

[10.1071/WR22110](https://doi.org/10.1071/WR22110)

Wildlife Research

Supplementary Material

Stable dingo population structure and purity over 11 years of lethal management

Danielle Stephens^{A,}, Malcolm S. Kennedy^B, and Tracey L. Kreplins^C*

^AZoological Genetics, C/PO, Inglewood, SA 5133, Australia.

^BThreatened Species Operations, Department of Environment and Science, 203 Tor Street, Toowoomba, Qld 4350, Australia.

^CDepartment of Primary Industries and Regional Development, 75 York Road, Northam, WA 6401, Australia.

*Correspondence to: Danielle Stephens Zoological Genetics, C/PO, Inglewood, SA 5133, Australia Email: stephens@zoologicalgenetics.com

S1: Isolation by distance

The continuous sampling scheme and use of a large mobile animal in this study means it is likely that there is a strong relationship between genetic and geographic distance. The presence of this relationship must be taken into account in the interpretation of population structure (Meirmans, 2012; Perez *et al.*, 2018), so we first tested for significant isolation by distance within the sample. We performed a Mantel test between genetic and Euclidean geographic distances using the R package *adegenet* v2.1.3 (Jombart, 2008; R core team 2020). The data was randomised for 999 permutations to test for deviation of the observed data from simulated data, i.e., deviation from the null hypothesis of no spatial pattern to the distribution of dingoes within the study area. Geographic and genetic distances were compared and plotted using a two-dimensional kernel estimation (function *kde2d* in the R package *MASS* v7.3-54; (Venables & Ripley, 2002).

Isolation by distance – Results

The observed isolation by distance data showed no overlap with the simulated data (Fig. 1a), indicating that there is a relationship between the genetic distance and the geographic distance ($P < 0.01$). This confirms that the study area is greater than the usual movement area of these dingoes, and they are not a panmictic population at this scale. The plot of geographic and genetic distance (Fig. 1b) shows a single population of points, with no obvious discontinuities. This indicates that the pattern of isolation by distance is the predominant cause of genetic differentiation between the individuals in the study area.

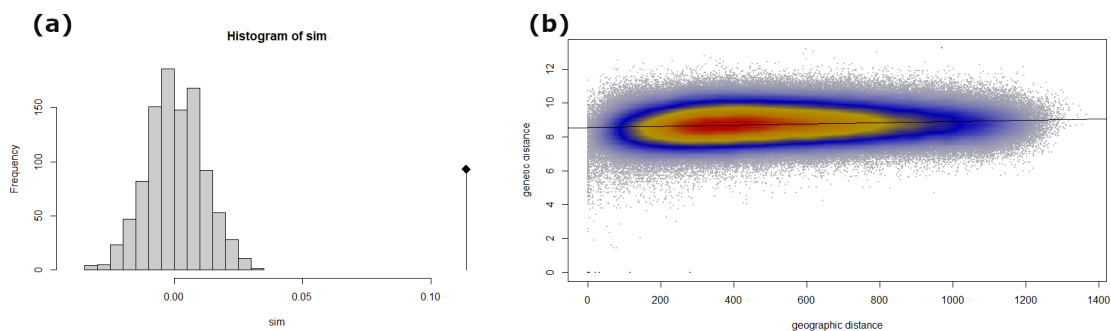


Figure S1: Tests for the presence of isolation by distance in the broader MRVC data set. (a) The frequency of simulated, randomised data (grey boxes) and the observed data (line with diamond), showing significant deviation of the observed data from random, and the null hypothesis of no correlation between genetic and geographic distance. (b) Local density plot of geographic vs. genetic distance, showing a positive correlation between genetic and geographic distance, with a regular distribution pattern more consistent with a cline of differentiation than strongly distinct clusters.

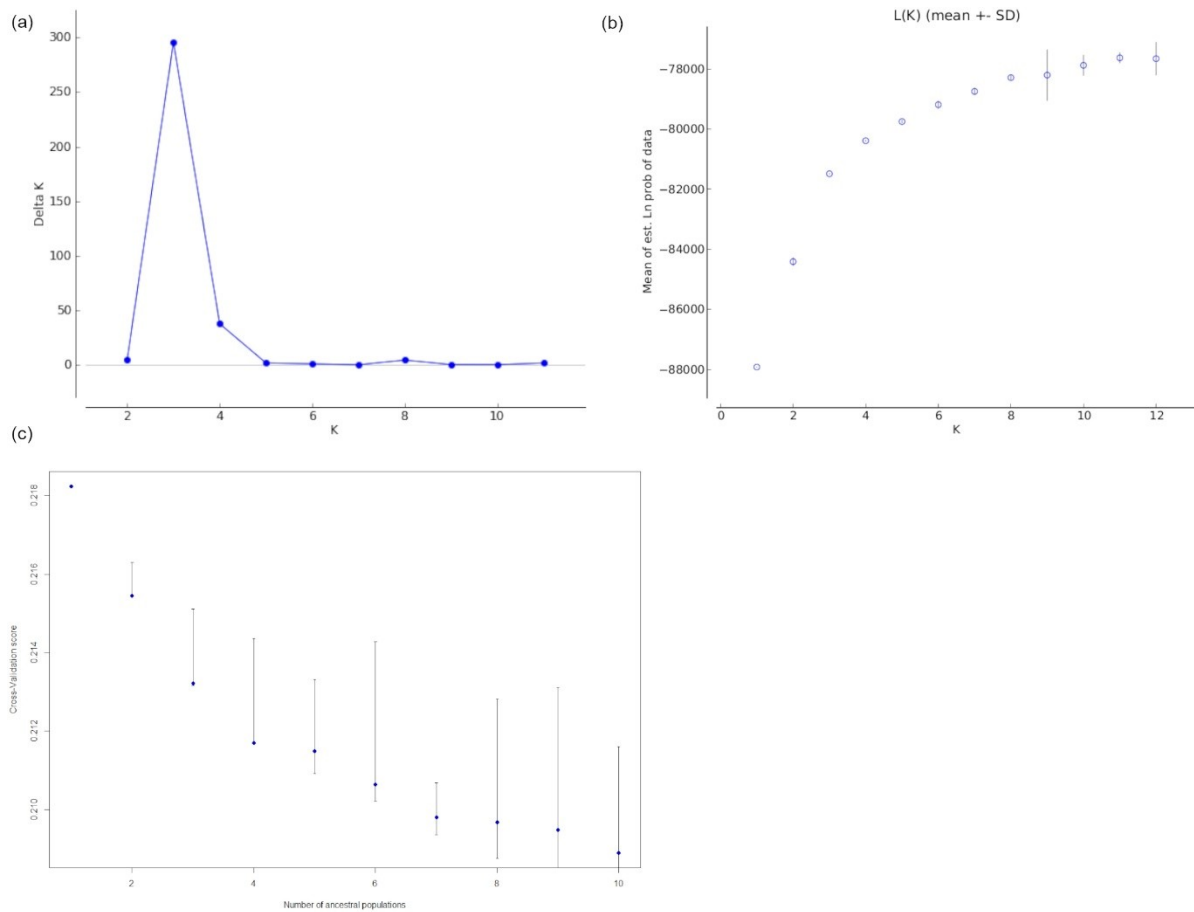


Figure S2: Assessment of the most likely number of populations (K). (a) Delta K assessment of the populations, showing a peak at K=3. (b) The mean of the estimated ln probability for each K, vertical lines are standard deviation from the mean. (c) Cross-validation plot of the likely number of populations determined in TESS3. The x-axis displays the number of ancestral populations (K) and the y-axis is the cross-validation score. The values show a flattening at K=3-4 once error bars are considered, indicating that this is the optimal number of populations determined by this method. K=3 was selected for display for comparison with the Structure results.

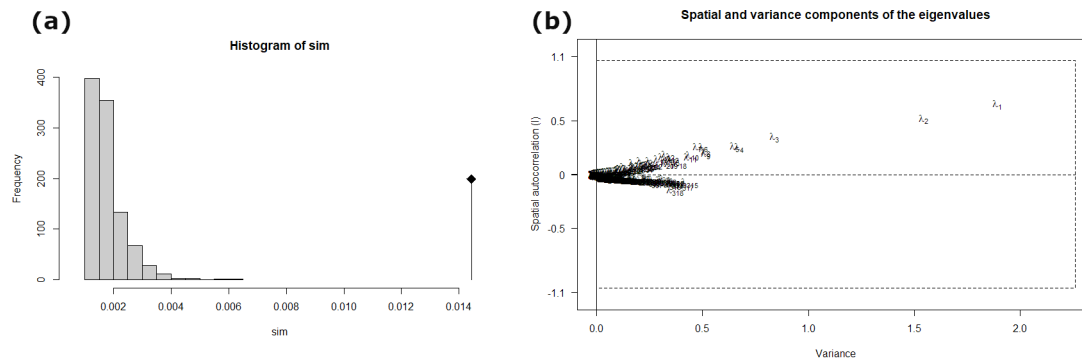


Figure S3: Results of spatial PCA on the broader MRVC data set. (a) Histogram of the randomised simulated data, showing the observed data (line and diamond) is significantly different from the null hypothesis of no spatial structure. The plot shown is for missing loci filled with the averages from allele frequencies, but the result was very similar when missing data were replaced with zeroes, indicating the missing data is not having a discernible effect on the results. (b) plot of the eigenvalues, the first two eigenvalues were the most informative and so were retained for the PCA.

References

- Jombart, T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Meirmans, P.G. (2012) The trouble with isolation by distance. *Molecular Ecology*, **21**, 2839-2846.
- Perez, M.F., Franco, F.F., Bombonato, J.R., Bonatelli, I.A.S., Khan, G., Romeiro-Brito, M., Fegies, A.C., Ribeiro, P.M., Silva, G.A.R. & Moraes, E.M. (2018) Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Biodiversity Review*, **24**, 1883– 1889.
- Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S*, Fourth edn. Springer, New York.