

**Supplementary Table 1. Sampling locations and sample sizes for mice and soil samples.**

Site name	Location	Latitude	Longitude	Habitat	N (pheno)	N (soil)	N (geno)
Colome	South Dakota: Tripp Co., cornfield near Colome, SD	N43°13.485'	W99°40.353'	Off	30	30	31
Dogear	South Dakota: Tripp Co., near Dog Ear Lake	N43°09.766'	W99°54.581'	Off	13	15	14
Sharps	Nebraska: Cherry Co., Sharp's Campground, Sparks, NE	N42°52.709'	W100°15.621'	Off	33	31	33
Ballard	Nebraska: Cherry Co., Ballards Marsh State Wildlife Management Area, Valentine, NE	N42°36.147'	W100°33.733'	SH	33	34	13
Gudmund.	Nebraska: Hooker Co., Gudmundsen Sand Hills laboratory, Whitman, NE	N42°04.881'	W101°28.154'	SH	24	23	24
SHGC	Nebraska: Hooker Co., Sand Hills Golf Club; Mullen, NE	N41°51.605'	W101°05.117'	SH	15	18	15
Arapaho	Nebraska: Arthur Co., Arapaho Prairie, Arthur, NE	N41°28.939'	W101°50.966'	SH	26	26	26
Lemoyne	Nebraska: Keith Co., Lemoyne, NE	N41°17.098'	W101°50.329'	SH	14	14	14
Ogallala	Nebraska: Keith Co., cornfield in Ogallala, NE	N41°09.824'	W101°39.640'	Off	24	24	24
Grant	Nebraska: Perkins Co., cornfield in Grant, NE	N40°54.612'	W101°43.631'	Off	30	30	30
Imperial	Nebraska: Chase Co., cornfield in Imperial, NE	N40°32.270'	W101°37.240'	Off	24	25	25

The name and location of each of 11 sampling sites are provided. “Habitat” refers to whether the site is on or off of the Sand Hills (“SH” and “Off”, respectively). “N” indicates the final sample sizes for analysis of phenotypic (“pheno”), environmental (“soil”), and genetic (“geno”) variation. Note: juveniles were genotyped but not scored for phenotype.

**Supplementary Table 2. Strength of phenotypic association for previously identified candidate SNPs.**

Candidate (2013)	Traits (2013)	Position in scaffold	Dorsal brightness	Dorsal hue	Ventral brightness	D-V boundary	Tail stripe
31327	Tail stripe	25175452	<b>0.082</b>	0.024	0.045	<b>0.046</b>	<b>0.175</b>
77047	D-V boundary	25128625	0.055	0.023	0.032	0.024	0.059
118834	Dorsal hue	25085120	<b>0.203</b>	<b>0.047</b>	<b>0.122</b>	0.026	<b>0.123</b>
$\Delta_{ser}$	Tail stripe; ventral color	25075943	<b>0.543</b>	<b>0.049</b>	<b>0.096</b>	<b>0.497</b>	<b>0.137</b>
147347	Tail stripe	25055240	0.051	0.035	0.037	0.026	0.058
147789	D-V boundary	25054805	<b>0.069</b>	0.023	0.044	0.037	0.049
Mean			0.055	0.026	0.038	0.033	0.058

Candidate (2013) and Traits (2013) refer to significant genotype-phenotype associations identified in Linnen *et al.* (2013); remaining columns pertain to this study. Values in cells correspond to the posterior inclusion probabilities (PIP) for each trait and each candidate SNP (location on scaffold 16 is in base pairs). Gray boxes indicate SNPs that exceed our PIP threshold (>0.1). Bolded numbers exceed background levels of association for that trait (mean + 0.01; mean PIP across all SNPs for each trait is given in the last row).

**Supplementary Table 3. Hyper-parameters describing the genetic architecture of five color traits.**

Trait	h	PVE	rho	PGE	n_gamma
Dorsal brightness	0.66 (0.37, 0.92)	0.79 (0.60, 0.95)	0.76 (0.43, 0.98)	0.87 (0.68, 0.99)	132.10 (7, 291)
Dorsal hue	0.32 (0.10, 0.59)	0.31 (0.11, 0.54)	0.42 (0.03, 0.92)	0.41 (0, 0.93)	73.44 (0, 276)
Ventral brightness	0.40 (0.20, 0.63)	0.45 (0.25, 0.66)	0.57 (0.18, 0.93)	0.62 (0.23, 0.95)	92.43 (3, 277)
D-V boundary	0.57 (0.30, 0.82)	0.68 (0.48, 0.87)	0.78 (0.47, 0.98)	0.87 (0.68, 0.99)	62.37 (6, 214)
Tail stripe	0.80 (0.53, 1.00)	0.88 (0.71, 1.00)	0.82 (0.60, 0.98)	0.88 (0.73, 0.99)	196.64 (60, 296)

All parameters were estimated using a BSLMM model in GEMMA, 8,616 SNPs, and a relatedness matrix estimated from genome-wide SNPs. Parameter means and 95% credible intervals (lower bound, upper bound) were calculated from ten independent runs per trait. Interpretation of hyper-parameters is as follows: *PVE*, the total proportion of phenotypic variance that is explained by genotype; *PGE*, the proportion of genetic variance explained by sparse (*i.e.*, major) effects; *h*, an approximation used for prior specification of the expected value of *PVE*; *rho*, an approximation used for prior specification of the expected value of *PGE*; *n\_gamma*, the expected number of sparse (major) effect SNPs.

**Supplementary Table 4. Posterior inclusion probabilities for candidate SNPs identified in association mapping analysis of all SNPs.**

SNP Position	Dorsal brightness	Dorsal hue	Ventral brightness	D-V boundary	Tail stripe	Maximum
scaffold00016:25175228	0.0470	0.0086	0.0279	0.7796	0.0412	0.7796
scaffold00016:25075943	0.4683	0.0196	0.0479	0.2306	0.0818	0.4683
NW_006502179.1:31437	0.0735	0.0073	0.0203	0.0036	0.3777	0.3777
scaffold00016:25075903	0.0237	0.0080	0.0107	0.0350	0.3404	0.3404
NW_006501090.1:4441929	0.0287	0.0077	0.0103	0.3168	0.0177	0.3168
NW_006501041.1:4744196	0.0160	0.0084	0.0091	0.2539	0.0248	0.2539
NW_006501185.1:901540	0.2517	0.0089	0.0121	0.0048	0.0214	0.2517
scaffold00016:25076212	0.0428	0.0076	0.0227	0.0156	0.2495	0.2495
NW_006501416.1:2319587	0.0435	0.0068	0.0249	0.0034	0.2466	0.2466
scaffold00016:25085405	0.0111	0.0145	0.0109	0.0295	0.2251	0.2251
NW_006501625.1:1977638	0.2184	0.0062	0.0072	0.0226	0.0137	0.2184
NW_006501710.1:808272	0.0122	0.0102	0.0124	0.0178	0.2157	0.2157
NW_006501970.1:52602	0.0122	0.0085	0.0104	0.2151	0.0535	0.2151
scaffold00016:25074418	0.0153	0.0077	0.0091	0.0067	0.2151	0.2151
scaffold00016:25173672	0.0735	0.0081	0.0925	0.0316	0.2147	0.2147
NW_006501363.1:1300597	0.0149	0.0094	0.0088	0.0038	0.2106	0.2106
NW_006501215.1:2183663	0.0109	0.0075	0.0118	0.0045	0.2096	0.2096
NW_006501164.1:562237	0.0149	0.0077	0.0084	0.0040	0.2042	0.2042
NW_006501505.1:1234861	0.0144	0.0082	0.0222	0.0036	0.2035	0.2035
NW_006501064.1:4448524	0.0104	0.1988	0.0079	0.0039	0.0279	0.1988
scaffold00016:25076355	0.0711	0.0072	0.0129	0.0121	0.1936	0.1936
NW_006501066.1:6401068	0.0096	0.0070	0.0080	0.0044	0.1926	0.1926
NW_006501363.1:1424874	0.0126	0.0099	0.0089	0.1915	0.0212	0.1915
scaffold00016:25108248	0.1841	0.0073	0.0332	0.0077	0.0318	0.1841
NW_006501066.1:6400593	0.0112	0.0092	0.0092	0.1833	0.0195	0.1833
NW_006501258.1:1259668	0.0531	0.0076	0.0156	0.0633	0.1811	0.1811
scaffold00016:25078686	0.1800	0.0070	0.0753	0.0152	0.0334	0.1800
scaffold00016:25175452	0.0577	0.0073	0.0178	0.0167	0.1751	0.1751
scaffold00016:25111157	0.1734	0.0079	0.0379	0.0133	0.0366	0.1734
scaffold00016:25175879	0.0137	0.0097	0.0112	0.1706	0.0166	0.1706
NW_006501149.1:6204758	0.0198	0.0083	0.0138	0.0088	0.1694	0.1694
scaffold00016:25195324	0.0113	0.0084	0.0122	0.1643	0.0750	0.1643
scaffold00016:25117322	0.0351	0.0072	0.0170	0.0041	0.1628	0.1628
scaffold00016:25033802	0.0143	0.0090	0.0196	0.0050	0.1598	0.1598
scaffold00016:25085108	0.1598	0.0210	0.0926	0.0050	0.0740	0.1598
scaffold00016:25058748	0.0115	0.0119	0.0114	0.0089	0.1587	0.1587
NW_006501112.1:2369721	0.0428	0.0071	0.1560	0.0034	0.0215	0.1560



scaffold00016:25103837	0.0402	0.0141	0.0142	0.0092	0.1547	0.1547
scaffold00016:25103836	0.0413	0.0143	0.0145	0.0094	0.1543	0.1543
scaffold00016:25066132	0.0134	0.0077	0.0124	0.1537	0.0149	0.1537
scaffold00016:25108191	0.1531	0.0078	0.0345	0.0074	0.0306	0.1531
scaffold00016:25085120	0.1502	0.0226	0.0793	0.0055	0.0678	0.1502
scaffold00016:25063343	0.0206	0.0080	0.0088	0.1491	0.0166	0.1491
NW_006501047.1:4679252	0.0257	0.0092	0.0079	0.1488	0.0225	0.1488
scaffold00016:25113309	0.1483	0.0109	0.0089	0.0058	0.0160	0.1483
NW_006501634.1:1603622	0.0192	0.0076	0.0112	0.0042	0.1478	0.1478
scaffold00016:25074446	0.0128	0.0167	0.0082	0.0062	0.1415	0.1415
NW_006501392.1:259982	0.0163	0.0078	0.0080	0.0041	0.1403	0.1403
scaffold00016:25076229	0.0242	0.0085	0.0160	0.0157	0.1320	0.1320
scaffold00016:25039123	0.0120	0.0086	0.0103	0.1317	0.0182	0.1317
NW_006501505.1:1621286	0.1308	0.0067	0.0116	0.0090	0.0157	0.1308
NW_006501476.1:2057524	0.0274	0.0095	0.0098	0.0051	0.1298	0.1298
NW_006501197.1:6731032	0.0129	0.0081	0.0092	0.0056	0.1281	0.1281
NW_006501271.1:2908119	0.0296	0.0083	0.0095	0.0148	0.1253	0.1253
NW_006501154.1:2893063	0.0111	0.0075	0.0074	0.0036	0.1249	0.1249
scaffold00016:25066476	0.0108	0.0079	0.0068	0.1240	0.0151	0.1240
NW_006501355.1:2485903	0.1238	0.0125	0.0161	0.0046	0.0198	0.1238
NW_006501156.1:4189217	0.0103	0.0147	0.0076	0.0043	0.1227	0.1227
scaffold00016:25098160	0.0145	0.0080	0.0117	0.1215	0.0349	0.1215
scaffold00016:25063986	0.0182	0.0080	0.0104	0.1214	0.0383	0.1214
NW_006502017.1:177456	0.1211	0.0106	0.0098	0.0037	0.0196	0.1211
NW_006501451.1:1085407	0.0134	0.0082	0.0103	0.0124	0.1205	0.1205
scaffold00016:25084934	0.1204	0.0118	0.0357	0.0087	0.0465	0.1204
NW_006504112.1:718	0.0130	0.0083	0.0077	0.0040	0.1197	0.1197
NW_006501124.1:2767575	0.0094	0.0082	0.0097	0.0043	0.1186	0.1186
NW_006501363.1:1301032	0.0196	0.0079	0.0099	0.1168	0.0190	0.1168
scaffold00016:25085230	0.0102	0.0069	0.0089	0.0038	0.1160	0.1160
scaffold00016:25173030	0.0315	0.0086	0.0131	0.0109	0.1158	0.1158
NW_006501178.1:430728	0.0144	0.0074	0.1152	0.0045	0.0171	0.1152
NW_006501342.1:346813	0.0334	0.0079	0.0205	0.0153	0.1152	0.1152
NW_006502126.1:565845	0.0105	0.0071	0.0075	0.0063	0.1150	0.1150
NW_006501818.1:1377012	0.1143	0.0074	0.0142	0.0043	0.0213	0.1143
NW_006501064.1:4448280	0.1132	0.0070	0.0071	0.0037	0.0262	0.1132
scaffold00016:25086542	0.0306	0.0072	0.0321	0.0074	0.1127	0.1127
scaffold00016:25075977	0.1126	0.0117	0.0368	0.0288	0.0160	0.1126
NW_006501049.1:2210807	0.0131	0.0108	0.0076	0.0047	0.1122	0.1122
NW_006502572.1:27506	0.0134	0.0094	0.0082	0.0561	0.1114	0.1114
scaffold00016:25126621	0.0149	0.0232	0.0084	0.0047	0.1109	0.1109
scaffold00016:25157370	0.0174	0.0083	0.0100	0.0057	0.1087	0.1087
NW_006501094.1:2379444	0.0116	0.0080	0.0227	0.1080	0.0184	0.1080

NW_006501818.1:1376978	0.1078	0.0076	0.0167	0.0048	0.0190	0.1078
scaffold00016:25175932	0.0605	0.0086	0.0349	0.1065	0.0921	0.1065
NW_006501066.1:6860433	0.0388	0.0295	0.0134	0.0051	0.1063	0.1063
NW_006501280.1:3657919	0.1059	0.0073	0.0085	0.0067	0.0162	0.1059
scaffold00016:25195321	0.0125	0.0097	0.0106	0.1053	0.0944	0.1053
NW_006501131.1:1733529	0.0123	0.0103	0.0091	0.0044	0.1043	0.1043
NW_006501237.1:5407969	0.0137	0.0095	0.0092	0.0070	0.1035	0.1035
scaffold00016:25107989	0.1035	0.0084	0.0492	0.0151	0.0521	0.1035
scaffold00016:25155955	0.0099	0.0071	0.0096	0.0050	0.1025	0.1025
NW_006501392.1:260035	0.1024	0.0085	0.0109	0.0051	0.0194	0.1024
scaffold00016:25064166	0.0142	0.0080	0.0102	0.1023	0.0239	0.1023
NW_006501363.1:1420928	0.0088	0.0070	0.0085	0.1018	0.0137	0.1018
NW_006501387.1:1348066	0.1016	0.0080	0.0568	0.0059	0.0211	0.1016
scaffold00016:25063484	0.0136	0.0078	0.0152	0.1011	0.0282	0.1011
NW_006501267.1:5399680	0.0315	0.0080	0.0095	0.0062	0.1011	0.1011

Position of each SNP is given as scaffold:position (in base pairs) in scaffold. *Agouti* is located on scaffold 16. All other scaffolds are “non-*Agouti*”. SNP posterior inclusion probabilities (PIPs) were estimated using a BSLMM model in GEMMA, 8,616 SNPs, and a relatedness matrix estimated from genome-wide SNPs. SNPs were included as candidates in this table if they had a PIP>0.10 for at least one pigmentation trait.

**Supplementary Table 5. Likelihood values for each model, obtained with the full dataset including linked SNPs (3D-MSFS).**

<b>Models</b>	<b>#p</b>	<b>Estimated Log10 (Lhood)</b>	<b>Max possible Log10 (Lhood)</b>	<b><math>\Delta</math>log10 (Lhood)</b>	<b><math>\Delta</math>best model log10 (Lhood)</b>
Topology Off North; (On, Off South)	16	- 287,775	- 287,025	-749	0
Topology (Off North, On); Off South	16	- 287,778	- 287,025	-753	-3

Number of parameters (#p) is given. “Estimated Log10(Lhood)” is the estimated likelihood values for each model, in log10 scale. “Max possible Log10(Lhood)” is the likelihood if there were a perfect fit between the expected and observed site frequency spectrum. “ $\Delta$ log10(Lhood)” is the difference between the estimated and maximal possible log10 likelihood. “ $\Delta$ best model Log10(Lhood)” is the difference in log10 units between the likelihood of each model and the likelihood of the best model; the best model is the model with a  $\Delta$ best model of 0.

**Supplementary Table 6. Parameter estimates obtained for the different demographic models.**

Parameters	Point estimates		95% Confidence intervals	
	Worst model Topology (Off N, On), Off S	Best model Topology Off N; (On, Off S)	Best model Topology Off N;(On, Off S)	
<b><i>Effective sizes (diploid Ne)</i></b>				
N ancestral	19,911	17,262	5,679	25,386
N Off North	38,343	39,847	27,890	58,658
N On	33,871	43,708	30,098	51,162
N Off South	60,166	54,568	39,936	68,150
N ancNorth	87,790	66,502	38,645	86,798
N ancSouth	55,416	56,768	30,953	84,090
N BotOn (100 generations)	32	17	6	209
N BotAncestral (100 generations)	13	884	18	1,379
<b><i>Time of events (years)</i></b>				
Time Split On	4,026	3,709	3,400	7,942
Time Split ancestral	39,410	45,471	37,073	57,782
<b><i>Scaled migration rates (2Nm, number of immigrants per generation)</i></b>				
On to Off N	13.93	12.46	2.6	18.2
Off N to On	6.97	6.40	1.4	13.1
Off S to On	17.82	18.28	12.5	24.0
On to Off S	0.37	3.62	0.0	11.3
Between Off S and Off N	4.99	4.87	1.9	8.6
Between Off N (or Off S) and ancestral On-Off S (or On-Off N)	5.6e-4	3.61e-4	0.0	0.3
<b><i>Migration rates (m, proportion of migrants per generation)</i></b>				
On to Off N	1.59e-04	1.56e-04	2.36e-05	3.27e-04
Off N to On	9.03e-05	7.32e-05	1.79e-05	1.99e-04
Off S to On	2.31e-04	2.09e-04	1.54e-04	3.35e-04
On to Off S	2.67e-06	3.32e-05	3.24e-10	1.39e-04
Between Off S and Off N	5.72e-05	6.11e-05	3.42e-05	8.08e-05
Between Off N (or Off S) and ancestral On-Off S (or On-Off N)	2.84e-09	2.71e-09	6.64e-11	2.57e-06

Description of the parameter tags is given in Supplementary Figure 2. The parameter estimates are similar and consistent across models, pointing to a recent divergence time of the Sand Hills population and to high migration rates. These estimates suggest: (i) a bottleneck associated with the split of the Sand Hills population (N BotOn); (ii) recent split times for the on Sand Hills population (3.4-9.7 kya); (iii) a very old split of Off North (37.1-57.8kya); (iv) asymmetric and high migration rates ( $2Nm > 5$ ), with higher gene flow in the direction south to north, consistent with a southern colonization.

**Supplementary Table 7.  $F_{ST}$  values in the *Agouti* locus.**

<b>Position</b>	<b><math>F_{ST}</math>: On vs Off North</b>	<b><math>F_{ST}</math>: On vs Off South</b>
25103904	0.442782	0.455174
25107989	0.475032	0.466737
25108113	0.452818	0.454937
25108191	0.451366	0.474141
25108248	0.451366	0.463734
25108273	0.422605	0.475471
25111157	0.457267	0.490538
25111390	0.402897	0.49633

Summary of  $F_{ST}$  values “On” the Sand Hills relative to “Off North” of the Sand Hills and separately to “Off South” of the Sand Hills. SNP positions given in basepairs on scaffold 16.

**Supplementary Table 8. Summary of the filtered datasets used for analysis.**

	Population structure (1 SNP per block)	Demographic analyses	Association mapping
Length blocks (kb)	1.5	1.5	1
Min. #accessible sites per block	150	150	NA
Min. median distance among SNPs within a block (in bp)	3	3	NA
Min. depth of coverage DP	4	4	2
Min. individual mean DP	8	8	2
Max. DP	$2*mDP$	$2*mDP$	$2*mDP$
Max. missing data per site allowed	0.1	0	0.1
Max. missing data per individual allowed	0.25	0	0.25
Minor allele frequency filter	$1/(2*n)$	0	0.01
Number of individuals	161	30	236
Number of SNPs	5412	140358	12920
Number of accessible sites	NA	2674174	NA

$mDP$  corresponds to the mean depth of coverage (DP) per individual, and  $n$  corresponds to the number of individuals. Three columns provide details for each analyses presented – the dataset used to assess population structure, demographic analyses, and association mapping.

**Supplementary Table 9. Summary of the dataset based on the random background genomic regions.**

	#Blocks	#SNPs	#Accessible sites	Median #SNPs/block	Median #CallableSites/block
All blocks	52566	649300	6224355	4	100
Filtered blocks	11770	284287	2814532	19	207

These data are used for the population structure and demographic analyses, dividing the scaffolds into contiguous blocks of 1.5kb size. Blocks with less than 150 accessible sites, and with a median distance among consecutive SNPs below 3bp, were discarded. Note: further filters were applied for the population structure and demographic analyses (see Supplementary Table 6).

**Supplementary Table 10. Inference of demographic history.**

Parameter tag	Range units	Parameter range		
		lower	upper	
N ancestral	uniform	10000	1000000	*
Time Split On	uniform	200	50000	
Time Split ancestral	uniform	1	50000	<sup>a</sup>
N North	uniform	1000	400000	
N On	uniform	1000	400000	
N South	uniform	1000	400000	
N ancNorth	uniform	1000	400000	
N ancSouth	uniform	1000	400000	
N BotOn	uniform	10	3000	*
2Nm On to Off N	logunif	1.00E-06	10	*
2Nm Off N to On	logunif	1.00E-06	10	*
2Nm Off S to On	logunif	1.00E-06	10	*
2Nm On to Off S	logunif	1.00E-06	10	*
2Nm Off S to Off N	logunif	1.00E-06	10	*
2Nm Off N (or Off S) to ancestral On-Off S (or On- Off N)	logunif	1.00E-06	10	*

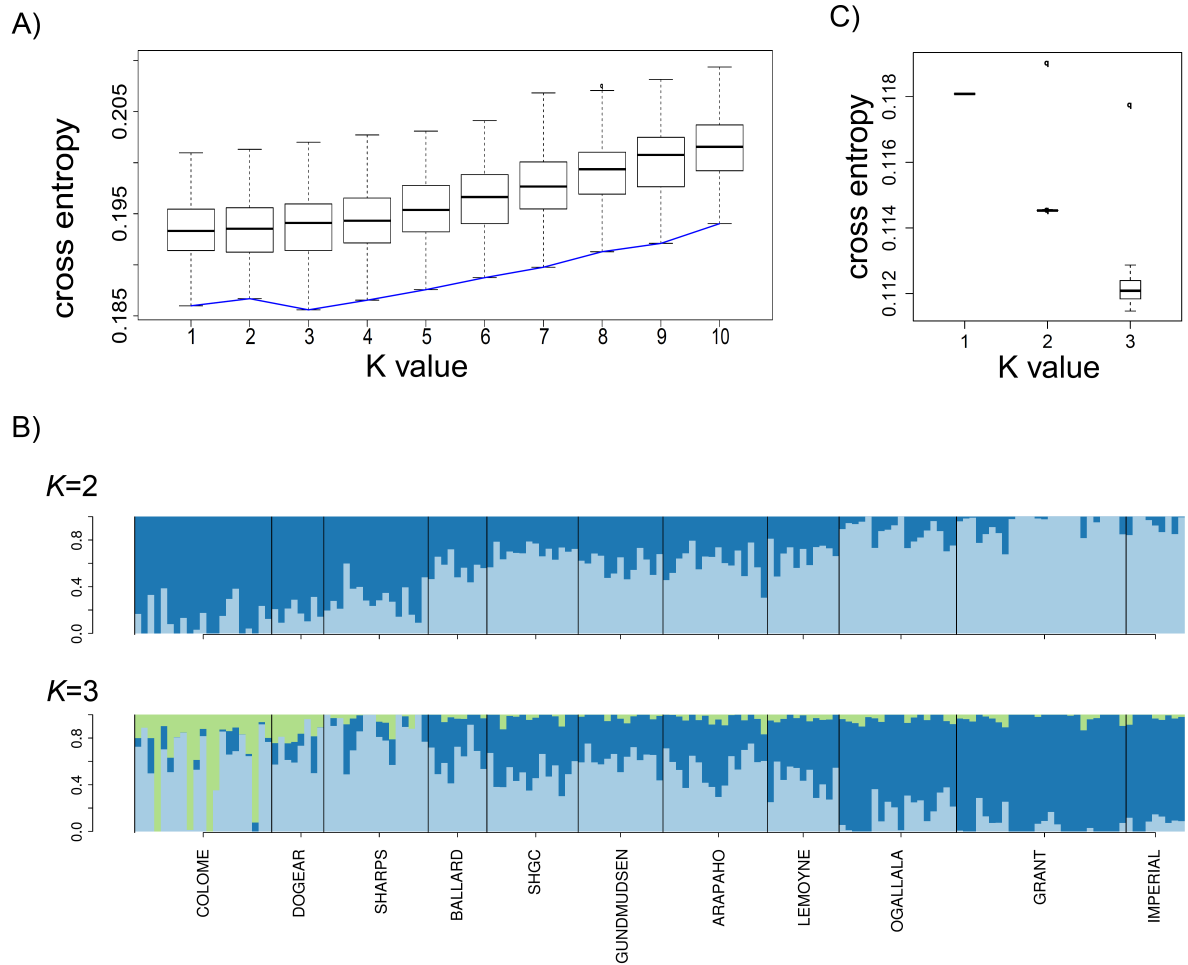
Search range for each parameter (see Supplementary Figure 2 for parameters description). The upper ranges are soft (*i.e.*, if the parameter approaches these values during the optimization, the upper range was increased), except for the parameters indicated with an asterisk (\*). <sup>a</sup>The time of split ancestral is given by adding the range reported to “Time Split On”.



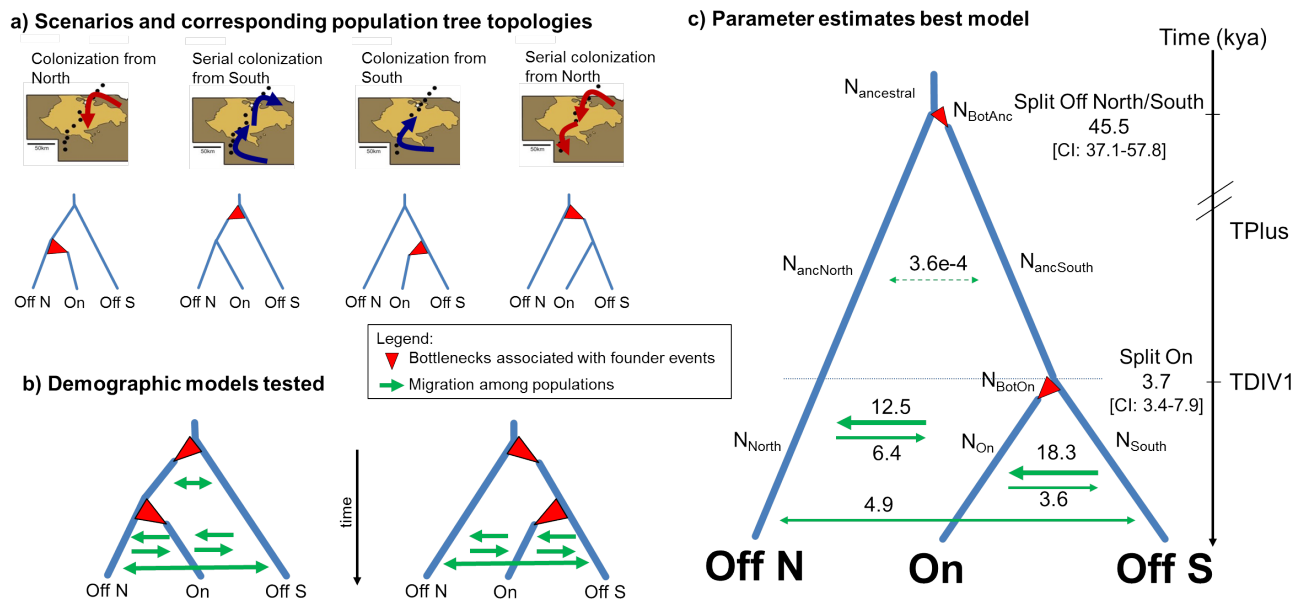
**Supplementary Table 11. Nested ANOVA**

		Source	DF	SS	MS	F	P	PVE
<b><u>Environment</u></b>	<b>Soil brightness</b>	Habitat	1	129.66	129.66	307.40	<b>&lt;2e-16</b>	0.48
		Habitat:Site	9	30.37	3.37	8.00	<b>1.89E-10</b>	0.11
		Residuals	260	109.68	0.42			0.41
<b><u>Color traits</u></b>	<b>PC1 (dorsal bright.)</b>	Habitat	1	87.01	87.01	218.56	<b>&lt;2e-16</b>	0.39
		Habitat:Site	9	37.68	4.19	10.52	<b>8.07E-14</b>	0.17
		Residuals	254	101.12	0.40			0.45
	<b>PC2 (dorsal hue)</b>	Habitat	1	17.60	17.60	18.41	<b>2.53E-05</b>	0.06
		Habitat:Site	9	28.86	3.21	3.35	<b>0.000671</b>	0.10
		Residuals	254	242.83	0.96			0.84
	<b>PC3 (ventral bright.)</b>	Habitat	1	40.76	40.76	60.95	<b>1.53E-13</b>	0.15
		Habitat:Site	9	62.16	6.91	10.33	<b>1.42E-13</b>	0.23
		Residuals	254	169.85	0.67			0.62
	<b>PC4 (ventral hue)</b>	Habitat	1	0.90	0.90	1.08	0.3	0.00
		Habitat:Site	9	8.91	0.99	1.19	0.304	0.04
		Residuals	254	212.12	0.84			0.96
	<b>D-V boundary</b>	Habitat	1	59.32	59.32	104.62	<b>&lt;2e-16</b>	0.25
		Habitat:Site	9	33.93	3.77	6.65	<b>1.54E-08</b>	0.14
		Residuals	255	144.60	0.57			0.61
<b>Tail stripe</b>	Habitat	1	80.65	80.65	170.21	<b>&lt;2e-16</b>	0.35	
	Habitat:Site	9	29.34	3.26	6.88	<b>7.54E-09</b>	0.13	
	Residuals	251	118.94	0.47			0.52	
<b><u>Non-color traits</u></b>	<b>Total length</b>	Habitat	1	0.09	0.09	0.11	0.74	0.00
		Habitat:Site	9	60.27	6.70	8.54	<b>3.89E-11</b>	0.24
		Residuals	250	196.08	0.78			0.76
	<b>Tail length</b>	Habitat	1	0.14	0.14	0.18	0.672	0.00
		Habitat:Site	9	49.12	5.46	7.19	<b>2.85E-09</b>	0.21
		Residuals	250	189.91	0.76			0.79
	<b>Hind foot length</b>	Habitat	1	2.10	2.10	3.01	0.084	0.01
		Habitat:Site	9	105.40	11.71	16.81	<b>&lt;2e-16</b>	0.38
		Residuals	248	172.70	0.70			0.62
	<b>Ear length</b>	Habitat	1	0.35	0.35	0.48	0.491	0.00
		Habitat:Site	9	44.39	4.93	6.70	<b>1.38E-08</b>	0.20
		Residuals	247	181.78	0.74			0.80

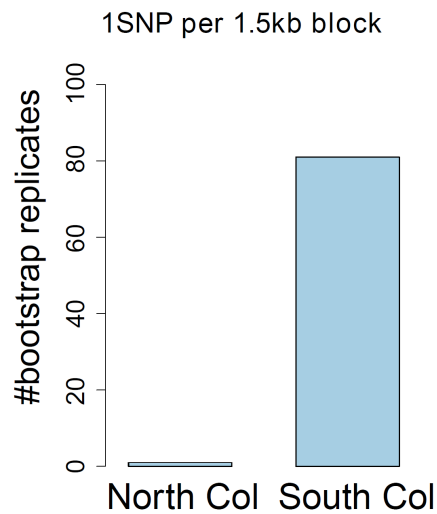
Nested ANOVA tables for environmental, color trait, and non-color trait measurements, with degrees of freedom (DF), sum of squares (SS), mean sum of squares (MS), *F*-statistic (F), *P*-value (P), and proportion variance explained (PVE). “Habitat” refers to whether samples were collected on or off the Sand Hills; “Habitat:Site” refers to collecting locations nested within habitat type. Significant *P*-values ( $\alpha = 0.05$ ) are bolded.



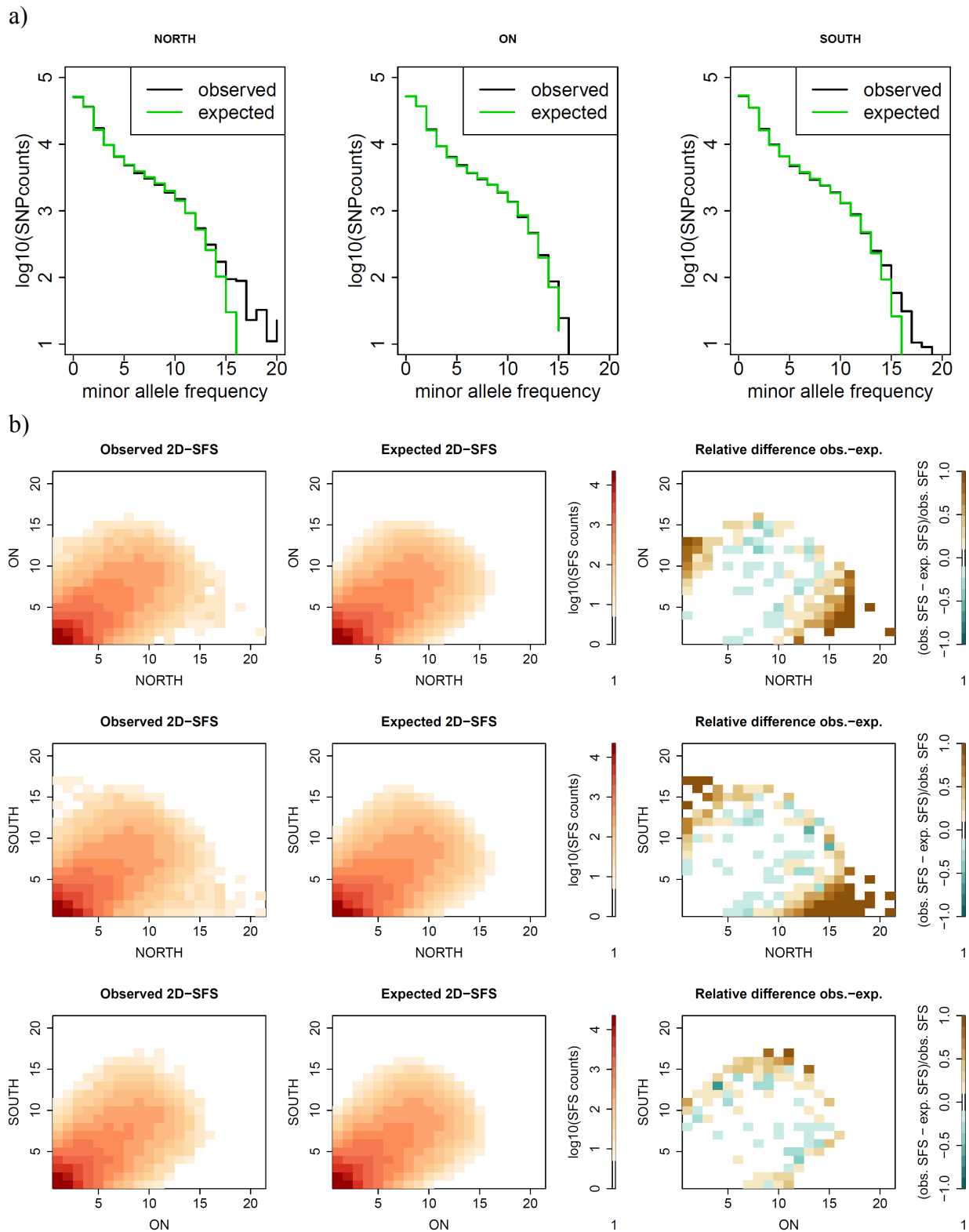
**Supplementary Figure 1. Population structure analyses with sNMF and TESS3.** A) Cross-entropy for each  $K$ -value. Boxplots correspond to the distribution of cross-entropy values obtained by performing 100 runs for each  $K$  value. Blue line represents the minimum cross-entropy for each  $K$  value. The distribution of cross-entropy is similar for values  $K=1-4$ , but the lowest  $K$  value (corresponding to the minimal cross-entropy) is  $K=3$ . B) Estimated ancestry proportions for each individual for  $K=2$  and  $K=3$  with sNMF, without accounting for geographic location of individuals. C) Cross-entropy for  $K$  values from 1 to 3 using the method accounting for geographic information implemented in TESS3. Boxplots correspond to the distribution of cross-entropy values obtained by performing 1000 runs for each  $K$  value. The blue line represents the minimum cross-entropy for each  $K$  value. When accounting for the geographic sampling location of individuals, the cross-entropy decreases for  $K$  from 1 to 3, with a minimum value at 3. Results reported with lambda value of 0.01, but similar results were obtained when changing the lambda values from 0.01 to 0.05 (values nearer to zero give less weight to the geographic location of individuals).



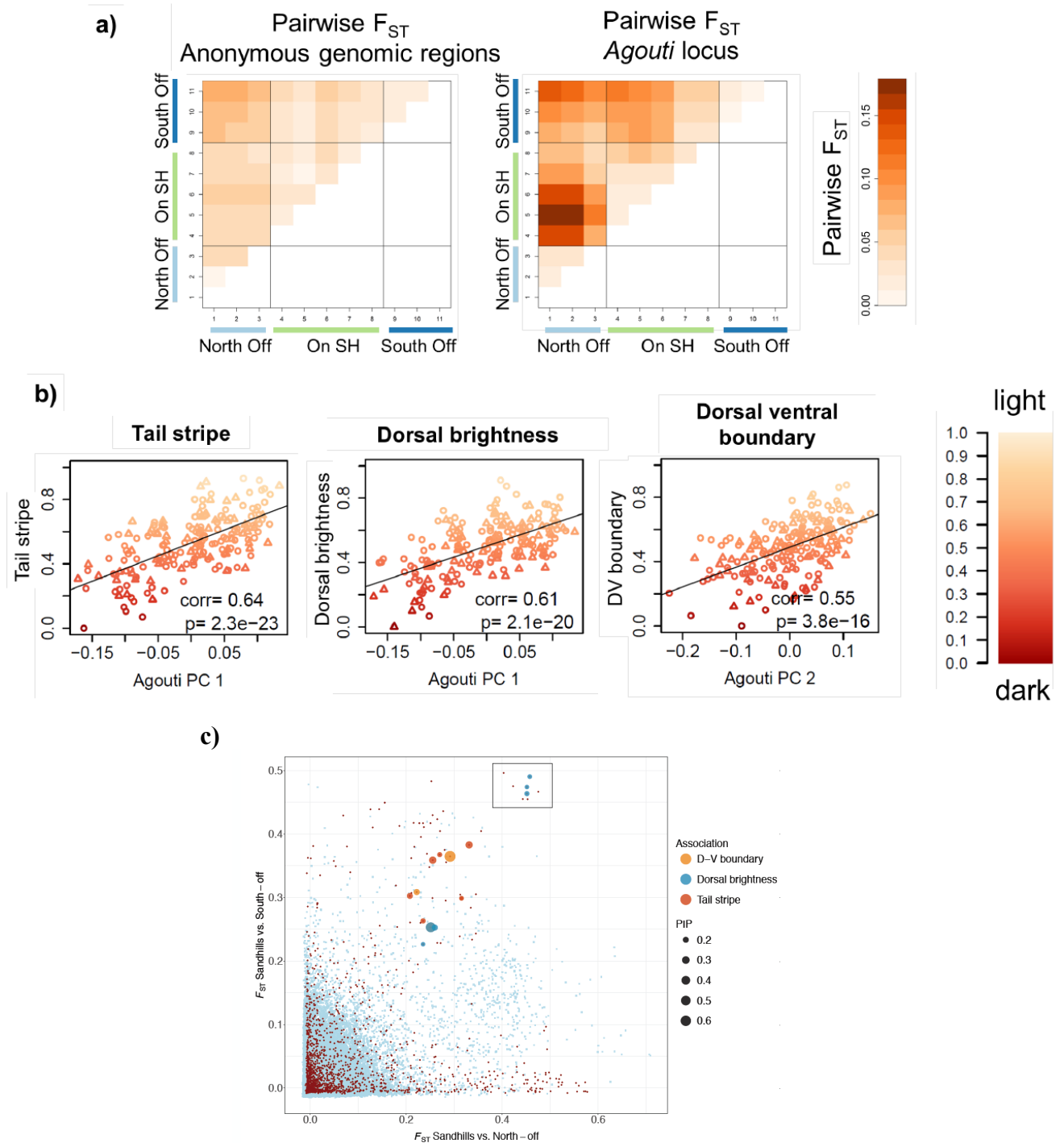
**Supplementary Figure 2. Demographic modeling favors colonization from the south.** (a) Four scenarios were tested and the corresponding three-population model tree topologies are provided. The three samples correspond to populations north of the Sand Hills (Off N), on the Sand Hills (On), and south of the Sand Hills (Off S). Although alternative scenarios might have identical population trees, they can be distinguished as founder events associated with the colonization of the Sand Hills are expected to occur on different lineages. (b) Demographic models tested with two topologies, allowing for up to two bottlenecks associated with the population split, and gene flow among populations. Under this modeling framework, we can distinguish among scenarios with the same tree topology by inferring the strength of the bottlenecks associated with the split of the Sand Hills populations (indicated as red arrows), representing potential founder events related with the colonization of the Sand Hills. For instance, for the tree topology ((Off N, On), Off S), if inferred parameters indicate a stronger bottleneck associated with the Sand Hills (“On”), it would support a colonization from the north scenario, whereas a stronger bottleneck associated with the split of the ancestor of “Off N” and “On” would suggest a serial colonization from the south. (c) Parameter estimates under the best model, supporting a southern colonization route. Parameter estimates for scaled immigration rates  $2Nm$  (i.e., average number of immigrants per generation) are shown above or below corresponding arrows, and the time of events (given in thousand years (kya)) is indicated at the nodes, assuming 2.5 generations per year and a mutation rate of  $3.67 \times 10^{-8}$  per site per generation. Parameter tags are indicated next to each corresponding parameter. Demographic modeling was based on the three-dimensional minor allele frequency spectrum (3D MSFS), considering models with three populations. The 3D folded SFS was generated by dividing the data set into blocks of 1.5kb, sampling ten individuals for each block from each population without missing data - resulting in a 3D folded SFS with 140,358 SNPs.



**Supplementary Figure 3. Model support based on Akaike's information criterion (AIC) across bootstrap replicates.** Only SNPs further than 1.5kb apart were considered, and hence each bootstrap replicate was obtained by sampling one SNP per 1.5kb block. Likelihoods were computed for each bootstrap replicate based on the expected site frequency spectrum for each model (approximated with  $10 \times 10^6$  coalescent simulations) using the parameter estimates that maximized the likelihood with a larger data set (140,358 SNPs), potentially including linked sites. The y-axis represents the number of bootstrap replicates supporting a given model. Each bootstrap replicate was assigned to the model with a relative likelihood (based on the AIC) larger than 0.95. On the y-axis, a value of 100 indicates that a given model was supported by all bootstrap replicates. Note: 14 bootstrap replicates could not be assigned to any model as the relative likelihood was below 0.95 for both models.

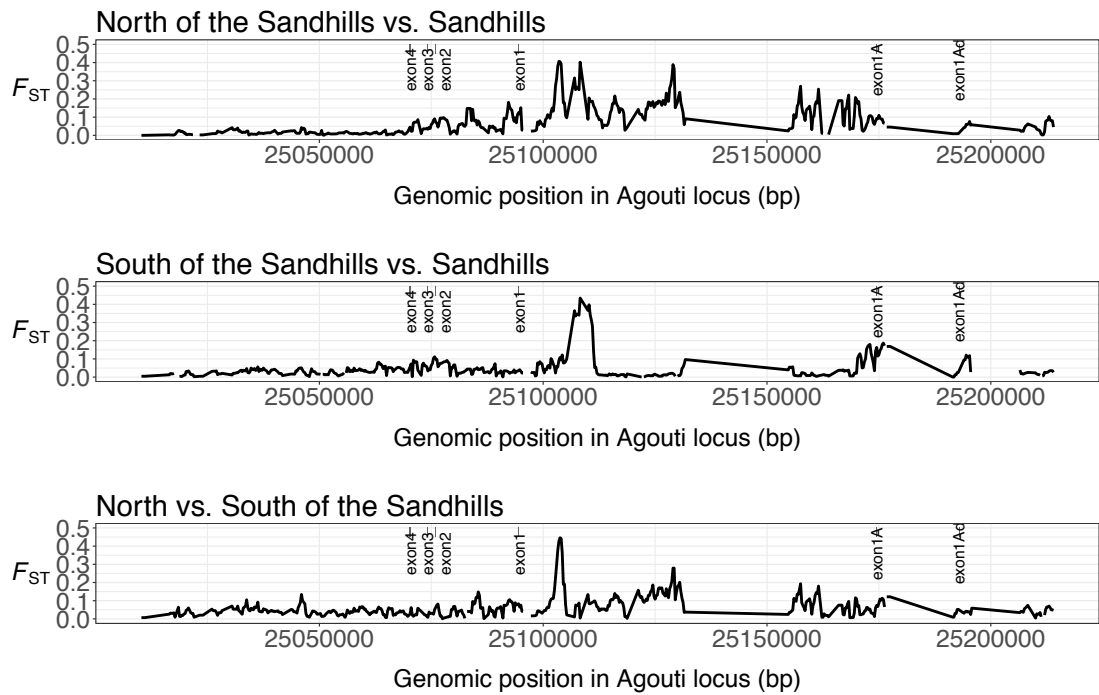


**Supplementary Figure 4: Observed marginal site frequency spectrum (SFS).** (a) 1D and (b) 2D SFS and fit of the marginal minor allele frequency (MAF) spectrum for the best model. The south colonization model fits the MAF spectrum for each sample well. In panel b, the x-and y-axis represent the MAF. As can be seen by the marginal 2D observed SFS, the sample “On” the Sand Hills has more SNPs along the diagonal with the sample south “Off” the Sand Hills, suggesting it is more closely related to the southern population.

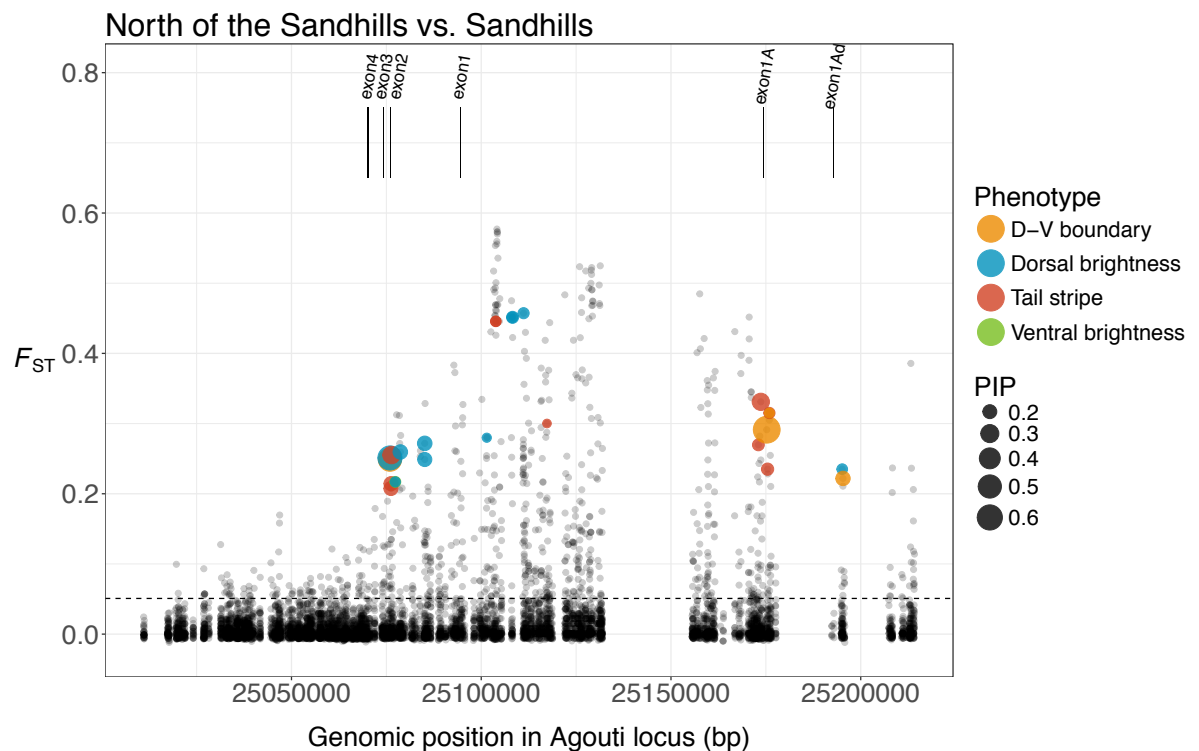


**Supplementary Figure 5. Comparison of genetic differentiation for random background genomic regions and *Agouti*.** (a)  $F_{ST}$  for samples along the transect. At the *Agouti* locus, higher  $F_{ST}$  values are observed, particularly for comparisons between samples On (green) and Off (blue) of the Sand Hills. (b) Genetic differentiation at *Agouti* is correlated with phenotypic trait values. PC1 correlates with tail stripe and dorsal brightness, whereas PC2 correlates with the dorsal ventral boundary, in agreement with Linnen *et al.* 2013 (color-coded with darker phenotypes in darker shades). Circles represent males, and triangles females. (c) The x-axis gives  $F_{ST}$  values for On vs. North, and the y-axis gives values for On vs. South, for *Agouti* (red) and background regions (blue). Points highlighted by a square correspond to the highest peak of differentiation. For SNPs with  $F_{ST}$  values above 0.2, their estimated contribution to the phenotypes considered in the association study are indicated by dot size (PIP > 0.15).

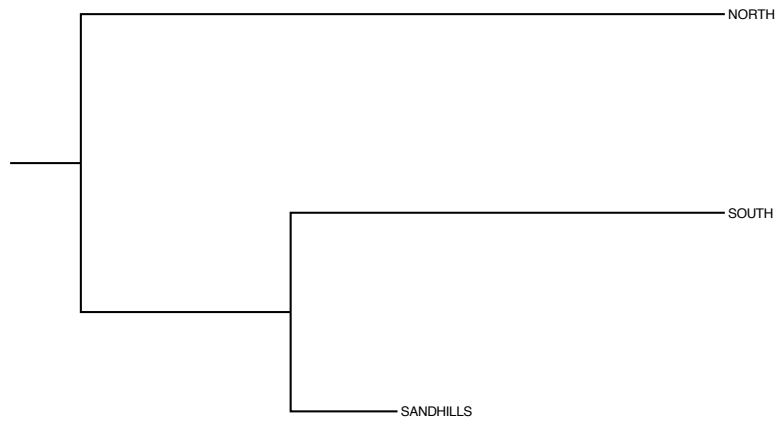
a)



b)

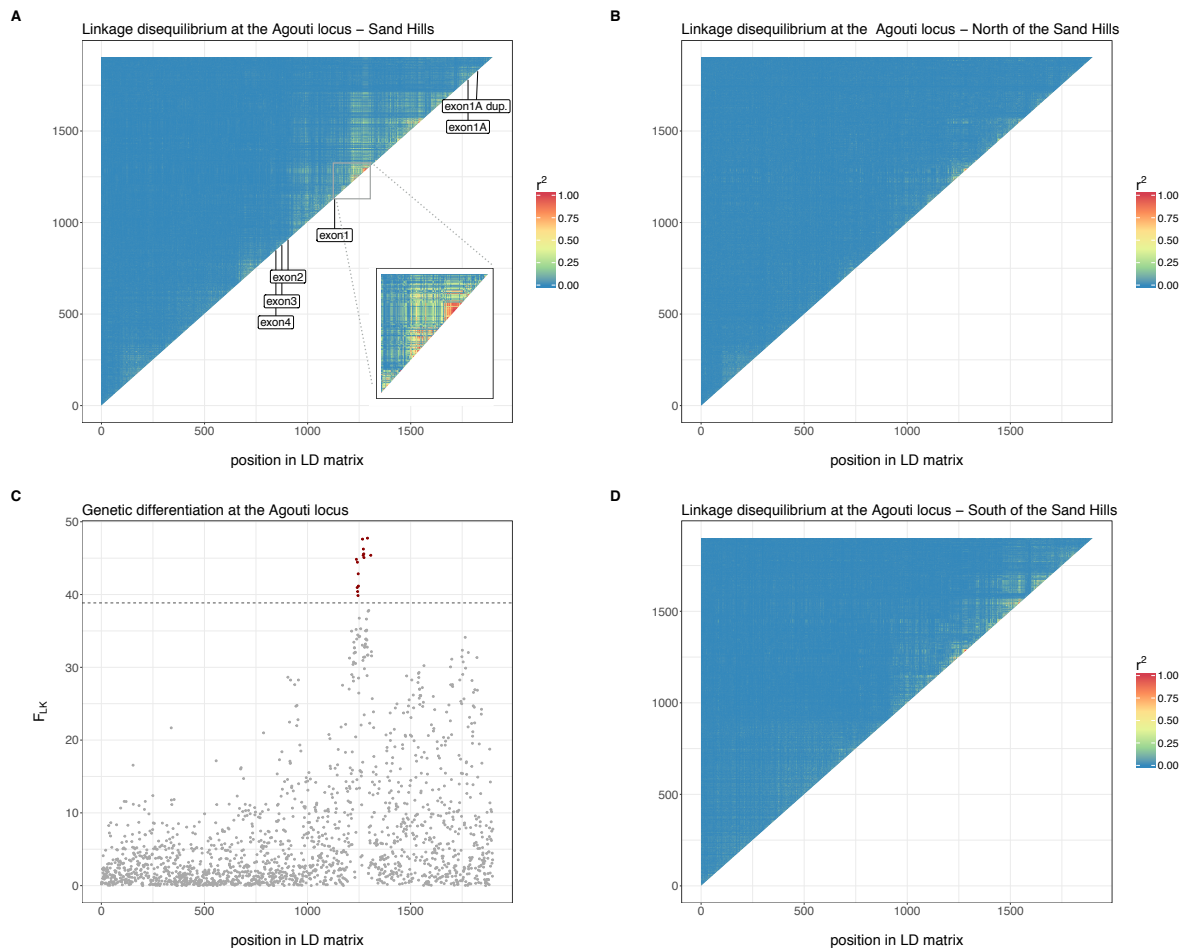


**Supplementary Figure 6. Genetic differentiation ( $F_{ST}$ ) between populations On and Off of the Sand Hills at *Agouti*.** Values are calculated (a) in sliding windows (window size: 1kb, step size: 100bp) and (b) per SNP using VCFtools v.0.1.12b. Vertical lines correspond to the locations of known exons in the *Agouti* locus.

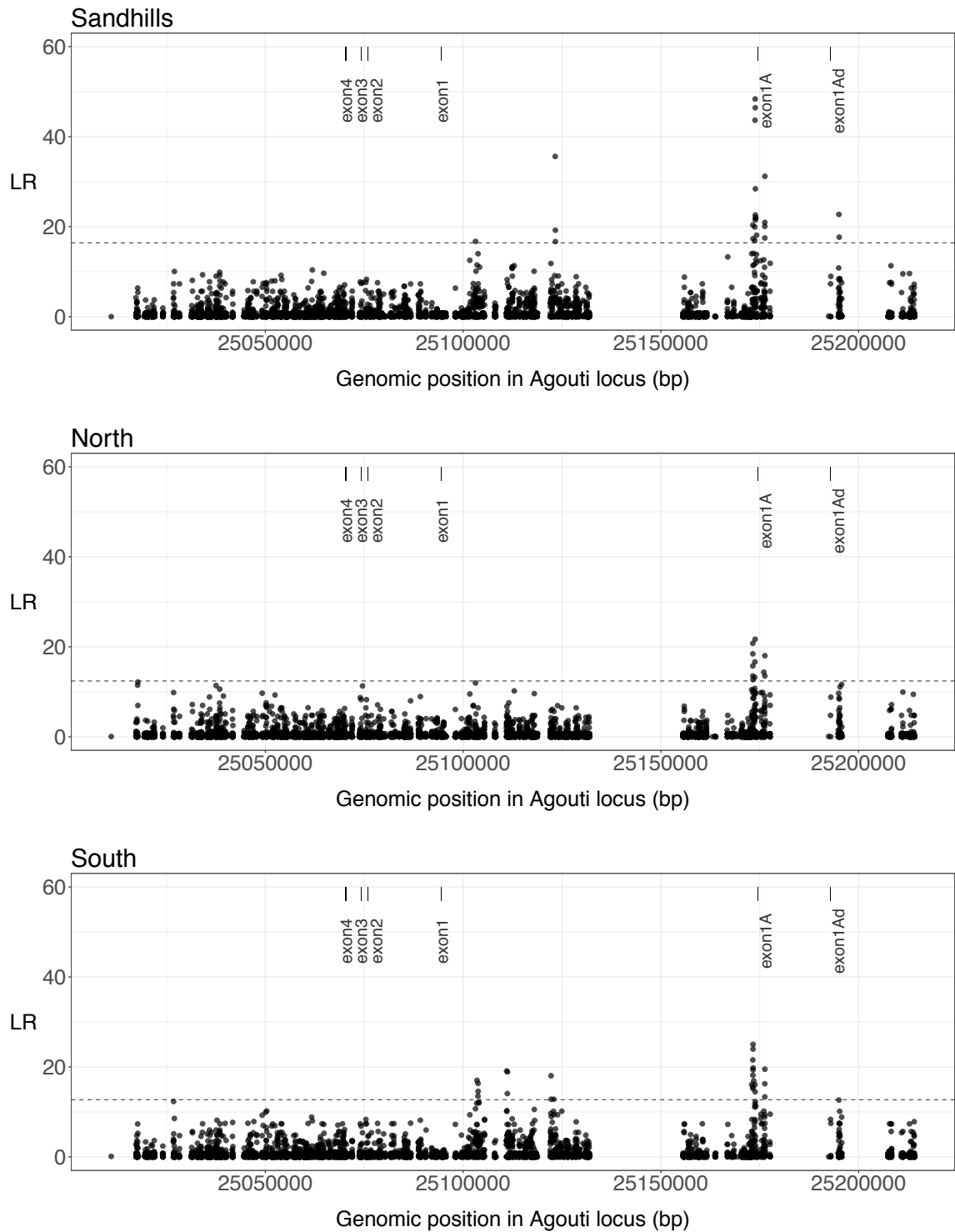


**Supplementary Figure 7. Graphical representation of the kinship matrix.** This tree is output produced by HapFLK, using all background genomic regions (excluding *Agouti*). Note shorter branch lengths represent more similar allele frequencies among populations.

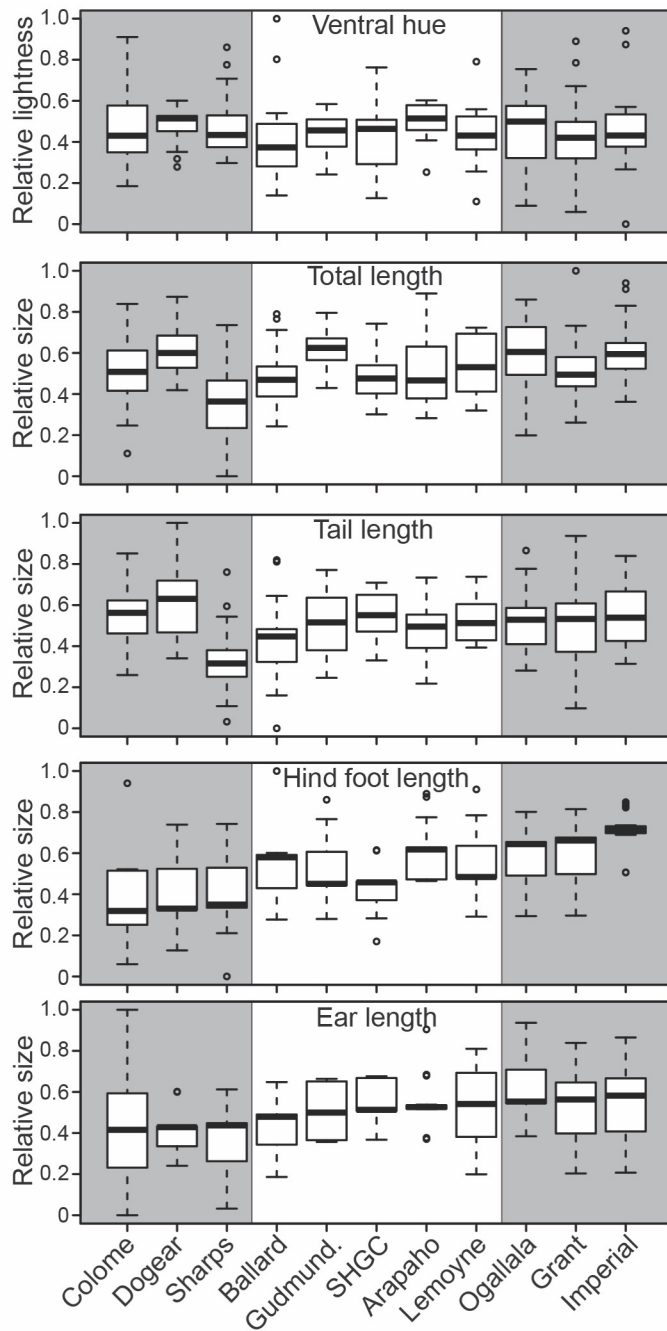




**Supplementary Figure 8. Patterns of linkage disequilibrium (LD) across the *Agouti* locus.**  $r^2$  values have been calculated with vcftools (--hap-r2) excluding all variants with a minor allele frequency strictly lower than 5%. The x-axis of the 4 panels are aligned. A) LD patterns in mice sampled at the Sand Hills. The inset highlights the fine-grained LD patterns around the significant peak of genetic differentiation (panel C). Approximate positions of all exons are indicated. B) LD patterns in mice sampled North of the Sand Hills. C) The FLK statistic represents variant-specific allelic differentiation across the 3 main geographic localities. The same filtering as in the LD analysis has been applied such that x-axes could be aligned. D) LD patterns in mice sampled South of the Sand Hills. For panels C and D, only variants present in the Sand Hills were used so that all three heat maps are constructed with an identical set of variants.



**Supplementary Figure 9. The composite likelihood ratio (CLR) test across the *Agouti* locus.** The CLR test identifies SNPs that deviate from neutral expectations. The CLR statistic was obtained using a pre-computed SFS obtained from the background dataset for the (a) Sand Hills, (b) North of the Sand Hills and (c) South of the Sand Hills populations. Dotted horizontal lines correspond to a theoretical significance threshold calculated for the Sand Hills region using neutral simulations.



**Supplementary Figure 10. Ventral hue and non-color traits are not significantly different on and off the Sand Hills.** Box plots for each sampling site were produced using scaled (range: 0-1), normal-quantile transformed values for: ventral hue (PC4), total body length, tail length, hind foot length, and ear length. In all plots, higher values correspond to lighter/brighter color or longer lengths. Grey shading indicates non-Sand Hills sites. Ventral hue did not differ significantly between habitats (on vs. off Sand Hills) or among sites ( $P > 0.05$ ). All four length traits differed significantly among sites nested within habitats, but not among habitats.