Supplementary material for:

# Genomic signatures of inbreeding in a critically endangered parrot, the kākāpō

Yasmin Foster<sup>\*</sup>, Ludovic Dutoit<sup>\*</sup>, Stefanie Grosser<sup>\*</sup>, Nicolas Dussex<sup>†, ‡, §</sup>, Brodie J. Foster<sup>\*</sup>, Ken G. Dodds<sup>\*\*</sup>, Rudiger Brauning<sup>\*\*</sup>, Tracey Van Stijn<sup>\*\*</sup>, Fiona Robertson<sup>\*</sup>, John C. McEwan<sup>\*\*</sup>, Jeanne M. E. Jacobs<sup>††</sup> & Bruce C. Robertson<sup>\*</sup>

\* Department of Zoology, University of Otago, PO Box 56, Dunedin 9054, New Zealand

† Centre for Palaeogenetics, Svante Arrhenius väg 20C, SE-106 91 Stockholm, Sweden

‡ Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden

§ Department of Zoology, Stockholm University, SE-106 91 Stockholm, Sweden

\*\* AgResearch Invermay Agricultural Centre, Private Bag 50034, Mosgiel 9053, New Zealand

†† AgResearch Lincoln Research Centre, Private Bag 4749, Christchurch 8140, New Zealand

### **Correspondence:**

Yasmin Foster

y.al.foster@gmail.com

# Contents

#### Page Item Supplementary Figure S1 ..... 2 Supplementary Figure S2 Supplementary Figure S3 Supplementary Figure S4 Supplementary Figure S5 Supplementary Figure S6 Supplementary Figure S7 Supplementary Figure S8 Supplementary Table S1



**Figure S1.** Principal component analysis of the total  $k\bar{a}k\bar{a}p\bar{o}$  dataset (n=161) using filtered SNP markers (n=12,241). Grey points represent Stewart Island founders and descendants, and green points for mainland descendants. Green star represents sole mainland founder, Richard-Henry.



**Figure S2.** Fin-plot for KGD-filtered SNPs (n=11,782) for the total kākāpō dataset (n=161) (Dodds et al. 2015). (a) Hardy-Weinberg disequilibrium plotted against minor allele frequency (MAF), with points shaded grey (low SNP depth) to blue (high SNP depth). (b)  $G_5$  self-relatedness estimates (diagonal elements of genomic relatedness matrix) plotted against log-transformed sample depth.



**Figure S3.** Density plot representing observed homozygosity for the total  $k\bar{a}k\bar{a}p\bar{o}$  dataset (n=161) using filtered SNP markers (n=12,241). Dashed line represents the mean, with blue shading for Stewart Island and red for mainland descendants.



b) Heterozygosity-heterozygosity correlation



**Figure S4** Molecular tests of identity disequilibrium for the total kākāpō dataset (n=161) using filtered SNP markers (n=12,241). (a) Distribution of the g2 statistic (p = 0.01). (b) Corresponding distribution of heterozygosity-heterozygosity correlation coefficients. For both plots, the dashed bar represents the mean and error bars are the 95% confidence interval produced after bootstrapping (n=1000).



**Figure S5.** Correlations between inbreeding estimates for dataset containing only Stewart Island descendants (n=151): (a)  $F_H$  and  $F_{RoH}$ , (b)  $F_H$  and  $F_{GRM}$ , and (c)  $F_{RoH}$  and  $F_{GRM}$ . Pearson's R and P-values are above each plot.



**Figure S6.** Density plots for (a)  $F_H$ , (b)  $F_{GRM}$ , (c)  $F_{RoH}$  for the total kākāpō dataset (n=161). Dashed line represents the mean.



**Figure S7.** Density plots for (a)  $F_{H}$ , (b)  $F_{GRM}$ , (c)  $F_{RoH}$ , and (d)  $F_{RoH10}$  for all individuals partitioned by Stewart Island or mainland descendant group. Dashed lines represent the mean for each group, with blue shading for Stewart Island and red for mainland descendants.



**Figure S8.** Genomic architecture of runs of homozygity (RoH) for the total  $k\bar{a}k\bar{a}p\bar{o}$  dataset (n=161) using filtered SNP markers (n=12,241). (a) Total number of RoH for all individuals on each chromosome. Density plots for: (b) number of RoH per individual, (c) total length of RoH between all individuals. Dashed lines represent the mean for each group, with blue shading for Stewart Island and red for mainland descendants.

**Table S1**. Summary of generalized linear model (GLM) with a binomial distribution and logit link function. Fixed predictor variables of ancestry and inbreeding ( $F_{RoH}$ ), and response variable of chick survivorship (survival ~ ancestry +  $F_{RoH}$ ).

					95% Confidence interval	
Variable	Estimate	SE	z-value	p-value	Lower bound	Upper bound
Intercept	-0.751	2.276	-0.330	0.741	-5.372	3.759
Ancestry	-0.442	1.634	-0.270	0.787	-3.751	2.826
FRoH	0.502	18.291	0.027	0.978	-36.394	37.065

## Reference

Dodds, K. G., McEwan, J. C., Brauning, R., Anderson, R. M., Stijn, T. C., Kristjánsson, T., & Clarke, S. M. (2015). Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics*, 16(1), 1–15.