



Fig. 1 Demographic models compared in approximate Bayesian computations. Model *a* assumes that three populations [western (W), south-eastern (SE), north-eastern (NE)] diverged and underwent simultaneous exponential growth; bidirectional gene flow was assumed between all pairs of populations. Model *b* is identical to model *a* but without gene flow. Model *c* assumes that the W and SE populations expanded simultaneously and that the NE population subsequently diverged from the SE population, with all populations undergoing exponential growth; bidirectional gene flow was assumed between all population pairs. Model *d* is identical to model *c* but without gene flow. N_x : effective population size of population x , m_{x-y} : gene flow per generation from population x to population y , T_{EXP} : divergence time at the beginning of the Holocene, T_{NE-SE} : time at which the NE population split from the SE population.

conducted a preliminary principal component analysis (PCA) with R software (MASS package, function prcomp), based on 3000 simulated data sets for *M. sylvestris*, for the four models, to reveal visual correlations between the main model parameters and the summary statistics (Tellier *et al.* 2011).

We assumed a generalized stepwise model of microsatellite evolution (Estoup *et al.* 2002). The mutation rate was allowed to vary across loci, with locus-specific mutation rates being drawn from a gamma distribution ($\alpha, \alpha/\mu$) in which μ is the mutation rate per generation and α is a shape parameter. We assumed a log-uniform prior distribution for μ [0.00001, 0.02] and a uniform distribution for α [1, 30].

We compared the four models by calculating their Bayes factors (Wegmann *et al.* 2010) and estimating their relative posterior probabilities, based on the 1% of simulated data sets most closely matching the observed data (i.e. 2000 simulated data sets). Once the best model had been chosen, we estimated demographic parameters under this scenario, using a general linear model (ABC-GLM) postsampling regression adjustment for the 2000 retained simulations (Leuenberger & Wegmann 2010; Wegmann *et al.* 2010). We report the mode and 95% highest posterior density (HPD) interval for each model parameter estimate.

The performance of the method for discriminating between competing historical models was assessed by analysing test data sets (called pseudo-observed data sets) simulated with the same number of loci and individuals as for the observed data sets. We simulated 2000 such data sets for each competing model, using parameter values drawn from the same prior distributions as for the original analyses. We determined the relative posterior probabilities of competing models for each pseudo-observed data set, using the model choice procedure, as described above (Wegmann *et al.* 2010). Confidence in model choice was then estimated from the percentage of times that a given scenario did not have the highest posterior probability of the competing scenarios when it was actually the true scenario (type I error) and the percentage of times that a given scenario had the highest posterior probability when it was not the true scenario (type II error).

Projections of *Malus sylvestris* distribution during LGM

We modelled the climatic niche of *M. sylvestris* from the current species distribution with the BIOMOD package (v 1.1-7.00, 2011-08.03; Thuiller *et al.* 2009) downloaded from WORLDCLIM data set v1.4 (<http://www.worldclim.org/>;