On the use of genome-wide data to model and date the time of anthropogenic hybridisation: an example from the Scottish wildcat

Howard-McCombe, J., Ward, D., Kitchener, A. C., Lawson, D., Senn, H. V., Beaumont, M.

**Supp. Table 1**. Sample information. Q6546 indicates the ADMIXUTRE Q values generated using the ddRAD-seq data. Q values are also given for the 35 SNP test, with the lower and upper bounds (LBQ/UBQ, 90% CI) also shown. Pelage scores (7PS) were available for a subset of samples. The final column indicates whether a sample was included in the final dataset, 21 samples were excluded due to high levels of missing data, or to limit relatedness in the data.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **INDV ID** | **SOURCE POP** | **YEAR SAMPLED** | **Q6546** | **Q35** | **LBQ35** | **UBQ35** | **7PS** | **Included in the final dataset? (Yes/No)** |
| WCQ0047 | Wild | 1996 | 0.892 | 0.931 | 0.854 | 0.987 | 18 | Y |
| WCQ0052 | Wild | 1999 | 0.853 | 0.91 | 0.838 | 0.966 | 18 | Y |
| WCQ0099 | Wild |  | 0 | 0.012 | 0 | 0.049 | 16 | Y |
| WCQ0100 | Wild |  | 0.315 | 0.289 | 0.184 | 0.402 | 10 | Y |
| WCQ0107 | Wild | 1996 | 0.522 | 0.599 | 0.476 | 0.718 |  | Y |
| WCQ0110 | Wild | 2008 | 0.78 | 0.851 | 0.754 | 0.932 |  | Y |
| WCQ0132 | Captive |  | NA | 0.822 | 0.709 | 0.922 | 14 | N |
| WCQ0158 | Wild | 1997 | 0.855 | 0.876 | 0.787 | 0.949 | 11 | Y |
| WCQ0165 | Wild |  | 0 | 0.035 | 0 | 0.102 | 9 | Y |
| WCQ0168 | Wild | 2007 | 0.412 | 0.402 | 0.284 | 0.523 | 10 | Y |
| WCQ0172 | Wild |  | 0.253 | 0.056 | 0 | 0.147 | 13 | Y |
| WCQ0208 | Wild | 2013 | 0.693 | 0.724 | 0.612 | 0.826 | 14 | Y |
| WCQ0209 | Wild | 2013 | NA | 0.666 | 0.552 | 0.773 | 16 | N |
| WCQ0210 | Wild | 2005 | 0.78 | 0.851 | 0.754 | 0.932 |  | Y |
| WCQ0211 | Wild | 1999 | 0.908 | 0.884 | 0.8 | 0.952 | 14.5 | Y |
| WCQ0212 | Wild | 2011 | 0.722 | 0.794 | 0.692 | 0.884 | 16 | Y |
| WCQ0213 | Wild | 2002 | 0.737 | 0.732 | 0.618 | 0.835 | 15 | Y |
| WCQ0214 | Wild | 2003 | 0.832 | 0.909 | 0.834 | 0.966 | 13 | Y |
| WCQ0216 | Wild | 2014 | 0.581 | 0.599 | 0.478 | 0.714 | 15 | Y |
| WCQ0217 | Captive | 2014 | 1 | 0.986 | 0.941 | 1 | 17 | Y |
| WCQ0218 | Wild | 2011 | 0.674 | 0.777 | 0.67 | 0.873 | 13 | Y |
| WCQ0222 | Wild | 2014 | NA | 0.734 | 0.625 | 0.833 | 12 | N |
| WCQ0223 | Wild | 2014 | NA | 0.556 | 0.436 | 0.673 | 15 | N |
| WCQ0224 | Wild | 2014 | NA | 0.796 | 0.691 | 0.89 | 8 | N |
| WCQ0227 | Wild | 2012 | 0.206 | 0.288 | 0.182 | 0.401 | 12.5 | Y |
| WCQ0229 | Wild | 2014 | 0.272 | 0.282 | 0.171 | 0.401 | 10 | Y |
| WCQ0230 | Wild | 2011 | 0.321 | 0.367 | 0.251 | 0.488 | 15 | Y |
| WCQ0231 | Wild | 2011 | 0.411 | 0.466 | 0.348 | 0.585 | 17 | Y |
| WCQ0234 | Wild | 2014 | 0.463 | 0.626 | 0.507 | 0.74 | 7 | Y |
| WCQ0236 | Wild | 2011 | 0.499 | 0.489 | 0.371 | 0.607 | 9 | Y |
| WCQ0243 | Captive | 2014 | 1 | 0.937 | 0.858 | 0.992 | 19 | Y |
| WCQ0245 | Captive | 2014 | NA | 0.959 | 0.877 | 1 | 19 | N |
| WCQ0246 | Wild | 2014 | 0.627 | 0.433 | 0.315 | 0.553 | 14 | Y |
| WCQ0247 | Wild | 2014 | 0.682 | 0.721 | 0.608 | 0.825 |  | Y |
| WCQ0248 | Wild | 2014 | 1 | 0.974 | 0.905 | 1 | 17 | Y |
| WCQ0249 | Wild | 2013 | 0.462 | 0.632 | 0.513 | 0.745 | 10 | Y |
| WCQ0252 | Wild | 2014 | NA | 0.528 | 0.406 | 0.648 | 15 | N |
| WCQ0255 | Wild | 2012 | 0.391 | 0.327 | 0.216 | 0.445 | 18 | Y |
| WCQ0340 | Captive | 2014 | 1 | 0.938 | 0.846 | 1 | 18.5 | Y |
| WCQ0343 | Captive | 2014 | 1 | 0.991 | 0.961 | 1 | 20 | Y |
| WCQ0344 | Captive | 2014 | 1 | 0.99 | 0.959 | 1 | 20 | Y |
| WCQ0358 | Captive |  | NA | 0.98 | 0.919 | 1 | 17 | N |
| WCQ0383 | Wild | 2014 | 0.541 | 0.431 | 0.315 | 0.55 | 8 | Y |
| WCQ0387 | Wild | 2014 | 0.688 | 0.548 | 0.427 | 0.666 | 15.5 | Y |
| WCQ0390 | Wild | 2014 | NA | 0.641 | 0.523 | 0.753 | 13 | N |
| WCQ0402 | Captive | 2017 | 1 | 0.967 | 0.881 | 1 | 16.5 | Y |
| WCQ0404 | Captive | 2017 | 1 | 0.974 | 0.901 | 1 | 15 | Y |
| WCQ0408 | Captive | 2017 | 0.495 | 0.537 | 0.418 | 0.654 | 21 | Y |
| WCQ0419 | Captive | 2006 | NA | 0.984 | 0.933 | 1 |  | N |
| WCQ0420 | Captive | 2013 | 0.992 | 0.934 | 0.856 | 0.99 | 16 | Y |
| WCQ0421 | Captive | 2013 | 0.978 | 0.962 | 0.901 | 0.999 |  | Y |
| WCQ0422 | Captive | 2013 | 1 | 0.993 | 0.97 | 1 |  | Y |
| WCQ0427 | Captive | 2012 | 0.829 | 0.865 | 0.773 | 0.942 |  | Y |
| WCQ0428 | Captive | 2015 | 1 | 0.967 | 0.881 | 1 | 14 | Y |
| WCQ0429 | Captive | 2015 | 1 | 0.989 | 0.952 | 1 |  | Y |
| WCQ0430 | Captive | 2015 | NA | 0.826 | 0.72 | 0.918 |  | N |
| WCQ0431 | Captive | 2015 | NA | 0.828 | 0.724 | 0.918 |  | N |
| WCQ0432 | Captive | 2015 | 0.92 | 0.856 | 0.752 | 0.944 |  | Y |
| WCQ0433 | Captive | 2015 | 0.919 | 0.867 | 0.77 | 0.948 |  | Y |
| WCQ0434 | Captive | 2015 | NA | 0.85 | 0.744 | 0.941 |  | N |
| WCQ0435 | Captive | 2015 | 1 | 0.993 | 0.968 | 1 |  | Y |
| WCQ0436 | Captive | 2015 | 1 | 0.991 | 0.962 | 1 |  | Y |
| WCQ0437 | Captive | 2015 | 0.73 | 0.745 | 0.633 | 0.847 |  | Y |
| WCQ0439 | Captive | 2015 | 1 | 0.976 | 0.906 | 1 |  | Y |
| WCQ0443 | Domestic | 2015 | NA | 0.01 | 0 | 0.041 |  | N |
| WCQ0485 | Wild | 2014 | 0.138 | 0.146 | 0.061 | 0.246 |  | Y |
| WCQ0486 | Wild | 2014 | 0.152 | 0.128 | 0.042 | 0.228 |  | Y |
| WCQ0487 | Captive |  | NA | 0.979 | 0.922 | 1 | 15 | N |
| WCQ0488 | Captive | ? | NA | 0.803 | 0.7 | 0.893 | 19 | N |
| WCQ0489 | Captive | 2017 | 1 | 0.985 | 0.939 | 1 |  | Y |
| WCQ0490 | Captive | 2015 | 1 | 0.989 | 0.954 | 1 |  | Y |
| WCQ0491 | Captive | 2015 | 1 | 0.983 | 0.931 | 1 |  | Y |
| WCQ0492 | Domestic | 2007 | 0 | 0.023 | 0 | 0.079 |  | Y |
| WCQ0494 | Domestic | 2010 | 0 | 0.012 | 0 | 0.051 |  | Y |
| WCQ0501 | Domestic | 2008 | 0 | 0.023 | 0 | 0.086 |  | Y |
| WCQ0504 | Domestic | 2011 | 0 | 0.043 | 0.001 | 0.112 |  | Y |
| WCQ0515 | Wild | 2014 | 1 | 0.992 | 0.966 | 1 | 14-16 | Y |
| WCQ0519 | Wild | 2010 | 0.283 | 0.271 | 0.164 | 0.387 |  | Y |
| WCQ0527 | Wild | 2015 | 0.229 | 0.294 | 0.183 | 0.412 | 11-14 | Y |
| WCQ0528 | Wild | 2009 | 0.732 | 0.727 | 0.613 | 0.831 |  | Y |
| WCQ0529 | Wild | 2015 | 0.783 | 0.733 | 0.621 | 0.834 |  | Y |
| WCQ0531 | Captive | 2015 | 1 | 0.979 | 0.926 | 1 | 16 | Y |
| WCQ0540 | Captive | 2015 | 1 | 0.988 | 0.95 | 1 |  | Y |
| WCQ0541 | Captive | 2015 | 1 | 0.986 | 0.942 | 1 |  | Y |
| WCQ0544 | Captive | 2015 | 1 | 0.987 | 0.944 | 1 | 17 | Y |
| WCQ0545 | Captive | 2015 | 1 | 0.984 | 0.934 | 1 | 18 | Y |
| WCQ0546 | Captive | 2015 | 1 | 0.985 | 0.941 | 1 | 19.5 | Y |
| WCQ0547 | Captive | 2015 | 1 | 0.977 | 0.917 | 1 | 18 | Y |
| WCQ0549 | Captive | 2015 | 1 | 0.985 | 0.936 | 1 |  | Y |
| WCQ0550 | Captive | 2015 | 1 | 0.977 | 0.923 | 1 | 19.5 | Y |
| WCQ0551 | Captive | 2015 | 1 | 0.975 | 0.918 | 1 | 20 | Y |
| WCQ0552 | Captive | 2015 | 1 | 0.968 | 0.89 | 1 | 19.5 | Y |
| WCQ0553 | Captive | 2015 | 1 | 0.981 | 0.935 | 1 | 17.5 | Y |
| WCQ0554 | Captive | 2016 | 1 | 0.955 | 0.861 | 1 |  | Y |
| WCQ0555 | Captive | 2016 | 1 | 0.943 | 0.84 | 1 | 18 | Y |
| WCQ0556 | Captive | 2016 | 1 | 0.975 | 0.908 | 1 |  | Y |
| WCQ0557 | Captive | 2016 | 0.938 | 0.843 | 0.74 | 0.931 | 13.5 | Y |
| WCQ0558 | Captive | 2016 | NA | 0.895 | 0.798 | 0.975 |  | N |
| WCQ0559 | Captive | 2016 | 0.912 | 0.825 | 0.721 | 0.915 | 13 | Y |
| WCQ0560 | Captive | 2016 | 1 | 0.968 | 0.891 | 1 | 18 | Y |
| WCQ0564 | Captive | 2016 | 1 | 0.931 | 0.822 | 1 | 17.5 | Y |
| WCQ0567 | Captive | 2016 | 1 | 0.99 | 0.957 | 1 |  | Y |
| WCQ0578 | Wild | 2016 | 0.189 | 0.144 | 0.065 | 0.236 | 15 | Y |
| WCQ0586 | Captive | 2016 | 0.93 | 0.93 | 0.849 | 0.989 |  | Y |
| WCQ0588 | Captive | 2016 | 1 | 0.982 | 0.931 | 1 |  | Y |
| WCQ0589 | Captive | 2016 | 0.467 | 0.328 | 0.213 | 0.448 |  | Y |
| WCQ0603 | Wild | 2015 | NA | 0.646 | 0.528 | 0.757 |  | N |
| WCQ0604 | Wild | 2015 | 0.546 | 0.641 | 0.523 | 0.751 | 10.5 | Y |
| WCQ0606 | Wild | 2015 | 0.458 | 0.424 | 0.306 | 0.544 |  | Y |
| WCQ0612 | Captive | 2016 | NA | 0.927 | 0.835 | 0.998 |  | N |
| WCQ0613 | Wild | 2016 | 0.47 | 0.445 | 0.326 | 0.565 | <14 | Y |
| WCQ0614 | Captive | 2016 | 0.947 | 0.89 | 0.796 | 0.967 |  | Y |
| WCQ0615 | Captive | 2016 | 1 | 0.985 | 0.938 | 1 |  | Y |
| WCQ0616 | Captive | 2016 | 1 | 0.986 | 0.941 | 1 |  | Y |
| WCQ0617 | Captive | 2016 | 1 | 0.979 | 0.916 | 1 |  | Y |
| WCQ0618 | Captive | 2016 | 0.971 | 0.88 | 0.774 | 0.967 |  | Y |
| WCQ0619 | Captive | 2016 | 0.961 | 0.854 | 0.754 | 0.938 |  | Y |
| WCQ0620 | Captive | 2016 | 0.969 | 0.882 | 0.786 | 0.96 |  | Y |
| WCQ0621 | Captive | 2016 | 0.974 | 0.938 | 0.859 | 0.996 |  | Y |
| WCQ0622 | Captive | 2016 | 0.977 | 0.886 | 0.789 | 0.963 |  | Y |
| WCQ0624 | Captive | 2016 | NA | 0.965 | 0.884 | 1 | 16 | N |
| WCQ0626 | Captive | 2016 | 1 | 0.956 | 0.87 | 1 |  | Y |
| WCQ0627 | Captive | 2016 | 1 | 0.956 | 0.884 | 1 | 20 | Y |
| WCQ0628 | Captive | 2016 | 0.981 | 0.974 | 0.907 | 1 | 20.5 | Y |
| WCQ0629 | Captive | 2016 | 1 | 0.988 | 0.949 | 1 | 15.5 | Y |
| WCQ0901 | Wild | 2015 | 0.389 | 0.418 | 0.3 | 0.539 | 12 | Y |
| WCQ0902 | Wild | 2015 | 0.255 | 0.249 | 0.141 | 0.365 |  | Y |
| WCQ0903 | Wild | 2013 | 0.545 | 0.644 | 0.523 | 0.759 | 13 (2 marking not scored) | Y |
| WCQ0904 | Wild | 2015 | NA | 0.344 | 0.233 | 0.461 | 13 | N |



**Supp. Figure 1.** A) Proportion of variance explained by the first 20 principal components of the PCA. B) PCA plot for PC3/PC2 (individuals coloured by source population)

**A**

**B**



**Supp. Figure 2.** There is a very strong correlation between PC1 position and proportion of ‘domestic’ ancestry at K=2: The higher the PC1 coordinate the more domestic cat ancestry an individual is likely to have. This supports a strong genetic differentiation between domestic cats and a group of putative wildcats.

**Supp. Figure 3.** For captive animals only, the relationship between PC2 or PC3 position and inbreeding coefficient (F), excluding three hybrid individuals with negative F estimates (i.e., higher than expected heterozygosity). Inbreeding does not seem to explain the distribution of captive individuals across PC2 and PC3.

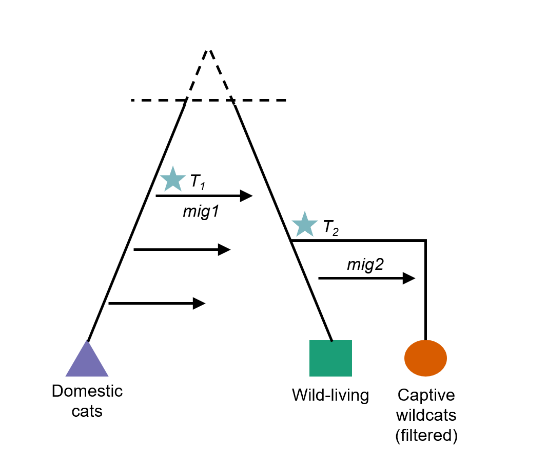
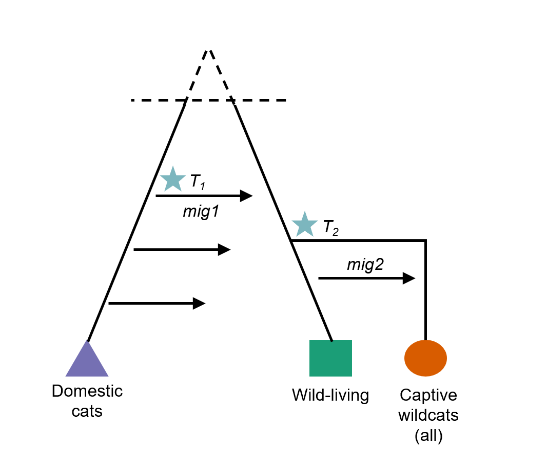
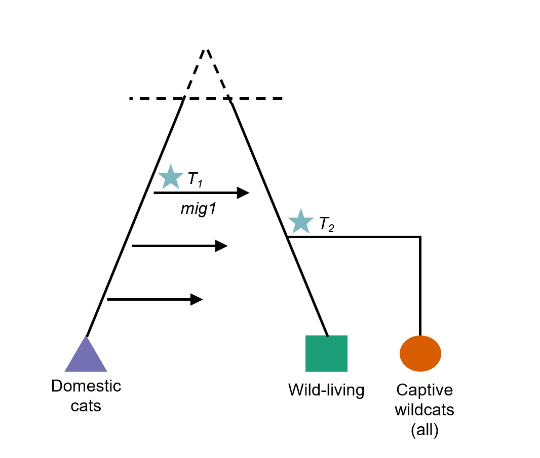
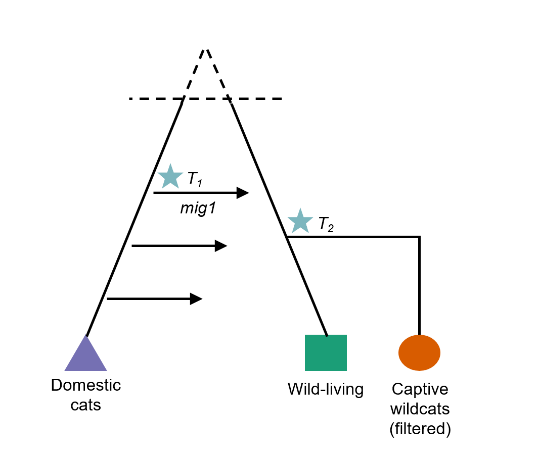


**Supp. Figure 4.** A) Cross-validation error for the ADMIXTURE analyses. The value of K with the lowest error is the best supported, though this does not always translate to the genuine number of ancestral populations. For these data cross-validation error supports K=5 as the most likely value of K. (B) K=5 gives a complex picture of wildcat population structure demonstrating the existence of as yet incompletely characterised population structure. Domestic cats form the only cluster that remains consistent at every value of K.

**B**

**A**

**Supp. Figure 5.** Data was initially simulated from four different models (A), with 22 summary statistics (see Supp. Table 2). These were used to identify the best combination of including/excluding an additional migration parameter (mig2, allowing gene flow from the wild-living population into the captive population) and including/excluding introgressed individuals in the captive population observed data. Approximately 30,000 simulations were generated using each model. In the filtered dataset individuals with a Q35 score less than 0.9 were excluded from the target data, these individuals are marked with a cross on the PCA (B).



**Model 1**

**Model 2**

**Model 3**

**Model 4**



**A**

**B**



**Supp. Figure 6.** As different target datasets were used across the four models (Q35 filtered vs. unfiltered, Supp. Fig. 4) a goodness-of-fit approach was used to examine model fit (using the R package abc; Csilléry, Francois & Blum 2012). For each model, the left-hand plot shows the first two principal components for the projected summary statistics, with target data shown in red. The right-hand plot shows the results of the goodness-of-fit test. The null distribution for the mean distance between the observed data and accepted summary statistics (tol=0.01) was generated using 100 replicates (each replicate using a simulated dataset as pseudo-observed data). The blue line shows the actual mean distance between the observed and simulated data (distance and associated p-value shown). For all models tested the observed data was outside the null distribution (p-value < 0.05). Models 3 and 4, with the mig2 parameter, appeared to fit better.

**Supp. Box 1**

To choose which summary statistics should be dropped from the model the following procedure was carried out. For the target summary statistics, and every simulated set of summary statistics (with parameters drawn from the prior), compute the Mahalanobis distance to the nearest neighbour of that point. The target and simulated summary statistics are scaled to have unit variance prior to PCA rotation. The nearest neighbour distance (nnd) is an estimate of a quantity proportional to density (Silverman, 2018), in this case the prior predictive density. The idea is to compare the nnd of the target to the nnd of all the simulated points. We can then define a highest prior predictive density (HPPD) band, e.g. HPPD0.95, such that 95% of all simulated points have nnd > HPPD0.95. Using this as a guide we iteratively dropped summary statistics, PCA rotated and scaled, and recomputed Mahalanobis nnd for all the points, until nnd of the target > HPPD0.95. The commented R script for achieving this is available in [github site].



**Supp. Figure 7.** A method of iteratively dropping summary statistics, based on nearest neighbour distances for the simulated and target data, was devised with the aim of improving model fit. Using this method the five ‘worst’ summary statistics were dropped for models 3 and 4, and goodness-of-fit re-examined. Dropping problem summary statistics improved fit for both models – mean distances were smaller than reported for the full set of summary statistics (Supp. Fig. 6), and these distances were not significantly different from the null distribution (p-value > 0.05). Model 4, using the second migration parameter (mig2) and filtering introgressed captive individuals from the target data, appeared to be the best fitting and was used for larger-scale simulations.

**Supp. Table 2.** Summary statistics used for ABC. Following a novel approach dropping summary statistics to improve model fit, 14 of the initially devised summary statistics were used in the final analysis.

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Name** | **Description** | **Included in final analysis?** |
| 1 | ss1 | Hierarchical clustering | No |
| 2 | ss2 | Yes |
| 3 | ss3.1 | Distribution of individuals across PC1 | Yes |
| 4 | ss3.2 | Yes |
| 5 | ss3.3 | Yes |
| 6 | ss3.4 | Yes |
| 7 | ss4.1 | Distribution of individuals across PC2 | Yes |
| 8 | ss4.2 | Yes |
| 9 | ss4.3 | Yes |
| 10 | ss4.4 | Yes |
| 11 | ss5 | Pairwise genetic distance between populations | Yes |
| 12 | ss6 | Yes |
| 13 | ss7 | Yes |
| 14 | ss8 | Linkage disequilibrium | Yes |
| 15 | ss9 | Yes |
| 16 | ss10.1 | Distribution of domestic individuals across PC1 | No |
| 17 | ss10.2 | Distribution of domestic individuals across PC2 | No |
| 18 | ss11.1 | Distribution of wild individuals across PC1 | No |
| 19 | ss11.2 | Distribution of wild individuals across PC2 | No |
| 20 | ss12.1 | Distribution of captive individuals across PC1 | No |
| 21 | ss12.2 | Distribution of captive individuals across PC2 | No |
| 22 | ss13 | Number of SNPs | No |

**Supp. Figure 8.** Correlation between summary statistics, with (B) showing the correlation between the projected summary statistics. Overall, it appears projection helped to reduce dimensionality and correlation between summary statistics. Plotted using corrplot (Wei & Simko, 2017)



**A**

**B**

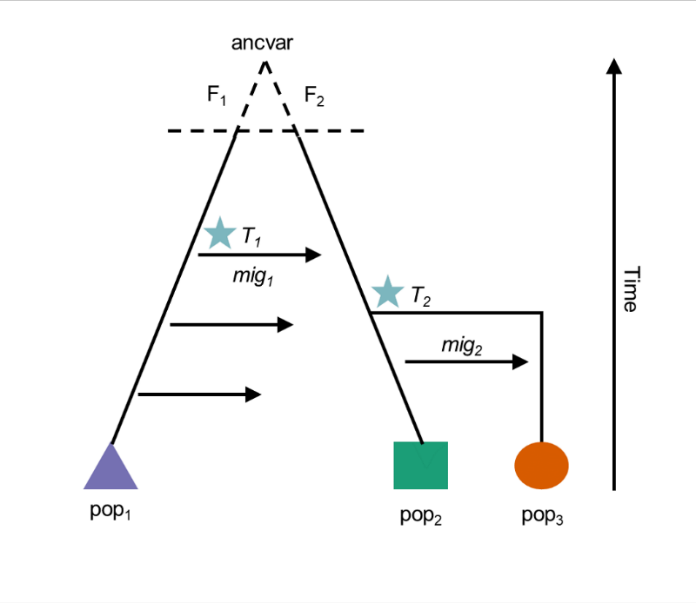


**Supp. Figure 9.** PCA plots for summary statistics and target data. Simulated data are shown in grey, with accepted points (tol=0.01) highlighted in blue, target data is shown in red.



**Supp. Figure 10.** Actual versus projected values of the prior for all model parameters.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Description** | **Prior distribution** | **Posterior mean (95% HPD)** |
| ancvar | Generates baseline ancestral variation | Exponential†  λ=0.1 | 4.155 (2.441 -5.801) |
| F1 | Drift from baseline (pop1) | Beta  α=2, β=10 | 0.211 (0.047 – 0.391) |
| F2 | Drift from baseline (pop2) | 0.183 (0.036 – 0.336) |
| log(pop1) | Log population size | Normal  µ=6.5, σ=0.5 | 6.429 (5.813 – 7.167) |
| log(pop2) | 6.580 (5.924 – 7.426) |
| log(pop3) | Normal‡  µ=4.6, σ=0.5 | 4.469 (3.986 – 5.099) |
| T1 | Onset of gene flow from pop1 to pop2 (number of generations) | Exponential  λ=0.02 | 3.326 (1.209 – 5.602) |
| T2 | Time pop3 is established from a sample of pop2 (number of generations) | Gamma  α=9, θ=0.5 | 19.272 (9.430 – 30) |
| mig1 | Migration (per generation) pop1 to pop2 | Beta  α=5, β=20 | 0.128 (0.067 – 0.192) |
| mig2 | Migration (every three generations) pop2 ­to pop3 | Gamma§  α=1, θ=1 | 0.012 (0 – 0.037) |



**A**

**Supp. Figure 11.** Prior and posterior distributions for the full set of parameters, shown on the model schematic (A). Details of each parameter, including prior and posterior distributions, are given in the table. The prior distributions for all parameters, fitted with locfit (Loader, 2013), can be seen in (B).

† (exponential distribution with rate parameter λ=0.1)+1 to avoid values of ancvar less than 1

‡The lower bound of this distribution was limited to 60 to avoid simulating a population of captive individuals smaller than the target data

§ (gamma distribution with shape parameter α=1 and scale parameter θ=1)/size of captive population



**B**

**Supp. Table 3.** SNPs associated with PC1 that were outliers with respect to population structure (p-value < 1x10-6)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | SNP No. | Chromosome | Position (bp) | Allele 1 | Allele 2 | Principal component | P-value |
| 1 | 2022 | B2 | 52989424 | G | T | 1 | 1.403 x 10-11 |
| 2 | 5147 | D4 | 75300817 | G | A | 1 | 1.991 x 10-7 |
| 3 | 5885 | E3 | 20260711 | A | G | 1 | 1.794 x 10-7 |

References

Csillery, K., Francois, O. and Blum, M. G. B. (2012), abc: an R package for approximate Bayesian computation (ABC). Methods in Ecology and Evolution 3: 475-479.

Loader, C., (2013) locfit: Local regression, likelihood and density estimation. R package version 1.5-9.1. <https://CRAN.R-project.org/package=locfit>

Silverman, B. W., Density estimation for statistics and data analysis. Routledge, 2018.

Wei, T., Simko, V. (2017). R package "corrplot": Visualization of a correlation matrix (version 0.84). Available from <https://github.com/taiyun/corrplot>