCTTCATANNTGACTAGACAGAAACATTCTCATAAACTTCTTTGTTATGTGTTCATTCAACTCACAGAGTTGAANCTTCCTTATGCTAGAGCAATTTTGAAACACTCTTTTTGGAG
4AM1	AATCTGCAAGTGGATATTTGGAGCNCTTTGAGGCCTATNGTGGAAAAGGAAATATCTTCACATAAAAACTACACAGAAGCATTCTGAGAAACTTCTTTGTGATGTGTGCATTCATCTCACAGAGTTGAACCTTTCTTTTGATTGAGCAGTTTTGAAACACTCTTTTTGTAG
5BR1	AATCTGCANGTGGATATTTGGAGCNCTTTGAGGCCTTCNTTGGAAACGGGAATATCTTCACATAAAAACTAGACAGAAGCATTCTCAGAAACTTCTTTGTGATGTGTGCATTCAACTCACAGAGTTGAACCTTCCTTTTGATAGAGCAGTTTTGAAACACTCTTTTTGTAG
5AR2	AATCTGCAAGTGGACATTTGGAGCGCTTTGAGGCCTATGGTGGAAAAGGAAATATCTTCACATAAAAACTAGACAGAAGCATTCTCAGAAACTTCTTTGTGATGTGTGCATTCAACTCACAGAGTTGAACCTTTCTTTTGATAGAGCAGTTTTGAAACACTCTTTTTGTAG
3BW1	AATCTGCAAGTGGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATAAACTTCACATAACTAAACAGAAGCATTCTCAGAAACTTCTTTGTGATGTTTGCATTCAACTCACAGAGTTGAACCTTCCTTTGATAGTTCAGGTTTGAAACACTCTTTTTGTAG
3BW2	AATCTGCAAGTGGATATTTGGACCACTTTGTGGCCTTCGTTCGAAACGGGTATATCTTCACATCAAACCTAGACAGAAGCATTCTCAGAANGTTTCTGTGATGACTGCATTCAACTCACAGAGTTGAACAATCCTNTTGATGGAGCAGTTTTGAAACTCTCTTTCTTTGG
3BW3	AATCTGCAAGTGGATATGTGGACCTCTTTGAAGATTTCGTTGGAAACGGGATCATCTTCACATAAAAACTAAACAGAAGCATTCTCAGAAACTNCTTTGTGATGTTTGTGTTCAACTCCCAGAGTTGAACTTTCCTTTTGANAGAGCAGCTNTGAAACACTCTTTTTCTAG
3AW4	AATCTGCAAGTGGACATTTGGAGGGCTTTGAGGCCTGTGGTGGAAAAGGAAATATCTTCACATAAAAACTAGATAGAAGCATTCTCAGAAACTACTTTGTGATGATTGCATTCAACTCACAGAGTTGAACATTCCTTTTGATAGAGCAGTTTGGAAACACTCTTTTTGTAG
3AW5	AATCTGCAAGTGGAGATTTGGACCGCTTTGAGGCCTANGGTAGTAAAGGAAATAACTTCATATAAAAACTAGACAGAAGCATTCTCAGAAAATTCTTTGTGATGATTGAGTTTAACTCACAGAGCTGAACATTCCTTTNGATGGAGCAGTTTCNAAACACACTTTTTGTAG
0Ax0	AATCTGCAAGTGANATTTGGAGCGCTTTGAGGCCTATGGTGGAAAAAGGAAATATCTTCACATAAAAACTAGACAGAAGCATTCTCAGAAACTTCTTTGTGATGTGTGCATTCAACTCACAGAGTTGAACCTTTCTTTTGATAGAGCAGTTTTGAAACACTCTTTTTGTAG
0Bx0	AATCTGCAAGTGGATATTTGGACCNCTTTGAGGCCTTCGTTGGAAACGGGAATATCTTCACATAAAANCTAGACAGAAGCATTCTCAGAAACTTCTTTGTGATGTNTGCATTCAACTCACAGAGTTGAACNTTCCTTTTGATAGAGCAGNTTTGAAACACTCTTTTTGTAG
HOR_W_11_1	AATCTGCAGGTGGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATAACGTCACCTAACTAAACAGAAGCTTTCGCAGAAACATCTTTCTGACGTTTGCATTCAAAGTCCAGAGTTGAACCTTCCTTTGATAGTTCACGTTTGAAACACTCTTGTTGGAG
HOR_W_11_2	GACCTGCAAGTGGATATTTGGAGCACTTTGTGGCCTTCGTTCGAAACGGGTATATCTTCACATAAAATCTAGACAGAAGCCTTCTCAGAAACTTCTCTGTGATGACTGCATTCAACTCACAGAGTTGAACATTCCTTTTGATAGAGCAGTTTTGAAACTCTCTTTTTCTAG
HOR_W_11_3	CATCTGCAAATGGATAGGTGGAAGTCTGTGAAGATTTCTTTGGAAACGGGAATATCTTCACGTAAAAAGTAAACAGAAGCATTCTCAGAAACTCCTTTGTGAGGCTTGTGTTCAACTCCCAGAGTATAACATTGCTTTTCATAGAGCAGTTTTGAAACATTCTTTTCGTAG
HOR_W_11_4	AGTCTCCAAGTGGACATTTGGAGCGCTTTCAGGCCTGTGGTGGAAAAGGAAATATCTTCACATAAAAACTAGAGAGAAGCATTGTCAGAAACTTCTTTGTGATGATTGCATTCAACTCACGGAGTTGAAGATTCCTTTTGATACAGCAGTTTGGAAACACTCTTTCGGTGG
HOR_W_11_5	AATCTGCAAGCGGATATGTGGACCTCTTTGAACATTTCGATGGAAAAGGGATAATCTTCCCATAAAAGCTAAACGGAAGCATGCTCAGGAACTTCTTTGTGATGTTTGCATTCAACTCACAGAGTTGTACTTTCCTTTTGATAGAGCAGCTTTGAAACCCTCTCTTTCTAG
HOR_W_11_6	CATCTGCAAGGGGACATTTGGAGGGCTTCGAGGCCTGGGGTGGAAAAGGAAATATCTGCTCATAAAAGCTACATGGAAGCATTCTCAGAAACTGCTTTGTGATGATTGCATTCAAGTCACAGAGTTGAACATTCCCTTTGATAGAGCCGTTTGGAAACACACTTTTGGTAG
HOR_W_11_7	AATCTGAAAGGGGAGATTTGGACCGCTTTGAGGCCTATGGCAGCAGAGGATATAACTGCCCATAAAAACTAGACAGTAGCATTCCCAGGAAACACTTTGTGACGATTGAGTTCAACTCACAGAGCTGAACATTCCTTTGGATGGAGCAGTTTCAAAACACACTTTCTGTAG
HOR_W_11_8	AATCTGCAAGTGGATATTTGGACCTCTCTGAGGATTTCGTTGGATACGGGAGAAAACTCACCTATCTAAACAGAAGCATTCTCAGAACCTTCTTCGTGATGCTTGCATTCAACTCACAGTGTTGAACCTTTCTCTGATAGTTCAGGTTTGAAACACTCCTTCTGCAG
HOR_W_11_9	AATCTGCAAGTGGAGATTTGGACCTCTTTGAGGCCTATCGTCGTAAAGGAAATAACTTCATCCTAAAACAAGACAGAAGCATTCTCAGAAAATTCTTTGTGATGATTGAGTTTAACTCACAGAGCTGAGCATATCTTTTGATGGAGCACTTTCAAAACACACTTTTTGTAG
HOR_W_11_10	AATATGCAAGTGGATATTTGTACTTCTCTGAGAATTTCGTTGGAAACGGGATAAAACTCACATAACTGAAGAGAAACATTCCCAGAACTTCTTTGTGATGTTGGCATTCAACTGACAGAGTTGAACCTTCCCTTGTGAGTTCAGGTTGAAACGCTCTTTTCGTAG
HOR_W_11_11	TATCTGCAAGTGGAGATTTGGAACGCTTTGAGGCCTACGGTAGTAAAGGAAACAGCTTCATGTAAAAACTGGACAGAAGCATTCTCAGAAAATACTTTGTGATGATTGAGTTTAACTCACAGAGCTGAACATGCCTTTGGGTGGAGCAGTTTGGAAACACACTTTTTGCAG
HOR_W_12_1	AATCTGCAAGGGGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATCAACTTCCCATAACTGAACGGAAGCAAACTCAGAACATTCTTTGTGATGTTTGTATTCAACTCACAGAGTTGAACCTTCCTTTGATAGTTCAGGTTTGCAACACCCTTGTAGTAG
HOR_W_12_2	AATCTGCAAGTGTATATTTTGACCACTTTGTAGCCTTCGTTTGAAACGTCTATATCTTCACATCAAACCTAGACAGAAGCATTCTCAGAAAGTTTTCTGCGATGACTGCATTCAACTCACAGAGTTGAACAATCCTTCTGATGGAGCAGTTTTGAAACCCTCTTTCTTTGG
HOR_W_12_3	AATCTGCAAGGGGATATGTGGACCTCTTTGAAGATTTCACTGGAAACGGGATCATCTTCACATAAAAACTAAACAGAAGCATTCTCGGAAACTACTTTGTGATGTTTGTATTCAACTCCCAGAGTTGAACTTTCCTTTTGAAAGAGCAGCTATGAAACACTCTTTTTCGAG
HOR_W_12_4	AATCTGCAAGTGGACGTTTGGAGGGCTTTGAGGCCTGTGGTGGAAAAGGAAATATCTTCACATAAAAACTAGATAGAAGCATTCTCAGAAACGACTTTGTGAGGATGGCATTCAACTCATGGAGTTGAACAATCCTATTGATAGAGCAGATTGGAATCACTCTTTTTGTAG
HOR_W_12_5	AATCTGCAAATGGAGATTTGGACTGCTTTGAGGCCTACGGTAGTATAGGAAGGAACTTCATATAAAAGGCAAACGGAAGCATTCTCAGAATATTCTTTGTGATGATGGAGTTTCACTCACAGAGCTGAACATGCCTTTTGATGGAGCAGTTTCCAAATACACTTTTGGTAG
HOR_W_12_6	AATCTGCAGGTGGATATTTGGAGCTCTCTGAGGATTTCGTTGGAAACGGGAATAATTTCCCATAACTAAACACAAACACTCTGAGAAAGTTCTTCATGATGAATGCATTTAACTCGCAGAGATGAACCTGCCTTTGAGAGTTCAGGTTCGAAACACTCTTTCTGTAG
HOR_W_12_7	AATCTGCAAGTGGATATTTGGACCACTGGGTGGCCTTCGTTCGAAACGGGTATATGTTCACGTAAAAACTAAAGAGAAGCATTCTCAGAAACTTCTGAGTGATGATTGCATTCAAGTCACACAGTTGAACCCTCCTTTTGATGGAGCAGTTTTGAAACTGTCTTTTTGTAG
HOR_W_12_8	AATCTGTAAGTGGATACGTGGACCTCTTTGAAGATTTCTTTGGAAACGGGAATATTTCCACAGAAAAACTAAACTGAAGCATTCTCAGAAACTGCTTTGTGATGTTTGTGTTCGAGCCACAGAGTTTAACATTGCTTTTCATAGAGCAGTTTTGAAATATTCTTTTGGCAG
HOR_W_12_9	AATCTGCAAGTGGACATTTGGAGCGCTTTCAGGCCTGTGGTGGAAAAGGCCTGAAAGCCTTTTCCTTTATCTTCACAGAAAGACGAGAGAGAAGCATTGTCAGAAACTTCTTTGTGATGATTGCATTCAACTCACAGAGTTGAAGATTCCTTTTGAAACAGCAGTTTCGAAACACTCTTTCTGTGGG
HOR_W_12_10	ATCCGCAAGGGGATATTTGGACCTCTTTGAAGGTTTCGTTGGAAACGGGATAATCTTCACCTAAAAGCTAAACGGAAGCATTCTCAGAAACTTCTTTGGGATGTTTGCATTCACCTCACAGAGTTGAACTTTCCCTTTGATAGCGCAGCTTTGACACACTTTTTCTAC
HOR_W_12_11	AATGTGCAAGTGGCTATTTAGCGGGCTTGGAGGACTGTGTTGGAAAAGGAAATATCTTCTCCTAAAAACGACATAGAAGCATTCTCAGAAACTGCTCTGTGATGATTGCATTCAACTCCCAGAGTTGAACATTCCTTTTGATAGAGCAGTTTGCAAACACTCTTTTTGTAG
HOR_W_12_12	AATCTGCAAGTGGAGATTTGGACCGCTTTGAGGCCTGTGGTAGTGAAGGAAAGAACTTCATATAAAAACCAGACGGTAGCACTCTCAGAAAATTCTTTGTGACGATGGAGTTTAACTCAGGGAGCTGAACATTCGTTATGATGGAGCAGTTTCCAAACACACGTTTGTAG
HOR_W_16_1	AATCTGCAAGTGCATATTTGGACCTCTGTGAGGAATTCGTTGGAAACGGGATAATTTCAGCTGACTAAACAGAAGCAGTCTCAGAATCTTCTTTGTGATGTTTGCATTCAAATCCCCGAGTTGAACTTTCCTTTCAAAGTTCACGTTTGAAACACTCTTTTTGCAG
HOR_W_16_2	GATCTACAAGTGGATATTTGGACCACTCTGTGTCCTTCGTTCGAAACGGGTATATCTTCACATGACATCTAGACAGAAGCATCCTCAGAAGCTTCTCTGTGATGACTGCATTCAACTCACGGAGTTGAACACTCCTTTTGAGAGCGCAGTTTTGAAACTCTCTTTCTGTGG
HOR_W_16_3	CATCTGCAAGGGGACATGTAGACCTCTTTGAAGATTTCGTTGGAAACGGAATCATCTTCACATAAAAACTATACAGATGCATTCTCAGGAACTTTTTGGTGATGTTTGTATTCAACTCCCAGAGTTGAACTTTCCTTTGGAAAGAGCAGCTATGAAACACTCTTTTTCTAG
HOR_W_16_4	AATCTGCAAGTGGACGTTTGGAGGGCTTTGTGGTTTGTGGTGGAAAAGGAAATATCTTCACCTAAATACTAGATAGAAGCATTCTCAGAAACTGCTTTGTGATGATTGCATTCACCTCACAGAGTTGAACATTCCTATTGATAGAGCAGTTTGGAAACACTCTTGTTGTGG
HOR_W_16_5	AATGTGCAAGTGGAGATTTGGAGCGCTTTGAGGCCTATGGTAGTAAAGGGAATAGCTTCATAGAAAAACTAGACAGAAGCATTCTCAGAAAATACTTTGTGATGATTGAGTTTAACTCACAGAGCTGAACATTCCTTTGGATGGAGCAGGTTTGAGACACACTTTTTGTAG
HOR_W_16_6	AATCTACAAGTGGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATAACTGCACCTAACTAAACGGAAGCATTCTCAGAAACTTCTTGGTGATGTTTGCATTCAAATCCCAGAGTTGAACCTTCCTTTGATAGTTCAGGTTTGAAACACTCTTTTTGTAG
HOR_W_16_7	GATCTGCAAGTGGATATTTGGACCACTCTGTGGCCTTCGTTCGAAACGGGTATATCTTCGCATAAAATCTAGACAGAAGCCTTCTCAGAAACTTCTCTGTGATGATTGCATTCAACTCACAGAGTTGAACCCTCCTATGGATAGAGCAGTGTTGAAACTCTCTTTTTGTGG
HOR_W_16_8	AATCTGCAAGTGGATATGTGGACCTCTCCGAAGATGTCTTTGGAAACGGGAATATCTTCACATAAAAACTAAACAGAAGCATTCTCAGAAACTTCTCTGTGATGTTTGTGTTCAACTCCCAGAGTTTCACATTGCTTTTCATAGAGTAGTTCTGAAACATGCTTTTCGTAG
HOR_W_16_9	TGTCTACAAGTGGACATTTGGAGCGCTTTCAGGCCTGTGGTGGAAAACGAATTATGGTCACATAAAAACTGGAGAGAAGCATTGTCAGAAACTTCTTTGTGATGATTGCATTCAACTCACAGAGTTGAAGGTTCCTTTTCAAAGAGCAGTTTCCAATCACTCTTTCTGTGG
HOR_W_16_10	AATCTGCAAGTGGATATTTGGACCTATTTTGAAGATTTCGTTGGAAACGGGAGAATCTTCACAGGAAAGCTAAACAGAAGCATTCTCAGAAACTTCTTTGTGATGCTTGCATTCAACTCACAGAGTTGAACTTTCCTTTCGAGAGAGAAGCTTTGAAACACTCTTTTTCCAG
HOR_W_16_11	AATCTGCAAGTGGACATTTGGAGGGCTTTGAGGCCTGTGGTGGAAAAGGAATTATCTTCCCGTAAAAGCTAGATAGAAGCATTCTCAGAAACTACTTTGTGATGATTGCATTCAAGTCACAGAGTTGAACATTCCCTTTGACAGAGCAGTTTGGAAACTCTCTTTGTGTAG
HOR_W_16_12	AATCTGCAAGTGGAGATATGGACCGCTTTGAGGCCTATGGTAGTAAAGGAAATAGCTTCATATAAAAGCTAGACAGTAGCATTCACAGAAAACTCTTGGTGACGACTGAGTTTAACTCACAGAGCTGAACATTCCTTTGGATGGAGCAGTTTCGAAACACACTATTTGTAG
HOR_W_16_13	AATGTGCAAGTGGATATTTAGGCCTCTCTGAGGATTTCGTTGGAAACGGGATAAACCGCACAGAACTAAACAGAAGCATTCTCAGAACCTTCTTCGTGATGTTTGCATTCAACTCACAGTGTTGAACCTTTCTTTGATAGTTCAGGTTTGAAACGGTCTTTCTGTAG
HOR_W_16_14	AAACTGCAAGTAGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATAACCCGCACAGAACTAAAACAGAAGCATTCTCAGAACCCTCTTCGTGATGTTTGCATTCAACTCACAGTGCTGAACCTTTCTTTGATAGTTCAGCTTTGAAACACTCTTTTTGTAG
HOR_W_16_15	AAACTGCAAGTGGATATTTGGTCCTCTCTGAGGATTTCGTTGGAAACGGGATAAACTGCACAGAACTAAACAGAAGCATTCTCAGAACCTTCTTCGTGATGTCTGCATTCAACTCACAGTGTGGAACCTTTCTTTGATAGTTCAGGTTTGAAACACTCTTTTTGTAG
HOR_W_16_16	AAACTGCAAGGGGATAATTGCACTCTTTGAGGAGTACCGTAGTAAAGGAAATAACTTCCTATAAAAAGAAGACAGAAGCTTTCTCAGAAAATTCTTTGGGATGATTGAGTTGAACTCACAGAGCTGAGCATTCCTTGCGATGTAGCAGTTTAGAAACACACTTTCTGCAG
HOR_W_5_1	AATCTGCAAGTGGATATTTGGACCTCTCTGAGGATTTCGTTGGAAACGGGATAAACTTCCCAGAACTACACGGAAGCATTGTGAGAAACTTCTTTGTGATGTTTGCATTCAACTCACAGAGTTGAACCTTGCTTTCATAGTTCAGCTTTCAAACACTCTTTTTGTAG
HOR_W_5_2	AATCTGCAAGTGGATATTTGGACCACTTTGTGGCCTTCCTTCGAAACGGGTATATCTTCACATCAAACCTAGACAGAAGCATTCTCAGAATGTTTCCTGTGATGACTGCATTCAACTCACAGAGGTGAACAATCCTGCTGATGGAGCAGTTTTGAAACTCTCTTTCTTTGG
HOR_W_5_3	ATTCTGCAAGTGGATATGTGGACCTCTGTGAAGATTTCGTTGGAAACGGGTTCATCTTCACAGAAAAACTAAACAGAAGCATTCTCAGAAACTGCTTTGTGATGTTTGTGTTCCACTTCAGGAATTGAACTTTCCTCTTGACAGAGCAGCTCTGAAACCCTCTTTTTCTAG
HOR_W_5_4	AATCTGCAAGTGGACATTTGGAGGGCTTTGAGGCCTGTGGTGGAAAAGGAAAATCTTCACATAAAAACTAGATGGAAGCATTCTCAGAAACTACTTTGTGATGATTGCATTCGACTCACAGAGTTGAACATTCCTATAGATAGAGCAGGTTGTAAACAATCTTTTTGTAG
HOR_W_5_5	AATCTGCGATTGGAGATTTGGACTGCTTTGAGGCCTACTGTAGTAAAGGAAATAACTTCATCTAAAAACCAAACGGAAGCATTCACAGACAATTCTTAGTGATCATTGGATTGAACTAACAGAGCTGAACATTCCTTTAGATGGAGCAGTTTCCAAACACACTTTCTGTAG

Table S2. List of alpohid monomer sequences used for monomer-encoding of long reads.
DNA sequence of 14 generic monomers and 44 chromosome-specific monomers used in this study.

Family	Group	Name	Accession
2	B	D1	AJ130751.1
2	A	D2	AJ130752.1
1	A	J1	AJ130753.1
1	B	J2	AJ130754.1
4	A	M1	AJ130755.1
5	B	R1	AJ130756.1
5	A	R2	AJ130757.1
3	B	W1	AJ130758.1
3	B	W2	AJ130759.1
3	B	W3	AJ130760.1
3	A	W4	AJ130761.1
3	A	W5	AJ130762.1
*	*	A	AJ131207.1
*	*	B	AJ131208.1

Table S3. List of 14 generic monomers and their Accession IDs used in this study. Accession IDs in the Sequence Read Archive of publicly available reference sequences for 14 generic monomers, along with their assignment to suprachromosomal families, evolutionary groups, and canonical names.

List of Supplementary Data:

Supplementary Data S1. Summary of variant HORs detected.
The number (before any normalization) of detected variant HORs (i.e., HOR-SVs) in each sample. Here are four sheets each corresponding to the chromosome 11, 1, X, and 17.

Supplementary Data S2. Local enrichment of variant HORs.
Frequencies of 2-mer, in which the variant appears sporadically (S) or tandemly (T), were tested against the null model where they existed randomly according to the frequency of the variant themselves. Comparisons where the statistical significance was attained are colored in red. Left: 13-mer variant (13m9-13) in chr17, Right: 6-mer variant (6m1) in chr11

Supplementary Data S3. Summary of SNVs observed.
The summary of SNVs among canonical HOR copies in each sample. The relative frequencies and basic statistics for each SNV (position and alternative base) are shown. Here are four sheets each corresponding to the chromosome 11, 1, X, and 17.

Supplementary Data S4. Analysis of paralogous SNVs on canonical HORs.
Here are two sheets according to how SNVs were counted.
Unique sites = SNVs are counted only once if they were detected in any of samples (with CLR reads).
Total sites = SNVs are counted according to the number of samples where it was detected.
For example, pSNV was counted as "10" sites if the SNV was detected in 10 samples.
Each sheet contains four parts for the chromosomes 1, 11, 17, and X.

Here, we assumed substitution error rate was 9% in total (i.e., 3% for each alternative base), and we used the observed number of alternative base to perform binomial test. The resulted p-values were adjusted by multiplying them with the approximate number of tests for the HOR: (3 alternatives) * (171 bp / monomer) * (length of HOR unit in monomers). Specifically, these adjusting factors were 5643, 2565, 8208, and 6156 for 11mer(chr1), 5mer(chr11), 16mer(chr17), and 12mer(chrX), respectively. Finally, we detected paralogous SNV if the adjusted p-value was less than 0.01.

In the tables of detected variants,
Rows = 2-mer contexts of the variant (the first base is to be substituted)
Columns = Alternative bases. The latter columns represent counts normalized by frequency of 2-mer in the original sequences. The transitions (A/G, G/A, C/T, or T/C) are boxed to be contrasted with the transversion (other types of substitutions).