**Supplementary Material S5 – Description of simulations**

Data was simulated using a combination of coalescent methods ([MaCS; Chen *et al.*, 2009](#_ENREF_1)) and AlphaDrop ([Hickey and Gorjanc, 2012](#_ENREF_2)). Using MaCS 100 founder haplotypes were simulated for one chromosome which was 100 cM in length comprising 1.0×108 base pairs, using a per site mutation rate of 2.5×10-8 and a varying effective population size (*Ne*) over time reflecting the estimates for the Holstein cattle population ([Villa-Angulo *et al.*, 2009](#_ENREF_3)). In the final generation of haplotype sequence simulation *Ne* was 100 with linear increases to a *Ne* of 1256 at 1000 years ago, a *Ne* of 4350 at 10 000 years ago, and a *Ne* of 43 500 at 100 000 years ago. Using AlphaDrop the simulated founder haplotypes were dropped through 6 generations of a simulated pedigree. In each generation there were 1000 individuals simulated (500 male, 500 female). The parents of each generation comprised all females and 50 randomly selected males from the previous generation. A random sample of 2000 segregating sites was selected to be used as SNP. The individuals in generation 6 were the target for imputation and thus had all but 200 of their markers masked. The training set of high-density genotyped individuals used for imputation comprised male ancestors from generations 1, 2, and 3 and all individuals in generations 4 and 5.

**References**

Chen GK, Marjoram P and Wall JD 2009. Fast and flexible simulation of DNA sequence data. Genome Research 19, 136-142.

Hickey JM and Gorjanc G 2012. Simulated data for genomic selection and genome-wide association studies using a combination of coalescent and gene drop methods. G3: Genes|Genomes|Genetics 2, 425-427.

Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP and Grefenstette JJ 2009. High-resolution haplotype block structure in the cattle genome. BMC Genetics 10, 19.