

SUPPLEMENTARY FIGURES AND TABLES

Figure S1

Repeat graph generated from a human WGS sequence library. We tested the RepeatNet algorithm using a fosmid end-sequence library (WIBR2). We retrieved the fosmid end sequences that contained the overrepresented k-mers shown in this figure and checked their repeat content using RepeatMasker. We found that the largest cluster (red) denotes the alpha-satellite repeats in the human genome, confirming the use of our RepeatNet algorithm. The components in the repeat graph are annotated by the repeat sequence they represent. The clusters from the largest to the smallest denote alpha-satellite (red), satellite II and III (green), LINE and SINE (yellow), homopolymer repeats (blue), and low complexity (orange).

Figure S2

Example of FISH results of CFA-specific fosmid clones. RepeatNet graph is reported in (C) showing four different repeat clusters. (A) FISH with the clone G630P8139H12_BI1-1903P24 corresponding to green cluster; (B) FISH with the clone G630P816490E11_BI1-411I22 corresponding to blue cluster; (D) FISH using the clone G630P84226G11_BI1-1356M22 corresponding to red cluster; and (E) FISH with the clone G630P88303G10_BI1-3749N20 corresponding to yellow cluster.

Figure S3

Amplification results obtained using primers designed on the specific consensus sequences of species-specific genomic DNA (Table 2). 1, 2log ladder; 2, ECAcons70; 4, ECAcons71; 6 ECAcons421+424; 8, ECA3cons221; 10, ECA4cons450; 12 ECA5cons451; 14, CFAcons244+246; 16, DNOcons173; 18, LAFcons842+936; 20, MDOcons528 and 22, OANcons144. 3, 5, 7, 9, 11, 13, 15, 17, 19, 21 empty wells.

Figure S4

FISH results using PCR products as probes obtained for each consensus sequences. (A) FISH results with ECAcons70 on ECA metaphases; (B) FISH with LAFcons842+936 on LAF metaphases; (C) FISH with DNOcons173 on DNO metaphases; (D) FISH with MDOcons528 on MDO metaphases; and (E) FISH with OANcons144.1+144.2 on OAN metaphases.

Figure S5

Cohybridization results using species-specific amplification products (green) and ChIP (red) DNA in horse (A), elephant (B), and opossum (C). In yellow, overlapping signals between immunoprecipitated DNA and PCR probes.

Table S1

Sequences extracted by PHRA and TFR (see text for details); FISH and ICHC results obtained on the six studied mammals. In brackets the positive chromosomes out of the full karyotype are reported.

Table S2

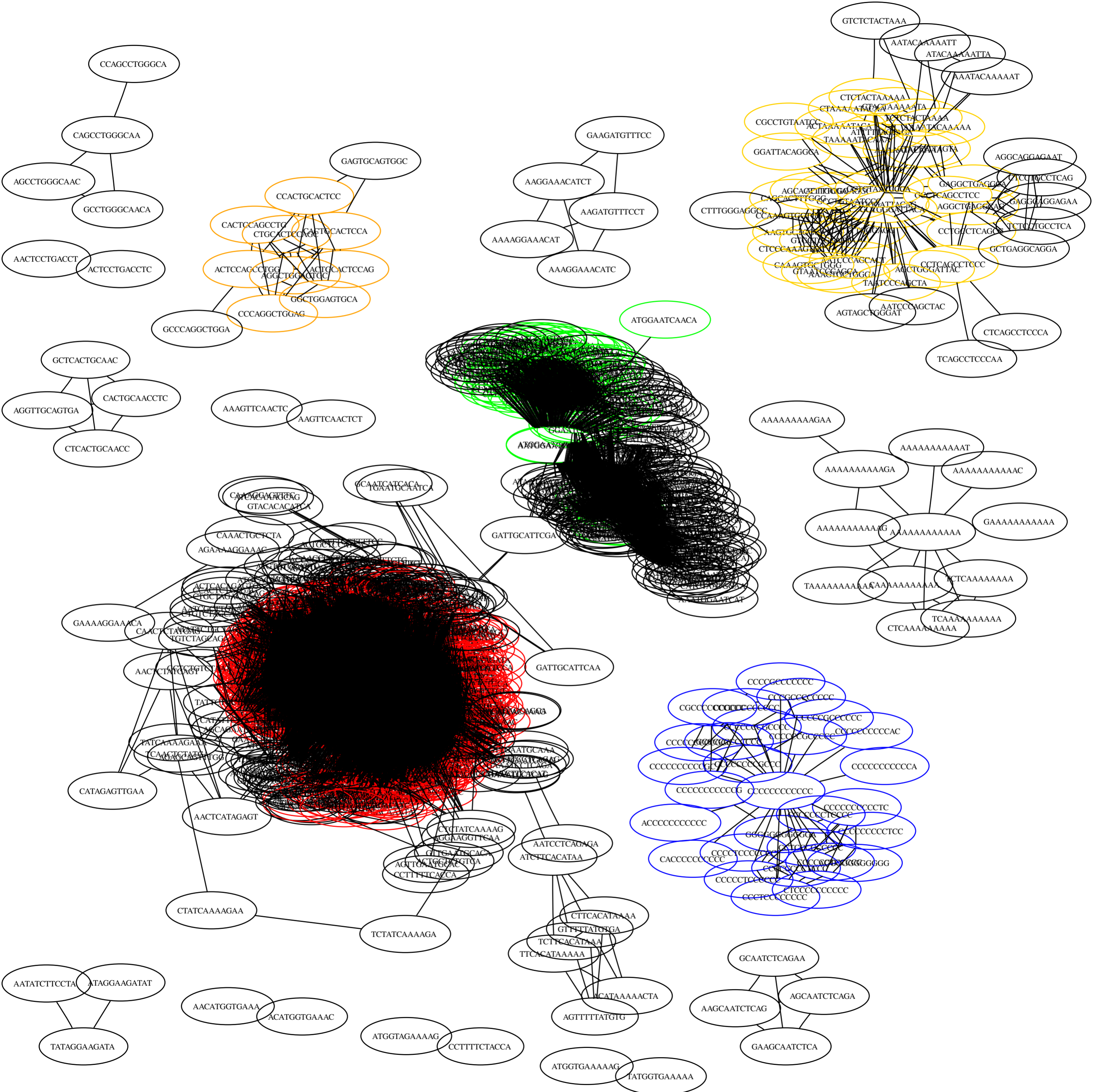
RepeatMasker analysis on the consensus sequences.

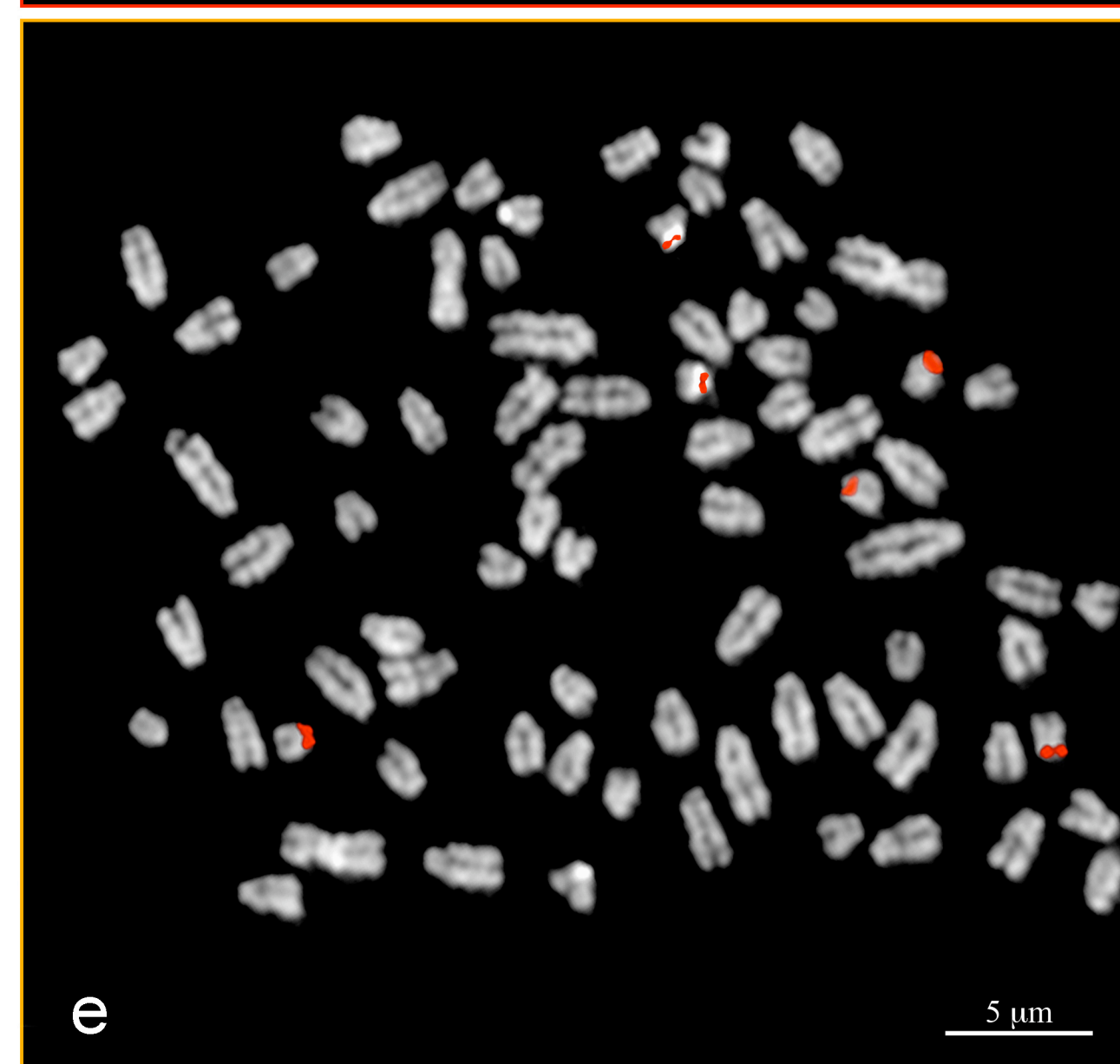
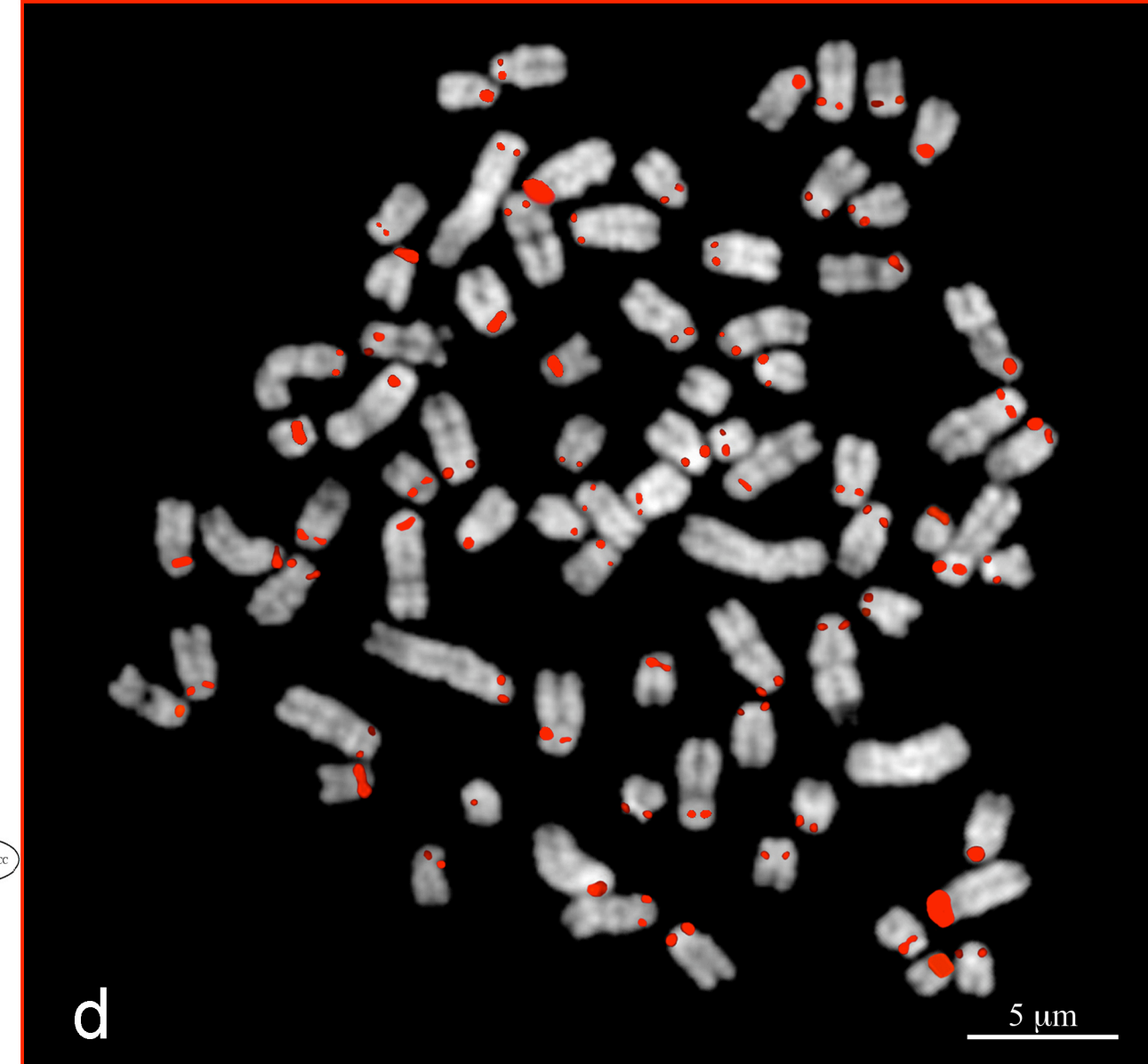
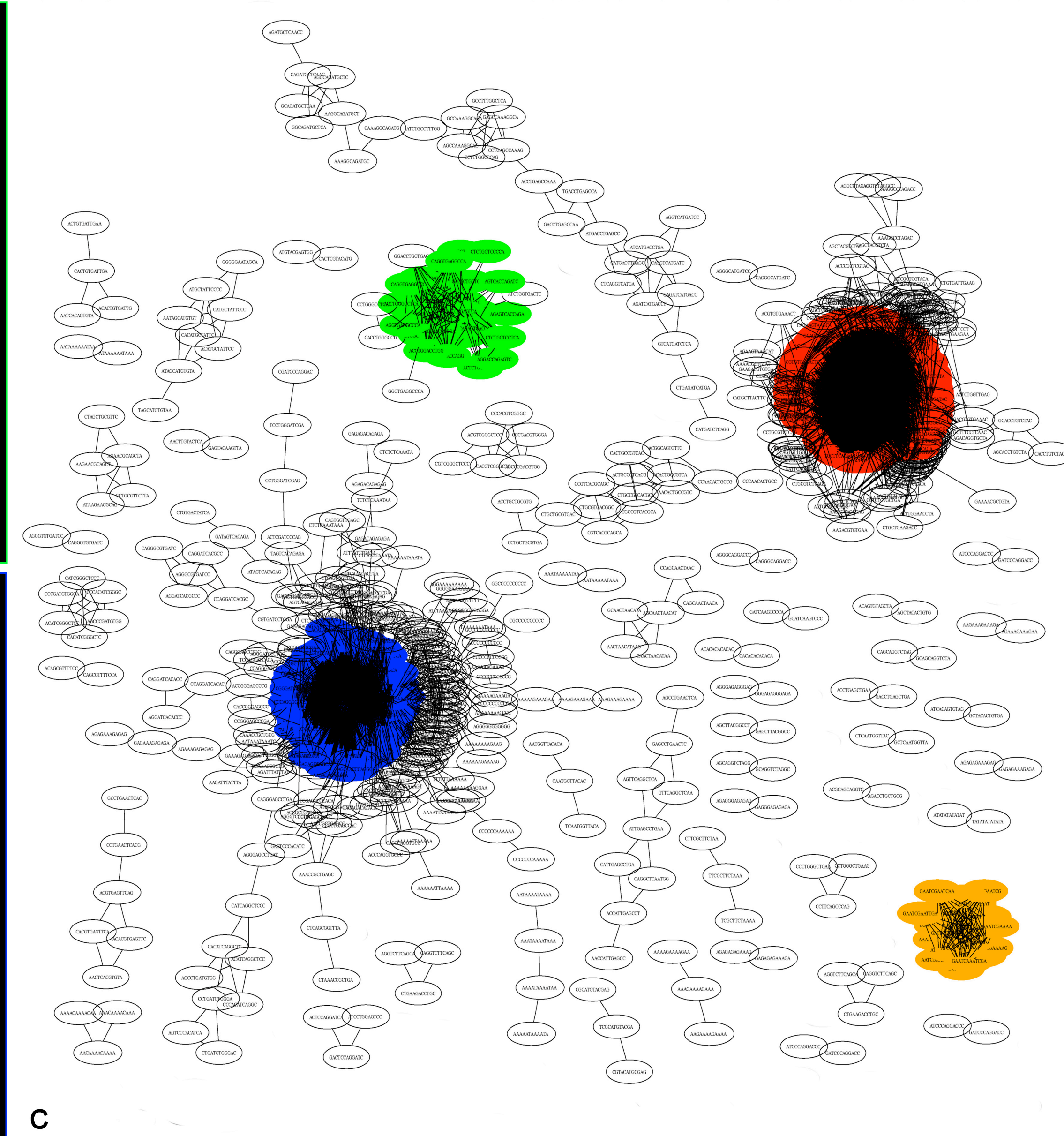
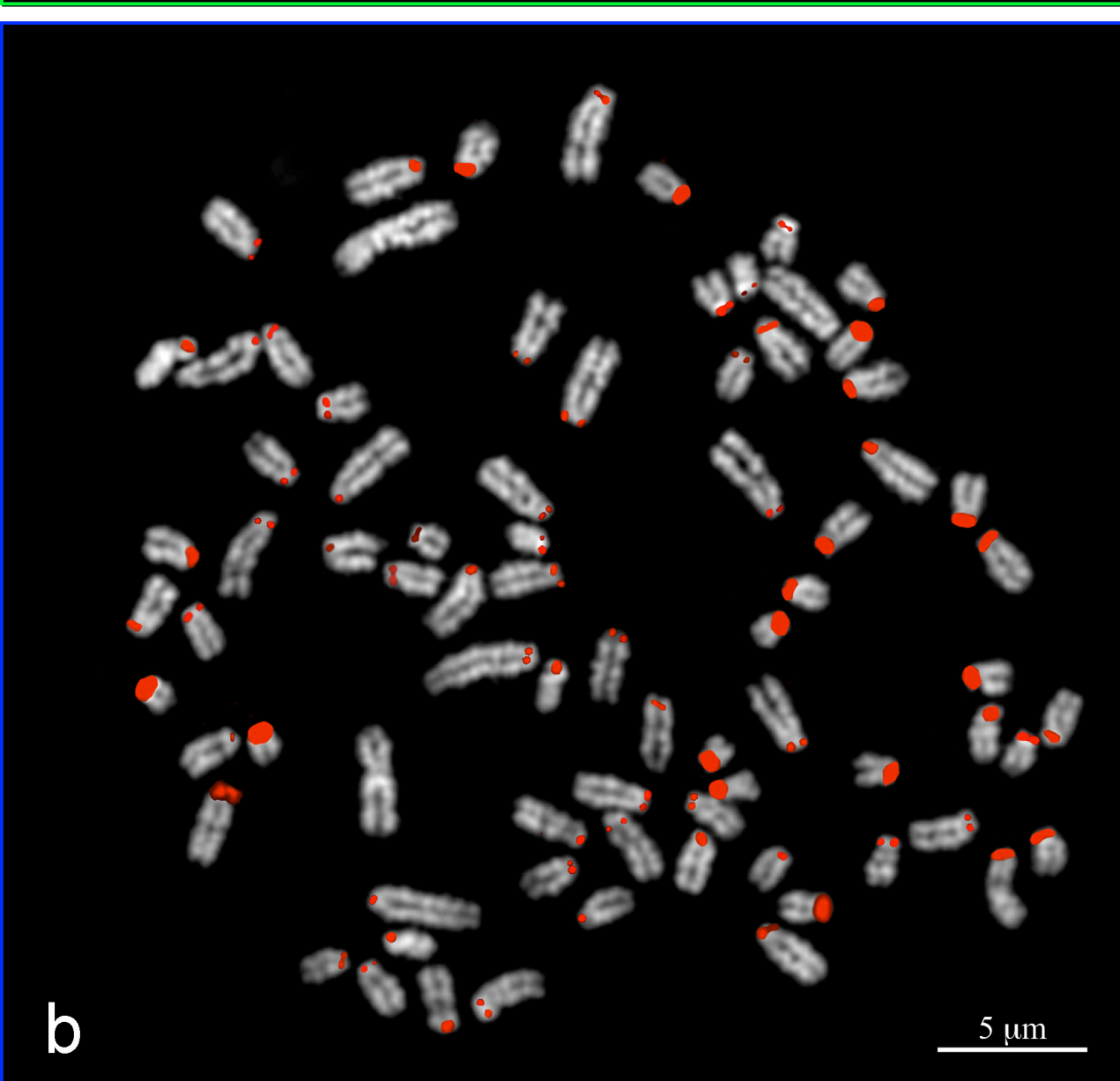
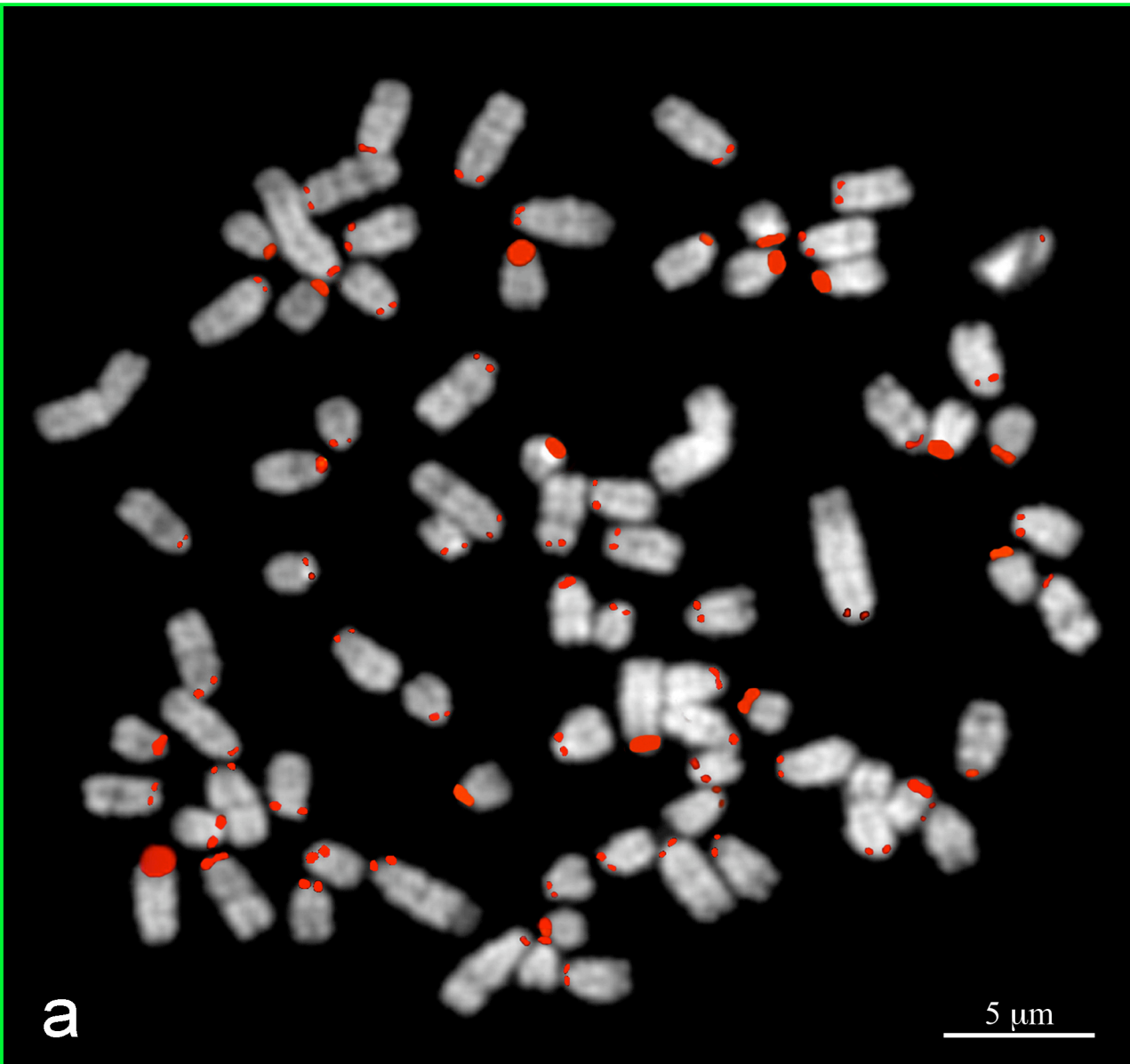
Table S3

FISH results with specific CFA centromeric fosmids (in parentheses the number of chromosomes according to (Breen et al. 1999)).

Table S4

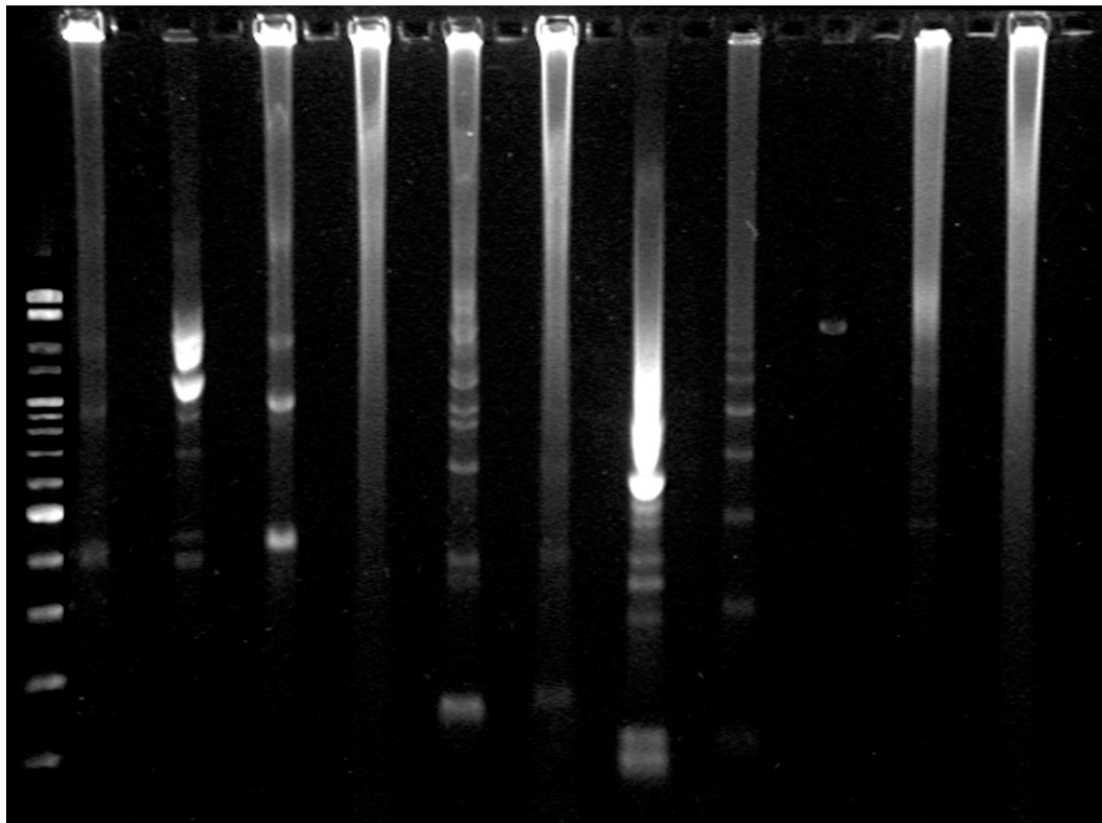
ICHC results.



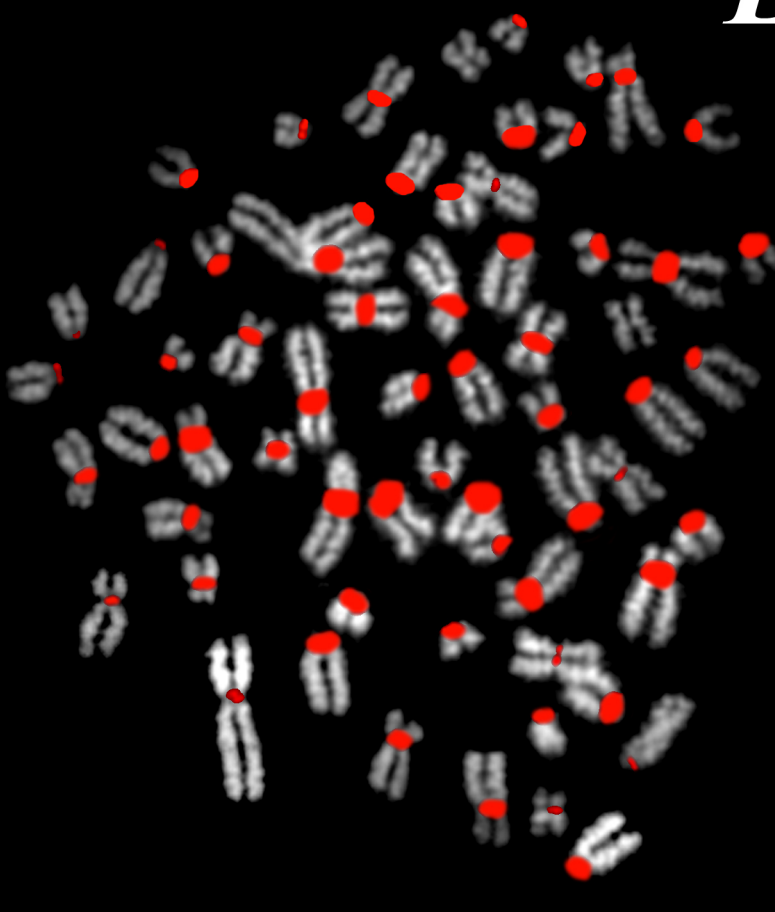


1 2 4 6 8 10 12 ECA CFA 14 DNO 16 LAF 18 MDO 20 OAN 22

3000
2000
1500
1200
1000
900
800
700
600
500
400
300
200
100



ECA



A

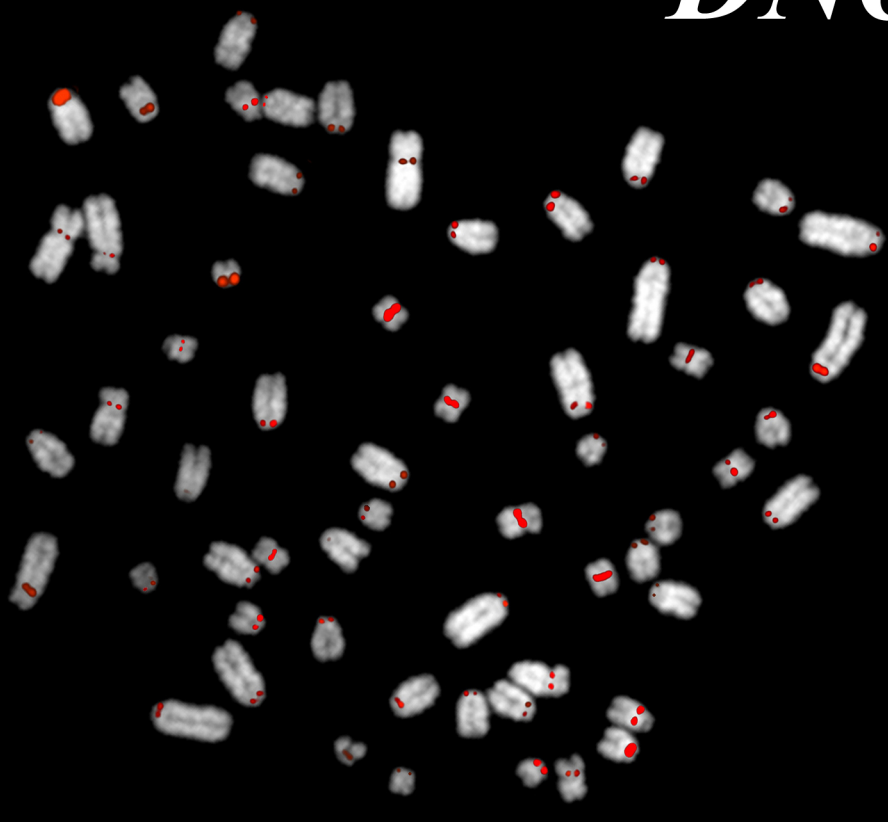
LAF



B

5 μ m

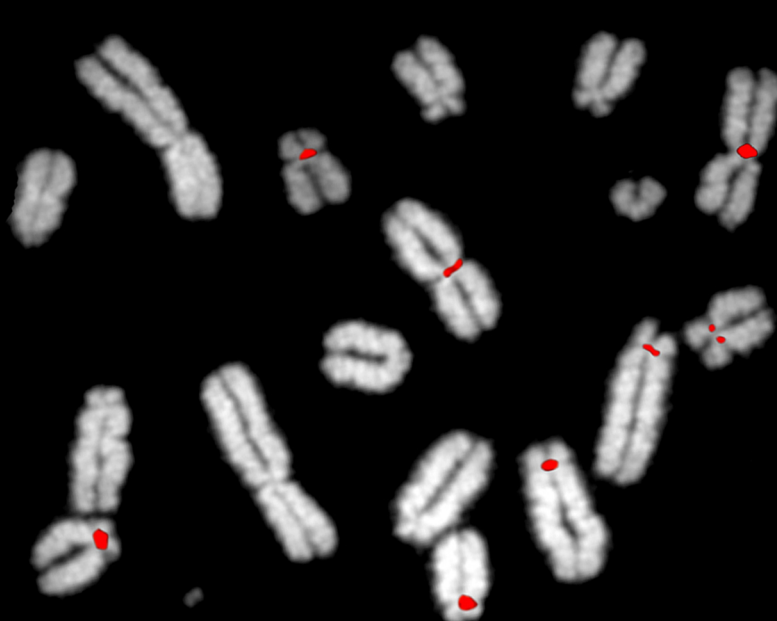
DNO



C

5 μ m

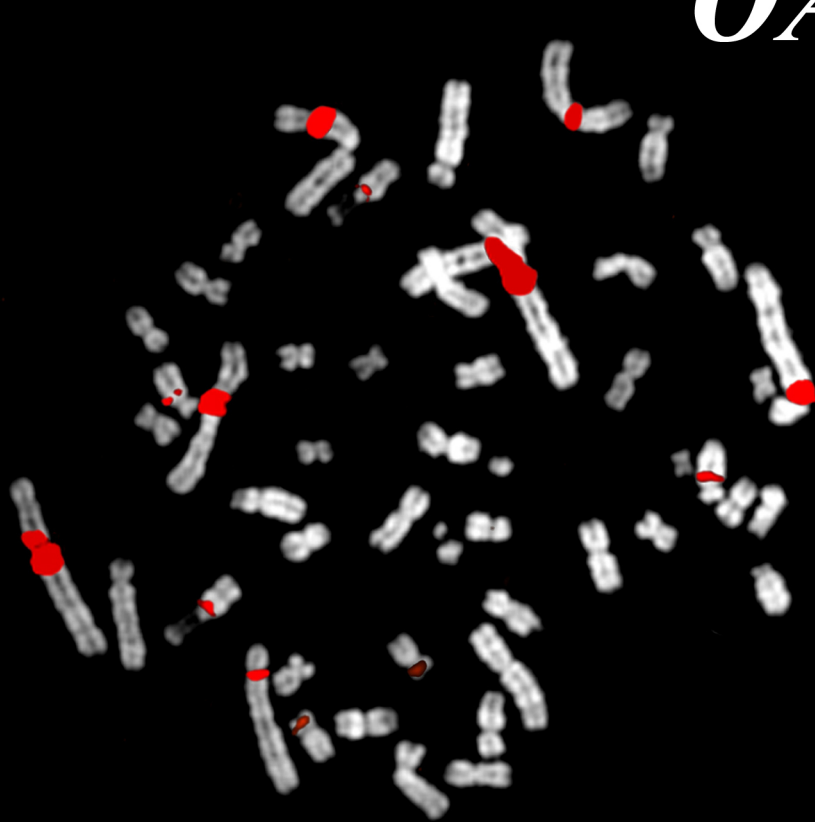
MDO



D

5 μ m

OAN



E

5 μ m

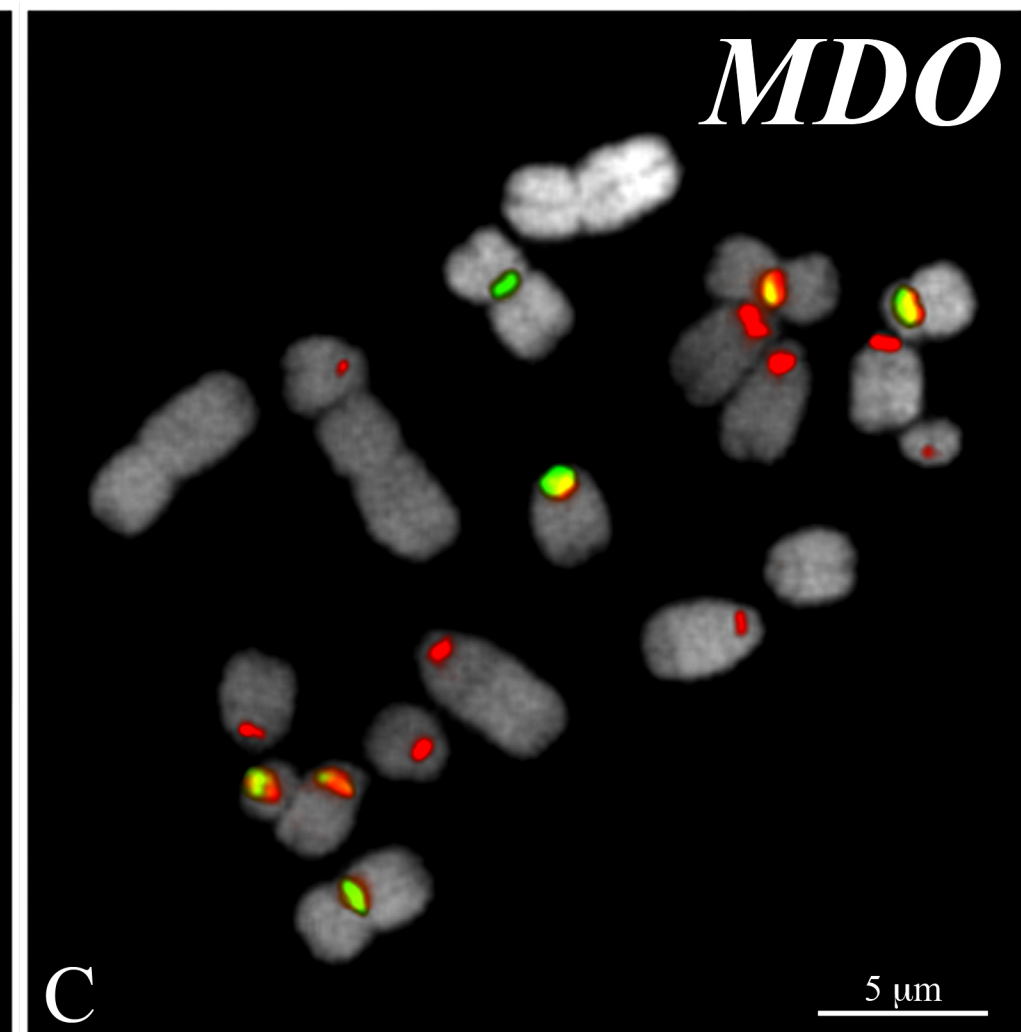
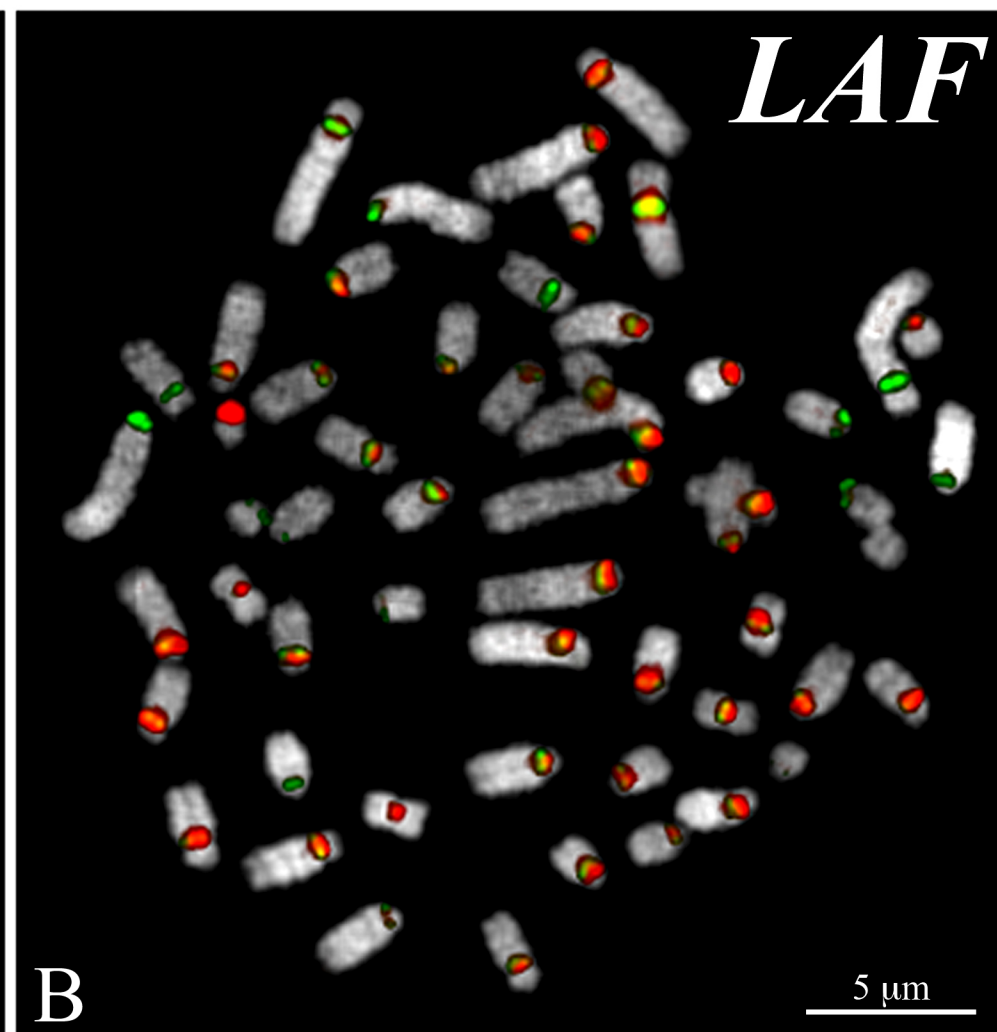
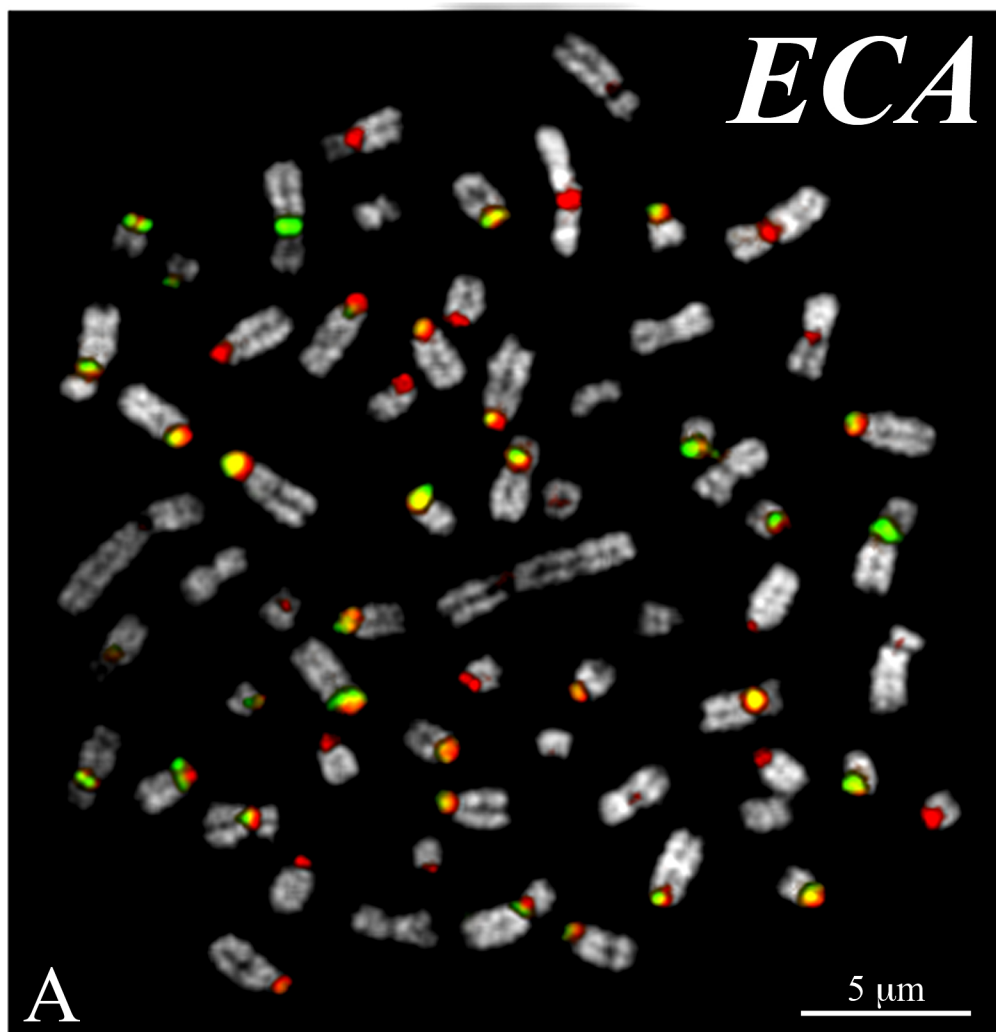


Table S2. RepeatMasker analysis on the consensus sequences

SPECIES	Sequence Code	Sequence	Size bp	Begin	End	Matching repeats	Repeat class/family
ECA	ECAcons70	TCAGCTGCCTTTGTATGAGAGTTTCCCAGGACGCTGTAAGAGCACTGTGGAA AGCGAGTTCTTTCCCAGCTTCCCTAAAGAGCTGGAAGCAAGACAGTTTATGGC TTGTCTCCAATTGAAGGATGGAGGCAGTGCTTTGTGCCTTCCACCTTAGAG CAATGGAGGGCACGGCTGAGAGCAAAGGGCCCTTCTGACATAGGTGCATT CTGACACTCTGACCCAGTGCACAATGGCGGTTTGTA AAAAGCCTATTGTC TGCTCCTCCTAAGCATGTGGAAGCACAATCATCTGGGCCCTCGTCCCCGTTGCT TAAGGGGAAGATGTAGGCATTTCTGCTGAGCCGGTTGGCAAGGTGAAAAT GTGGCAGAAGCAGAAGTCCAAAGCGGGAGTTAAGCGCTGAAAAGAAATG GCATT	418	7	187	SAT_EC	Satellite
	ECAcons71	TGTTTCTTAGCTTCCCAAAGAGCTGGAAGCAAGATGCTGTGGGGCCCAACTC GCCCTTTGGAAGAAGCCTGCACGTTGTGCCTTTCAGCTCTAGGGCAAAGTA GCACACCCAGAGCAGAAGTTCTACTTCAGCCAGATCAGGCCTGCAAAGAAA CTGCGTTTACAGGCCCTTTGGAAGAGATGTTCCCGGTAGGCTGTAAGAGCAC TGTGCAGAGCGAGT	221	1	221	SAT_EC	Satellite
	ECAcons421+424	TATGTCAGAAAGGGCCCTTTGCTCTCAGCCGTGCCTCCATTGCTCTAGAGGT GGAAGGCACAAAGCACTGCCTCCATCCTCAATGGGAGACAAGCCATAAAC TGTCTTGCTCCAGCTCTTTAGGAAGCTGAGAAAGAACTCGCTTCCACAGT GCTCTTACAGCATAACGGGAAAACATCTCATACAAAGGCAGCTCAAATGCCATT TCTTCCCAGCGCTTAACTCGCCGCTTTAGGACTTCCGCTTCTGCCACAATTCC ACCTTGCCAAACCCGGGCTCAGACGAAATGCTACATCTTCCCTTAAGCAAC GGGGGATGAGGCCCAAATGATTGTGCTTCCACATGCTTAGGAGGAGCAGAC AATAGGCTTTTTACAAACCACCATTGTGCACTGGGTGAGAGAGTGTGAGAAA GCACCTATGTCAGAAAGAGCCCTTTGCTCTCAGCCGTGCCCTCCATTGCTCT AGAGGT	475	13	217	SAT_EC	Satellite
	ECA3cons221	TCTTGCTTCCAGCTCTTTGGGAAGCTAAGAAAAAAAGTCGCTCTGCACAGTG CTCTTACAGCCCTACCGGGAACATCGTCTTACAAGGCCTGTGAAACGCAATT TCTTTGCAGGCCTGATCTGGCTGAAATAGAACTTCTGCTCTGCGTGTGCTACA TTGCCCTAGAGCTGAAAGGCACAAAAGTGTGCTTCTTTCCAAAGGGGCAGA TGGCCCCAAAGCA	221	1	221	SAT_EC	Satellite
	ECA4cons450	CTCATACAAAGCCTGTGAAACGCAGTTTACTTGGAAGGCCTGCATCTGGCT GAAATAGAACTTCTGCTCTGCCATGTCCTACATTGCCCTAGAGCTGAAAGCA CAAAGTGGAAGTCTTCTTCCAAAGGGCAGTAGGGCCCCAAAGCATCTTGC GTTCCAGCTCTTTGGGAAGCTAAGAAAAAACTCGCTCGGGCCAGTGCTCTT ACAGCCTACCGGAGAACATCGTCTTACAAGCCCTGTGAAACGCAATTTCTTT GCAGGCCTGATCTGGCTGAAGTAGAACTTCCGCGCTGAGTGTCTTACATTGC CATAACGCTGCAAGGCACAAAAGTGTGCTTCTTTAGAAAGGGGCAGATGA GCCCCAAAAGCATCTTACTTCCAGCTCTTTGGGAAGCTAAGAAAACAAATCG CTCTGCACAGTGCCTTAAATCCACCGGGGACATC	450	1	450	SAT_EC	Satellite

	ECA5cons451	CCCCAAAGCATCTTTCGCTTCCAGCTCTTTGGGAAGCTAAGAAAAAACTCGC TCGGGCCAGTGCTCTTACAGCCTACCGGAGAACATCGTCTTACAAGCCCTGT GAAACGCAATTTCTTTGAGGCCTGATCTGGCTGAAGTAGAACTTCCGCGCT GAGTGTCCATACATTGCCATACAGCTGCAAGGCACAAAGTGCTGCCTTCTTA GAAAGGGGCAGATGAGCCACAAAGCATCTTACTTCCAGCTCTTTGGGAAG CTAAGAAACAAATCGCTCTGCACAGTGCCTTAAATCCCAACGGGGAAATCC TCATACAAAGGTCCTGTGAACGGCATTACTTGCAAGGCCTGATCTGGCTGA AATATGACATTCTGGCTCTGCCATGGGCTACATTGCACCTAAAGCTGAAAGGC ACAAAATGCTGACTTCTTTCCATACGGGCAGAGGGT	451	1	444	SAT_EC	Satellite
CFA	CFAcons244+246	AGACACAAACCTTCCAGGCCAGCAGCTTGCCTTGCATCTAAGACAGCTTCT GACAATCATCTAAAAGAAGACGTTAGAACTAAATCAACAGGCAGTGTGGA ACTGCCAAGGACAGCAACTGTGTCAACACTCAGAAGTATACCTTTAAGCCA CCTTACAACTATAACCAGCTCCAATGGGAGTTACAGGCTAACAAAGATTCTACA GTAGGAATTTCCATTGCAAAAAGTACTAACTCAGAAAACACAAGCCTTCC AAGCCTGAAGCTTGACACTTGCATCTAATACAGCTTATAACAATCATCTAAAA TGAGGATGTTGGAAGTAAATCCCCAGGCAGTGTGGAAGTGTGAG	357	1	338	CarSat1	Satellite
LAF	LAFcons842+936	AACAGATACTTTAAGCACAGTGAGGCCAGGGAAAACGCAGTGACCGGTAA AATCCTAGGAAAACCCATGTCTTTCCCACTTGAATGCATGTCAATTTAGCTA TTTGCCAAGAACAGAGCTCAGAGGCAACATCAGATTTAGTGCATCTCGGTGA AAGAAAAGAACACTCTGACTTCCATTCCTCTGACGAAAGCAGTTAGAAA GAGAAAACATCAGTTGGTTTCACTTCCCTGCCTGAGCAGGCAAACTGCC CATCGGAAGGGAAGGTTTCCCTAGAGAAGTGCCTTGGCGTAGCCTGTTCT GGAAGTGGAAACGAGTATGCAGTGTAGCTTAAAGCTATTGTGAAGGCTCGTT TTGCCTCGGAAAACAAAGTGCCGGTGAATCCTTACAAGCACTGGGGCGT TTCCACAGTTCAAAGTGTGGAAGTACAGACTCTTTCCGATGAAGACAGCGC AGAGACAACAGCAGCCTGCATGTTTCTCAGTGAAAGGGAGAACTCCCTCTC TTCATCCAGTCCCTGAGCAAGGGAAGTGAATGAGAGAACTTTTCTTTCC TTCTAGTTCCCTGCAATGAATTCCTCATTTGCTCTTGCAAACGCAAGACTTCT TCTGAAAAACCACATCGCAGTGGATTTATCCTTCCAGAGGAGAACTACATAC TCTTAGTTTAAGAGATTGTGAAAGCACCCAGGGAGAAAAAGGCCTTTTGTAT TCTCTTTCCTGCCTGAAGTCTGTGAATTGGCACTTGCGAAGGCAAGGTTTCC AAGCACAACACTGCAAAGCAGTAACCTCTCTCTGGCAGAGTCAATCCTATGT AGTGTTCGCGTGACGAGATAGTGTGGGTACGCTGAGGCCAGGGAAAACA CACTGCGCAGAGAAATCCTAGCAAAGCCGAGGGCTTCCCCACAGGAACCCA TGTCCGTTTCAGTCTTTGCCGAGAACGGAGCTCGAGGCAACATCAGATTG TGCGTATCTCAGGGAGAGAAAAGAACTGTGACTTCCCCGTTCTCTGACG AAAGCAGTTAGAAGGAGAAAAGCGTCTTTGCTTTCCGTTTCCCGCAGGGAG	1220			No repeat detected	No repeat detected
DNO	DNAcons173	TATAGAAAAGCATAACGGCAGGTATAGCAGCTAACTCTTTAAACCTAAATGA AACAAATGGCACATATGATTTACTTGTAGAACTAAGCATTGTGTGCAGAGA TTTTGCAGCTATCTCTTCTTCCAGTTTTTGTGCTAACATTCCATCTGTAAA ACTAGCATTACAC	173			No repeat detected	No repeat detected

MDO	MDOcons528	GATCGTCGTAAGGACGCGTGGCGTAGCGGTCTCCCCGTAGCGAGGATTTTT CTCCCCGAAAGCCAGCCGTCTGAAGTACGCGTGGAGTAGAGGTCTCCATT TCTGATCCTGGCGCGTTCTCACATCGAAATGGGGGAGTGTTCACAGTAGGG CATTGGTTCCCAGGGAGGGATTGATTCCTCGATGTGGACATTTTCACATTCA CAAGTCTGAGCAATTTAGATTTACAATCCCCTTGAGGGATCATTGGGAGA CCAGTCTCCCCACTGGTCATTTACATGATCAATTTGTAATCCTCAATGCCCT TCTAACTCGAGAGAGACACGCTCTTCCACTTTCGAAGAGGAATGAGAACCG TGAGGGAGCAAATCGGCCAGAAAAGAAAGCCAAACCAATGTTTTGCTGGGC GCATCGCGAGAAAGCCAAATTCGCGACAGTCCCTTTTGGCCTAAATGTGTA CATTTACTCCACACCGGAAGACGCTTTCATTGCGTAGCAAGGAGTTTTCT CCCCTAT	528	5	235	ERV37_MD_LTR	LTR/ERV1
				236	477	ERV37_MD_I-int	LTR/ERV1
				495	527	ERV37_MD_LTR	LTR/ERV1
OAN	OANcons144.1+144.2	GTAAACCTCTGCCCCGCCACCCTTTGCTCCCGCCTGATCTTTCTCC CGGCTAAACCTCTGCTCCCGGCTAAACCTCTGCTCCCGGCCGCCCTTTGCT CCCGCCTGATCTTTCTCCCGGCTAAACCTCTGCTCCCGCTAAACCTCTGC TCCCGCCTG	168	6	168	MonoRep114	Satellite

Table S3. FISH results with specific CFA centromeric fosmids (in parenthesis the number of chromosomes according to Breen et al. 1999 (Chromosome Research 7: 401;406, 1999))

CLONE NAME	FISH RESULTS
G630P88580D11 BI1-3897G21	chr 37
G630P816490E11 BI1-411I22	all except chr 1 (strong signals on chrs: 29; 32; 34; 35; 36; 37; 38)
G630P89020G11 BI1-3524M21	all except chr 1 (strong signals on chrs: 21, 31, 36, 37)
G630P84226G11 BI1-1356M22	all except chr 1 (strong signals on chrs: 21, 31, 34, 36)
G630P88303G10 BI1-3749N20	chrs 36; 37; 38
G630P8903H12 BI1-1850P24	chr 36; 37; 38
G630P8139H12 BI1-1903P24	all except chr 1 (strong signals on chrs: 21, 31, 34, 35, 36, 37)
G630P8317C12 BI1-1922F23	chr1 centromeric heterochromatic block+ others heterochromatic regions

Table S4. ICHC results using CENPB/CENPC antibodies combination

SPECIES	RESULTS
ECA	all chromosomes except ECA 1;3;6;7;9;11;12;13;15;X
CFA	no results
LAF	no results
DNO	no results
MDO	MDO 2, 6
OAN	OAN 3, 4, 12, 14, 18, X5