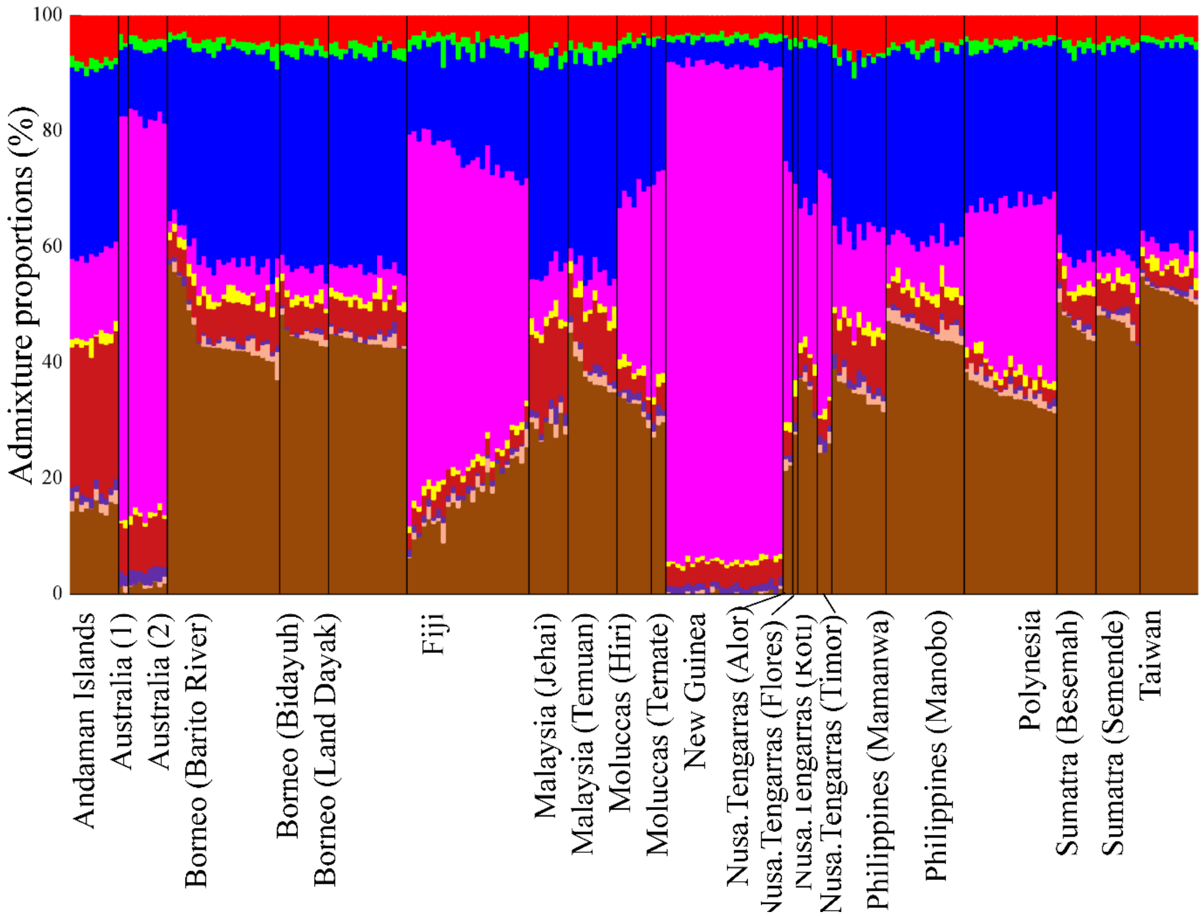
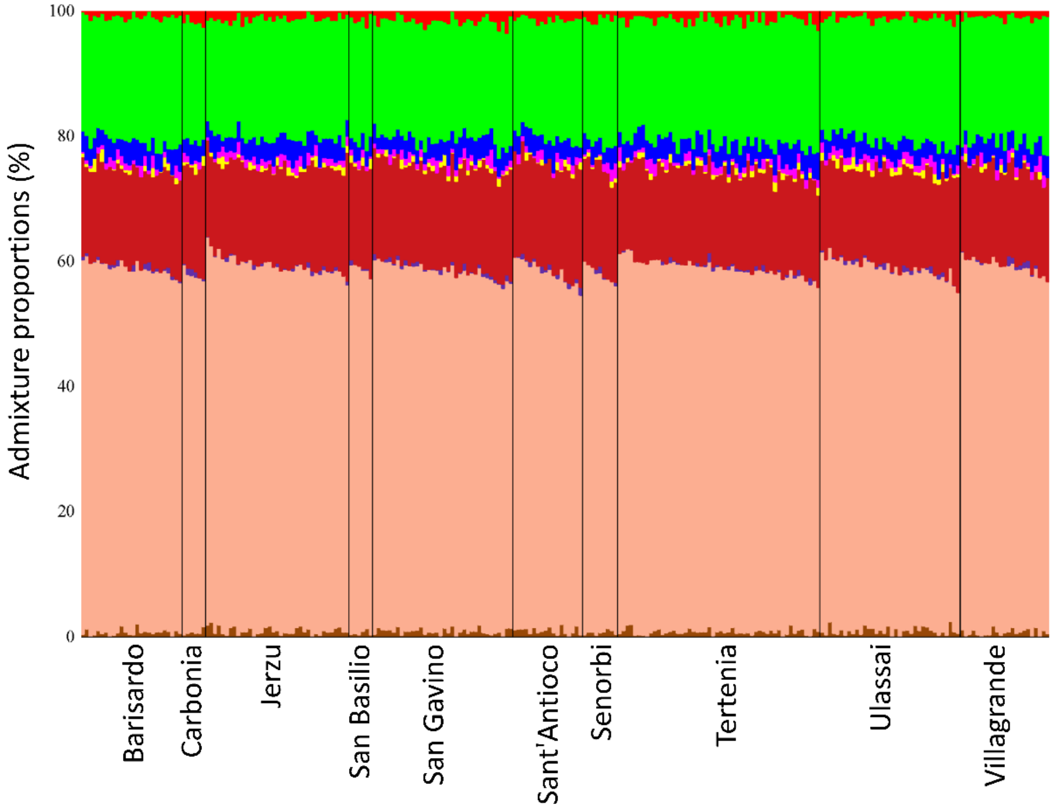
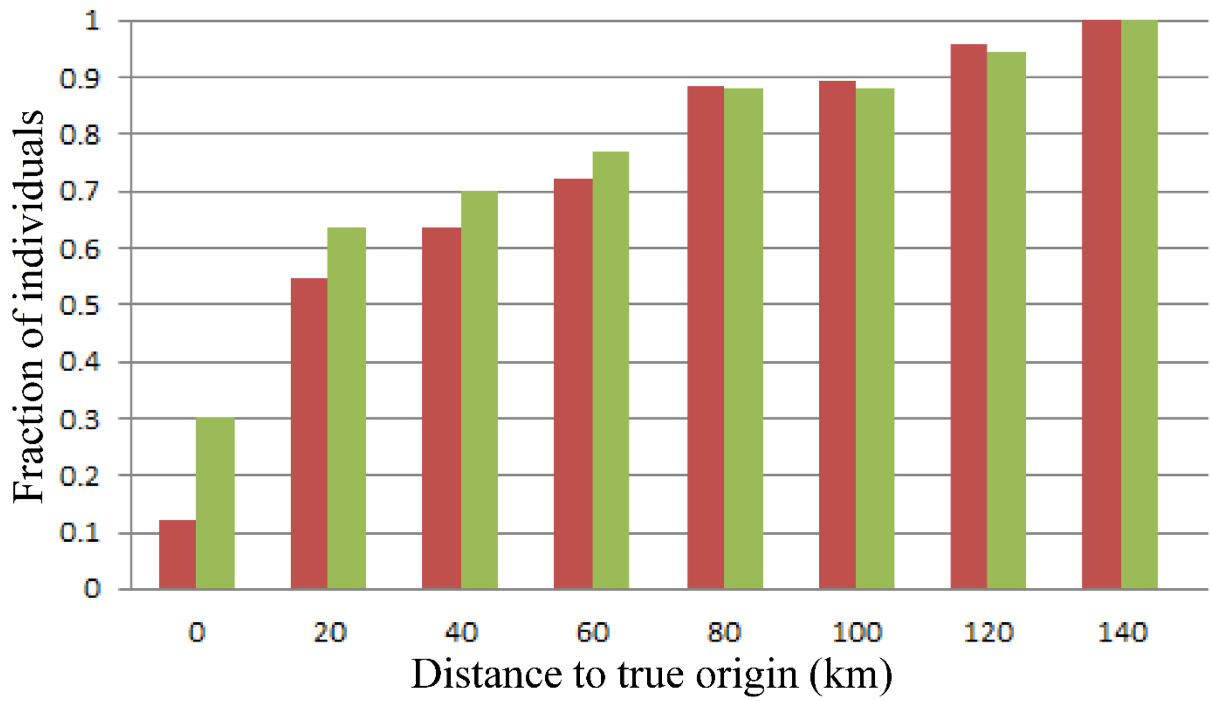
Supplementary Figures



Supplementary Figure 1. Admixture analysis (*K*=9) of Asian and Oceania populations and subpopulations. The *x*-axis represents individuals from populations sorted by their ancestries. Each individual is represented by a vertical stacked column of color-coded admixture proportions of the putative ancestral populations.



Supplementary Figure 2. Admixture proportions for Sardinians. The *x*-axis represents individuals from populations sorted according to their ancestries. Each individual is represented by a vertical stacked column of color-coded admixture proportions of the ancestral populations.



Supplementary Figure 3. Prediction accuracy for 249 Sardinian male (red) and females (green) individuals calculated as the distance between the predicted point and the village of origin.

Supplementary Tables

Supplementary Table 1 – Summary of population details

Populations*	Countries and inner-countries regions	Nick	n	Latitude	Longitude
Bermudian	Bermuda	BE	16	32.17	-64.47
Bulgarian	Bulgaria	BU	15	42.41	23.19
Chinese	China	CHB	12	39.55	116.2
Danish	Denmark	DA	15	56	10
Egyptian	Egypt	EG	15	30.03	31.15
Finnish	Finland	FIN	12	60.1	24.56
Georgian	Georgia	GO	4	41.43	44.47
German	Germany	GR	15	52.51	13.38
Greek	Greece	GK	15	37.96	23.71
Southern Indian	Andhra Pradesh, Karnataka, Kerala, Tamil Nadu	IS	18	10.4	78.13
Northern Indian	Jammu and Kashmir, Himachal Pradesh, Uttaranchal, Haryana, Punjab, Rajasthan, Uttar Pradesh, Jharkhand, Chhattisgarh, Madhya Pradesh	IN	20	28.23	78.35
Western Indian	Maharashtra, Gujarat, Goa	IW	16	22.21	72
Eastern Indian	West Bengal, Orissa, Bihar	IE	10	22	86
Iranian	Iran	IR	16	32.42	53.68
Italy	Sardinia	ID	15	40	9
	Toscana	TSI	15	43	11
Japanese	Japan	JPT	12	38	138
Luhya (Kenya)	Kenya	LWK	12	-0.02	37.9
Kuwaiti	Kuwait	KU	18	29.22	47.58
Lebanese	Lebanon	LE	22	33.85	35.86
Madagascar	Antananarivo	MO	3	-18.91	47.53
(Ambilobe)	Antsiranana	MB	4	-13.12	49.3
(Manakara)	Fianarantsoa	MN	4	-22.14	48
(Andilambe)	Mahajanga	MD	3	-15.75	47.98
(Toliara)	Toliary	MT	6	-23.35	43.66

Mexico	Guanajuato	XG	3	21	-101.15
	Hidalgo	XH	8	20.28	-98.51
	Morelos	XM	4	18.44	-99.04
Mongolian	Mongolia	MG	11	45	111
Namibia (Southeastern)	Erongo, Hardap	NS	5	-28.1	19.53
(Kaokoveld)	Kunene	NK	5	-19.4	13.91
(Hereroland)	Omaheke	NH	5	-21	19.5
(Tsumkwe)	Otjozondjupa	NT	2	-19.59	20.5
Yoruban (West African)	Nigeria	YRI	12	8	4
Papuan (Papua New Guinea)	Central, Western, Southern Highlands, Western Highlands, Eastern Highlands, Northern, Enga, Sandaun, Madang, Chimbu, Morobe, Gulf	PC	16	-9.3	147.1
Bougainville-Nasioi (Oceania)	North Solomons	PN	10	-6.05	155.19
Peruvian (Highland)	Cusco, Puno	PH	12	-13.5	-71.97
Peruvians (Lima)	Lima, Callao	PEL	12	-12.04	-77.06
Puerto Rican	Puerto Rico	PR	15	18.28	-66.07
Romanian	Romania	RO	15	44.8	26.06
Russia	Altay	RA	15	50.61	86.21
(Northern Caucasian)	Chechnya, Ingush	CA	12	43.8	45.71
	Moscow	RM	16	55.75	37.62
	Tatarstan	RT	15	55.18	50.72
RSA	Johannesburg	SJ	2	-26	28
	Underberg	SU	2	-29.47	29.3
	Schmidtsdrift, Northern Cape	SS	9	-29	21.85
	Kroonstad , Orange Free State	SK	5	-27	27
Iberian	Spain, Portugal	IBS	12	40.3	-3.72
Pamiri	Tajikistan	PT	13	38.35	68.48
Tunisia	Tunisia	TU	12	36.48	10.11
British	United Kingdom	UK	13	52	-3.18
Ni-Vanuatu	Vanuatu	VA	10	-17.44	168.1
Kinh (Vietnam)	Vietnam	KHV	12	21.02	105.5

* Four populations (African and Mexican Americans, Brahmin Indians, and Romanian gypsies) were omitted from later analyses.

Supplementary Table 2 - Prediction accuracy for subpopulations

			Prediction accuracy (%)	
Population	**Subpopulation**	**N**	**Region**	**Country**
Mexican	Guanajuato	3	33	67
	Morelos	4	75	75
	Hidalgo	8	75	88
Peruvian	Highland	12	83	100
	Lima	12	58	75
Italian	Sardinian	15	100	100
	Tuscan	15	60	60
Russian	Tatarÿ	15	100	100
	Moscow	16	88	88
	Altaian (Siberian)	15	93	100
Namibian	Kaokoveld	15	100	100
	Tsumkwe	5	80	100
	Hereroland	2	0	100
	Southeastern	5	20	80
RSA	Kroonstad, Free State	5	100	100
	Underberg	5	60	60
	Johannesburg	2	100	100
	Schmidtsdrift, Northern Cape	2	0	50
Madagascar	Ambilobe	9	11	22
	Antananarivo	4	50	100
	Toliara	3	67	100
	Manakara	6	50	100
	Andilambe	4	50	100
Indian	Northern	2	50	100
	Eastern	21	67	67
	Southern	10	100	100
	Western	17	82	82
Papuan	Papua New Guinea, Center	16	81	81
	Bougainville-Nasioi	17	88	88

Supplementary Table 3. Details on the Asian and Oceania populations and subpopulations analyzed. For brevity, we adhered to the population names and nicknames used by Reich et al.[1]

Broader group	Detailed	Nick	N	Latitude	Lonitude
Australian	Northern Territories	AU-1	8	-17.3	133.24
	Cell Cultures	AU-2	2	-20.3	141.85
Borneo	Barito River	BO-BR	23	-1.1	114.14
	Bidayuh	BO-BI	10	1.62	110
	Land Dayak	BO-DY	16	0	109.46
Chinese	Beijing (Han)	CHB	87	39.55	116.2
Fiji	Fiji	FI	25	-17.71	178.06
Japanese	Japanese	JPT	88	34	136
Malaysia	Jehai	ML-JE	8	4.21	101.97
	Temuan	ML-TM	10	4.21	101.97
Moluccas	Hiri	MO-HI	7	1	127
	Ternate	MO-TE	3	0.3	127.3
Nusa Tenggaras	Alor	NT-AL	2	-8.4	124.2
	Flores	NT-FL	1	-8.7	120.62
	Roti	NT-RO	4	-10.91	123.26
	Timor	NT-TI	3	-9.31	125.55
Andamanese	Onge	AN	10	11.74	92.65
Philippines	Manobo	PH-MN	16	16	129
	Mamanwa	PH-MA	11	10	129
New Guinea	Highlander	SH	24	-9.3	147.1
Polynesia	Western Polynesia	PO	19	-16.83	-148.37
Sumatra	Besemah	SU-BE	8	-3.03	102.39
	Semende	SU-SM	9	-4.3	103.88
Taiwan	Taiwan	TA	12	23.69	120.96

Supplementary Table 4 - Average predicted distances for males and females from ten Sardinian villages

Village	Elevation (m)	Lat	Lon	Females		Males	
				$\overline{Dist}$	*N*	$\overline{Dist}$	*N*
Barisardo*	41	39.84	9.64	22.442	15	19.86	11
Carbonia*	98	39.17	8.52	42.425	4	24.21	2
Jerzu*	439	39.79	9.52	33.541	17	28.71	20
San Basilio	407	39.54	9.20	0	2	38.23	4
San Gavino*	52	39.55	8.79	36.962	16	44.05	20
Sant'Antioco	193	39.06	8.44	47.883	10	72.83	8
Senorbi	201	39.54	9.13	50.455	7	36.23	2
Tertenia*	130	39.69	9.58	34.490	20	43.46	32
Ulassai*	720	39.81	9.50	22.032	25	16.64	11
Villagrande	900	39.96	9.51	23.443	10	31.92	13

**Coastal villages*

Supplementary Table 5 - Pearson correlation between altitude and distance from predicted origin to village of origin for females (F) and males (M). Results are reported as: (correlation coefficient, Student's *t*-test *p*-value, and sample size). Significant results (p<0.05) are bolded.

Gender	All villages	Coastal villages	Inland villages
F	(-0.137, 0.064, 125)	(-0.11, 0.142, 96)	(-0.266, 0.082, 29)
M	**(-0.174, 0.028, 122)**	**(-0.214, 0.019, 95)**	**(-0.382, 0.024, 27)**
F*	(-0.132, 0.076, 119)	(-0.105, 0.161, 92)	(-0.283, 0.076, 27)
M*	**(-0.181, 0.026, 116)**	**(-0.22, 0.017, 93)**	**(-0.406, 0.027, 23)**

*Results for the eight largest villages with *N*>9.

Supplementary Notes

Supplementary Note 1

```
GPS<-function(outfile_name='GPS_results.txt',N_best=1,fname="data.csv",
directory_name="C://GPS//") {

  setwd(directory_name)    #set directory
  GEO=read.csv("GEO.csv", header=TRUE,row.names=1)
  GEO=GEO[,1:2]
  GEN=(read.csv("GEN.csv", header=TRUE,row.names=1))  #as.numeric
  TRAINING_DATA=read.csv(fname, header=TRUE, row.names=1)
  y=dist(GEO)
  x=dist(GEN)

  LL=length(y)
  for(l in 1:LL){
    if(y[l]>=70 || x[l] >=0.8) {y[l]=0; x[l]=0;}
  }

  eq1<-lm(y~x);

  GROUPS=unique(TRAINING_DATA$GROUP)
  write("Population\tSample_no\tSample_id\tPrediction\tLat\tLon",outfile_name,
append=FALSE)
  N_best<-min(N_best,length(GEO[,1]))

  for(GROUP in GROUPS){
    Y=subset( TRAINING_DATA, TRAINING_DATA$GROUP_ID==GROUP)
    K=length(Y[,1])
    for(a in 1:    K)
    {
      X<-Y[a,1:9]
      E<-rep(0, length(GEO[,1])  )  ;
      minE=10000; minG=-1; second_minG=-1;
      for(g  in 1: length(GEO[,1])){
        ethnic<-attributes(GEO[g,])$row.names;
        gene<-as.numeric(GEN[ethnic,1:9])
        E[g]<-   sqrt(sum((gene-X)^2))
      }
```

```
    minE=c();minE<-c(minE,sort(E,FALSE)[1:N_best])
    minG=c();
    for(g  in 1: length(GEO[,1])){
      for(j in 1:N_best){
        if( isTRUE(all.equal(minE[j], E[g]))){minG[j]=g;}
      }
    }
    radius<-E[minG];
    best_ethnic<- attributes(GEO[minG,])$row.names; #best_ethnic;
    radius_geo=(eq1[[1]][2]*radius[1])
    W<- (minE[1]/minE)^4;
    W=W/(sum(W));
    delta_lat<-(GEO[minG,][[1]]-GEO[minG[1],][[1]])
    delta_lon<-(GEO[minG,][[2]]-GEO[minG[1],][[2]])
    new_lon<-sum(W*delta_lon)
    new_lat<-sum(W*delta_lat)
    lo1<-new_lon*min(1,radius_geo/sqrt(new_lon^2+new_lat^2) )
    la1<-new_lat*min(1,radius_geo/sqrt(new_lon^2+new_lat^2))
    write(paste(GROUP,                a,                row.names(Y[a,]),
best_ethnic[1],GEO[minG[1],1]+la1,GEO[minG[1],2]+lo1,          sep="\t"),outfile_name,
append=TRUE)
   }
  }
  return ("GPS is done!");
}
```

Supplementary References

1 Reich, D. *et al.* Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am. J. Hum. Genet.* **89**, 516-528 (2011).