

Current Biology, Volume 27

Supplemental Information

**Parallel Trajectories
of Genetic and Linguistic Admixture
in a Genetically Admixed Creole Population**

Paul Verdu, Ethan M. Jewett, Trevor J. Pemberton, Noah A. Rosenberg, and Marlyse Baptista

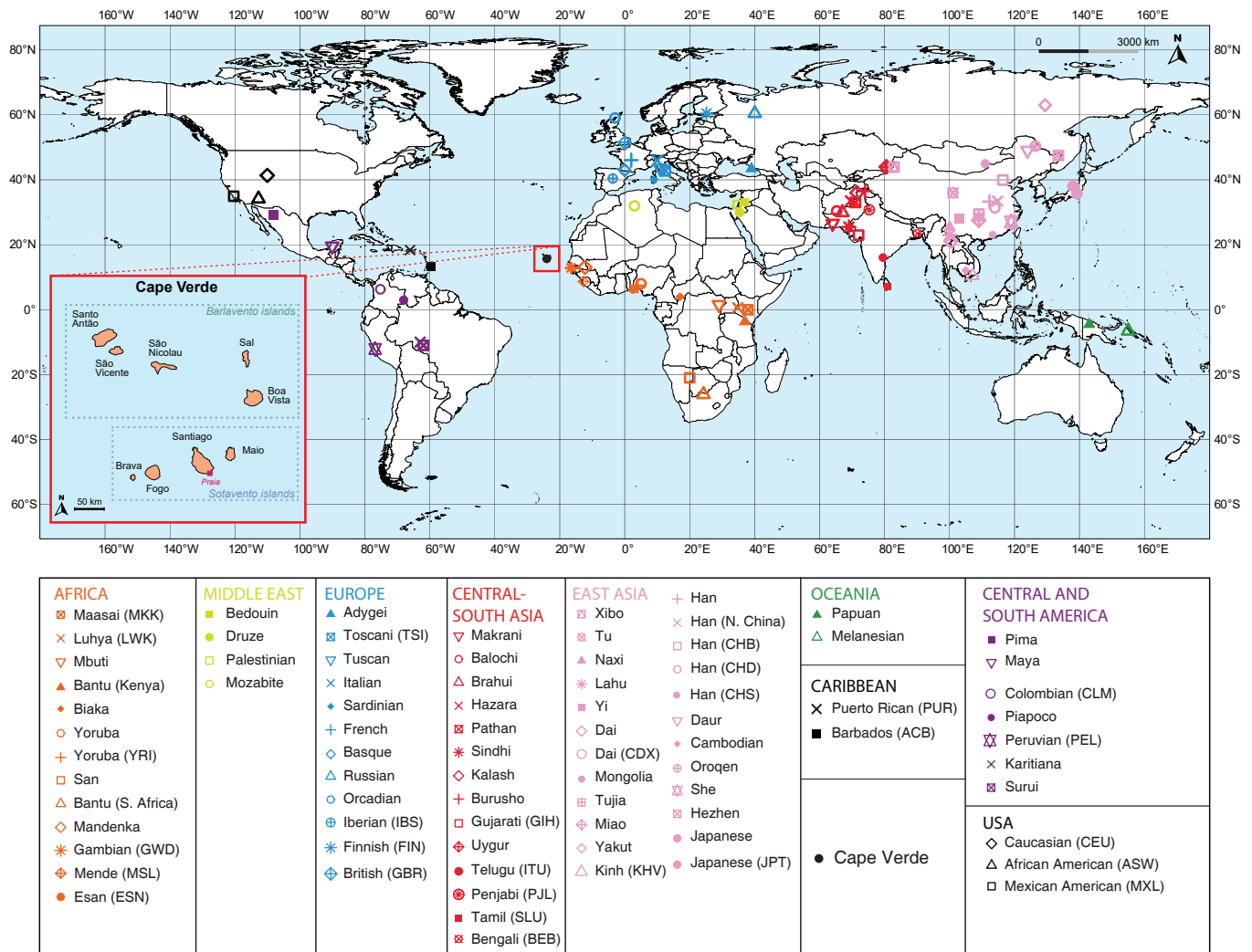


Figure S1. Locations of 79 populations that were combined with the Cape Verde data. Genetic data from these populations were used to produce the dataset used in Figure 1. Portuguese settlement of Cape Verde began in the 1400s, beginning on Santiago around 1460. The remaining islands were settled over the following 400 years.

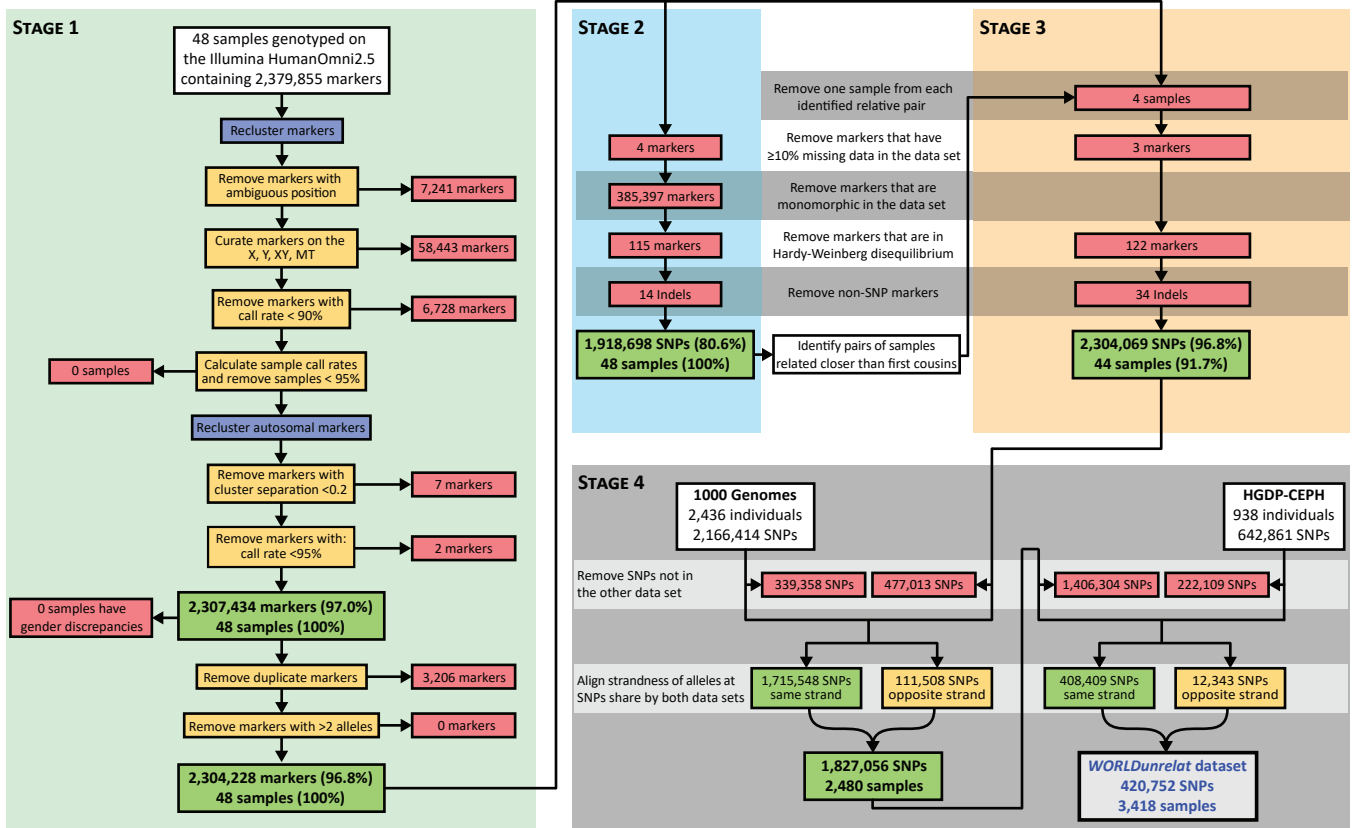


Figure S2. Pipeline for preparation of the genotype dataset used in the study. The data resulting from this pipeline were used in the genetic data analysis in [Figure 1](#).

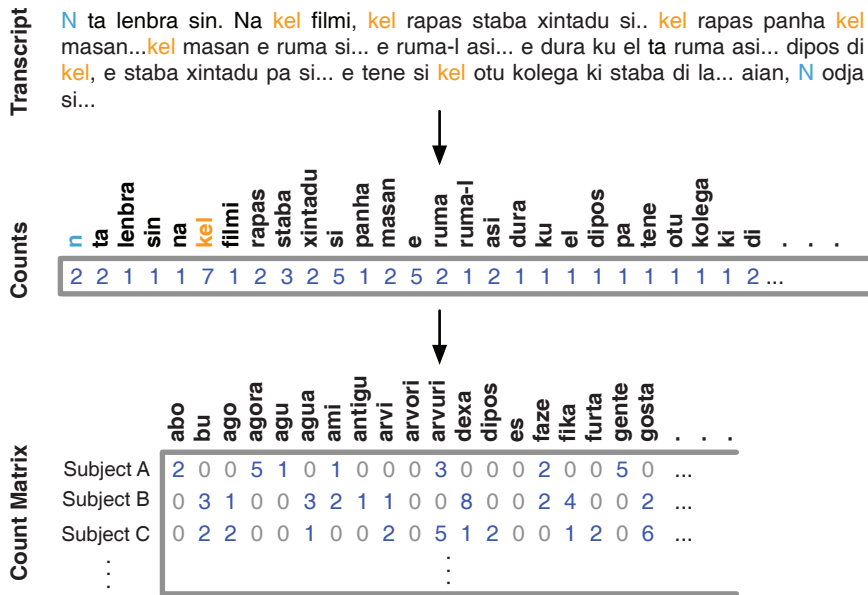


Figure S3. Word frequency profiles. The number of times each unique word appeared in a given transcript was recorded. These counts were then compiled into a matrix with rows corresponding to individual subjects and columns to words. This matrix was used in the linguistic data analyses in [Figure 3](#).

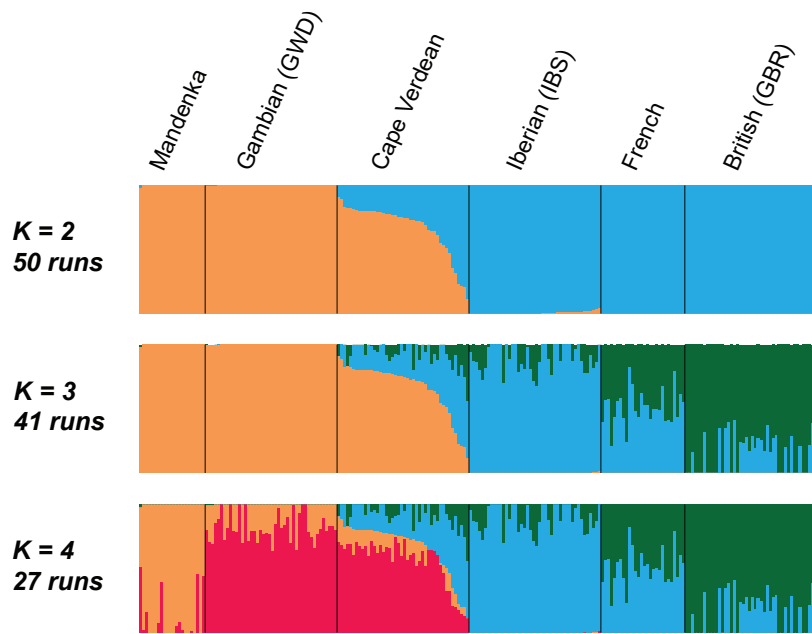


Figure S4. Downsampled ADMIXTURE analysis. A sample of 226 individuals is considered. The analysis and figure design follow [Figure 1C](#).

Table S1. Identification of words of African origin. The table shows the list of words considered in the linguistic analysis in Figure 3.

Word	Category	Description
al	**	Marker signaling desire and probability. From Portuguese <i>há de</i> ‘must do something’ and Wolof <i>yálla</i> , expression invoking God ([S1]: 98; [S6]: 134).
banbudu	*	Adj. “carried on one’s back”. From <i>banbu</i> ‘to carry a child on one’s nback’ on Manding, Manjaku, and Mankañ ([S9]: 78).
djobe/i	*	‘To look’. From Portuguese <i>olhar</i> and Mandinga <i>juubee</i> ([S4]: 150; [S6]: 218; [S9]).
Es	**	Third-person plural pronoun. From Wolof old form lees < li + ees ([S6]: 140). Wolof <i>mëneesu ko</i> ‘one can’t do it.’ Brito in 1888 writes the Cape Verdean Creole pronoun in <i>ēss</i> with a long vowel is reminiscent of the Wolof form ([S3]: 343, 359). Lang proposes a double etymology from Portuguese <i>eles</i> and Wolof <i>ees</i> ([S6]: 141).
fepu	*	Quantifier ‘all/entirely’. From Wolof <i>-épp</i> ‘all/entirely’, Mandinka <i>few</i> ‘completely’ ([S2]: 122).
Ka	**	Negative marker. From Portuguese <i>nunca</i> and different negative markers from Manjaku, in addition to other African languages in the region ([S1]: 95; [S6]: 129).
kabesa	***	Portuguese noun ‘head’. Used as a reflexive in the same way that Wolof head <i>bopp</i> is used as a marker of reflexivization ([S6]: 184). [In our corpus, <i>roba-l riba kabesa</i>].
korpu	***	Portuguese noun used as a reflexive. From Mankañ’s expression with same meaning <i>u-leef</i> ‘the body’ ([S11]: 80).
kotxi	*	Verb ‘to crush (corn or rice)’. Comes from <i>kócci</i> in Mandinka ([S2]: 125; [S8]: 76).
ku	***	Derived from Portuguese preposition <i>com</i> ‘with’ but used as a conjunction of coordination meaning ‘and/with’ and coordinating two nouns much in the same way that Wolof <i>ak</i> does ([S6]: 198).
kunpanheru	***	Portuguese noun meaning ‘companion’ but marking reciprocity in the same way that Bambara does ([S6]: 184).
la	*	Deictic marker signaling distance. From Wolof classifier –l- combined with deictic morpheme –a ([S6]: 147).
li	*	Deictic marker signaling proximity. From Wolof classifier –l- combined with deictic morpheme –i ([S6]: 147).
ma	**	Complementizer. From the older form <i>kuma</i> < Old Portuguese <i>coma</i> and Mandinka verb <i>kuma</i> ‘to talk’ and Mandinka morpheme <i>kó</i> combining the function of a verb ‘to say’ and a subordinate conjunction ([S6]: 129, 159).
moku	**	From Portuguese <i>mouco</i> ‘drunk’ and Wolof <i>mokk</i> ‘to be pounded’ ([S4]: 453) and meaning ‘drunk’ in Cape Verdean, Guinea Bissau, and Casamance Creoles, where the word has the same meaning ([S6]: 218; [S9]).
n	*	First person singular pronoun. From Balanta and Mandinka ([S6]: 129).
nbonji	***†	Type of beans.
nkontra	***†	To meet.
nsoda	***†	The meaning is unclear. <i>Nsoda</i> is not listed in the dictionaries that were consulted.
ntende	***†	To understand.
ntende	***†	To delegate.
ntrega	***†	To fill up.
nu	*	First-person plural pronoun. From identical Wolof pronoun <i>nu</i> with the same person and number ([S6]: 139). Also influenced by Portuguese <i>nos</i> , pronounced nus at the time of creolization ([S6]: 140).
ta	***	Aspectual marker. Functionally (semantically) equivalent to Wolof imperfective <i>-y</i> . <i>Ta</i> is viewed as an imperfective marker ([S6]: 161) that can be morphologically linked to Portuguese <i>estar</i> ([S6]: 162) but functionally behaves like the Wolof imperfective marker.

ten	***	Verb 'to have'. Semantically and morphologically associated with Wolof <i>am/ame</i> to express permanent versus temporary possession ([S6]: 132; [S7]: 55).
tene	***	Verb 'to have.' See 'ten.'
-ba	**	Anteriority marker. From Portuguese <i>-va</i> and Manjaku verb <i>ba</i> 'to finish, to complete' ([S1]: 94; [S6]: 129).
-du	**	Adjectival or past participle morpheme. Derived from Portuguese suffix <i>-do</i> and from the Wolof derivative suffix <i>-u</i> ([S6]: 183).
-m	*	First-person object pronoun. From Wolof or Bambara ([S6]: 140).

*Words and morphemes that trace directly back to an African language. The 9 items marked with a single asterisk are the items included in an African-origin set of particularly high confidence.

**Words that have a dual etymology, corresponding to cases of conflation between the European and the African languages and attesting to the intense contact between the two sets of languages.

***Words or expressions that are clearly Portuguese, but whose semantics or phonology reveal substratal influences.

†Words characterized by prenasal consonants that have been argued to originate from African substratal phonology ([S8]: 81). Though the words themselves are Portuguese, their phonological realization reveals traces of African substrates.

SUPPLEMENTAL REFERENCES

- [S1] Baptista M (2006). When substrates meet superstrate: the case of Cape Verdean Creole. In *Cabo Verde: Origens da sua Sociedade e do seu Crioulo*, eds Lang J, Holm J, Rougé J-L, Soares MJ (Gunter Narr, Tübingen), pp. 91-116.
- [S2] Bartens A (2006) A contribuição do substrato africano para a génese dos crioulos caboverdianos: o caso dos ideofones. In *Cabo Verde: Origens da sua Sociedade e do seu Crioulo*, eds Lang J, Holm J, Rougé J-L, Soares MJ (Gunter Narr, Tübingen), pp. 117-131.
- [S3] Brito A (1967) Apontamentos para a gramática do crioulo que se fala na ilha de Santiago de Cabo Verde. *Estudos linguísticos crioulos: reedição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa*, ed Morais-Barbosa, J (Academia Internacional da Cultura Portuguesa, Lisbon), pp. 329-404.
- [S4] Brüser M, dos Reis Santos A, Dengler E, Blum A, Lang J (2002). *Dicionário do Crioulo da Ilha de Santiago (Cabo Verde)*. Tübingen: Gunter Narr.
- [S5] Lang J (2006). L'influence des Wolof et du wolof sur la formation du créole santiagoais. In *Cabo Verde: Origens da sua Sociedade e do seu Crioulo*, eds J Lang, J Holm, J-L Rougé, MJ Soares (Gunter Narr, Tübingen), pp. 53-62.
- [S6] Lang J (2009). *Les Langues des Autres dans la Créolisation*. Tübingen: Gunter Narr.
- [S7] Quint N (2000). *Le Cap Verdien: Origines et Devenir d'une Langue Métisse*. Paris: L'Harmattan.
- [S8] Quint N (2006) Un bref aperçu des racines Africaines de la langue cap-verdienne. In *Cabo Verde: Origens da sua Sociedade e do seu Crioulo*, eds Lang J, Holm J, Rougé J-L, Soares MJ (Gunter Narr, Tübingen), pp. 75-90.
- [S9] Rougé J-L (1988) *Petit Dictionnaire Étymologique de Kriol de Guiné Bissau et de Casamance*. Bissau: INEP.
- [S10] Rougé J-L (1994) A propos de la formation des créoles du Cap Vert et de Guinée. *Papia* 3:137-149.
- [S11] Santos R (1979). *Le Créole des Iles du Cap-Vert: Comparaison avec les Langues Africaines*. Thesis, Université de Dakar, Sénégal.