**APPENDIX 1.**

Data vectors used to model breeding and conditional success probability jointly using frailty models.

- **BREED:** breeding events. 0 = nonbreeding, 1 = breeding.
- **IDB:** individual ID’s corresponding to breeding events.
- **DISPB:** dummy variables corresponding to breeding events. 0 = before dispersal, 1 = after dispersal.
- **YEARB:** years corresponding to breeding events.
- **SUCCESS:** success events conditional on BREED = 1. 0 = failure, 1 = success.
- **IDS:** individual ID’s corresponding to success events.
- **DISPS:** dummy variables corresponding to success events. 0 = before dispersal, 1 = after dispersal.
- **YEARS:** years corresponding to success events.

BREED=(1,1,1,1,1,0,0,0,0,1,1,1,1,1,1,1,1,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)  
IDB=(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
APPENDIX 2

A. As in classic approaches, the data \((Y)\) are regarded as sampled from a sampling distribution \(f(Y|\theta)\) governed by the parameter \(\theta\). However, in a Bayesian analysis, two probability distributions are used to make statements about \(\theta\): the prior and the posterior distributions. The prior summarizes what is known about the ranges and associated probability distribution for \(\theta\) without reference to the data \(Y\) (Link et al. 2002 a). The posterior provides the same summaries but informed by the data \(Y\). Inference about \(\theta\) is based on the posterior distribution \(f(Y|\theta)\), which can be expressed as a function of the prior \(\pi(\theta)\) and the sampling distribution \(f(Y|\theta)\) (see Bayes theorem).

The posterior distribution was assessed using a Monte Carlo approach (e.g., Spiegelhalter et al. 1996, Gelman et al. 1997, Link et al. 2002 a). This approach generates a first-order Markov chain of values sampled from the posterior. The transition matrix of the Markov process takes values in the range-space of \(\theta\), and its stationary distribution is equal to the posterior distribution \(f(Y|\theta)\). Given noninformative priors, the posterior mode of the chain approximates the maximum likelihood estimator of \(\theta\) (Gelman et al. 1997). The central 95\% range of values of the chain approximate the central 95\% region of the posterior distribution, which forms a Bayesian confidence interval for \(\theta\) (credible interval).

B. The DIC is defined as the sum of the Bayesian deviance (a measure of model fit) and the (effective) number of parameters (a measure of model complexity). As such, it is analogous in form and interpretation to the AIC (Burnham and Anderson 1998). The deviance, to within an additive quantity that does not depend on the parameters, is simply minus twice the log likelihood. The Bayesian deviance is the posterior mean of the deviance. Counting parameters in models containing random effects is difficult. For the DIC, Spiegelhalter et al. (2002) define the “effective number of parameters”, \(p\), as the posterior mean deviance minus the deviance evaluated at the posterior mean of the model parameters. This quantity may not be (and indeed seldom is) an integer, a phenomenon which is not unique to either Bayesian analyses, or random effects models in general (Burnham in review). For example, many nonparametric regression and smoothing methods involve non-integer measures of model complexity which are interpreted as “effective number of parameters” (e.g., see Hastie and Tibshirani 1990). In essence, non-integer values of “effective parameters” in a model are due to dependencies among the model parameters. In our case, individual effects are related to one another through specification of a normal prior distribution.