Supporting Information

Pierron et al. 10.1073/pnas.1321860111

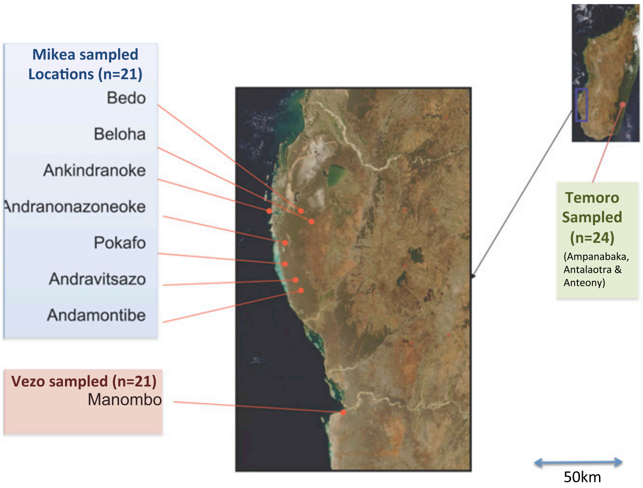
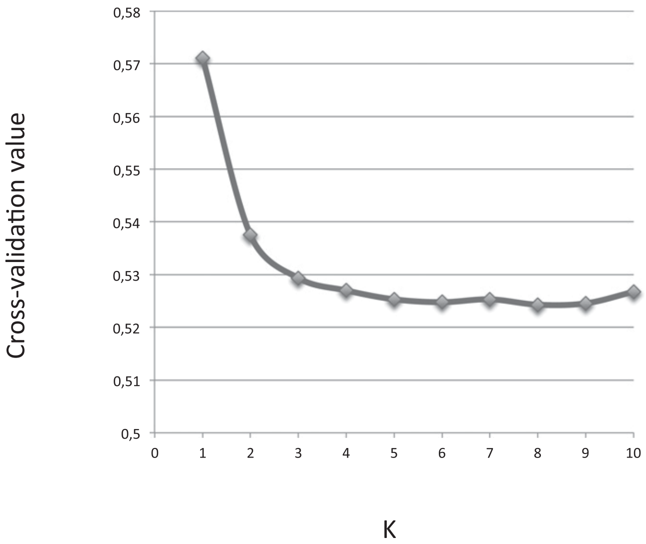


Fig. S1. Location of the genetic sampling in Madagascar.

Cross-validation Austronesian component



Austronesian component STRUCTURE-like chart for several K values

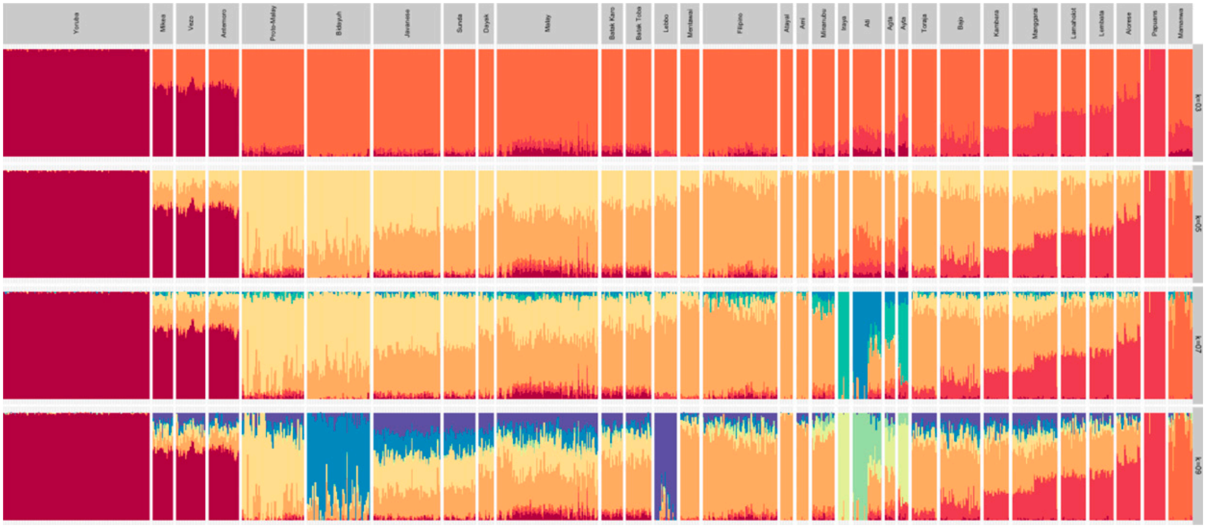
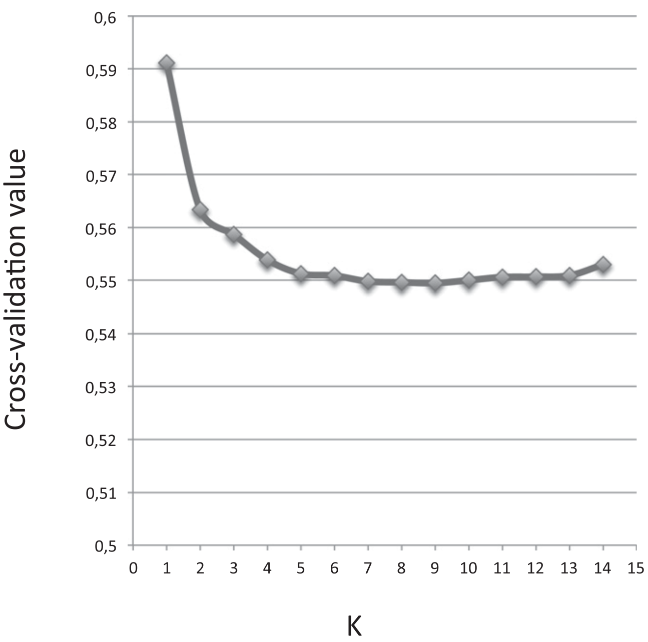


Fig. S2. (Continued)

Cross-validation African component



African component STRUCTURE-like chart for several K values

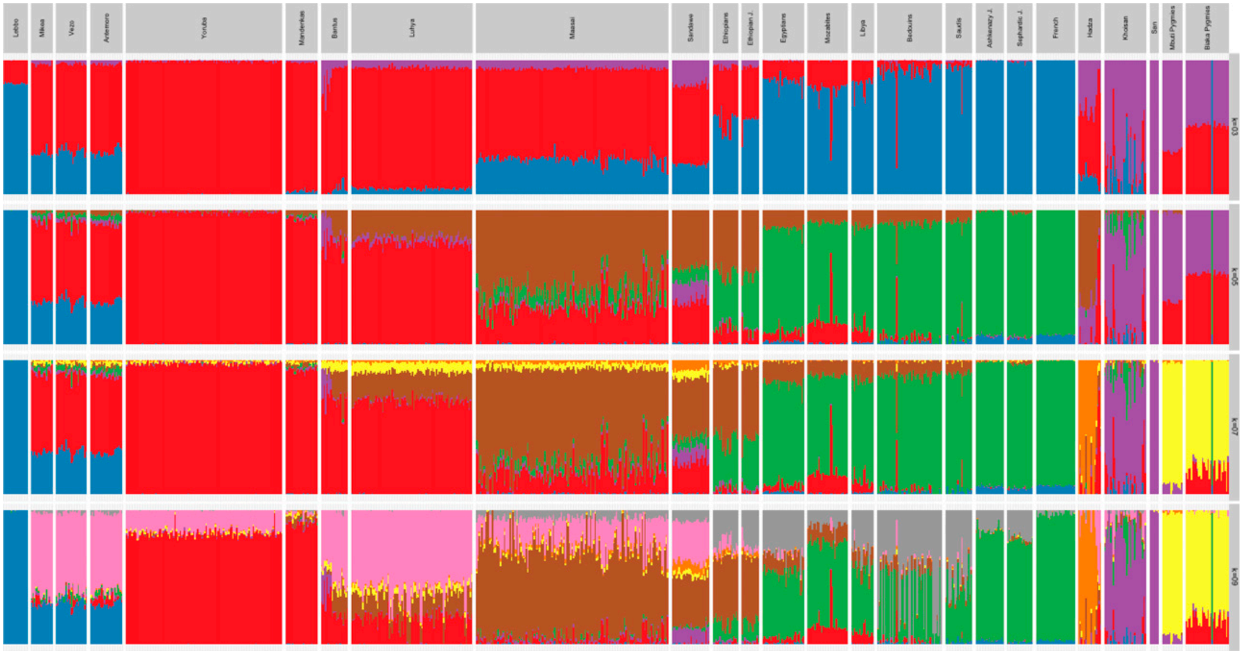


Fig. S2. ADMIXTURE results on Austronesian and African components.

Table S1. Estimates of genetic diversity and African ancestry in three Malagasy populations

		Genetic diversity	% of African ancestry (±SD) (ADMIXTURE analysis, K = 02)	
Populations	*n*	1-IBS (±SD)	Test A	Test B
Mikea	21	25.3% ± 1.25	68.21 ± 2.31	68.71 ± 2.10
Temoro	24	25.6% ± 0.17	67.15 ± 3.86	67.81 ± 0.34
Vezo	24	25.7% ± 0.32	66.28 ± 4.54	66.47 ± 2.10
All	69	25.7% ± 0.45	67.19 ± 3.76	67.50 ± 3.23

Estimates of genetic diversity from IBS analysis and African ancestry in three Malagasy populations from the two distinct analyses with ADMIXTURE, test A was focused on the Indonesian origin and test B on the African origin. IBS, identity-by-state distance.

Table S2. Overlaps of iHS signals among populations

	Overlap with top 5% iHS windows in second population						
Malagasy	Temoro	Mikea	Vezo	Bajo	Lebbo	CEU (Europe)	YRI (Yoruba)
Top 1% iHS							
Temoro		38.3	39.3	34.7	14.1	14.9	25.0
Mikea	29.3		24.1	19.8	10.9	13.9	19.3
Vezo	41.3	31.6		28.1	12.4	12.3	24.4
Bajo	27.3	24.3	24.1		23.4	11.2	14.0
Lebbo	13.5	8.7	10.6	18.1		7.8	12.4
CEU (Europe)	12.9	13.4	13.8	10.3	11.1		7.8
YRI (Yoruba)	23.9	15.4	20.3	11.0	7.6	9.0	

iHS, integrated haplotype score.

Table S3. Biological processes with less than 100 genes that were specifically enriched (EASE score <0.05) in Mikea top 5% iHS signals but not in Temoro and Vezo

GO ID	Gene ontology term	Windows in top 5% with this term	Windows in top 5% without this term	Overall windows with this term	Overall windows without this term	Individual genes in Mikea top 5%	Individual genes in 5%	
							Temoro	Vezo
GO:0014013	Regulation of gliogenesis	11	398	58	6,025	*NOTCH1*, *HES5*, *SOX10*, *CLCF1*, *ID4*, *DRD3*, *CNTN2*, *PPARG*, *GFAP*, *RNF10*	*HES5*, *SOX10*, *CLCF1*,	*NOTCH1*, *HES5*, *SOX10*,
GO:0001561	Fatty acid α-oxidation	4	405	5	6,078	*HAO1*, *PHYH*, *HACL1*, *PEX13*		*HAO1*
GO:0090330	Regulation of platelet aggregation	4	405	7	6,076	*SYK*, *PRKG1*, *PRKCD*, *PRKCQ*	*PRKCD*	
GO:0051584	Regulation of dopamine uptake	4	405	8	6,075	*GDNF*, *DRD3*, *DRD4*		*GDNF*
GO:0002819	Regulation of adaptive immune response	11	398	74	6,009	*CLCF1*, *DUSP10*, *FCER1A*, *IL20RB*, *IRF7*, *LTA*, *PRKCQ*, *RC3H1*, *TBX21*, *TNF*, *TNFSF13*, *TRAF2*	*CLCF1*, *TRAF2*, *IL20RB*	*TRAF2*, *IL20RB*
GO:0007215	Glutamate receptor signaling pathway	6	403	25	6,058	*GRIN1*, *GRIN2A*, *GRM5*, *GRM6*, *NRXN1*, *HOMER2*	*GRM5*	*GRIN1*
GO:0006941	Striated muscle contraction	8	401	46	6,037	*ADRBK1*, *CHRNA1*, *DTNA*, *PIK3CA*, *RPS6KB1*, *SLC8A1*, *TNF*, *TNNC2*, *TNNI1*, *TTN*	*ADRBK1*, *DTNA*, *CHRNA,1*	*CHRNA1*, *TNNC2*
GO:0071901	Negative regulation of protein serine/threonine kinase activity	11	398	78	6,005	*ABL1*, *DUSP10*, *DUSP16*, *DUSP5*, *DUSP7*, *GBA*, *GSTP1*, *LATS2*, *PRKCD*, *SPRED1*, *SPRED2*	*GSTP1*, *PRKCD*, *SPRED2*	*PRKCD*

Redundant terms excluded: GO:0042537, GO:0018958, GO:0051580, GO:0048710, GO:0007612, GO:0051588, GO:0051940, GO:0007216, GO:0045685, GO:0042417, GO:0045214, GO:0019395. (Underlined genes are present at least twice in the table.)

Table S4. Top 1% signals of haplotype homozygosity (iHS) shared by all three Malagasy populations

Chr:Loc (Mb)	Genes contained in the 200-kb window	YRI	CEU	Lebbo	Bajo
7_142.4	*TRBJ2-5*, *TRBJ2-3*, *TRBJ2-2P*, *TRBJ2-2*, *TRBJ2-1*, *TRBJ2-4*, *TRBJ2-6*, *TRBJ2-7*, *TRBC2*, *TRBV29-1*, *TRBV27*, *AC104597.1*, *TRBV28*, U66061.42, *TRBVB*, *TRBV26*, *PRSS2*, *U66061.31*, *TRPV6*, *EPHB6*, *TRBV30*, *U66061.35*, *U66061.38*, *RP11-114L10.2*, *PRSS1*				
2_27	*MAPRE3*, *C2orf18*, *DPYSL5*, *CDKN2AIPNLP2*, *AC013472.4*, *CENPA*			**	*
9_100.6	*HEMGN*, *C9orf156*, *ANP32B*, *FOXE1*, *RP11-535C21.3*, *Y_RNA U6*	**			
8_72	*EYA1*, *RP11-326E22.1*			*	*
2_177.6	*FUCA1P1*, *AC092162.2*, *AC073636.1*, *AC092162.1*, *U6*			*	**
2_125.6	*CNTNAP5*, *RN5S102*, *U6*				
2_72.4	*EXOC6B*, *AC006461.2*				
13_61.4	*RP11-196P2.1*				*

Signals that appear also at top 1% or top 5% in Yoruba (YRI), European (CEU), Lebbo, or Bajo populations are shown with ** or *, respectively.