Supporting Information Appendix

Multiple independent introductions of *Plasmodium falciparum* in South America

Yalcindag *et al.*

Index

SI Material and Methods

Microsatellite and SNP Genotyping. DNA from blood samples was extracted using the DNeasy Blood and Tissue kit (Qiagen, France) according to the manufacturer's recommendations, and eluted in 100 µl of elution buffer per 200 µl of whole blood or per filter blot. The extracted DNA was used as template for the PCR-based amplification of 12 polymorphic microsatellites distributed among eight of the 14 chromosomes of *P. falciparum* (listed in Table S2), as well as SNPs from coding and non-coding regions (Table S3). For microsatellite amplification, a two-step semi-nested PCR strategy with fluorescent end-labeled primers was used as in (1), following the methodology in (2). The amplified microsatellite repeats were resolved and sized relative to an internal size standard using a Genetic Analyzer 3130xl and the GeneMapper software. For SNPs, genotyping was performed by the Integragen Company (Evry, France) using the Illumina Golden Gate technology. We obtained interpretable genotypes for 354 of the 384 SNPs initially defined. To minimize missing data while maximizing information, we retained only the isolates that were genotyped for at least 80% of the 354 SNPs and when SNPs were successfully typed in at least 90% of the isolates. The resulting sample consisted of a dataset of 210 isolates genotyped at 272 SNPs (Table S3 for a list of SNPs) from 17 populations in 13 countries (Table S1).

Markers under natural selection. Natural selection, either divergent or convergent, may impact genetic frequencies differently in different populations. We searched for loci (microsatellites and SNPs) that could be under positive or balancing selection, using the method developed in (3) and implemented in the program LOSITAN (4). Outlier values of genetic differentiation at specific loci were detected by running coalescent simulations to generate the expected distribution of Wright's inbreeding coefficient F_{ST} as a function of heterozygosity (*He*). The distribution of F_{ST} was obtained by simulating an island model with a distribution centered on the empirical estimates averaged over loci. This average F_{ST} cannot be assumed to be neutral because (initially unknown) selected loci may be included in the computation. We first ran LOSITAN to determine a subset of candidate selected loci and then removed them for the computation of the neutral F_{ST}. The value obtained is likely to be a better approximation of the neutral F_{ST} (3). The approach is expected to be robust with respect to variation in mutation rate among loci, sample size, and departure from mutation/drift equilibrium (3), but see (5). In any case, we tested for selection on

a reduced dataset (microsatellites or SNPs) that included only the African and Asian populations so as to limit as much as possible departures from the assumptions of the test.

Statistical, Population Genetic, and Phylogenetic Data Analyses. Statistical calculations were performed with the R software (6), unless stated otherwise.

Population and individual-based genetic trees. We constructed Neighbor-joining trees (from microsatellite and SNP data) to examine the relationships between South American populations and those from the rest of the world. The trees were constructed using Cavalli-Sforza and Edwards' chord distance (7) computed between each pair of populations. This distance is typically used to reconstruct trees from population data (e.g., ref.(8)). All analyses were run using the software R-package APE (9). Matrices of pairwise shared allelic distances computed between isolates were also used to analyze the genetic relationship between SNP haplotypes (10).

Principal Component Analyses. Principal Component Analyses (PCA) were used to gain additional insights on the genetic structure of the studied isolates. The PCA were performed on the matrix of binary allele profiles using the adegenet R-package (11). Missing genotypes were replaced by the mean of the allelic frequency observed over all populations.

STRUCTURE analyses. We used the Bayesian clustering method implemented in STRUCTURE v.2.1 (12) to identify population structure. We ran models allowing for admixture, with the number of clusters K ranging from 1 to the number of populations (24, 33 or 15, depending on the dataset). All simulations used 100,000 Markov-chain Monte Carlo (MCMC) generations in the burn-in phase and 100,000 generations in the data collection phase. Ten independent runs were performed for each specified *K* to verify convergence in the estimates of posterior probabilities. The optimal number of clusters was estimated using the method proposed in (13) .

Approximate Bayesian Computations (ABC). For the microsatellite dataset, we used the Approximate Bayesian Computation (ABC) method as implemented in the package "DIY ABC" (14) to trace the introduction history of *P. falciparum* in South America. Three scenarios were considered and tested against each other. They differed in the order of introduction events or the presence of an unsampled population (Fig. S13).

In the first scenario, populations in South America were independently introduced from Africa ("Independent Introduction Scenario"). In the second scenario ("Serial Introduction Scenario"), only one South American population originated from the African source population, while a second South American population originated from the first. The third scenario ("Unsampled Population Scenario") assumed that the two introduced populations derived from an unsampled population that itself derived from the African source. In all scenarios considered, the divergence time of population 1 is constrained to be older than the divergence time of population 2.

We used a subset of the West African populations (S) as the source population. We chose these populations because the transatlantic slave trade was preying primarily on populations along the west coast of Africa (Fig. 1). We created one large pooled source population because the *P. falciparum* populations located along the African coast are genetically very little differentiated (F_{ST} =0.009 for "MS" and F_{ST} = 0.01 for MS+). For the introduced populations (Pop1 and Pop2), we successively took all possible pairs of South American populations (Camopi, Trois Sauts, Maripasoula, Colombia, Bolivia and Brazil) but excluded Venezuela and Peru, which displayed a clear pattern of recent admixture. We also created a "southern cluster" comprising the three French Guianan, Brazilian and Bolivian populations, and a "northern cluster" comprising the two Colombian populations.

Analyses were performed using the demographic parameter prior distributions described in Table S7. We used a generalized stepwise model to simulate mutation at the markers of interest (i.e. microsatellites). In order to determine which scenario was most likely, we simulated each scenario one million times, and estimated for each scenario several summary statistics (mean number of alleles, mean gene diversity, F_{ST} and shared allelic distance between pairs of populations). From these, we assessed the posterior probability of each scenario using a polychotomous weighted logistic regression. The posterior probabilities for divergence times were also retrieved.

SI References

1. Razakandrainibe FG, et al. (2005) "Clonal" population structure of the malaria agent *Plasmodium falciparum* in high-infection regions. *Proc Natl Acad Sci USA* 102:17388-17393.
2. Anderson TJC, Su ZX, Bockarie M, Lagog M, Day KP (1999) Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology* 119**:**113-125.
3. Beaumont MA, Nichols RA (1996) Evaluation loci for the use in the genetic analysis of population structure. *Proc Roy Soc Lond B* 263:1619-1626.
4. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics* 9:323.
5. Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity* 103:285-298.
6. R Development Core Team (2009), *Computing*. R Foundation for Statistical Computing, Vienna, Austria.
7. Cavalli-Sforza LL, Edwards AW (1967) Phylogenetic analysis. Models and estimation procedures. *Am J Hum Genet* 19**:**233-257.
8. Bowcock AM, et al. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368:455-457.
9. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20:289-290.
10. Jin L, Chakraborty R (1993) Estimation of Genetic Distance and Coefficient of Gene Diversity from Single-Probe Multilocus DNA Fingerprinting Data. *Mol Biol Evol* 11:120-127.
11. Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403-1405.
12. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
13. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.

14. Cornuet JM, et al. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24:2713-2719.
15. Anderson TJC, et al. (2000) Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular Biology and Evolution* 17:1467-1482.
16. Aubouy A, et al. (2007) Dramatically decreased therapeutic efficacy of chloroquine and sulfadoxine-pyrimethamine, but not mefloquine, in southern Benin. *Tropical Medicine & International Health* 12:886-894.
17. Bogreau H, et al. (2006) Genetic diversity and structure of African *Plasmodium falciparum* populations in urban and rural areas. *Am J Trop Med Hyg* 74:953-959.
18. Andriantsoanirina V, et al. (2009) *Plasmodium falciparum* drug resistance in Madagascar: facing the spread of unusual pfdhfr and pfmdr-1 haplotypes and the decrease of dihydroartemisinin susceptibility. *Antimicrob Agents Chemother* 53:4588-4597.
19. Pumpaibool T, et al. (2009) Genetic diversity and population structure of *Plasmodium falciparum* in Thailand, a low transmission country. *Malaria Journal* 8**:**96.
20. Nair S, et al. (2003) A selective sweep driven by pyrimethamine treatment in southeast Asian malaria parasites. *Molecular Biology and Evolution* 20:1526-1536.
21. Legrand E, Volney B, Meynard JB, Mercereau-Puijalon O, Esterre P (2008) In vitro monitoring of *Plasmodium falciparum* drug resistance in French Guiana: a synopsis of continuous assessment from 1994 to 2005. *Antimicrobial Agents and Chemotherapy* 52:288-298.
22. Restrepo E, Carmona-Fonseca J, Maestre A. (2008) *Plasmodium falciparum*: high frequency of pfcrt point mutations and emergence of new mutant haplotypes in Colombia. *Biomedica* 28**:**523-530.
23. Grande T, et al. (2007) A Randomised Controlled Trial to Assess the Efficacy of Dihydroartemisinin-Piperaquine for the Treatment of Uncomplicated Falciparum Malaria in Peru. *Plos One* 2(10):e1101.
24. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38**:**1358-1370.

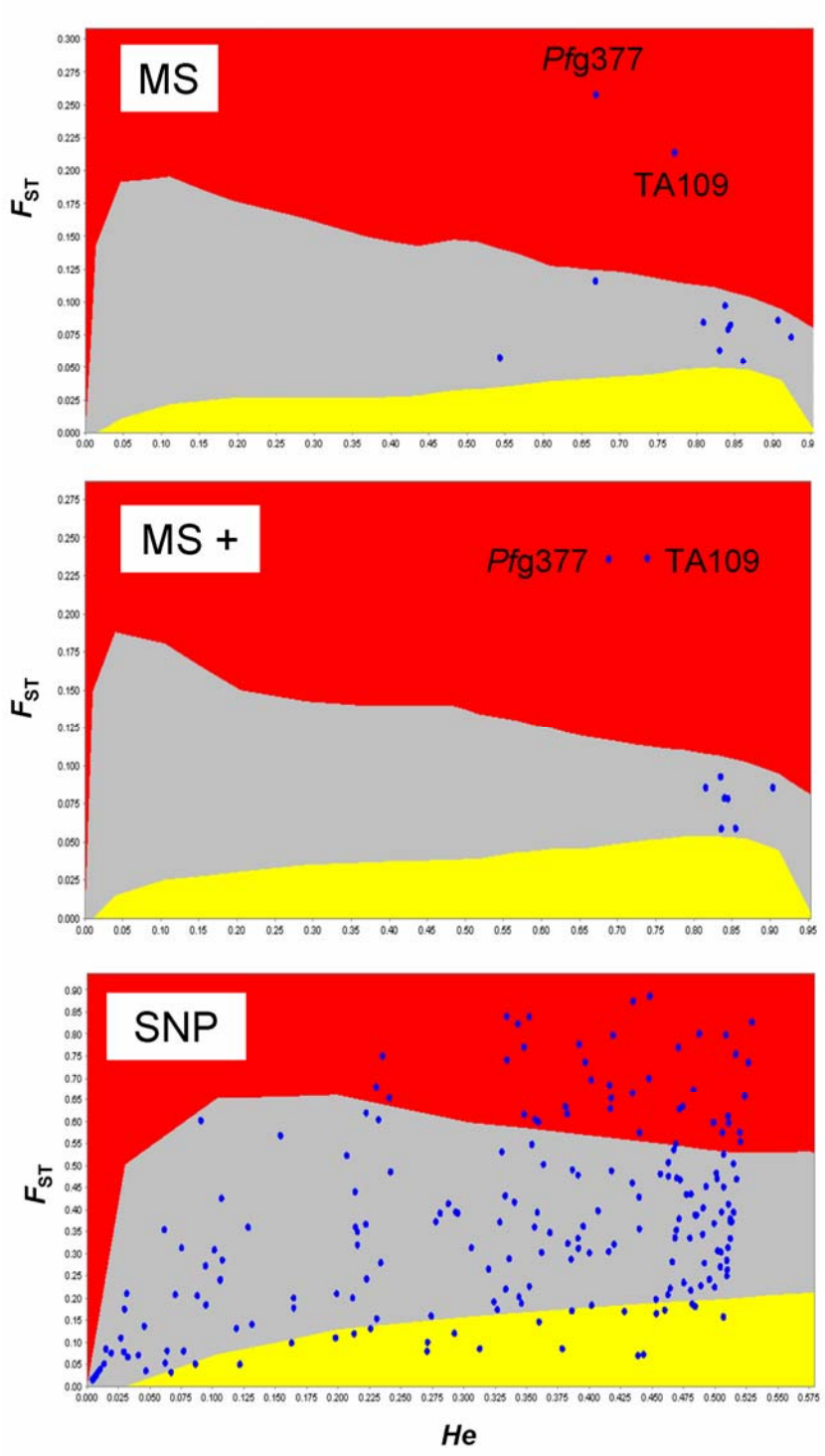

Fig. S1. F_{ST} values plotted against heterozygosity (*He*) and computed using **a**, the MS dataset; **b**, the MS+ dataset; and **c**, the "SNP" dataset (see Materials and Methods and Table S1 for details). For MS and MS+, the microsatellite markers suspected to be under selection (*Pfg*377 and TA109) are displayed. Grey: area including 95% of the neutral F_{ST} computed using an island model. Red: area including the highest 5% neutral F_{ST}. Yellow: area including the lowest 5% neutral F_{ST}. Detailed information is given in Tables S2 (microsatellites) and S3 (SNPs).

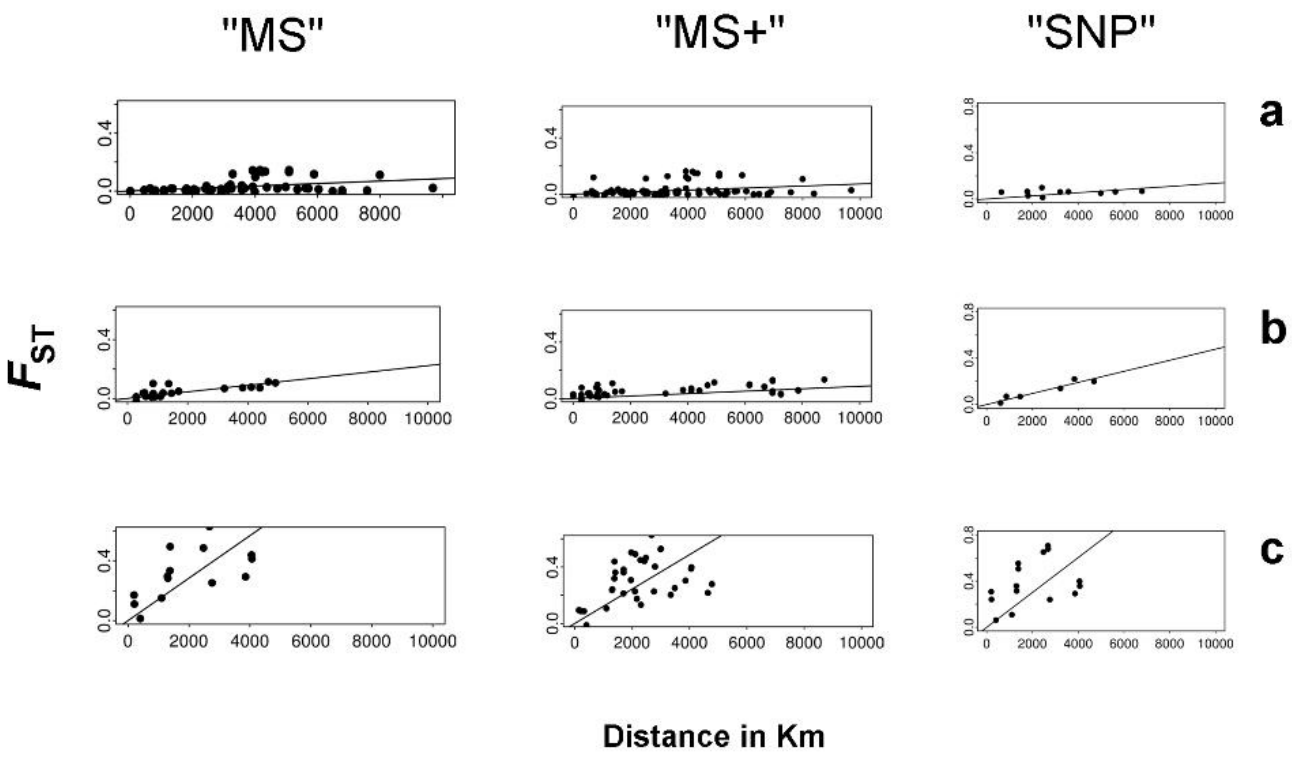

Fig. S2. Isolation by distance (IBD) within each continent. Pairwise geographic distances along landmasses plotted against pairwise genetic differentiation (F_{ST}) computed for each dataset: MS, MS+ and SNPs. Loci suspected to be under selection were removed. **a,** African populations; **b,** Asian Populations; **c,** South American populations. All Mantel tests gave *P-values* <0.05 except for Africa (all datasets) and South America (SNP dataset).

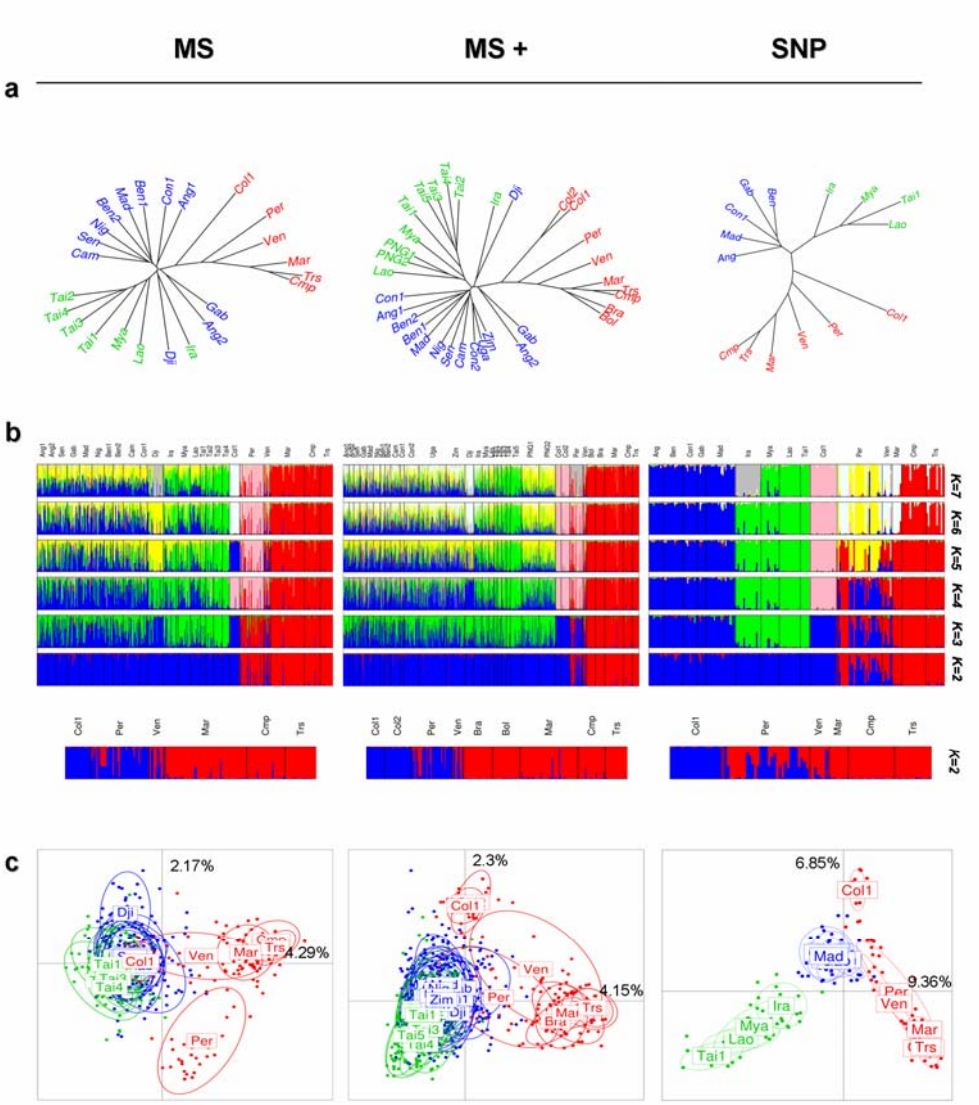

Fig. S3. Genetic relationship between South American populations and those from Africa and Asia, based on the microsatellite (MS and MS +) and SNP datasets. In all datasets, loci under selection were removed. **a**. Neighbor-joining trees of populations. **b.** Worldwide and South American population structure inferred by Bayesian clustering. Each isolate is shown as a thin vertical line partitioned into *K* colored components representing inferred membership in *K* genetic clusters. *K* varies from 2 to 7 for the worldwide analysis (upper graph). Bottom row provides the inferred population structure in South America for $K = 2$. **c.** First and second principal components of the Principal Component Analysis (PCA). Dots represent isolates. Ellipses represent 95% of the genetic variation within populations. Percentages of inertia are displayed directly along the respective axes (first axis: horizontal; second axis: vertical). The code name of each population is at the centroid of the ellipse. For trees and PCA: blue, African populations; green, Asian populations; red, South American populations. Ang1 and Ang2 : Angola, Ben1 and Ben2: Benin, Bra: Brazil, Bol: Bolivia, Cam: Cameroon, Cmp: Camopi (French Guiana), Col1 and Col2: Colombia, Con1: Republic of the Congo, Con2: Democratic Republic of the Congo, Dji: Djibouti, Gab: Gabon, Ira: Iran, Lao: Laos, Mad: Madagascar, Mar: Maripassoula (French Guiana), Mya: Myanmar, Nig: Niger, Per: Peru, PNG1 and PNG2: Papua New Guinea, Sen: Senegal, Tai1-Tai5: Thailand, Trs: Trois Sauts (French Guiana), Uga: Uganda, Ven: Venezuela, Zim: Zimbabwe. Details about the sampled populations are reported in Table S1.

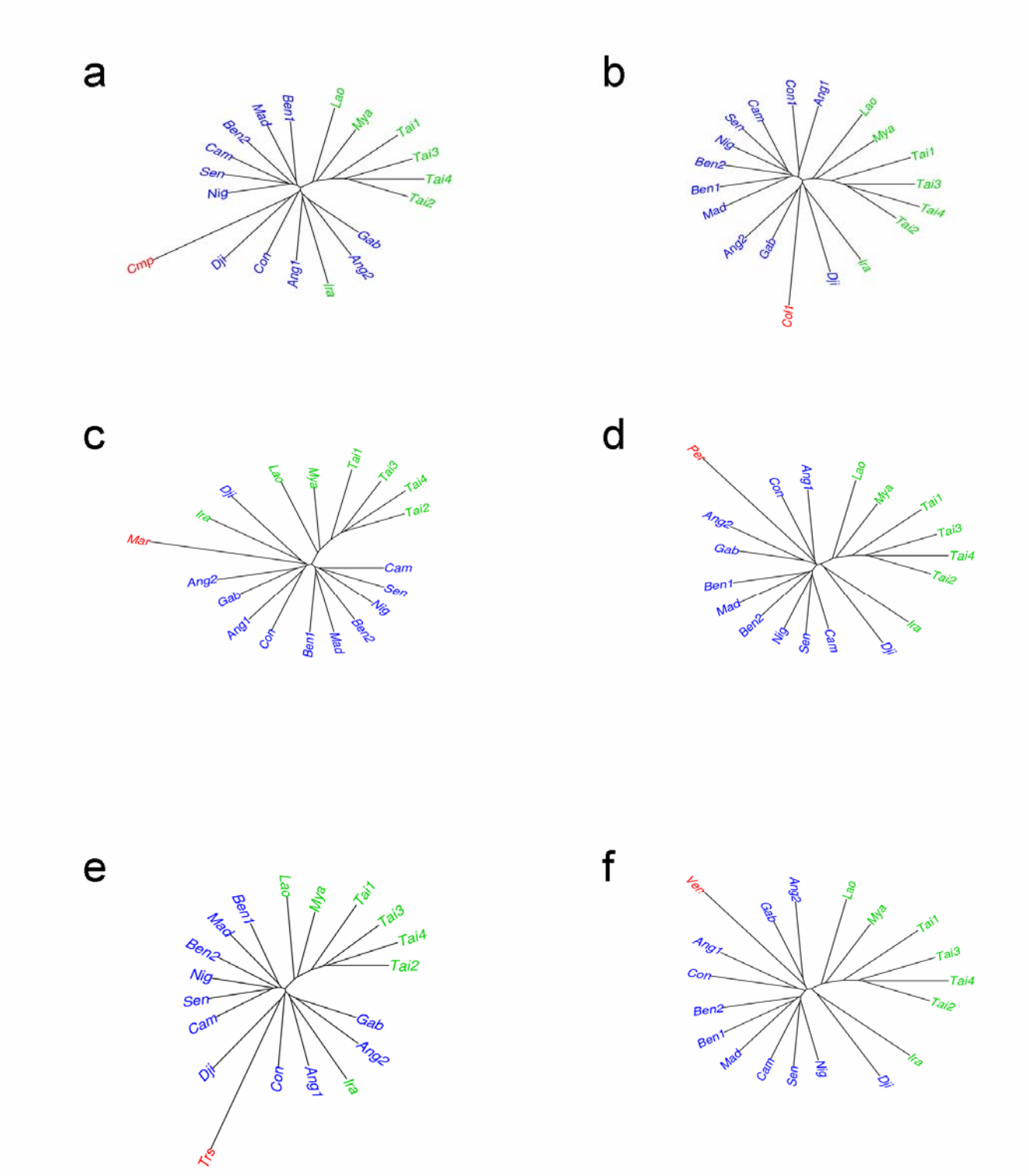

Fig. S4. Neighbor-joining trees of populations obtained by including the South American populations one by one (MS dataset excluding loci suspected to be under selection). **a,** Camopi (Cmp: French Guiana)**. b,** Colombia (Col1). **c,** Maripasoula (Mar: French Guiana). **d,** Peru (Per). **e,** Trois Sauts (Trs: French Guiana). **f,** Venezuela (Ven). Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

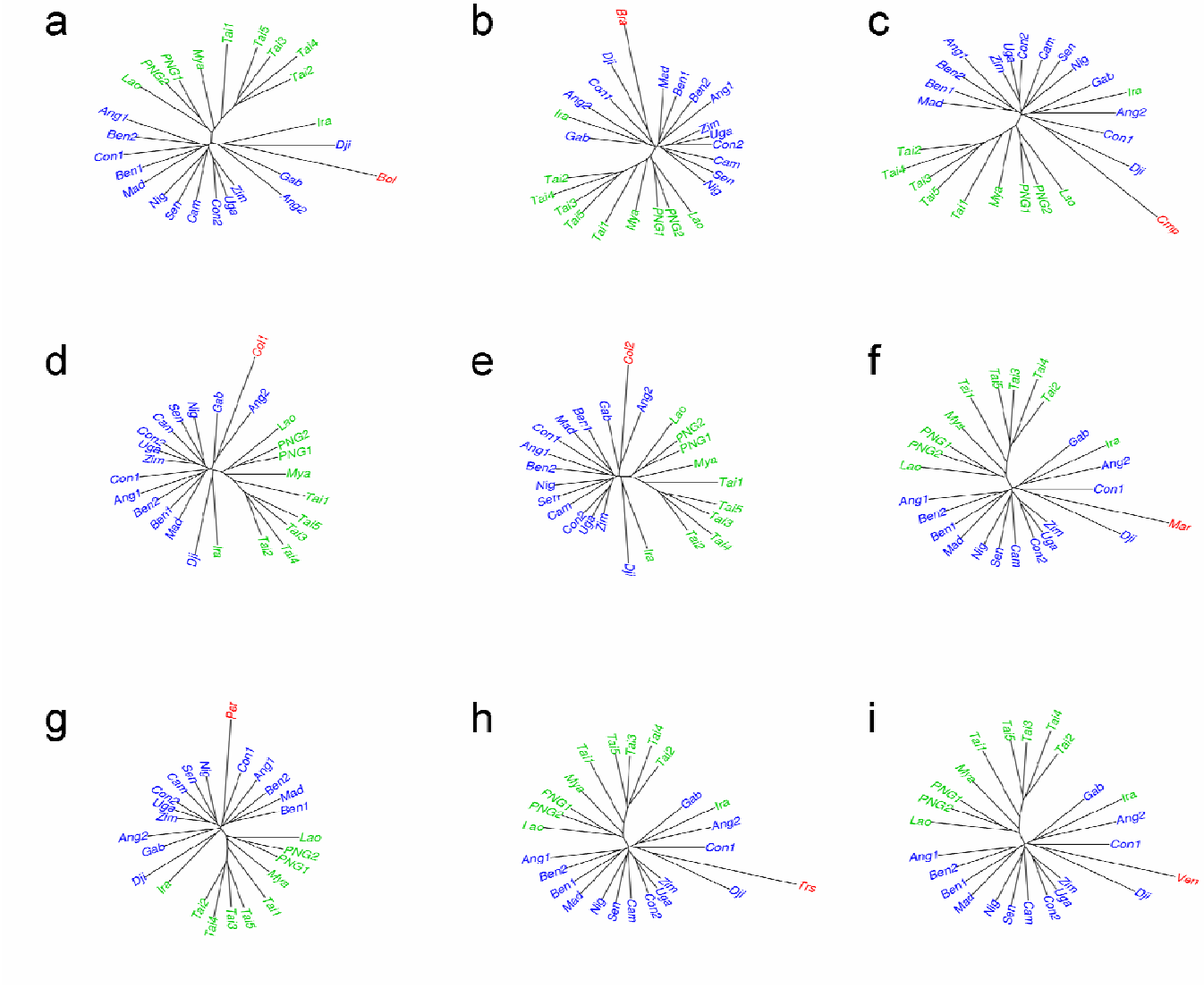

Fig. S5. Neighbor-joining trees of populations obtained by including the South American populations one by one (MS+ dataset including the data from ref. 1 and excluding the loci suspected to be under selection). **a,** Bolivia (Bol). **b,** Brazil (Bra). **c,** Camopi (Cmp: French Guiana). **d,** Colombia 1 (Col1). **e,** Colombia (Col2). **f,** Maripasoula (Mar: French Guiana). **g,** Peru (Per). **h,** Trois Sauts (Trs: French Guiana). **I,** Venezuela (Ven). Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

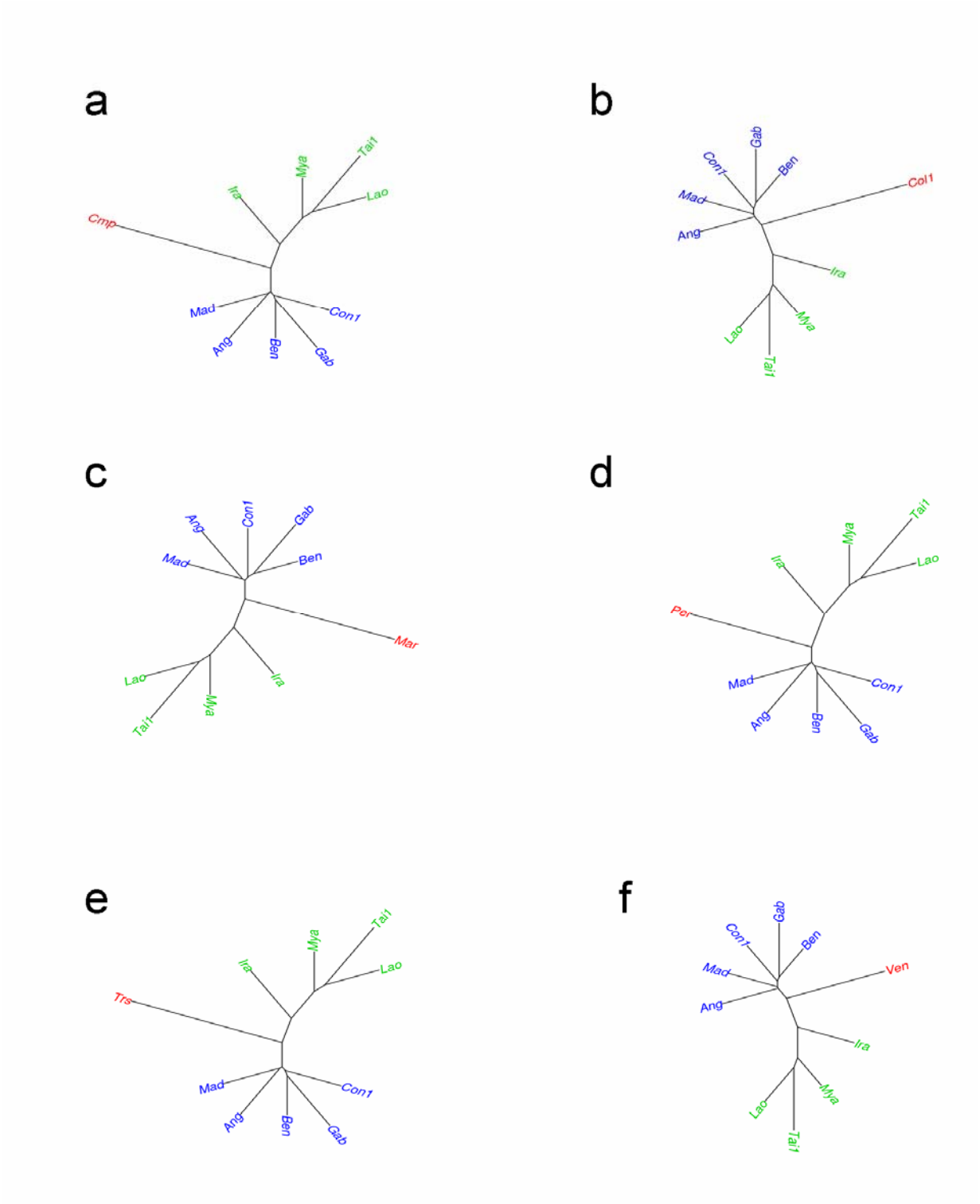

Fig. S6. Neighbor-joining trees of populations obtained by including the South American populations one by one (SNP dataset excluding loci suspected to be under selection). **a,** Camopi (Cmp: French Guiana)**. b,** Colombia (Col1). **c,** Maripasoula (Mar: French Guiana). **d,** Peru (Per). **e,** Trois Sauts (Trs: French Guiana). **f,** Venezuela (Ven). Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

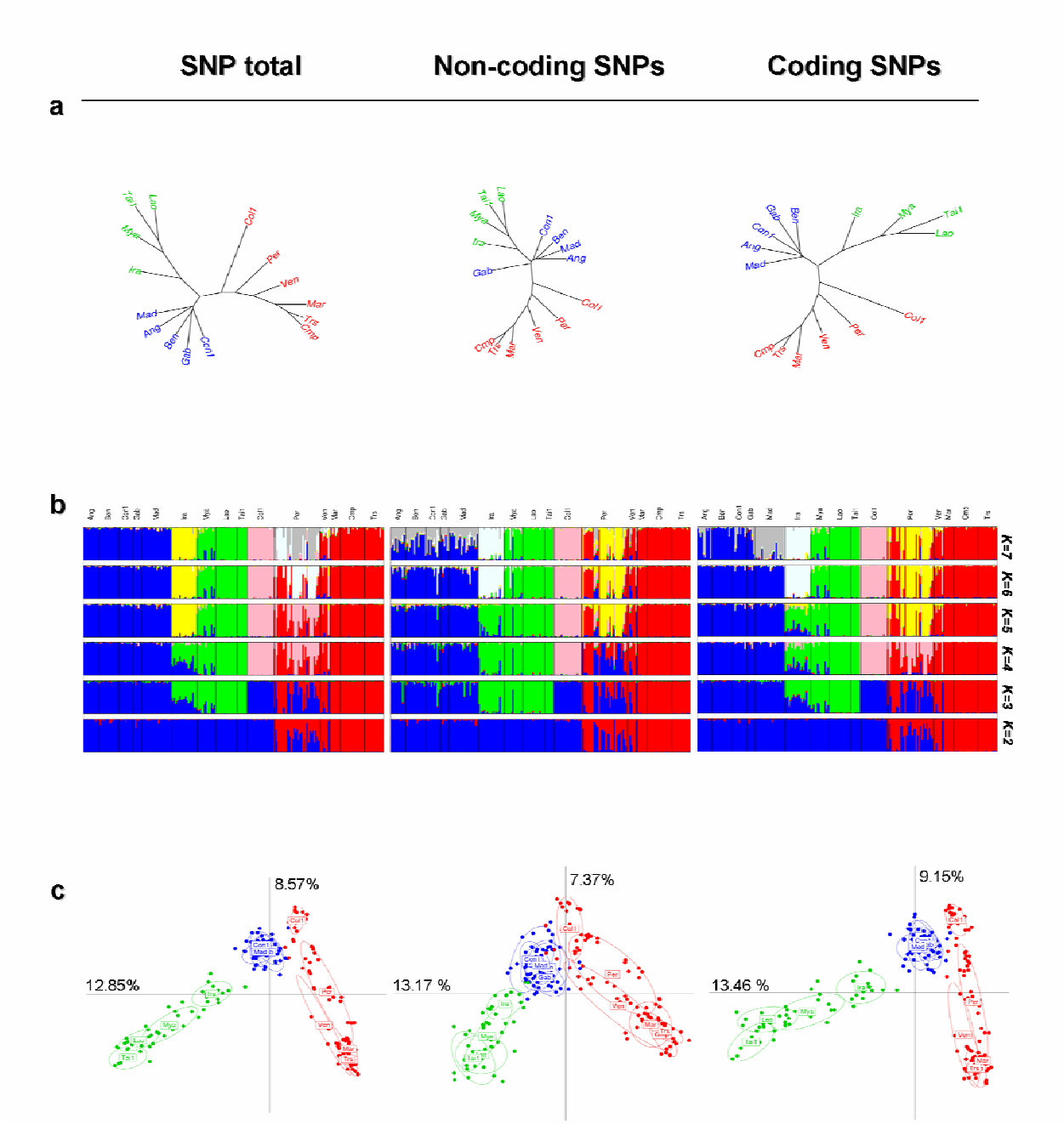

Fig. S7. Genetic relationship between South American populations vs. Africa and Asia based on all SNPs (SNP total), only non-coding SNPs, or only coding-SNPs. All loci are herein considered, even those suspected to be under selection. **a,** Neighbor-joining trees of populations. **b,** Population structure inferred by Bayesian clustering. Each individual is shown as a thin vertical line partitioned into *K* colored components representing inferred membership in *K* genetic clusters (*K* varying from 2 to 7). **c,** Projection of the Principal Component Analysis (PCA) on axes 1 (horizontal) and 2 (vertical). Dots represent individuals. Ellipses are approximate 95% confidence limits for populations. The name of each population is located at the centroid of the ellipse. Percentages of inertia are displayed directly along the respective axes (first axis: horizontal; second axis: vertical). For trees and PCA, Blue: African populations; Green: Asian Populations, Red: South American populations. For abbreviations, please refer to Table S1.

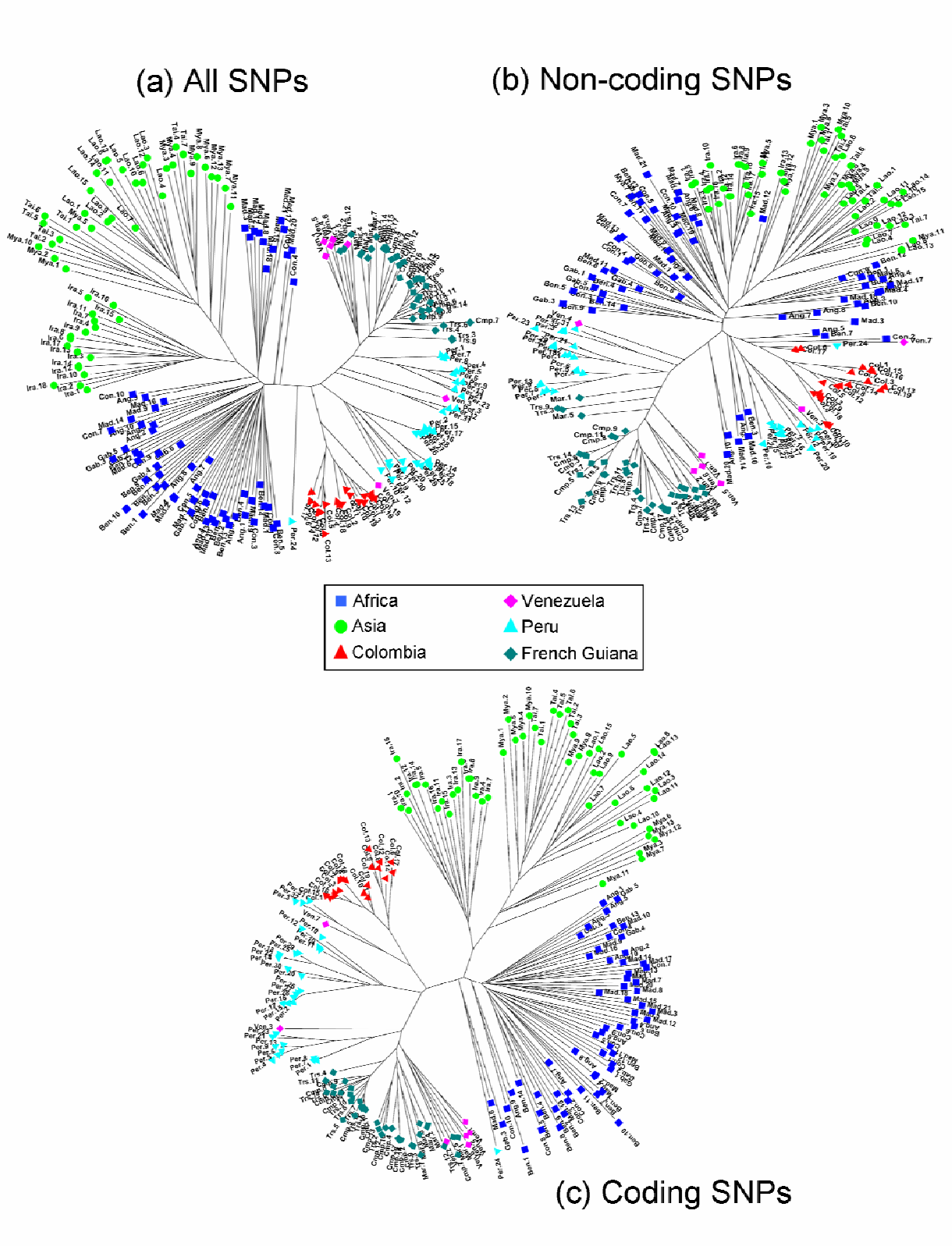

Fig. S8. Neighbor-joining trees of individuals, using all SNPs (**a**); only non-coding SNPs (**b**); or only coding SNPs (**c**). All loci are considered, including those suspected to be under selection.

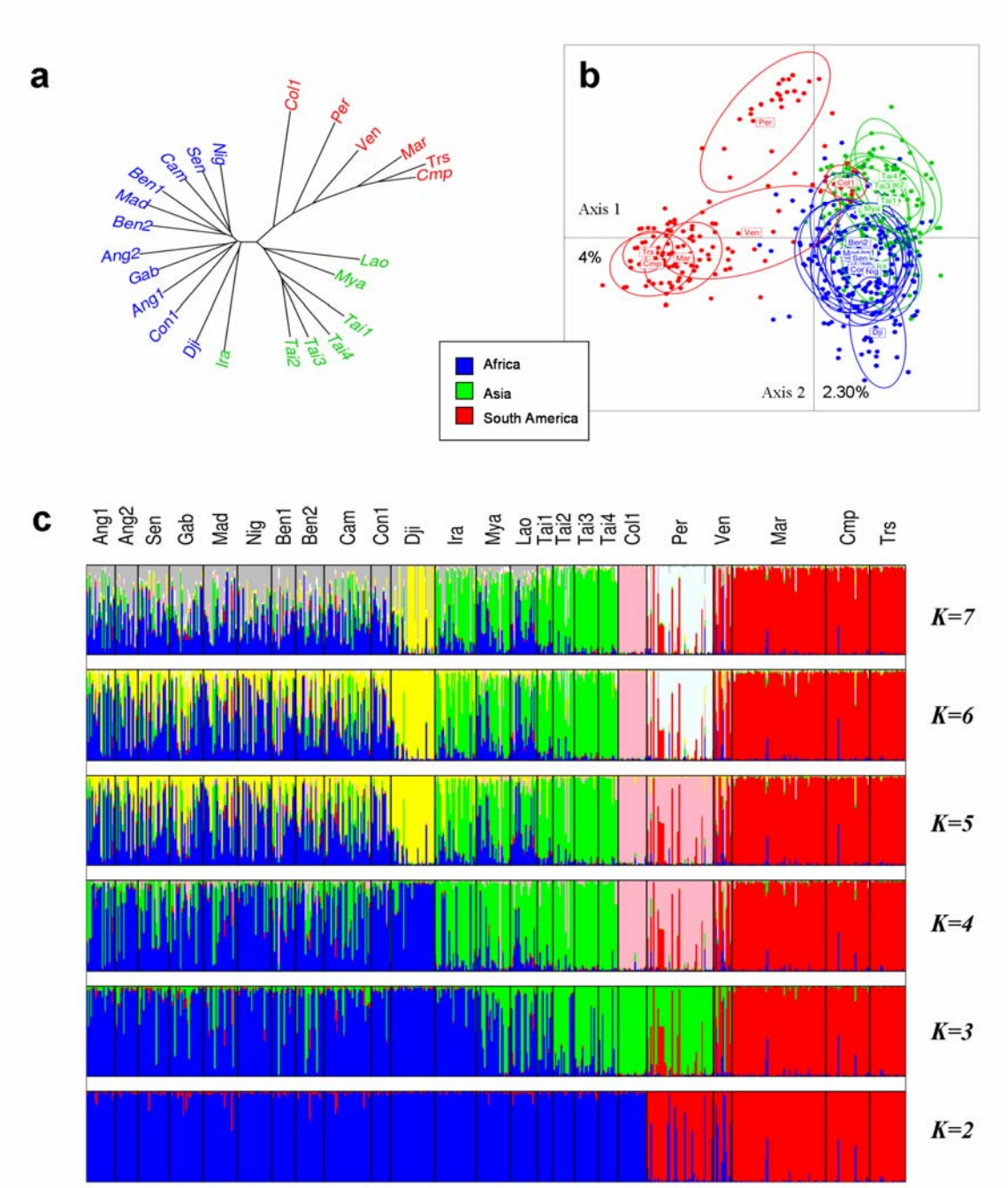

Fig. S9. Genetic relationship between South American populations vs. African and Asian populations, based on all microsatellite markers (MS dataset), including loci suspected of being under selection. **a,** Neighbor-joining trees of populations. **b,** Projection of the Principal Component Analysis (PCA) on axes 1 and 2. Dots represent individuals. Ellipses are approximate 95% confidence limits for populations. The name of each population is located at the centroid of the ellipse. Percentages of inertia are displayed directly along the respective axes (first axis: horizontal; second axis: vertical). **c,** Population structure inferred by Bayesian clustering. Each individual is shown as a thin vertical line partitioned into *K* colored components representing inferred membership in *K* genetic clusters (*K* varying from 2 to 7). For trees and PCA, Blue: African populations; Green: Asian Populations, Red: South American populations. For abbreviations, please refer to Table S1.

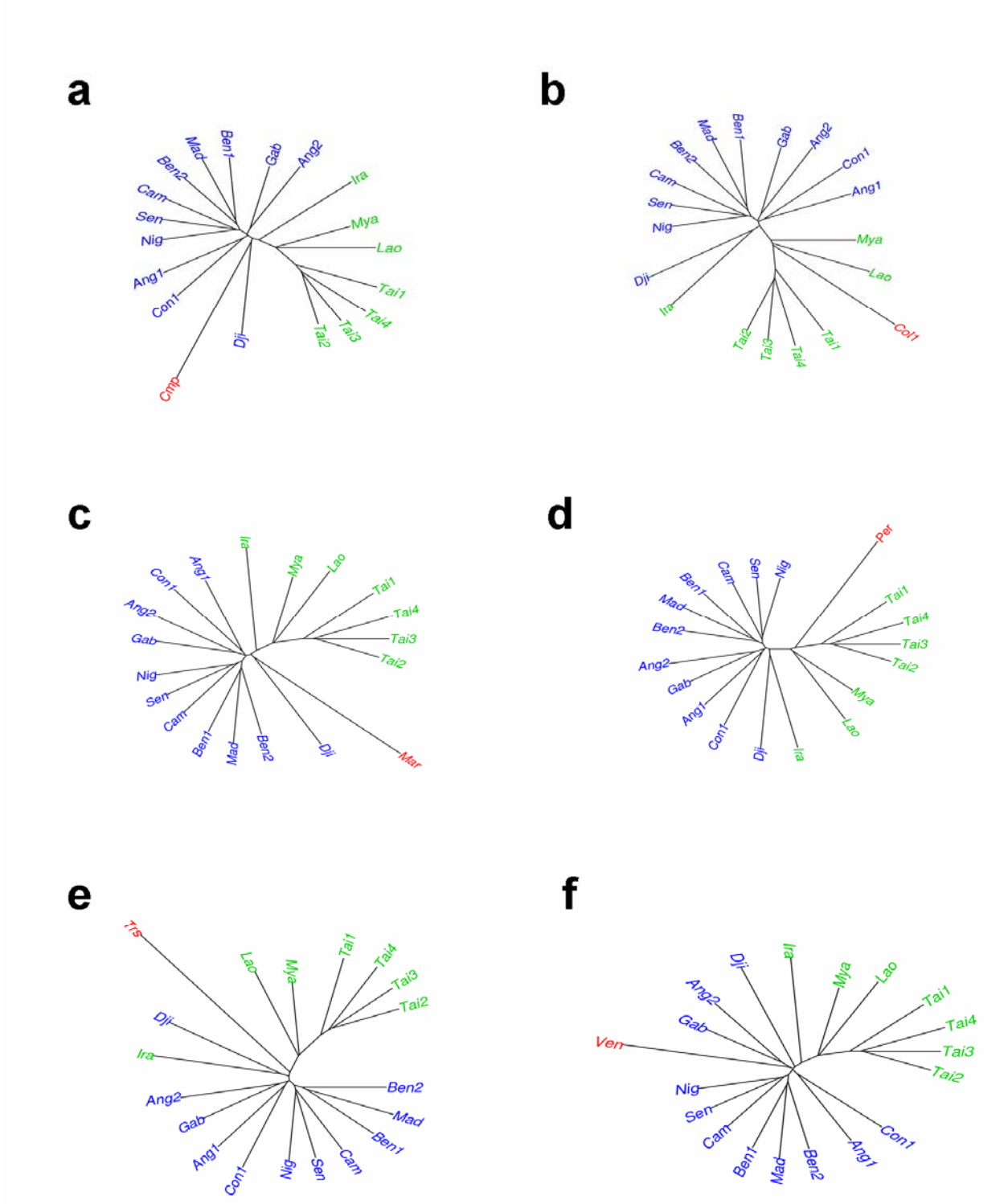

Fig. S10. Neighbor-joining trees of populations obtained by adding the South American populations one by one for the MS dataset including all loci even those suspected of being under selection. **a,** Camopi (Cmp: French Guiana)**. b,** Colombia (Col1). **c,** Maripasoula (Mar: French Guiana). **d,** Peru (Per)**. e,** Trois Sauts (Trs: French Guiana). **f,** Venezuela (Ven). Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

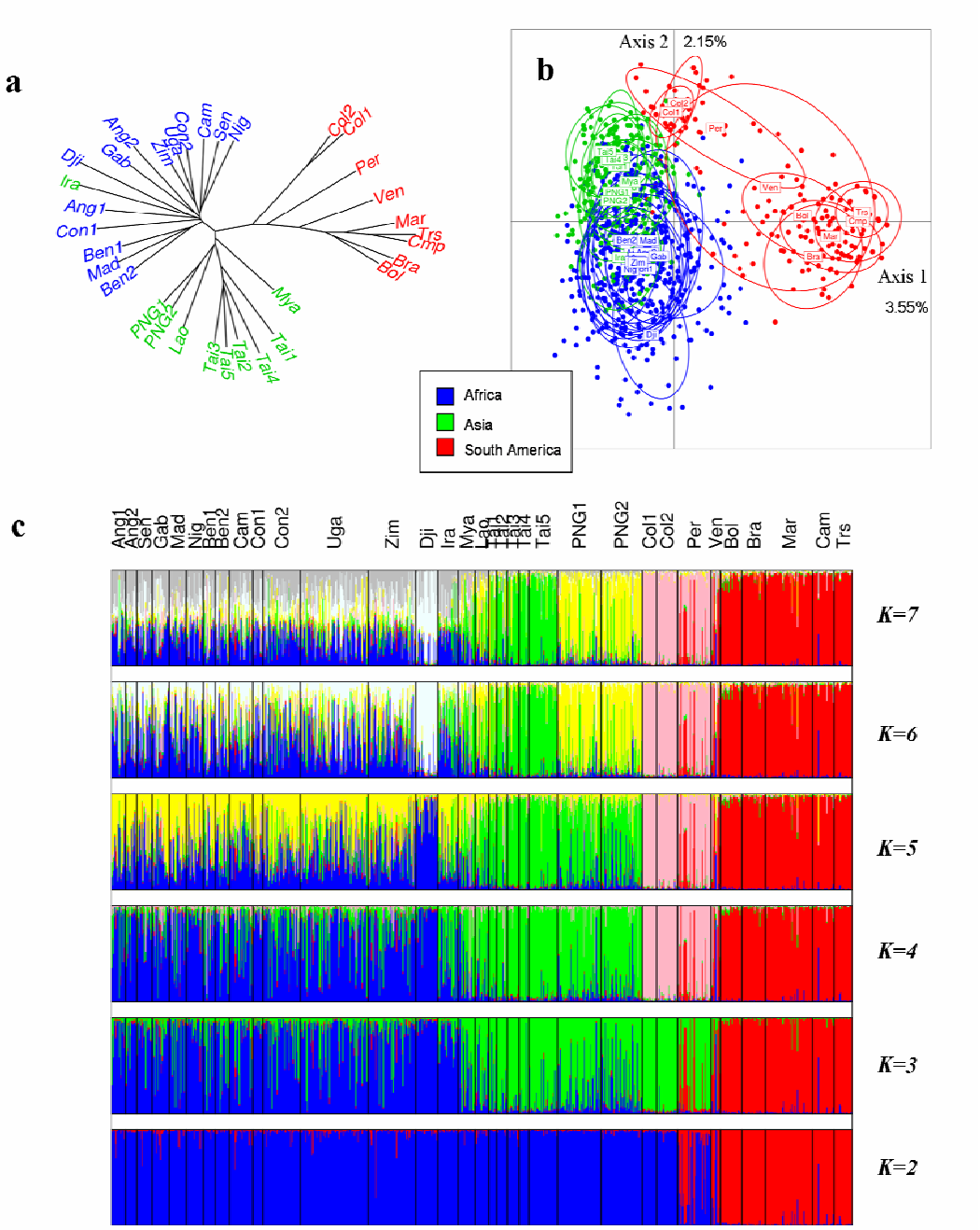

Fig. S11. Genetic relationship between South American populations vs. African and Asian populations using the MS + dataset including the data from (1) and considering all loci even those suspected to be under selection. **a,** Neighbor-joining trees of population relationships. **b,** Projection of the Principal Component Analysis (PCA) on axis 1 and 2. Dots represent individuals. Ellipses are approximate 90% confidence limits for populations. The name of each population is located at the barycentre of the ellipse. Percentages of inertia are displayed directly along the respective axes (first axis: horizontal; second axis: vertical). **c,** Population structure inferred by Bayesian clustering. Each individual is shown as a thin vertical line partitioned into *K* colored components representing inferred membership in *K* genetic clusters (*K* varying from 2 to 7). For trees and PCA, Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

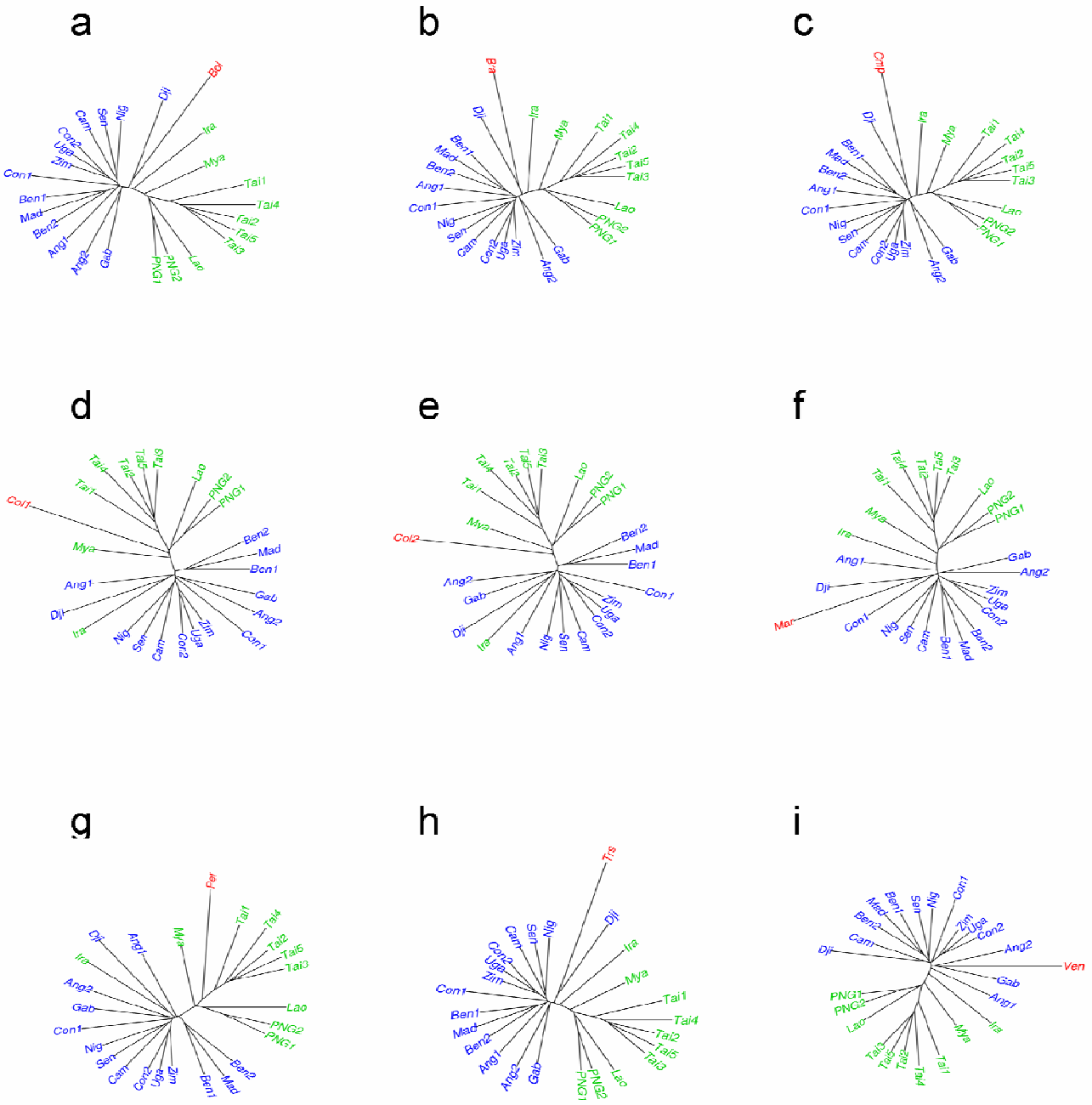

Fig. S12. Neighbor-joining trees of population relationships obtained by adding the South American populations one by one, using the MS+ dataset including the data from (1) and considering all loci even those suspected to be under selection. **a,** Bolivia (Bol). **b,** Brazil (Bra). **c,** Camopi (Cmp: French Guiana). **d,** Colombia 1 (Col). **e,** Colombia 2 (Col2). **f,** Maripasoula (Mar: French Guiana). **g,** Peru (Per). **h,** Trois Sauts (Trs: French Guiana). **I,** Venezuela (Ven). Blue: African populations; Green: Asian Populations, Red: South American populations. All other abbreviations as in Table S1.

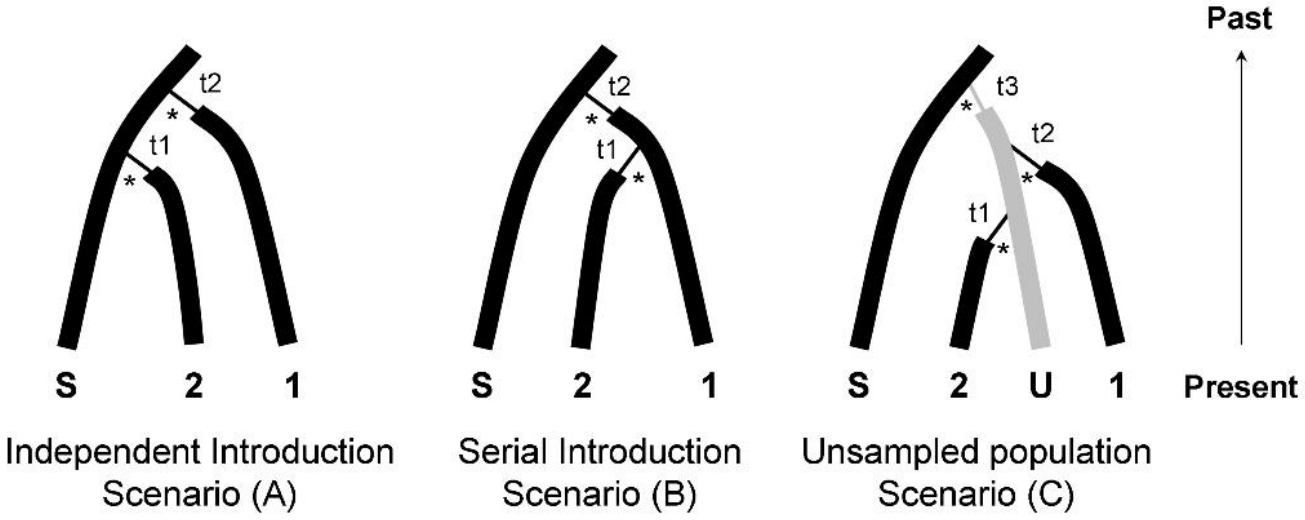

Fig. S13. Three scenarios tested in the ABC analysis. Table S7 gives the prior distribution for each of the demographic parameters included in the analysis. S: Source: 1: Population 1: 2: Population 2: U: Unsampled population: t3, t2, t1: Divergence times. *: period of bottleneck.

Table S1. Characteristics of the *Plasmodium falciparum* samples with geographical coordinates, total sample size, and number of mono-infected isolates. The abbreviations shown here are those used in the figures. *****, populations with multi-infected samples removed before our study. **#,** populations from (15) included in the MS+ dataset. §, from populations included in the SNP dataset. For mono-infected samples, the number between parentheses indicates the number of isolates genotyped at 272 SNPs. All isolates, except for the populations from (15) were genotyped by us.

Country	Site	Short Name	Latitude	Longitude	Sample size	Mono-infected sample size	Years of collection	Reference
AFRICA								
Angola	Chisseque	Ang1§	12° 42' 00" S	14° 10' 00" E	47	20 (2)	2008	Present study
	Libata	Ang2§	15° 39' 58" S	17° 26' 29" E	61	16 (8)	2008	Present study
Benin	Kpasse (Ouidah)	Ben1§	06° 22' 00" N	02° 05' 00'' E	42	17 (4)	2005	(16)
	Topka Dome (Ouidah)	Ben2§	06° 30' 00'' N	02° 00' 05' E	53	20 (10)	2005	(16)
Cameroon	Simbock/Tibati	Cam	05° 15' 20'' N	11° 41' 55'' E	33*	33	2003	Present study
Republic of the Congo	Brazzaville, Gamboma, Pointe Noire, Oyo	Con1§	00° 13' 44'' S	15° 49' 45'' E	85	14 (10)	2005	Present study
Djibouti	Djibouti City	Dji	11° 35' 00'' N	43° 08' 00'' E	43	31	2002	(17)
Gabon	Near Kango	Gab§	00° 08' 49" N	10° 04' 50" E	49	24 (6)	2008	Present study
Madagascar	All over the country	Mad§	18° 54' 53'' S	47° 31' 54'' E	60	24 (21)	2006-2007	(18)
Niger	Niamey	Nig	13° 30' 45" N	02° 06' 45" E	38	24	2001	(17)
Senegal	Dakar	Sen	14° 45' 00'' N	17° 20' 00'' W	40	22	2002	(17)
Democratic Republic of the Congo	Kimpese	Con2#	05° 33' 46" S	14° 26' 06" E	NA	53	1993	(15)
Uganda	Entebbe	Uga#	00° 03' 10'' N	32° 27' 54'' E	NA	96	1996-1997	(15)
Zimbabwe	Mutare/Mutasa	Zim#	18° 58' 56'' S	32° 40' 14'' E	NA	67	1998	(15)
ASIA								
Iran	Sistan-Baluchistan	Ira§	26° 37' 60'' N	61° 16' 32'' E	50	29 (18)	2000-2003	Present study
Thailand	Maehongsorn	Tai1§	19° 51' 25'' N	97° 57' 60'' E	11*	11 (7)	2005	(19)
	Tak	Tai2	16° 54' 07'' N	99° 00' 46'' E	15*	15	2007	(19)
	Kanchanaburi	Tai3	14° 41' 07'' N	99° 01' 56'' E	17*	17	2006	(19)
	Ranong	Tai4	10° 00' 41'' N	98° 41' 57'' E	14*	14	2006	(19)
	Tak (Shoklo region)	Tai5#	16° 54' 07'' N	99° 00' 46'' E	NA	40	1997-1998	(15)
Myanmar	Rakhine	Mya§	19° 48' 36'' N	93° 59' 16'' E	30	24 (13)	2000	(20)
Laos	Sekong	Lao§	15° 20' 21'' N	106° 43' 21'' E	30	19 (15)	1999-2000	(20)
Papua New Guinea	Buksak	PNG1#	NA	NA	NA	62	1997	(15)
	Mebat	PNG2 #	05° 04' 40'' S	145° 46' 52'' E	NA	58	1997	(15)

SOUTH AMERICA								
French Guiana	Camopi	Cmp[§]	03° 09' 49'' N	52° 20' 32'' W	59	31 (17)	2006-2008	Present study
	Trois Sauts	Trs[§]	02° 20' 44'' N	52° 42' 14'' W	49	26 (14)	2006-2008	Present study
	Maripasoula	Mar[§]	03° 37' 54'' N	54° 02' 10'' W	66*	66 (7)	2002-2006	(21)
Colombia	Turbo/Bajo Cauca	Col1[§]	08° 24' 53'' N	75° 53' 52'' W	25	20 (19)	2002-2004	(22)
	El Bagre (Bajo Cauca region)	Col2[#]	07° 33' 36'' N	74° 58' 01'' W	NA	30	1998	(15)
Peru	Province of Iquitos	Per[§]	03° 43' 54" N	73° 16' 39" W	70	47 (32)	2003-2004	(23)
Venezuela	Bolivar and Amazonas departments (provinces)	Ven[§]	06° 21' 44'' N	63° 34' 50'' W	15	13 (7)	2003-2007	Present study
Bolivia	Guayaramerin	Bol[#]	10° 50' 44'' S	65° 23' 08'' W	NA	31	1994	(15)
Brazil	Porto Velho/Rondonia	Bra[#]	08° 45' 43'' S	63° 54' 07 '' W	NA	33	1997-1998	(15)

Table S2. The 12 *P. falciparum* microsatellite loci with marker code, chromosome location, GenBank accession number, number of alleles detected per locus, and size range of the amplified alleles in base pairs. The primer sequences for each pair of PCR primers and PCR conditions are as in (19). * Loci in common with the study by (15).

Microsatellite marker	**Chromosome location**	**GenBank access no.**	**Number of alleles**	**Size range (Base pairs)**
TA80	10	G38857	5	142-154
ARP2	13	G37793	12	150-192
TA1 *	6	AF010507	13	159-195
Poly ά *	4	L18785	21	131-194
TA60 *	13	AF010556	12	66-91
ARA2 *	11	X17484	12	56-89
Pfg377 *	12	L04161	7	92-110
PfPK2 *	12	X63648	14	163-208
TA87 *	6	AF010571	15	82-122
TA109 *	6	AF010508	21	144-210
TA81 *	5	AF010510	11	108-138
CIM8	1	G38013	19	156-210

Table S3. Characteristics of the 272 Single Nucleotide Polymorphisms (SNPs) genotyped in *P. falciparum* isolates. [a], "Yes" means that the mutation induced an amino acid change (non-synonymous SNP); "No" means that it does not, either because it is not located in a coding region or because the mutation corresponds to a synonymous polymorphism.

SNP name	Chromosome location	Chromosome position (bp)	Gene ID	Coding Region	Non-synonymous SNP[a]
MAL1.1303	1	161748	PFA0110w	Yes	No
MAL1.1698	1	248536	Null	No	No
MAL1.1775	1	269371	PFA0310c	Yes	No
MAL1.2165	1	375657	PFA0460c	Yes	Yes
MAL1.2547	1	454537	Null	No	No
MAL1.2718	1	497763	PFA0625w	Yes	No
MAL1.3532	1	552009	Null	No	No
MAL2.1214	2	69773	Null	No	No
MAL2.1464	2	109248	Null	No	No
MAL2.1601	2	133655	Null	No	No
MAL2.1674	2	161537	Null	No	No
MAL2.1692	2	165520	Null	No	No
MAL2.1769	2	185349	PFB0190c	Yes	yes
MAL2.1949	2	233720	Null	No	No
MAL2.2488	2	373912	PFB0405w	Yes	Yes
MAL2.2514	2	380021	Null	No	No
MAL2.2914	2	501802	PFB0555c	No	No
MAL2.3062	2	549049	PFB0190c	Yes	No
MAL2.3274	2	585780	PFB0650w	Yes	No
MAL2.3532	2	650531	Null	No	No
MAL2.3551	2	660994	PFB0703w	Yes	Yes
MAL2.3840	2	725641	PFB0820c	Yes	Yes
MAL2.4312	2	818134	Null	No	No
MAL2.4313	2	819315	Null	No	No
MAL2.4543	2	846392	Null	No	No
MAL2.4573	2	848536	PFB0950w	Yes	Yes
MAL3.1838	3	213832	Null	No	No
MAL3.2220	3	308854	Null	No	No
MAL3.2648	3	392712	Null	No	No
MAL3.2912	3	447298	PFC0435w	Yes	Yes
MAL3.3175	3	507959	PFC0505c	Yes	Yes
MAL3.3213	3	511564	PFC0510w	Yes	Yes

MAL3.3652	3	619437	PFC0640w	Yes	yes
MAL3.3796	3	656356	PFC0710w-a	No	No
MAL3.3815	3	662233	Null	No	No
MAL3.4155	3	748662	PFC0050c	Yes	Yes
MAL3.4358	3	816349	Null	No	No
MAL3.4439	3	848225	PFC0905c	Yes	Yes
MAL3.4573	3	889290	Null	No	No
MAL3.4644	3	908297	Null	No	No
MAL3.4961	3	982177	Null	No	No
MAL3.4987	3	995640	PFC1065w	Yes	Yes
MAL4.1520	4	110331	PFD0080c	Yes	Yes
MAL4.1525	4	110712	PFC0380w	Yes	No
MAL4.1685	4	134538	Null	No	No
MAL4.2067	4	158452	Null	No	No
MAL4.2520	4	234468	PFC0200c	Yes	No
MAL4.2588	4	245963	Null	No	No
MAL4.3297	4	399294	PFD0390c	Yes	No
MAL4.3544	4	472279	PFD0495c	Yes	No
MAL4.4003	4	551570	PFD0610w	Yes	Yes
MAL4.4969	4	686026	PFD0735c	Yes	Yes
MAL4.5256	4	755115	PFD0830w	Yes	Yes
MAL4.5260	4	755558	PFD0830w	Yes	Yes
MAL4.5386	4	782790	PFD0840w	Yes	Yes
MAL4.5679	4	843527	PFD0900w	Yes	No
MAL4.5987	4	905957	PFD0970c	Yes	Yes
MAL4.7161	4	1053149	Null	No	No
MAL4.7367	4	1101563	PFD1155w	Yes	Yes
MAL4.8110	4	1142144	Null	No	No
MAL5.526	5	65071	PFE0060w	Yes	No
MAL5.891	5	150460	Null	No	No
MAL5.1224	5	219220	Null	No	No
MAL5.1380	5	269459	PFE0320w	Yes	Yes
MAL5.1449	5	278013	PFE0320w	Yes	Yes
MAL5.1505	5	289918	PFE0340c	Yes	Yes
MAL5.1790	5	343384	Null	No	No
MAL5.1889	5	364200	PFE0435c	Yes	No
MAL5.2034	5	382974	PFE0450w	Yes	Yes
MAL5.2081	5	396039	PFE0465c	Yes	Yes
MAL5.2309	5	482703	Null	No	No
MAL5.2353	5	504226	PFE0570w	Yes	No
MAL5.2470	5	546121	Null	No	No

MAL5.3572	5	796712	PFE0970w	Yes	No
MAL5.3637	5	810754	Null	No	No
MAL5.3953	5	958435	PFE1150w	Yes	Yes
MAL5.4016	5	980524	Null	No	No
MAL5.4167	5	1040130	Null	No	No
MAL5.4615	5	1113504	PFE1325w	Yes	Yes
MAL5.4913	5	1181711	PFE1435c	Yes	No
MAL5.5469	5	1308985	PFE1600w	Yes	Yes
MAL6.595	6	74312	Null	No	No
MAL6.627	6	88829	Null	No	No
MAL6.847	6	148357	PFF0175	Yes	Yes
MAL6.1334	6	301723	Null	No	No
MAL6.1877	6	490306	Null	No	No
MAL6.1882	6	493449	PFF0575c	Yes	Yes
MAL6.2294	6	577117	PFF0670w	Yes	Yes
MAL6.2803	6	682442	PFF0795w	Yes	No
MAL6.3497	6	769375	PFF0885w	Yes	No
MAL6.3628	6	799787	PFF0930w	Yes	Yes
MAL6.3801	6	860702	Null	No	No
MAL6.4018	6	943311	PFF1110c	Yes	Yes
MAL6.4098	6	967410	PFF1145c	Yes	Yes
MAL6.4538	6	1116171	PFF1350c	Yes	No
MAL6.4668	6	1183651	Null	No	No
MAL6.5107	6	1289574	PFF1495w	Yes	Yes
MAL7.1561	7	145763	MAL7P1.7	Yes	No
MAL7.1980	7	225281	MAL7P1.16	Yes	Yes
MAL7.2219	7	285085	Null	No	No
MAL7.2220	7	287581	Null	No	No
MAL7.2307	7	307517	Null	No	No
MAL7.2405	7	339933	Null	No	No
MAL7.2631	7	417697	MAL7P1.23	Yes	Yes
MAL7.2888	7	470921	PF07_0037	Yes	Yes
MAL7.5309	7	665075	MAL7P1.61	Yes	No
MAL7.5506	7	692568	PF07_0053	Yes	Yes
MAL7.5593	7	723752	PF07_0058	Yes	Yes
MAL7.5886	7	783281	Null	No	No
MAL7.6134	7	837487	MAL7P1.89	Yes	Yes
MAL7.6292	7	899602	Null	No	No
MAL7.6500	7	940109	PF07_0085	Yes	No
MAL7.1416410	7	1416418	MAL7P1.176	yes	No
MAL7.7454	7	1193678	Null	No	No

MAL7.7642	7	1225730	MAL7P1.152	Yes	Yes
MAL7.7865	7	1289641	MAL7P1.157a	Yes	Yes
MAL7.8102	7	1363687	MAL7P1.187	Yes	Yes
MAL7.8344	7	1440599	PF07TR005	No	No
MAL7.8353	7	1414328	MAL7P1.176	Yes	No
MAL7.8356	7	1414598	MAL7P1.176	Yes	Yes
MAL7.8360	7	1414872	MAL7P1.176	Yes	Yes
MAL7.8369	7	1416729	MAL7P1.176	Yes	Yes
MAL8.1085	8	86747	PF08_0137	Yes	Yes
MAL8.1224	8	99179	Null	No	No
MAL8.1325	8	126761	Null	No	No
MAL8.1448	8	151154	PF08_0132	Yes	No
MAL8.1620	8	197031	Null	No	No
MAL8.1757	8	233936	RNAzID:4265	No	No
MAL8.1900	8	265780	MAL8P1.139	Yes	Yes
MAL8.2100	8	337607	MAL8P1.135	Yes	No
MAL8.2209	8	370855	Null	No	No
MAL8.2514	8	423288	MAL8P1.122	Yes	No
MAL8.3752	8	547962	MAL8P1.107	Yes	Yes
MAL8.3768	8	551333	PF08_0095	Yes	Yes
MAL8.4068	8	682727	PF08_0075	Yes	Yes
MAL8.4553	8	846885	MAL8P1.66	Yes	Yes
MAL8.4660	8	884477	MAL8P1.63	Yes	Yes
MAL8.5093	8	1006923	MAL8P1.43	Yes	Yes
MAL8.6236	8	1311881	PF08_0002	Yes	Yes
MAL9.1364	9	150689	Null	No	No
MAL9.1631	9	254304	PFI0260c	Yes	Yes
MAL9.1733	9	288133	Null	No	No
MAL9.1875	9	334347	PFI0330c	Yes	Yes
MAL9.2306	9	466622	PFI0495w	Yes	Yes
MAL9.2415	9	512730	PFI0550w	Yes	Yes
MAL9.2768	9	604333	PFI069c	Yes	Yes
MAL9.2977	9	689167	PFI0805w	Yes	Yes
MAL9.3207	9	767141	Null	No	No
MAL9.3582	9	973974	Null	Yes	Yes
MAL9.3705	9	929691	PFI1120c	Yes	Yes
MAL9.4188	9	1077048	PFI1300c	Yes	No
MAL9.4521	9	1178953	PFI1445w	Yes	Yes
MAL9.4543	9	1183700	Null	No	No
MAL9.4600	9	1199374	PF09TR005	No	No
MAL9.4694	9	1203641	PFI1475w	Yes	No

MAL9.4717	9	1203920	PFI1475w	Yes	Yes
MAL9.4752	9	1205108	PFI1475w	Yes	Yes
MAL9.4775	9	1205725	PFI1475w	Yes	No
MAL9.4807	9	1206288	PFI1475w	Yes	Yes
MAL9.4825	9	1206801	PFI1475w	Yes	Yes
MAL9.5592	9	1428002	PFI1740c-a	Yes	Yes
MAL10.1341	10	82376	Null	No	No
MAL10.1470	10	121129	PF10_0028-a	Yes	Yes
MAL10.1743	10	222496	PF10_0051	Yes	No
MAL10.2084	10	320664	PF10TR002	No	No
MAL10.2263	10	380845	Null	No	No
MAL10.2475	10	455426	PF10_0115	Yes	Yes
MAL10.4821	10	1182257	PF10_0281	Yes	Yes
MAL10.5105	10	1266136	PF10_0308	Yes	Yes
MAL10.5509	10	1377619	PF10_0336	Yes	No
MAL10.6506	10	1440352	Null	No	No
MAL10.6660	10	1509425	PF10_0373	Yes	No
MAL11.1470	11	120056	PF11_0037	Yes	Yes
MAL11.1490	11	123964	PF11_0038	Yes	Yes
MAL11.1524	11	138084	Null	No	No
MAL11.1653	11	185698	PF11_0053	Yes	No
MAL11.1727	11	214600	Null	No	No
MAL11.3166	11	659734	PF11_0178	Yes	Yes
MAL11.3627	11	768606	RNAzID:784	No	No
MAL11.4786	11	1179874	Null	No	No
MAL11.4816	11	1191858	PF11_0319	Yes	Yes
MAL11.4921	11	1221389	PF11_0326	Yes	Yes
MAL11.5107	11	1274174	Null	No	No
MAL11.5262	11	1295157	PF11_0344	Yes	Yes
MAL11.5485	11	1384948	PF11_0364	Yes	Yes
MAL11.5918	11	1546730	PF11TR006	No	No
MAL11.6405	11	1715136	Null	No	No
MAL11.6421	11	1722533	PF11_0442	Yes	Yes
MAL11.6680	11	1791173	PF11_0464	Yes	Yes
MAL12.2148	12	83558	PFL0055c	Yes	No
MAL12.2346	12	127973	PFL0115w	Yes	Yes
MAL12.2609	12	217434	PFL0245w	Yes	Yes
MAL12.3146	12	428463	Null	No	No
MAL12.3171	12	432868	PFL0465c	Yes	Yes
MAL12.3458	12	527375	PFL0585w	Yes	No
MAL12.3752	12	646259	PFL0770w	Yes	No

MAL12.4596	12	776377	PF12TR008	No	No
MAL12.4761	12	797285	Null	No	No
MAL12.4817	12	817723	PFL0980w	Yes	No
MAL12.5266	12	971730	PFL1145w	Yes	No
MAL12.5512	12	1074422	PFL1280w	Yes	Yes
MAL12.5831	12	1162017	PFL1385c	Yes	Yes
MAL12.6458	12	1325996	PFL1540c	Yes	Yes
MAL12.6805	12	1425744	PFL1645w	Yes	No
MAL12.6980	12	1486964	Null	No	No
MAL12.7216	12	1552073	PFL1800w	Yes	No
MAL12.7353	12	1589318	PFL1855w	Yes	yes
MAL12.7708	12	1691359	Null	No	No
MAL12.8354	12	1755389	PFL1980c	Yes	No
MAL12.8381	12	1770575	Null	No	No
MAL12.8583	12	1834449	PFL2110c	No	No
MAL12.8666	12	1850403	PFL2120w	Yes	Yes
MAL12.9166	12	1995203	PFL2335w	Yes	Yes
MAL12.9383	12	2054073	PFL2405c	Yes	Yes
MAL12.9386	12	2054196	PFL2405c	Yes	Yes
MAL12.9638	12	2138722	PFL2520c	Yes	Yes
MAL12.9690	12	2162863	PFL2545c	Yes	Yes
MAL13.776	13	92674	MAL13P1.60	Yes	Yes
MAL13.808	13	103054	PF13_0075	Yes	No
MAL13.1014	13	109962	MAL13P1.62	Yes	Yes
MAL13.1228	13	169982	Null	No	No
MAL13.1246	13	175971	PF13_0019	Yes	No
MAL13.1518	13	256929	MAL13P1.26	No	No
MAL13.1745	13	351823	Null	No	No
MAL13.1784	13	376068	MAL13P1.39	Yes	Yes
MAL13.1978	13	430054	Null	No	No
MAL13.2283	13	512792	PF13_0066	Yes	Yes
MAL13.2457	13	596533	MAL13P1.70	Yes	Yes
MAL13.2681	13	714924	PF13_0095	Yes	No
MAL13.2756	13	743761	MAL13P1.93	Yes	Yes
MAL13.2992	13	815532	Null	No	No
MAL13.3001	13	816155	MAL13P1.105	Yes	No
MAL13.3095	13	854947	Null	No	No
MAL13.3251	13	901094	MAL13P1.119	Yes	Yes
MAL13.3651	13	1013875	MAL13P1.113	Yes	Yes
MAL13.4132	13	1174520	MAL13P1.151	Yes	Yes
MAL13.4972	13	1465989	PF13_0201	Yes	yes

MAL13.5481	13	1624489	PF13_0223	Yes	Yes
MAL13.5689	13	1696966	MAL13P1.214	Yes	No
MAL13.5805	13	1742617	MAL13P1.380	Yes	Yes
MAL13.5880	13	1760453	MAL13P1.390	Yes	Yes
MAL13.6075	13	1809977	PF13_0241	Yes	No
MAL13.6321	13	1872783	PF13_0247	Yes	Yes
MAL13.6616	13	1957902	MAL13P1.246	Yes	Yes
MAL13.6890	13	2034181	MAL13P1.256	Yes	Yes
MAL13.7128	13	2111329	MAL13P1.268	Yes	Yes
MAL13.7384	13	2236537	Null	No	No
MAL13.7598	13	2306330	MAL13P1.286	Yes	Yes
MAL13.7767	13	2362067	Null	No	No
MAL13.7922	13	2419099	MAL13P1.301	Yes	Yes
MAL13.8065	13	2465780	MAL13P1.307	Yes	Yes
MAL13.8177	13	2502209	Null	No	No
MAL13.8453	13	2567063	PF13_0338	Yes	Yes
MAL13.9059	13	2745068	Null	No	No
MAL14.693	14	39478	Null	No	No
MAL14.697	14	40095	Null	No	No
MAL14.742	14	44103	Null	No	No
MAL14.1056	14	126288	PF14_0032	Yes	Yes
MAL14.1066	14	130979	Null	No	No
MAL14.1236	14	192760	Null	No	No
MAL14.1446	14	290368	Null	No	No
MAL14.1774	14	417069	Null	No	No
MAL14.1943	14	480286	Null	No	No
MAL14.2522	14	666649	Null	No	No
MAL14.3677	14	1043837	PF14_0247	No	No
MAL14.3812	14	1106081	PF14_0261	Yes	No
MAL14.4145	14	1199184	PF14_0282	Yes	No
MAL14.4985	14	1484089	Null	No	No
MAL14.6443	14	1995720	Null	No	No
MAL14.6695	14	2077736	Null	No	No
MAL14.9187	14	2981922	PF14_0700	Yes	Yes
MAL14.9817	14	3154289	PF14_0736	Yes	Yes

Table S4. SNPs under selection: F_{ST} estimates lie outside the 95% F_{ST} expected under neutrality, as determined using the program LOSITAN**.** *He*: expected heterozygosity. F_{ST}: Weir and Cockerham's estimate of F_{ST} (24).

SNP name	*He*	F_{ST}	*Probability (neutral F_{ST} <Observed F_{ST})*
MAL1.1303	0.121387	0.051391	1.60E-05
MAL7.7642	0.43845	0.070871	1.22E-04
MAL4.7367	0.442788	0.074168	1.59E-04
MAL4.2067	0.312366	0.086747	4.03E-04
MAL2.3062	0.378112	0.086751	4.03E-04
MAL4.8110	0.270479	0.081374	0.001733
MAL13.7767	0.162529	0.09972	0.007178
MAL4.5987	0.270868	0.101402	0.00797
MAL2.2488	0.506473	0.158745	0.008416
MAL10.5509	0.359566	0.147461	0.010125
MAL6.4018	0.197629	0.111366	0.013895
MAL8.6236	0.483926	0.182324	0.018503
MAL11.1490	0.453142	0.166348	0.020373
MAL12.2148	0.212425	0.120795	0.021636
MAL10.1341	0.48136	0.188129	0.022154
MAL12.9638	0.29224	0.121619	0.022417
MAL7.8356	0.427718	0.171058	0.023856
MAL2.1769	0.385734	0.172551	0.025048
MAL2.3551	0.459693	0.174227	0.026438
MAL5.4913	0.326363	0.175245	0.02731
MAL4.7161	0.225579	0.13183	0.033559
MAL8.2209	0.118678	0.132731	0.034674
MAL13.776	0.401434	0.185533	0.037348
MAL3.3815	0.34558	0.189817	0.042224
MAL11.1470	0.323725	0.193431	0.046676
MAL7.1980	0.130963	0.141978	0.047343
MAL13.5689	0.221963	0.620363	0.951603
MAL12.9166	0.353901	0.548829	0.953478
MAL4.5260	0.240366	0.654517	0.967414
MAL12.2609	0.506565	0.526607	0.968698
MAL4.5679	0.439729	0.575387	0.972076
MAL12.7216	0.229995	0.680309	0.976347
MAL7.6134	0.468506	0.549626	0.983171
MAL3.3652	0.35902	0.600497	0.983527
MAL12.5831	0.356427	0.605508	0.985255
MAL9.2306	0.519905	0.555157	0.985625
MAL12.5266	0.347714	0.616919	0.988621
MAL9.1733	0.38197	0.618437	0.989015
MAL9.4543	0.090447	0.603199	0.990506
MAL3.2912	0.234921	0.751362	0.991254
MAL13.3651	0.416391	0.630392	0.991727
MAL4.4969	0.505643	0.575442	0.99216
MAL9.2768	0.51926	0.576422	0.992395
MAL9.2415	0.380653	0.635136	0.992631

MAL5.3953	0.416987	0.654534	0.995492
MAL8.2100	0.510676	0.597562	0.996149
MAL10.2475	0.498597	0.598757	0.996299
MAL9.4752	0.434049	0.666855	0.996754
MAL6.1877	0.510061	0.613501	0.997765
MAL8.1085	0.41573	0.684454	0.998016
MAL8.3752	0.401235	0.696965	0.998625
MAL5.2034	0.471365	0.629083	0.99872
MAL6.4538	0.447026	0.700233	0.998754
MAL7.5886	0.474045	0.635379	0.998985
MAL13.3095	0.523445	0.658726	0.999586
MAL12.4761	0.396389	0.737227	0.99962
MAL13.2457	0.334079	0.742435	0.999682
MAL7.2888	0.482507	0.674483	0.999781
MAL4.3544	0.347656	0.771762	0.999891
MAL7.6500	0.39161	0.777817	0.999913
MAL3.4439	0.418535	0.798566	0.999962
MAL8.5093	0.52619	0.736648	0.999986
MAL13.6321	0.342857	0.824353	0.999988
MAL13.2992	0.351889	0.840393	0.999994
MAL8.4068	0.516092	0.755806	0.999994
MAL13.6075	0.333754	0.841598	0.999995
MAL6.1882	0.470389	0.77092	0.999997
MAL5.4016	0.434497	0.875834	0.999999
MAL8.3768	0.487084	0.801964	0.999999
MAL13.7922	0.508365	0.799308	0.999999
MAL6.4668	0.447751	0.88731	1
MAL5.1449	0.52904	0.828028	1

Table S5. Expected heterozygosity (*Hs*) computed for each of three datasets: MS, MS + and SNP. (Loci under selection are not included.) NA: Not available. *The two Angolan (Ang1 and Ang2) and also the two Benin (Ben1 and Ben2) populations were combined in the SNP analysis. P-values of Student *t*-test on logarithms of heterozygosity between continents are given at the bottom of the table. Ang1 and Ang2: Angola, Ben1 and Ben2 : Benin, Bra: Brazil, Bol: Bolivia, Cam: Cameroon, Cmp: Camopi (French Guiana), Col1 and Col2: Colombia, Con1: Republic of Congo, Con2: Democratic Republic of Congo, Dji: Djibouti, Gab: Gabon, Ira: Iran, Lao: Laos, Mad: Madagascar, Mar: Maripasoula (French Guiana), Mya: Myanmar, Nig: Niger, Per: Peru, PNG1 and PNG2: Papua New Guinea, Sen: Senegal, Tai1-Tai5: Thailand, Trs: Trois Sauts (French Guiana), Uga: Uganda, Ven: Venezuela, Zim: Zimbabwe.

Location	*Hs* "MS"	*Hs* "MS +"	*Hs* "SNP"
Ang1	0.73	0.79	0.20*
Ang2	0.73	0.75	NA
Ben1	0.72	0.76	0.23*
Ben2	0.73	0.76	NA
Cam	0.77	0.80	NA
Con1	0.70	0.73	0.22
Con2	NA	0.85	NA
Dji	0.54	0.57	NA
Gab	0.75	0.79	0.21
Mad	0.75	0.77	0.24
Nig	0.75	0.79	NA
Sen	0.74	0.78	NA
Uga	NA	0.83	NA
Zim	NA	0.84	NA
Mean (Africa)	**0.719**	**0.77**	**0.220**
Ira	0.63	0.73	0.21
Lao	0.71	0.74	0.20
Mya	0.70	0.74	0.26
PNG1	NA	0.75	NA
PNG2	NA	0.76	NA
Tai1	0.63	0.70	NA
Tai2	0.60	0.68	0.19
Tai3	0.57	0.66	NA
Tai4	0.50	0.56	NA
Tai5	NA	0.63	NA
Mean (Asia)	**0.620**	**0.70**	**0.215**
Bol	NA	0.36	NA
Bra	NA	0.42	NA
Cmp	0.19	0.17	0.10
Col1	0.22	0.26	0.07
Col2	NA	0.42	NA
Mar	0.41	0.42	0.12
Per	0.40	0.48	0.18
Trs	0.15	0.11	0.10
Ven	0.46	0.40	0.18
Mean (South America)	**0.305**	**0.32**	**0.13**
P-value (Africa vs Asia)	**0.018**	**0.015**	**0.66**
P-value (Africa vs America)	**0.0042**	**0.00054**	**0.0088**
P-value (Asia vs America)	**0.0083**	**0.0012**	**0.011**

Table S6. Probabilities (with 95% confidence intervals in brackets) for three invasion scenarios inferred from ABC analyses. For each simulation, the selected scenario (shaded cell) was the one with the highest probability value. Confidence in scenario choice was evaluated from 500 randomly constructed simulated datasets. Median divergence time estimates (in number of generations with 95% confidence intervals) are also given except for cases where no scenario could be chosen with confidence. *Indicates populations where ABC analyses were performed on seven microsatellite loci (MS+ dataset excluding loci suspected to be under selection). For the other pairs of populations, analyses were performed on ten microsatellite markers (MS dataset excluding loci suspected to be under selection). Bol: Bolivia, Bra: Brazil, Cmp: Camopi (French Guiana), Col1, Col2: Colombia, Mar: Maripassoula (French Guiana), Trs: Trois Sauts (French Guiana). The "northern cluster" contains the two populations Col1 and Col2, whereas the "southern cluster" comprises the populations Bol, Bra, Cmp, Mar and Trs.

Population 1	Population 2	Scenario A	Scenario B	Scenario C	Divergence time estimates in generations
Northern Cluster *	Southern Cluster *	0.5835 [0.5689 : 0.5980]	0.08245 [0.0804 : 0.0845]	0.3340 [0.3257 : 0.3424]	t1:3710 [756 : 8450] t2:5450 [2400 : 9860]
Southern Cluster *	Northern Cluster *	0.6922 [0.6749 : 0.7094]	0.0658 [0.0641 : 0.0675]	0.2420 [0.2360 : 0.2480]	t1:2380 [392 : 7530] t2:6860 [2170 : 9860]
Col 1	Cmp	0.3218 [0.3138 : 0.3299]	0.2284 [0.2227 : 0.2341]	0.4498 [0.4385 : 0.4610]	t1:2220 [153 : 6930] t2:5450 [1790 : 8970] t3 :6710 [2210 :9830]
	Trs	0.5510 [0.5372 : 0.5648]	0.1500 [0.1462 : 0.1537]	0.2990 [0.2915 : 0.3065]	t1:3070 [280 : 8300] t2:6090 [1810 : 9700]
	Mar	0.7436 [0.7250 : 0.7630]	0.0886 [0.0864 : 0.0908]	0.1678 [0.1636 : 0.1720]	t1:2970 [251 : 7820] t2:5390 [1540 : 9640]
	Col 2 *	0	0.9252 [0.9021 : 0.9484]	0.0747 [0.0728 : 0.0766]	t1:608 [73 : 3020] t2:4520 [545 : 9620]
	Bol *	0.4076 [0.3974 : 0.4178]	0.0797 [0.0777 : 0.0817]	0.5126 [0.4998 : 0.5254]	t1:2260 [184 : 6960] t2:5570 [1840 : 9090] t3:7390 [2870 : 9880]
	Bra *	0.4988 [0.4863 : 0.5113]	0.0690 [0.0672 : 0.0707]	0.4322 [0.4213 : 0.4430]	t1:2760 [200 : 7900] t2:5910 [1730 : 9720]
Col 2 *	Cmp *	0.3468 [0.3381 : 0.3554]	0.2026 [0.1975 : 0.2076]	0.4506 [0.4393 : 0.4618]	t1:3050 [335 : 7480] t2:5750 [2020 : 9110] t3:7580 [3020 :9880]
	Trs *	0.4272 [0.4165 : 0.4379]	0.1403 [0.1368 : 0.1438]	0.4325 [0.4218 : 0.4433]	--
	Mar *	0.4368 [0.4259 : 0.4478]	0.1228 [0.1197 : 0.1259]	0.4404 [0.4293 : 0.4514]	--
	Col 1 *	0	0.9136 [0.8907 : 0.9364]	0.0864 [0.0842 : 0.0885]	t1:416 [53 : 2270] t2:4660 [561 : 9670]
	Bol *	0.4965 [0.4841 : 0.5089]	0.1238 [0.1207 : 0.1269]	0.3797 [0.3702 : 0.3891]	t1:2870 [293 : 7940] t2:6160 [1930 : 9770]
	Bra *	0.5403 [0.5267 : 0.5538]	0.0873 [0.0851 : 0.0895]	0.3723 [0.3630 : 0.3817]	t1:2470 [201 : 7710] t2:5920 [1740 : 9740]
Cmp	Col 1	0.6531 [0.6367 : 0.6694]	0.0890 [0.0868 : 0.0912]	0.2580 [0.2514 : 0.2643]	t1:2340 [172 : 7410] t2:5980 [1820 : 9720]
	Trs	0	0.9994 [0.9744 : 1.024]	0.0006 [0.0005 : 0.0007]	t1:612 [74 : 2770] t2:5830 [771 : 9800]
	Mar	0	0.7048 [0.6871 : 0.7224]	0.2952 [0.2878 : 0.3026]	t1:985 [118 : 3860] t2:3950 [470 : 9560]
	Col 2 *	0.6275 [0.6118 : 0.6432]	0.0383 [0.0373:0.0756]	0.3342 [0.3258 : 0.3426]	t1:1910 [156 : 7180] t2:6150 [1870 : 9740]

	Bol *	0	0.78072 [0.7612:0.8002]	0.2193 [0.2138 : 0.2248]	t1:486 [51 : 2720] t2:4400 [424 : 9670]
	Bra *	0	0.5408 [0.5273:0.5543]	0.4592 [0.4477 : 0.4706]	t1:511 [49 : 2930] t2:3740 [347 : 9580]
Trs	Col 1	0.8314 [0.8107 : 0.8523]	0.0364 [0.0355 : 0.0373]	0.1321 [0.1288 : 0.1354]	t1:2550 [196 : 7580] t2:6340 [2050 : 9760]
	Cmp	0	0.9983 [0.9734 : 1.023]	0.0017 [0.0015 : 0.0017]	t1:663 [80 : 2950] t2:5750 [750 : 9790]
	Mar	0	0.8691 [0.8473 : 0.8908]	0.1309 [0.1276 : 0.1342]	t1:537 [70 : 2430] t2:4330 [489 : 9620]
	Col 2 *	0.5278 [0.5148 : 0.5412]	0.0692 [0.0675 : 0.0710]	0.4028 [0.3927 : 0.4185]	t1:3050 [231 : 8140] t2:6350 [2030 : 9770]
	Bol *	0	0.9561 [0.9083 : 1.00]	0.04388 [0.0417:0.0460]	t1:408 [55 : 2160] t2:5520 [741: 9770]
	Bra *	0	0.9008 [0.8728 : 0.9233]	0.0992 [0.0967 : 0.1017]	t1:709 [83 : 3320] t2:5600 [781 : 9780]
Mar	Col 1	0.7882 [0.7685 : 0.8079]	0.0536 [0.0523 : 0.0550]	0.1582 [0.1542 : 0.1621]	t1:2210 [187 : 7160] t2:5630 [1630 : 9690]
	Cmp	0	0.7089 [0.6912 : 0.7266]	0.2911 [0.2838 : 0.2984]	t1:702 [85 : 3020] t2:3960 [504 : 9570]
	Trs	0	0.8393 [0.8183 : 0.8603]	0.1607 [0.1567 : 0.1647]	t1:332 [42 : 1730] t2:4220 [491 : 9610]
	Col 2 *	0.4895 [0.4772 : 0.5017]	0.1089 [0.1061 : 0.1116]	0.4019 [0.3916 : 0.4117]	t1:1760 [107 : 7320] t2:4250 [243 : 9680]
	Bol *	0	0.9079 [0.8852 : 0.9306]	0.0921 [0.0898 : 0.0944]	t1:336 [47 : 1800] t2:5050 [400 : 9720]
	Bra *	0	0.9005 [0.8780 : 0.9230]	0.0995 [0.0970 : 0.1020]	t1:396 [49 : 2070] t2:5060 [397 : 9730]
Bol *	Col 1 *	0.4951 [0.4827 : 0.5075]	0.1016 [0.0990 : 0.1041]	0.4033 [0.3932 : 0.4139]	t1:3180 [284 : 8170] t2:6210 [1890 : 9770]
	Cmp *	0	0.8834 [0.8613 : 0.9055]	0.1166 [0.1137 : 0.1195]	t1:819 [98 : 3910] t2:5710 [709 : 9800]
	Trs *	0	0.9813 [0.9568 : 1.005]	0.0187 [0.0182 : 0.0192]	t1:351 [50 : 1900] t2:6130 [926 : 9810]
	Mar *	0	0.9470 [0.9233 : 0.9707]	0.0530 [0.0517 : 0.0543]	t1:505 [72 : 2520] t2:5260 [650 : 9740]
	Col 2 *	0.6574 [0.6409 : 0.6738]	0.0529 [0.0514 : 0.0540]	0.2900 [0.2827 : 0.2972]	t1:2330 [191 : 7590] t2:6400 [2040 : 9780]
	Bra *	0	0.7575 [0.7385 : 0.7764]	0.2425 [0.2364 : 0.2485]	t1:348 [35 : 2080] t2:4640 [471 : 9690]
Bra *	Col 1 *	0.6281 [0.6124 : 0.6438]	0.0855 [0.0833 : 0.0876]	0.2864 [0.2792 : 0.2935]	t1:3040 [265 : 8070] t2:6350 [1960 : 9790]
	Cmp *	0	0.8799 [0.8578 : 0.9019]	0.1202 [0.1172 : 0.1231]	t1:909 [115 : 4210] t2:6030 [855 : 9810]
	Trs *	0	0.9647 [0.9405 : 0.9887]	0.0353 [0.0344 : 0.0362]	t1:448 [61 : 2370] t2:6100 [936 : 9810]
	Mar *	0	0.9400 [0.9163 : 0.9633]	0.0602 [0.0587 : 0.0617]	t1:441 [62 : 2330] t2:5320 [655 : 9740]
	Col 2 *	0.6468 [0.6306 : 0.6630]	0.0681 [0.0664 : 0.0698]	0.2850 [0.2780 : 0.2921]	t1:2660 [234 : 7860] t2:6100 [936 : 9810]
	Bol *	0	0.9064 [0.8837 : 0.9290]	0.0936 [0.0912 : 0.0960]	t1:352 [41 : 2020] t2:5730 [716 : 9790]

Table S7. Range of prior distributions of the parameters used in the ABC analyses. *Ne S* is the effective population size of the source. *Nebt1*, *Nebt2* and *Nebt 3* are the effective population sizes during bottlenecks. *Ne 1* and *Ne 2* are the effective population sizes after population expansion. *t1*, *t2* and *t3* are divergence times expressed in number of generations before present. μ is the single locus mutation rate. *db* corresponds to the duration of the bottleneck.

Parameters	***Prior (lower bound)***	***Prior (upper bound)***
Ne S	10	100000
Ne 1	10	10000
Ne 2	10	10000
Nebt1	10	1000
Nebt2	10	1000
Nebt3	10	1000
t1	1	10000
t2	1	10000
t3	1	10000
db	1	100
Mutation rate (*μ*)	5.10^{-5}	5.10^{-3}