Supplemental Information

Modeling recent human evolution in mice: the case of *EDARV370A*

Table of Contents

Supplemental Figure Legends

Figure S1: Linkage disequilibrium patterns of the genomic region surrounding *370A*, related to Figure 1A. Figure obtained by HaploView (Barrett et al., 2005). The entire region of the analysis included 280 SNPs over a ~1,400 kb span. The *370A* variant was pointed out by an arrow. The ~139kb block of linkage disequilibrium covering the *370A* variant is highlighted in yellow. Of note, in many individuals examined, this haplotype remained nearly unbroken for more than 1Mb.

Figure S2: Approximate posterior density estimates of demographic and evolutionary parameters, related to figure 1B. ABC was performed retaining the top 5,000 simulations among a total of 1,083,966 simulations (tolerance level 0.46%). The posterior density estimates shown in dash lines are from the top 1,074 simulations (tolerance level 0.1%).

Figure S3: Data used for demic forward simulation, related to Figure 1B. A) Demic carrying capacity of the geographic region simulated. B) Climatic factor of the geographic region simulated. Data deducted by Koppen's climate classification (Kottek et al., 2006). C) 157 locations with archaeobotanical evidence for rice cultivation (Fuller *et al*. (2011).

Figure S4: Ranked Euclidean distances of the 0.55% of the best simulations, related to Figure 1B. The plateau ends after the 5,000th best simulation (red), corresponding to a threshold value of 0.46%. The smallest 0.1% Euclidean distances are in blue.

Figure S5: Inferring the age of *370A* using coalescence-based allele frequency spectrum, related to Figure 1. A) The allele frequency spectrum (AFS) from the region of length 850kb flanking *370A*. 48 haplotypes carrying the putative selected allele at *370A* were randomly selected from CHB samples of the 1000 Genomes Project. The SNPs with known polarity (the ancestral allele is identifiable by comparing to the Chimpanzee genome) are pooled together to construct the AFS. B) and C) The marginal likelihood curves for selection intensity (*s*) and selection time (*T*) when the other parameter took the value of MLE. The curves are fitted using cubic B-spline with 2 evenly spaced knots.

Figure S6: Validation of homologous recombination into *Edar* locus, related to Figure 2. A) Southern blot showing homologous recombination of targeting construct using 5' probe binding upstream of the 5' arm of homology. Genomic DNA was isolated from correctly (49 and 139) and incorrectly targeted (control) ES clones and digested with SexA1 (cuts upstream of 5' arm) and AfeI (cuts in 3' arm after Neo cassette) restriction enzymes. After digest, correctly targeted clones contain a 10.446kb (resulting from integration of the targeting construct, denoted by *) and an 8.460kb band (endogenous locus, denoted by Δ). Note that intervening lanes were cut out from image of Southern blot to consolidate lane order so that control, clone 49 and clone 139 lanes could be displayed in the shown order. B) Southern blot using a probe against the Neo cassette shows correct 3' integration into *Edar* locus and single Neo integration. Genomic ES DNA was digested with AflII enzyme (cuts inside Neo cassette and downstream, outside of the 3' targeting arm). A single band at the expected size (6.142kb, *) was detected in correctly targeted clones (49 and 139), whereas an incorrectly targeted clone (control) shows non-homologous recombination in the ES cell genome. C) Eag1 diagnostic digest validates presence

of point mutation T58065848C in correctly targeted ES clones 49 and 139. A 411bp PCR product was amplified from targeted ES cell genomic DNA using (Primers: 5'CACCGAGTTGCCGTTTGACTG3' and 5'CCATGCATCCTGCAGGCATAAG3'). Digestion with Eag1 yields the expected size bands of 71bp (**) and 340bp (***) in clones 49 and 139 resulting from the introduction of a unique Eag1 site as a result of the T58065848C mutation. Uncut control DNA is shown.

Figure S7: Frequency trajectories of alleles under positive selection, related to Figure 1. The trajectory of the selected *EDAR* allele was simulated using a Wright-Fisher model with an effective population size of N=10,000 diploid individuals, and an initial allele frequency $1/2N$ (representing a newly-arisen allele). Three scenarios were simulated with selection coefficient of 0.05, 0.10, and 0.15. In each generation, the allele frequency used for the binomial sampling was based on the allele frequency in the previous generation and the effect of positive selection on the allele. We conditioned our trajectories on eventual fixation of the allele, ignoring any simulations where the allele was lost from the population. For each genetic model (dominant, semi-dominant, recessive), we generated 500 simulations where the allele eventually became fixed, and took the average of these simulations as the expected trajectory of the allele conditional on fixation. In all three scenarios, the allele in semi-dominant model rose to near fixation in a shorter period of time than that in dominant model.

Supplemental Tables

Table S1: Populations used as observed data points in ABC simulations, related to Figure 1B

Ethnicity	Geographic Origin	Sample Size (2n)	Derived allele frequency	Latitude	Longitude
Makrani	Pakistan	50	0.000	27	62-66
Balochi	Pakistan	46	0.000	30-31	66-67
Brahui	Pakistan	48	0.000	30-31	66-67
Sindhi	Pakistan	50	0.000	24-27	68-70
Pathan	Pakistan	50	0.000	32-35	69-72
Hazara	Pakistan	40	0.400	33-34	70
Kalash	Pakistan	50	0.000	35-37	71-72
Burusho	Pakistan	50	0.060	36-37	73
Xibo	China	18	0.889	43-44	81-82
Uygur	China	20	0.500	44	81
Bengali	Bangladesh	144	0.049	24	90
Dai	China	18	0.889	21	100
Lahu	China	18	0.778	22	100
Naxi	China	20	0.700	26	100
Tu	China	18	0.833	36	101
Yizu	China	18	0.889	28	103
Cambodian	Cambodia	20	0.550	12	105
Han	China	88	0.955	26-39	108-120
Miaozu	China	20	1.000	28	109
Tujia	China	20	0.950	29	109
Mongola	China	20	0.750	48-49	118-120
She	China	18	0.722	27	119
Oroqen	China	20	1.000	48-53	122-131
Daur	China	20	0.900	48-49	124
Yakut	Siberia	48	0.854	62-64	129-130
Hezhen	China	20	1.000	47-48	132-135
Japanese	Japan	62	0.742	38	138
Papuan	New Guinea	34	0.000	-4	143
Melanesian	Bougainville	42	0.095	-6	155

Data of Bengali from Bangladesh is from unpublished data from the Sabeti Lab at Harvard. Other data are from Sabeti *et al*. (Sabeti et al., 2007), of which samples were collected by HGDP-CEPH. For data points occupying a range of latitude and longitude, we averaged the summary statistics (derived allele frequency) of the demes covered by the range and deemed it as one data point in the analysis.

Table S2: Prior and posterior distributions of the parameters in ABC simulations, related to Figure 1B

Parameters	Prior		Posterior			
	Lower	Upper	Mode	Median	2.5%	97.5%
Start generation	0	1450	188	363	17	1073
Start latitude (°N)	-6	62	34	36	61	23
Start longitude (°E)	62	155	113	114	102	137
Intra-demic gene-flow	0	0.2	0.142	0.106	0.007	0.195
Inter-demic gene-flow	0	0.2	0.143	0.104	0.006	0.195
Cultural diffusion	0	0.2	0.140	0.103	0.005	0.195
Selection coefficient	0	0.2	0.122	0.114	0.030	0.186
Long-distance migration	0	0.2	0.167	0.135	0.029	0.196
Farmer mobility	0	3	0.928	1.485	0.074	2.914
Hunter-gatherer mobility	0	3	2.099	1.516	0.074	2.933
Farming potential of HG	0.2	0.333	0.239	0.266	0.203	0.330

We performed a total of 1,497,844 simulations, of which 1,083,966 simulations had the simulated allele still present after 1,600 generations. Analysis was based on these 1,083,966 simulations. The posterior distributions of the parameters were estimated by using the top 5,000 (0.46%) simulations fitting to observed data.

Table S3: Posterior distributions of the parameters estimated using the top 1,074 (0.1%) simulations fitting to observed data, related to Figure 1B

Parameters	Posterior			
	Mode	Median	2.5%	97.5%
Start generation	181	342	17	1091
Start latitude (°N)	34	37	61	25
Start longitute (°E)	113	115	103	137
intra-demic geneflow	0.144	0.113	0.009	0.195
inter-demic geneflow	0.062	0.103	0.005	0.195
cultural diffusion	0.140	0.107	0.005	0.196
selection coefficient	0.043	0.080	0.020	0.187
Long-distance migration	0.170	0.134	0.024	0.196
Farmer mobility	1.103	1.524	0.087	2.921
Hunter-gatherer mobility	2.076	1.602	0.080	2.943

Farming potential of HG	0.239	0.265	0.204	0.329

Table S4: Correlation between the parameters in ABC simulations, related to Figure 1B

	StGen	IntraGF	InterGF	CulDiff	Selection	LongMig	F_Mobil	HG_Mobil	F_Pot
StGen	-	0.0001	0.0002	0.0000	0.0443	0.0002	0.0002	0.0002	0.0000
IntraGF	**0.6110**	-	0.0021	0.0000	0.0000	0.0000	0.0002	0.0001	0.0002
InterGF	**0.3055**	0.0014	-	0.0001	0.0005	0.0002	0.0001	0.0001	0.0000
CulDiff	**0.8662**	0.7724	0.4349	-	0.0000	0.0005	0.0002	0.0003	0.0003
Selection	**0.0000**	0.8236	0.1288	0.8945	-	0.0134	0.0000	0.0000	0.0000
LongMig	**0.3688**	0.9612	0.2853	0.1033	0.0000	-	0.0002	0.0002	0.0022
F_Mobil	**0.3239**	0.3009	0.5754	0.3460	0.7751	0.3829	-	0.0001	0.0001
HG_Mobil	**0.3415**	0.4463	0.5099	0.2431	0.7715	0.3520	0.5944	-	0.0000
F_Pot	**0.9985**	0.3701	0.7845	0.2081	0.8358	0.0009	0.6040	0.6252	-

R^2 values are above the diagonal of the table, while *P*-values are reported under the diagonal. Significant *P*-values are in bold, and the corresponding R^2 values are in bold and shaded. StGen – mutation start generation, IntraGF – intra-demic gene-flow, InterGF – inter-demic gene-flow, CulDiff – cultural diffusion, Selection – selection coefficient, LongMig – long-range migration, F_Mobil – farmer mobility, HG_Mobil – hunter-gatherer mobility, F_Pot – farming conversion rate of hunter-gatherer. The parameters were estimated by using the top 5,000 (0.46%) simulations fitting to observed data.

Table S5: Multivariate Analyses of Variance – Mouse phenotypes, related to Figure 3-6

Trait	Wilks' λ	*F*	df_{hyp}	df_{error}	*P*	Partial Eta^2	Obs. Power	Pairwise Comparison *P*			Phenotypic means								
											mean	SD	n	mean	SD	n	mean	SD	n
Awl Hair	.726	2.77	4	64	**.034**	.148	.731	**370V - 370V/370A**	**370V - 370A**	**370A - 370V/370A**	**370V**			**370V/370A**			**370A**		
4 cell (% of total)		4.42	2	33	**.020**	.211	.721	0.440	**0.007**	0.052	37.51	19.92	12	43.91	24.31	11	60.11	14.25	13
3 cell (% of total)		4.65	2	33	**.017**	.220	.744	0.326	**0.005**	0.065	62.11	19.67	12	54.37	21.64	11	39.83	14.32	13
2 cell (% of total)		1.20	2	33	.313	.068	.244	0.248	0.773	0.148	0.38	0.61	12	1.73	4.92	11	0.07	0.24	13
Auchene Hair	.641	2.57	6	62	**.027**	.199	.811	**370V - 370V/370A**	**370V - 370A**	**370A - 370V/370A**	**370V**			**370V/370A**			**370A**		
4 cell (% of total)		1.14	2	33	.333	.065	.233	0.908	.182	.234	0.19	0.67	12	0.34	0.82	11	1.86	4.96	13
3 cell (% of total)		5.66	2	33	**.008**	.255	.828	0.983	**0.007**	**0.007**	64.06	12.19	12	63.94	18.76	11	80.05	9.51	13
2 cell (% of total)		5.21	2	33	**.011**	.240	.794	0.887	**0.007**	**0.012**	34.27	12.75	12	33.41	19.27	11	17.86	10.19	13
1 cell (% of total)		2.54	2	33	.094	.133	.472	0.388	0.182	**0.033**	1.48	1.57	12	2.31	3.74	11	0.23	0.60	13
Meibomian Glands	.944	0.78	4	108	.538	.028	.244	**370V - 370V/370A**	**370V - 370A**	**370A - 370V/370A**	**370V**			**370V/370A**			**370A**		
suborbital glands area (pixels)		1.53	2	55	.225	.053	.312	n/a	n/a	n/a	8.08E4	1.2E4	13	8.69E4	1.1E4	25	8.75E4	1.1E4	20
superorbital glands area (pixels)		0.43	2	55	.655	.015	.116	n/a	n/a	n/a	1.47E5	1.4E4	13	1.50E5	1.2E4	25	1.51E5	1.2E4	20
Mammary Glands	.557	2.24	2	37	**.026**	.253	.883	**370V - 370V/370A**	**370V - 370A**	**370A - 370V/370A**	**370V**			**370V/370A**			**370A**		
total branches		1.99	10	66	.151	.097	.385	.861	.095	.081	569.55	152.56	10	580.37	155.64	19	687.00	162.90	11
branch density (per cm)		3.41	10	66	**.044**	.156	.605	0.446	**0.018**	**0.047**	353.66	67.61	10	381.90	94.26	19	455.01	111.55	11
total length (cm)		1.67	10	66	.202	.083	.329	.076	.251	.596	1.59	0.20	10	1.49	0.13	19	1.52	0.08	11
fat pad area(cm^2)		4.46	10	66	**.018**	.194	.730	0.315	**0.007**	**0.030**	1.02	0.13	10	0.98	0.09	19	0.89	0.11	11
gland area (cm^2)		1.74	10	66	.189	.086	.342	.079	.159	.835	0.49	0.07	10	0.44	0.08	19	0.45	0.05	11
Eccrine Glands	.382	6.95	8	90	**4.3E-07**	.382	1.000	**370V - 370V/370A**	**370V - 370A**	**370A - 370V/370A**	**370V**			**370V/370A**			**370A**		
FP-3 (glands)		29.82	2	48	**3.8E-09**	.554	1.000	**.001**	**6.3E-10**	**6.9E-05**	25.62	1.48	17	22.67	2.00	18	20.25	2.44	16
FP-4 (glands)		17.74	2	48	**1.7E-06**	.425	1.000	**.005**	**3.0E-07**	**.003**	23.26	1.84	17	21.39	1.61	18	19.59	1.86	16
FP-5 (glands)		20.87	2	48	**3.0E-07**	.465	1.000	**.004**	**5.2E-08**	**.001**	16.15	0.90	17	14.61	1.55	18	13.28	1.28	16
FP-6 (glands)		3.88	2	48	**.027**	.139	.674	.096	**.008**	.265	20.38	1.96	17	19.64	2.38	18	18.50	1.29	16
Eccrine Glands 379K rescue	.464	3.506	8	60	**.002**	.319	.966	**370V/379E - 370V/379K**	**370V/379E - 370A/379K**	**370V/379K - 370A/379K**	**370V/379E**			**370V/379K**			**370A/379K**		
FP-3 (glands)		7.90	2	20	**.002**	.324	.935	**.002**	.991	**.001**	23.12	2.21	12	19.91	1.92	11	23.11	2.48	13
FP-4 (glands)		7.91	2	20	**.002**	.324	.935	**.001**	.329	**.006**	21.04	2.31	12	17.77	2.27	11	20.23	1.51	13
FP-5 (glands)		10.97	2	20	**2.2E-4**	.399	.985	**8.0E-5**	.247	**.001**	14.41	1.63	12	11.36	1.32	11	13.69	1.83	13
FP-6 (glands)		9.98	2	20	**4.0E-4**	.377	.976	**9.0E-5**	.054	**.014**	20.71	1.79	12	16.72	2.19	11	19.00	2.37	13

Pairwise comparisons were computed using Least Significant Difference (LSD) posthoc tests
SD= standard deviation; Bold face values indicate $P<0.05$

Table S6. Complete genotype-phenotype association list of *370A* in a Han Chinese population, related to Figure 7

Phenotypic Traits	*370A*	Covariate				
		Age	Sex	Cluster	Age&Sex	Age&Sex&Cluster
Sweat Glands						
Density (fingertips)	**0.0047**	**0.0045**	**0.0035**	**0.0052**	**0.0033**	**0.0024**
Hair						
Density (axilla)	0.3947	0.2539	0.3882	0.2711	0.2468	0.2823
Density (back)	0.9186	0.8292	0.9086	0.5703	0.8400	0.6244
Density (chest)	0.7047	0.7682	0.7345	0.8516	0.8049	0.8104
Density (forearm)	0.4558	0.2981	0.4250	0.2595	0.2653	0.2576
Density (forehead)	0.0646	**0.0239**	0.0694	**0.0189**	**0.0260**	**0.0214**
Density (leg)	0.8838	0.7321	0.8451	0.6914	0.6845	0.7105
Density (scalp)	0.9332	0.9977	0.9420	0.9580	0.9885	0.9882
Density (thigh)	0.2724	0.0984	0.2256	0.0837	0.0609	0.0582
Teeth						
Shoveling	**0.0077**	**0.0024**	**0.0063**	**0.0067**	**0.0016**	**0.0054**
Double shoveling	**0.0004**	**0.0003**	**0.0003**	**0.0005**	**0.0002**	**0.0004**
Canine mesial ridge	0.7801	0.7898	0.8022	0.7666	0.8160	0.8928
Carabelli's cusp	0.1729	0.1403	0.1932	0.2017	0.1523	0.1748
Cusp 4	0.7781	0.7481	0.7007	0.7979	0.6669	0.6188
Cusp 5	0.7253	0.6358	0.5989	0.7753	0.5025	0.5649
Cusp 6	0.3551	0.3284	0.3528	0.4155	0.3190	0.3132
Cusp 7	0.7717	0.7738	0.6173	0.7746	0.6191	0.6713
Cusp number	0.2084	0.1892	0.2066	0.2407	0.1831	0.1509
Deflecting wrinkle	0.9429	0.9244	0.9668	0.9444	0.9904	0.8778
Groove pattern	0.6138	0.6109	0.6715	0.6182	0.6677	0.6421
Interruption groove	0.2053	0.2266	0.1917	0.1846	0.2168	0.2177
Odontome (mandible)	0.5949	0.5705	0.6040	0.6099	0.5763	0.5851
Odontome (maxilla)	0.4615	0.4812	0.4078	0.4257	0.4300	0.4566
Parastyle	0.4250	0.4394	0.4553	0.4114	0.4733	0.4731
Peg-shaped molar	0.9083	0.9862	0.7597	0.9860	0.8371	0.8475
Premolar lingual cusp variation	0.7010	0.7415	0.5925	0.6178	0.6363	0.5595
Protostylid	**0.0079**	**0.0065**	**0.0092**	**0.0087**	**0.0075**	**0.0082**
Third molar missing (mandible)	**0.0123**	**0.0053**	**0.0099**	**0.0124**	**0.0039**	**0.0079**
Third molar missing (maxilla)	0.6041	0.5086	0.4317	0.6881	0.3378	0.4412
Tuberculum dentale	0.2342	0.1636	0.2326	0.2503	0.1556	0.2244
Winging	0.8102	0.8546	0.7107	0.7685	0.7583	0.7506

P-values of a linear regression analysis using an additive model are shown in the table. Significant *P*-values are in bold. Samples collected from different groups (427 from villages and 196 from a college) were assigned distinctive a cluster number, which was used as a covariate to control potential substructure.

Supplemental Experimental Procedures

S1 Modeling the origins and spread of *370A* in an approximate Bayesian computation framework

The demic simulation model

We employed spatially explicit forward simulations to model the origin and the subsequent spread of the derived *370A* allele in East Asia and surrounding regions. This demic model was previously used to investigate the co-evolution of dairying and lactase persistence in Europe (Itan et al., 2009).

Geographic region and demes

The demic grid of this simulation model encompasses the geographic region between 6ºS and 62ºN and 62ºE and 155ºE, covering most of the available sampled populations where the *370A* allele has been observed (Table S1). This geographic region was modelled as a series of 6486 rectangular demes, with each being one degree in latitude and one degree in longitude. There are 3328 land demes, arranged to approximate the East Asian landmass and 3158 sea demes in total. Sea demes do not hold populations, but can be passed through by long-range migrations.

Length of the simulation

Modern humans arrived in East Asia as early as 60,000 years BP (Shi et al., 2008), and spread throughout East Asia by 40,000 years BP (Stanyon et al., 2009). Our simulations therefore start from 40,000 years BP (1,600 generations ago, assuming 25 years per generation). At the start of the simulation, hunter-gatherers occupy all demes with an initial population size set at 1/20 of their hunter-gatherer carrying capacity (that is 1/1,000 of the maximum carrying capacity of the deme when farmers appear; see below for more details). This initial population size was arbitrarily set for computational efficiency, so that most of the simulations could run forward without the *370A* allele going extinct after it had first appeared. Sensitivity tests showed that this setting had little impact on the results. Among the total of 1,497,844 simulations performed, the derived allele arose and did not go extinct after 1,600 generations in 1,083,966, or 72% of the simulations. Those simulations where the derived allele frequency fell to zero were excluded from further analysis, as they definitely would not fit the observed data.

The appearance of the *370A* allele

The *370A* allele is at high frequency in East Asian and Native American populations, while almost absent in European and African populations (Sabeti et al., 2007). It was repeatedly reported as showing a signature of positive selection (Bryk et al., 2008; Grossman et al., 2010; Sabeti et al., 2007; Xue et al., 2009). Assuming that the *370A* allele arose as a new mutation in East Asia where it had a selective advantage over the ancestral allele, thereby reaching high frequencies in those locations, we allowed the allele to appear at any spatial point within our simulated geographic region, at any time between 40,000 years BP and 3,750 years BP (prior range of [0-1,450] of the total 1,600 generations). The lower boundary of this range (40,000 years BP) was chosen at the same time as the simulation starts, so that a wide enough range

would be explored. The upper boundary (3,750 years BP) was chosen because an ancient DNA study on a Palaeo-Eskimo dated from 4,000 years BP reported the *370A* allele was present in this individual (Rasmussen et al., 2010). Once the appearance time of the allele has been drawn from the prior range, an origin deme is selected at random from those where the population size is at least 10, so that at least one individual can carry the new allele. After the deme has been chosen, the allele appears either in hunter-gatherers or farmers with a probability proportional to the respective population sizes.

Carrying capacity

Each simulated deme has a carrying capacity, K_{deme} – the maximum population size the deme can support. We followed the equation used in Itan et al (2009): $K_{deme} = (0.2cl + 0.8el)D_{max}A_{deme}$, for determining the carrying capacity (K_{deme}) of each deme (Figure S3A). This carrying capacity is mainly affected by three factors; the relative climate (*cl*), the relative elevation (*el*), and the area of the deme (A_{deme} in km^2, varying for each deme to account for the curvature of Earth). D_{max} is the maximum population density, which was fixed at 5 individuals per km^2. We used Koppen's climate definitions to determine the relative climatic factor (*cl*; Figure S3B). We used values of 0.25, 0.5, 0.75 and 1 for polar, cold, dry and tropical, and temperate climates, respectively. These values reflect the fact that the mild climate normally results in a high potential population size while the extreme climate results in a small potential population size. The relative elevation (*el*) was directly retrieved from Geographic Mapping Technologies (GMT) software (http://gmt.soest.hawaii.edu/) and the ETOPO1 one arc-minute global relief model (Amante and Eakins, 2009).

Incorporating the impact of farming

The appearance and spread of farming had a large impact on the demography of human populations, and therefore needs to be taken into account in the model (Bocquet-Appel and Naji, 2006; Currat and Excoffier, 2005; Luca et al., 2010). First, a farming subsistence strategy enables human populations to reach higher population densities than hunter-gathering. This difference was accommodated as follows: While the maximum population density (D_{max}) was fixed at 5 individuals per km^2, consistent with the population density of farmers (Hassan, 1975), hunter-gatherers had a maximum population density of 0.1 individuals per km^2. Hunter-gatherers were thereby attributed 1/50 of the carrying capacity of each deme.

We also considered these two subsistence strategies as distinct cultural choices and modelled farmers and hunter-gatherers as two separate cultural groups (Currat and Excoffier, 2005; Itan et al., 2009). While hunter-gatherers initially occupied the world 40,000 years BP (i.e. the time when the simulation started), farmers appeared with the origin of agriculture. Following archaeobotanical evidence, which suggests in East Asian agriculture is likely to have started independently in distinct locations (Fuller et al., 2011), in our model, we allowed farmers to appear independently in 157 demes. These demes correspond to locations that showed archaeobotanical evidence for early rice cultivation (Fuller et al., 2011); Figure S3C). We used the median date if there was a time range. For these 157 demes, a number of hunter-gatherers convert to farmers once the simulation generation reached the time for which we have evidence of rice cultivation. The farming conversion rate – the proportion of the hunter-gatherers population of the deme that converts to farmers – has a prior range of [0.20-0.333]. This prior range was chosen after preliminary testing showed that values lower than 0.20 prevented farmers

from outcompeting hunter-gatherers, while values higher than 0.34 made farmers outcompete hunter-gatherers too quickly – both of which contradict the real scenario and therefore were avoided in the simulation.

Farmers were either converted within a deme from the local hunter-gatherers – as explained above – or introduced through other simulated population movements. More specifically, in demes for which there are no dates for farming to start (based on archaeobotanical data), farmers come in either by cultural diffusion or by long-range migration (called sporadic unidirectional migration in Itan *et al*. (2009); see below).

Population movements within and between demes

There are four types of population movements within and between demes: 1) intrademic bidirectional gene-flow between farmers and hunter-gatherers; 2) interdemic bidirectional gene-flow within the same cultural group to one of the eight surrounding demes (*e.g.* farmers to farmers, hunter-gatherers to hunter-gatherers); 3) long-range migration within the same cultural group; and 4) cultural diffusion – individuals of one cultural group convert to the other (*e.g.* hunter-gatherers convert to farmers) based on the presence of that cultural group in surrounding demes. We briefly detail these four types of population movements here, but technical details of how they are simulated can be found in Itan *et al.* (2009).

While the two types of bidirectional gene-flow affect allele frequencies but not population sizes; both long-range migration and cultural diffusion can change population sizes of different cultural groups and their allele frequencies. In particular, the long-range migration rate (or sporadic

unidirectional migration in Itan *et al*. (2009) can take any values within the prior range [0-0.20]. The long-range migration process considers the unfilled carrying capacity in the destination deme as well as the number of demes a cultural group can move through. The standard deviation of the number of demes a cultural group can move through, i.e. the mobility of a cultural group, can take any values within the prior range [0-3]. Cultural diffusion is the process through which individuals from a cultural group within a deme can convert to the other cultural group within the same deme. It occurs with a rate ranging between [0-0.20]. The process accounts for the relative importance of the two cultures in the eight neighbouring demes.

Population growth

We fixed the intrinsic growth rate to 0.3 per generation, a value estimated from data of world population growth rate over the last 10,000 years, excluding the post-Industrial Revolution population boom (US Census Bureau:www.census.gov). We applied this growth rate to both farmers and hunter-gatherers for simplicity, although hunter-gatherers growth rates remain an ongoing debate. The estimated growth rates vary widely in the literature – from 0.001 per annum at the lower end (Hassan, 1975) (~0.03 per generation, assuming 25 years per generation) to 0.03 per annum at the higher end (Young and Bettinger, 1995) (population size triples per generation, assuming 25 years per generation). Because of this uncertainty, we specifically performed sensitivity tests to confirm that a lower growth rate for the hunter-gatherers (*e.g*. 0.04 per generation) would not affect the results.

In each deme, the two cultural groups undergo population growth following a competitive growth density-dependent model. The rationale behind this is that the space occupied by one

cultural group (*i*) in a deme affects how much space is available for the other (*j*). As farmers can reach higher carrying capacities than hunter-gatherers, the former ultimately outcompetes the latter (Currat and Excoffier, 2004). This was modelled as follows:

$N_i^{t+1} = N_i^t(1+r_i(K_{deme}-N_i^t - \alpha_{ij}{}^t N_j^t)/K_{deme})$

Where N_i^t *and* N_i^{t+1} are the population sizes of the cultural group *i* at generation *t* and *t+1*, respectively; r_i is the cultural group growth rate; N_j^t is the other cultural group population size and $\alpha_{ij}{}^t$ is the competitive growth density dependent factor at generation *t*. This factor is obtained as follows:

$\alpha_{ij}{}^t = N_j^t/(N_i^t + N_j^t)$

Note that when N_j^t =0 the remaining population undergoes basic logistic growth. This happens at the beginning of a simulation, as hunter-gatherers occupy the world prior to the appearance of farming, and at the end of a simulation, when farmers have outcompeted hunter-gatherers.

Selection coefficient

Selection on the *370A* allele changes the frequency as follows (Maynard Smith, 1998):

$p_{t+1} = (p_t^2(1+s) + p_tq_t(1+hs))/(1+s(p_t^2+2hp_tq_t))$

where p_{t+1} and p_t are the frequencies of the allele at generation *t* and *t+1*, respectively; q_t is the frequency of the ancestral allele ($q_t = 1-p_t$); *h* is equal to 0.5 here as *370A* is a semi dominant allele; *s* is the selection coefficient, drawn randomly from the range [0-0.20]. In addition, selection on the derived allele increases the number of individuals in the population concerned as follows:

$N^{t+1} = N^t (1+s(p_t^2+2hp_tq_t))$

Selection is effective regardless of the cultural group (hunter-gatherer or farmer), as the subsistence strategy is not known to have driven selection on this allele.

At the end of each simulation, we recorded the parameter values that generated the simulation, including the generation and the location where the allele originated, the cultural group from which it originated, and the derived allele frequency in 29 locations where observed allele frequency data is available (Table S1). Among the total of 1,497,844 simulations performed, 1,083,966 had the derived allele present in the last generation. Only these "successful" 1,083,966 simulations were included in the analysis described below.

Choice of cut-off in the approximate Bayesian computation framework

As there is no specific rule for determining the number of simulations that should be retained for parameter estimation (Wegmann et al., 2009; Bertorelle et al., 2010), we examined the distribution of ranked Euclidian distances based on differences between simulated and observed allele frequencies, and choose the first plateau as the best acceptance. Our reasoning here is that over this plateau Euclidian distances do not differ greatly from one another so there is little justification for accepting one but rejecting another. However, in the range where the Euclidian distances increase sharply we applied rejection.

We examined the ranked 5.5% Euclidean distances for defining a reasonable threshold value that could be used for obtaining posterior parameter estimates. Figure S4 shows the smallest 5.5% Euclidean distances plotted against their rank (ordered by the smallest). The threshold chosen was the larger number of simulations that could be included before there was a noticeable

increase (determined by the end of a plateau) of Euclidean distances. The rationale behind this was that various combinations of parameter can explain the data equally well. Therefore, there is no reason to exclude those simulations whose Euclidean distances are very similar. As the plateau of Euclidean distance values ends after the 5,000th smallest Euclidean distance, we decided to perform the retaining by setting the threshold right after these top 5,000 simulations (0.46%). In order to compare the posterior estimates of parameters using a more stringent threshold we also performed the retaining on the top 1,074 simulations (0.1%). In theory, the smaller the threshold value is, the closer the retained simulated statistics are to the observed and consequently the better the retained parameter sets are for explaining the observed data. However, retaining too few simulations may exclude parameter sets that would explain the observed data well, even though those parameter sets are not within the number of closest Euclidean distances accepted. This over-fitting may skew the posterior estimates of the parameters. Even though we present parameters estimates using the best 0.1% of the simulations, we chose to base our inferences on the best 0.46% in order to avoid over-fitting.

The posterior distributions of the parameters were very similar amongst those two threshold rejection values (Figure S2). The posterior ranges are of the same order of magnitude, but the mode values for the selection coefficient are different between the two thresholds (Tables S2, S3). This illustrates that various combination of parameter values are compatible with the data. It also demonstrates the importance of accounting for the stochasticity of population demography when investigating evolutionary histories.

We performed a non-parametric correlation test (Spearman's rank correlation) to check the independence of the parameters in our model. Among the 36 possible pairs of the 9 parameters, only 4 pairs showed significant but weak correlation (Table S4; 3 pairs if using the top 0.1% simulations (data not shown)), suggesting that the parameters used in our model were generally driving different aspects of the data.

S2 Estimating selection time of *370A* using the coalescent-based allele frequency spectrum

The allele frequency spectrum

The allele frequency spectrum (AFS) is defined as "the sampling distribution of allele frequency at any random polymorphic locus in the genome". In practice, the AFS categorizes polymorphic loci based on their frequencies of a specific allele (*i.e.*, the derived allele) for all loci discovered in a finite random sample collected from one population, and collects the proportions or numbers for every possible frequency category (Chen, 2012). Assume the sample is of size of *n*, the AFS can be summarized using the set of statistics $\{\mathbb{E}(S_i(n)), 0 < i < n\}$, where *i* denotes that there are *i* copies of derived alleles and $\mathbb{E}(\cdot)$ stands for expectation. The AFS is extensively used in population genetic inference, and is informative for estimating selection intensity and the time of selection (Bustamante et al., 2001). The AFS was usually modelled through diffusion equations, and constructed in a Poisson random field framework (Sawyer and Hartl, 1992). Here we use coalescent theory to analytically model the allele frequency spectrum of SNPs in the vicinity of the selected mutant after a selective sweep. To construct the allele frequency spectrum after a selective sweep, we divided segregating sites into two groups: "ancestral sites" ($\mathbb{E}(S_i^a(n))$), that

existed in the population before selection and "new sites" ($\mathbb{E}(S_i^n(n))$), which arose in the population since the start of selection. The over-all AFS is the sum of these two groups:

$$\mathbb{E}(S_i(n)) = \mathbb{E}(S_i^a(n)) + \mathbb{E}(S_i^n(n)). \quad (1)$$

The part of AFS contributed by "ancestral sites" can be written as

$$\mathbb{E}(S_i^a) = \sum_{m=1}^{n} g_{n,m}(\tau) \times \sum_{\{0\le k\le m, k\le i\}} \mathbb{P}(k \to i|n,m)\mathbb{E}(S_k^a(m)). \quad (2)$$

Here $\mathbb{P}(k \to i|n,m) = \binom{n-m}{i-k}\frac{k_{(i-k)}(m-k)_{(n-m-i+k)}}{m_{(n-m)}}$, is the Polya-Eggenberger distribution (Johnson and Kotz, 1977), $E_k^a(m)$ is the AFS of ancient sites at the time of selective sweep, which was derived by using the Etheridge sampling formula to model the hitch-hiking effect (Etheridge et al., 2006), and can be found in Eqn.15 of Chen (Chen, 2012). And

$$g_{n,m}(T) = \mathbb{P}(B(T) = m) = \sum_{i=m}^{n} \frac{(-1)^{i-m}(2i-1)m_{(i-1)}n_{[i]}}{m!(i-m)!n_{(i)}} e^{-i(i-1)T/4N}, \quad (3)$$

where $n_{(k)} = n(n+1)\dots(n+k-1), k \ge 1; n_{(0)} = 1$, and

$n_{[k]} = n(n-1)\dots(n-k+1), k \ge 1; n_{[0]} = 1$ are the rising and falling factorial functions.

The part of AFS contributed by "new sites" is given by (see Appendix 2 of Chen (2012) for the detailed derivation):

$$\begin{aligned} & \mathbb{E}(S_i^n(n)) \\ = & \sum_{m=1}^{n} g_{n,m}(T)\frac{(n-i-1)!(i-1)!}{(n-1)!}\sum_{k=m}^{n} k(k-1)\binom{n-k}{i-1}\mathbb{E}(T_k)|B(T)\mu,\ 0 < i \le n, \end{aligned} \quad (4)$$

Parameter inference

From the allele frequency spectrum obtained above, we can estimate the probability that a particular SNP is at frequency i out of n is

$$P(i|n,r,T,s) = \frac{\mathbb{E}(S_i(n))}{\sum_{i=1}^{n}\mathbb{E}(S_i(n))}, \quad (5)$$

where *r* is the recombination fraction between the selected mutant and the SNP, and *T* and *s* are the selection occurring time and selection intensity respectively. The full likelihood is constructed by taking the product of the marginal likelihood of all SNPs.

Using publicly available data from the 1000 Genomes Project, we randomly picked 50 haplotypes from the CHB population. We excluded two haplotypes that carry the ancestral allele at the *V370A* position and finally kept 48 haplotypes for further analysis. A region of length 850kb was used in the analysis. The derived alleles of all SNPs within this region were determined by comparing with the Chimpanzee sequence as an out-group. After removing SNPs with unknown ancestral states, there are 1,677 SNPs remaining in the analysis.

We first constructed the AFS as shown in Figure S5A. The maximum likelihood estimate (MLE) was achieved by optimizing the likelihood function over *s* and *T*. The likelihood surface is smoothed by B-spline with 2 evenly spaced knots. The variation of the MLE is estimated by taking the negative inverse of the second derivate of the smoothed likelihood curve at MLE. For a clear illustration, the marginal likelihood curves for *s* and *T* were presented in Figure S5B-C when the other parameter took the value of MLE.

S3 Generation and statistical analysis of the *370A* knock-in mouse

370A targeting vector construction

A C57BL/6 BAC clone (RP2394P5) spanning chr10: 57959771-58131014 (assembly mm9) was obtained from the BACPAC Resource Center at the Children's Hospital Oakland Research Institute (Oakland, California, USA). After PCR amplification from the BAC, the 3' homology arm (chr10:58062548-58065589 mm9) was cloned into the multiple cloning site of the TV targeting vector using PacI/EcoR1 digest (PCR primers: 5'GCGCTTAATTAATCCCTGGGACCAGCTGGGGATC3' and 5'AATTGAATTCACCACCAGCAACCTGAATCC3'). The 5' homology arm (chr10: 58065590-58069817 mm9) was subcloned in two steps. First, a fragment (chr10: chr10:58,065,590-58,065,851 mm9) was PCR amplified from the BAC to introduce the T58065848C point mutation (PCR primers: 5'AATTGAATTCAGTACGGCCGTGAAAACATGGCGCCACC3'and 5'AATTGAGCTCTGCAGTGTACAAGACAGTCCAAGACAACTC3'). The resulting point mutation (underlined) also generated a new Eag1 restriction site. This fragment was assembled with a second BAC PCR-amplified fragment (chr10: chr10:58065852-58069817, PCR primers:5'AATTGAATTCGGTACCATCTAAGCAGTTACACACTC3' and 5'AATTGCGGCCGCCTTCTCAGAGTTGTATGTAG3') into pBS vector, before shuttling the entire 5'arm into TV using Kpn1/Bsrg1 digest. The TV targeting vector contains a LoxP-pGK-Neo-LoxP cassette for positive and a Diphtheria toxin A (DTA) negative selection cassette. Complete TV vector sequence available upon request. In the final *Edar* targeting vector the LoxP-pGK-NEO-LoxP cassette was inserted in the 3'UTR 30bp downstream of the *Edar* stop codon in exon12. The final vector was sequence verified.

Generation of *370A* knock-in mice

The *370A* targeting construct was linearized and electroporated into J1(129/SV) ES cells (Brigham and Women's Hospital Transgenic Mouse Facility, Boston, USA). Genomic DNA was isolated from 186 G418-resistant and DTA negative ES cell clones. Two targeted ES cell clones (49 and 139) were identified as having been correctly targeted as confirmed by Southern blotting using a SexA1/Afe1 digest and 5' probe (chr10:58071825-58072223), and an AflII digest and Neo probe (amplified from 370A targeting vector using primers 5'CAACAGATGGCTGGCAACTA3'and 5'TGCTCCTGCCGAGAAAGTAT3') to check for correct 3'and unique Neo integration (Figure S6). Presence of the point mutation was confirmed by sequencing and confirmation of the presence of the Eag1 site introduced by the 370A mutation. Clones 49 and 139 were injected into C57BL/6 blastocysts (Genome Modification Facility, Harvard Stem Cell Institute, Cambridge USA) but only clone 139 was found to have generated chimeras with germline transmission. Chimeras were crossed to β-Actin Cre strain (Susan Dymecki, Harvard Medical School, Boston USA) on FVB background to excise the Neo cassette. Subsequently, offspring were backcrossed to FVB mice for a further four generations during which the Cre transgene was outcrossed before phenotyping analysis was carried out. The following primer pairs were used for genotyping *370A* animals (5'CACCGAGTTGCCGTTTGACTG3' and 5'ACAGCTGTCGGTCGTGACTTC3'). Primers used to genotype *379K* were described previously (Charles et al., 2009).

Statistical analysis of mouse phenotypes

All statistical tests for mouse data were performed using IBM SPSS Statistics Software (v.19, IBM). Multivariate analysis of variance (MANOVA) was used to assess the effect of genotype on each ectodermal appendage in the current study due to the fact that each appendage was assessed using multiple measured factors. Results and descriptive statistics are summarized in Supplemental Table S5. For each ectodermal appendage, the multivariate effect of genotype was assessed by Wilks' lambda and is reported with a corresponding F statistic, significance level (*P*), estimate of effect size (partial eta squared) and observed power. The effect of genotype on each measured factor of a given appendage was assessed by an individual F statistic and a corresponding *P*-value. For each significant univariate effect, Least Significant Difference (LSD) posthoc tests were used to assess pairwise comparisons between genotype classes.

Univariate analysis of variance (ANOVA) reveals that *Edar* genotype does not affect body weight at six week of age ($F_{(2,38)} = 0.795$, $P = 0.459$). However, linear regression analysis of 84 individual mammary glands from all genotype classes revealed a small effect of body weight on gland area and fat pad area (other mammary measurements were not affected by weight). Fat pad area was positively correlated with body size (unstandardized $\beta = 0.044$ cm^2/g, $t_{(39)} = 2.714$, $P = 0.01$) although the percent variance explained was low ($R^2 = 0.159$; $F_{(1,39)} = 7.639$, $P = 0.01$). Similarly, gland area was positively correlated with body size (unstandardized $\beta = 0.023$ cm^2/g, $t_{(39)} = 2.393$, $P = 0.02$) although the percent variance explained was low ($R^2 = 0.128$; $F_{(1,39)} = 5.725$, $P = 0.02$). To ensure that body weight did not confound our analysis of the effect of genotype on mammary fat pad and gland area, we excluded animals whose body weights were more than two standard deviations away from the population mean (mean = 20.7 g; SD = 1.1 g). Two outliers were excluded by this criterion. We then used covariance analysis (ANCOVA) to

reanalyze the effect of *Edar* genotype on gland and fat pad area using body weight as a covariate in order to control for the effect of body weight. By ANCOVA, fat pad area is significantly affected by genotype ($F_{(2,35)} = 3.397$, $P = 0.045$) but not body weight ($F_{(1,35)} = 1.823$, $P = 0.186$) whereas gland area is significantly affected by body weight ($F_{(1,35)} = 4.381$, $P = 0.044$) and not by genotype ($F_{(2,35)} = 0.817$, $P = 0.450$).

S4 Association study of *370A* in a Han Chinese population

Phenotype collection and calling

Written consent was requested before the participation in the study. Body temperature, room temperature and humidity were recorded. Individuals with a body temperature higher than 38°C were not admitted into the study.

Sweat gland density

We employed a widely used non-invasive method based on starch/iodine reactions to measure activated sweat glands (Juniper et al., 1964). We studied both thumb and index finger of the less preferred hand of the subjects. After cleaning the subjects' fingers with 70% Ethanol, we applied an iodine/Ethanol solution to the fingertips and allowed approximately one minute for the surface to air-dry. We then applied a thin film of a mixture of starch/castor-oil (1:1 solution). Sweat secreted from the sweat glands broke the oil barrier and initiated the starch-iodine reaction, producing visible black dots that demarcated activated sweat glands. After waiting approximately 2.5 minutes, we took 50× microscope pictures of the centers of each finger, using the skin analyzer Bose View 6100U. The skin analyzer was connected to a laptop, on which we saved the picture files on site. In the collection of the college samples, we took pictures of the

fingers with a high-resolution Nikon D70 digital camera. A label with the subjects' de-identified code was attached to the finger and included in the picture, so that we could track the sample code and correct the size of the pictures using the label as a size standard.

For the village samples whose images were collected by the microscope skin analyzer, an image size of 6mm × 8mm was captured. We manually counted the number of active sweat glands. To ensure the reliability of the counting, we performed a double-scoring scheme. This scheme involved independent assessment of the same images from two individuals (A and B), as well as an additional assessment by a third individual C, when the discrepancy between the two initial counts (A and B) was over a preset threshold X (20% in this study). The deduction of the final count was dependent on the consistency of scoring among A, B and C. The threshold measuring the consistency between C and the mean of A and B and was set as 5% in this study. This scheme was consistently performed whenever double-counting was involved in the study.

For the college samples whose images were collected by the digital camera, a square region of size 11.4 mm × 11.4 mm was cropped from the center of the finger in each photo (using the label as a size standard). The number of activated sweat glands in each cropped image was then counted by a combination of an automatic image-analyzing pipeline and subsequent manual corrections. We developed the image-analyzing pipeline in CellProfiler (Carpenter et al., 2006). In the pipeline, we first converted the image into grayscale and corrected for its illumination. We then subtracted the background by the robust-background global thresholding method. We identified the sweat gland objects – by the two-classes Otsu global thresholding method – as objects whose diameters range from 4 to 12 pixels and whose eccentricities are below 0.8.

Afterwards, the pipeline overlaid the identified sweat glands on the original image for manual correction. The parameter threshold was fine-tuned in this pipeline, so as to yield approximately 30% false negatives while minimize the false positives, which made the subsequent manual correction a relatively light job. We manually corrected for both false negatives and false positives, and recorded the final number of sweat glands in each image.

For each individual, the number of activated glands for both the thumb and index fingers were scored and tallied, unless the image quality was too low to process. We then used the average of the density of glands scored for the two fingers in the analysis. For samples with only one processed image, the final density was taken directly from the available image.

Hair density

We used the microscope skin analyzer (Bose View 6100U) to take 50× pictures of the skin in different body areas including axilla, back, chest, forearm, forehead, leg, scalp, and thigh. The hair density on the forehead and scalp is high, so that we only took one picture for each of these two areas. For the other areas where hair density is relatively lower, we took 2-6 pictures before taking the mean to increase the reliability of the data. For areas like axilla and scalp where hairs are normally long, we trimmed the hair in the area prior to imaging.

We manually counted the number of hairs in each pictures. We performed the double counting scheme described above to ensure the reliability and consistency of the manual counting.

Dental traits

Dental molds were made for the participants whose dentition was generally complete. Many of the participants from the villages, particularly the aged participants had relatively incomplete dentition. Their dental molds were not collected in this study. In total, we collected 396 dental molds (200 from the village samples and 196 from the college samples). We analyzed the dental molds by categorizing the traits using the ASU Dental Anthropology System (Turner et al., 1991).

Statistical analysis

We applied a t-test to determine if the means of the two groups (of the genotypes *370A* and *370V/370A*) were significantly different. The sample size of this study is large enough that it can be well-approximated by a normal distribution by the central limit theorem. In fact, the *P*-values generally remained similar when we used the non-parametric Mann-Whitney test (Wilcoxon rank sum test).

We used linear regression to test the association between the *370A* genotypes and the phenotypes by controlling other possible covariates. The empirical *P*-values were based on the Wald test. We adopted the additive model, as supported by the knock-in mouse results in this study as well as previous *370A* studies in humans (Fujimoto et al., 2008; Kimura et al., 2009). It is believed that age and sex are both important determinants of traits related to epidermal appendages (Scobbie and Sofaer, 1987), so we applied both age and sex as covariates in the linear regression. Since the data were collected in two groups of samples from villages and a college, we introduced another covariate, the collection location, to compensate for potential substructure in the

samples. We further tested the association while using room temperature and humidity as covariates. The significant *P*-values all remained significant after controlling these factors.

Finally, to test if the *P*-values of the linear regression analysis hold if the normality assumption is violated, we also used a generalized linear model. Similar *P*-values were obtained.

We used R, PLINK (Purcell et al., 2007), and PRISM to implement the above statistical analysis.

Supplemental References

Amante, C., and Eakins, B.W. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis (19 p.). NOAA Technical Memorandum NESDIS NGDC-24.

Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics *21*, 263-265.

Bertorelle, G., Benazzo, A., and Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. Mol. Ecol. 19, 2609–2625.

Bocquet-Appel, J., and Naji, S. (2006). Testing the Hypothesis of a Worldwide Neolithic Demographic Transition - Corroboration from American Cemeteries. Current Anthropology *47*, 341-365.

Bustamante, C.D., Wakeley, J., Sawyer, S., and Hartl, D.L. (2001). Directional selection and the site-frequency spectrum. Genetics *159*, 1779-1788.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., *et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol *7*, R100.

Charles, C., Pantalacci, S., Tafforeau, P., Headon, D., Laudet, V., and Viriot, L. (2009). Distinct impacts of Eda and Edar loss of function on the mouse dentition. PloS one *4*, e4985.

Chen, H. (2012). The joint allele frequency spectrum of multiple populations: a coalescent theory approach. Theoretical population biology *81*, 179-195.

Currat, M., and Excoffier, L. (2004). Modern humans did not admix with Neanderthals during their range expansion into Europe. PLoSBiol *2*, e421.

Currat, M., and Excoffier, L. (2005). The effect of the Neolithic expansion on European molecular diversity. ProcBiolSci *272*, 679-688.

Etheridge, A., Pfaffelhuber, P., and Wakolbinger, A. (2006). An approximate sampling formula under genetic hitchhiking. The Annals of Applied Probability *16*, 685-729.

Fuller, D.Q., van Etten, J., Manning, K., Castillo, C., Kingwell-Banham, E., Weisskopf, A., Qin, L., Sato, Y., and Hijmans, R.J. (2011). The contribution of rice agriculture and livestock pastoralism to prehistoric methane levels: An archaeological assessment. The Holocene *21*, 743-759.

Hassan, F.A. (1975). Determination of the Size, Density, and Growth Rate of Hunting-Gathering Populations. In Population, Ecology, and Social Evolution, S. Polgar, ed. (De Gruyter Mouton), pp. 27-52.

Johnson, N.L., and Kotz, S. (1977). Urn models and their application: An Approach to Modern Discrete Probability Theory. (New York: Wiley).

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World Map of the Koppen-Geiger climate classification updated. Meteorologische Zeitschrift *15*, 259-263.

Luca, F., Perry, G.H., and Di, R.A. (2010). Evolutionary adaptations to dietary changes. AnnuRevNutr *30*, 291-314.

Maynard Smith, J. (1998). Evolutionary genetics, 2nd edn (USA: Oxford University Press).

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., *et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. AmJHumGenet *81*, 559-575.

Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., *et al.* (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature *463*, 757-762.

Sawyer, S.A., and Hartl, D.L. (1992). Population genetics of polymorphism and divergence. Genetics *132*, 1161-1176.

Scobbie, R.B., and Sofaer, J.A. (1987). Sweat pore count, hair density and tooth size: heritability and genetic correlation. HumHered *37*, 349-353.

Shi, H., Zhong, H., Peng, Y., Dong, Y.L., Qi, X.B., Zhang, F., Liu, L.F., Tan, S.J., Ma, R.Z., Xiao, C.J., *et al.* (2008). Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations. BMCBiol *6*, 45.

Stanyon, R., Sazzini, M., and Luiselli, D. (2009). Timing the first human migration into eastern Asia. JBiol *8*, 18.

Turner, C.G., II, Nichol, C., and Scott, G. (1991). Scoring procedures for key morphological traits of the permanent dentition: The Arizona State University Dental Anthropology System. In Advances in Dental Anthropology, M. Kelley, and C. Larsen, eds. (New York: Wiley-Liss), pp. 13-31.

Wegmann, D., Leuenberger, C., and Excoffier, L. (2009). Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. Genetics 182, 1207–1218.

Xue, Y., Zhang, X., Huang, N., Daly, A., Gillson, C.J., Macarthur, D.G., Yngvadottir, B., Nica, A.C., Woodwark, C., Chen, Y., *et al.* (2009). Population differentiation as an indicator of recent positive selection in humans: an empirical evaluation. Genetics *183*, 1065-1077.

Young, D.A., and Bettinger, R.L. (1995). Simulating the global human expansion in the Late Pleistocene. Journal of Archaeological Science *22*, 89-92.