Supplementary Note 2: Details of the simulation design

We used simulation to test the performance of BiForce in binary traits in comparison with BOOST (Wan, *et al.*, 2010) and quantitative traits in comparison with PLINK (Purcell, *et al.*, 2007). Since BOOST has been tested against PLINK in a simulation study (Wan, *et al.*, 2010), to keep simple we choose to adopt the BOOST simulation design in this study. We used 500 replicates (in contrast to 100 replicates used in the BOOST study) for all simulation scenarios. Details of the adopted simulation design and variations are briefly described below.

Epistasis models

The four two-locus interaction models used to generate epistatic scenarios with marginal effects in disease loci in the BOOST study were applied here (Table_S2_1). Considering two loci A (disease risk allele a) and B (disease risk allele b), Model 1 is a multiplicative model where the odds of each joint genotype have a baseline value unless both loci have at least one disease-associated allele and their odds increase multiplicatively within and between genotypes (Marchini, *et al.*, 2005). Both Models 2 and 3 have the missing lethal genotype (i.e. the double homozygous genotype of disease alleles **aa-bb** does not lead to disease) (Li and Reich, 2000). Model 2 differs from Model 3 mainly in the double heterozygous genotype **Aa-Bb** that does not lead to disease and has been used to describe the genetics of handedness (Levy and Nagylaki, 1972; Neuman and Rice, 1992). In addition, Model 3 has the **aa-Bb** genotype leads to disease in contrast to **Aa-BB** in Model 2. Model 4 is a well known XOR (exclusive OR) model where only four single heterozygous genotypes (**AA-Bb**, **Aa-BB**, **Aa-bb** and **aa-Bb**) lead to disease (Li and Reich, 2000; Moore and Williams, 2009). The logical XOR operation on two binary variables is defined as: 0 XOR 0 = 0, 0 XOR 1 = 1, 1 XOR 0 =1, 1 XOR 1 = 0, where the last operation makes XOR extremely nonlinear. These models were used in this study to generate epistatic scenarios for binary and quantitative traits.

Epistatic scenarios

For each epistasis model six simulation scenarios were used to cover a sample size of 800 or 1600 (with balanced design) and a minor allele frequency (MAF) of 0.1 or 0.2 or 0.4 for the simulated epistatic SNPs (assumed equal MAF for both loci). In the binary trait case, the disease prevalence was fixed as 0.1 for all epistatic scenarios, whereas the disease heritability was set to 0.03 for Model 1 and 0.02 for Models 2 to 4. With these settings, the parameters α and θ are calculated and showed in Table_S2_2.

Table_S2_1. Genotype values of the simulated epistatic SNP pair in four epistasis models used in binary trait simulation scenarios*

	SNP_2		
SNP_1	BB	Bb	bb
Model 1			
AA	α	α	α
Aa	α	$\alpha(1+\theta)$	$\alpha(1+\theta)^2$
aa	α	$\alpha(1+\theta)^2$	$\alpha(1+\theta)^4$
Model 2			
AA	α	$\alpha(1+\theta)$	$\alpha(1+\theta)$
Aa	$\alpha(1+\theta)$	α	α
aa	$\alpha(1+\theta)$	α	α
Model 3			
AA	α	α	$\alpha(1+\theta)$
Aa	α	$\alpha(1+\theta)$	α
aa	$\alpha(1+\theta)$	$\alpha(1+\theta)$	α
Model 4			
AA	α	$\alpha(1+\theta)$	α
Aa	$\alpha(1+\theta)$	α	$\alpha(1+\theta)$
aa	α	$\alpha(1+\theta)$	α

*: The parameters α and θ control the disease prevalence and heritability based on Equations 15 to 18 given in the BOOST paper (Wan, *et al.*, 2010). They define the interaction pattern in each model together with minor allele frequency of disease loci (identical for SNP_1 and SNP_2)

Table_S2_2. Parameter values used to generate epistatic scenarios for binary traits

Model	MAF	Heritability	Prevalence	α	θ
1	0.1	0.03	0.1	0.100	3.448
1	0.2	0.03	0.1	0.091	1.300
1	0.4	0.03	0.1	0.075	0.624
2	0.1	0.02	0.1	0.077	1.533
2	0.2	0.02	0.1	0.065	1.640
2	0.4	0.02	0.1	0.065	1.640
3	0.1	0.02	0.1	0.099	3.003
3	0.2	0.02	0.1	0.088	1.746
3	0.4	0.02	0.1	0.068	1.585
4	0.1	0.02	0.1	0.078	1.537
4	0.2	0.02	0.1	0.067	1.601
4	0.4	0.02	0.1	0.061	1.717

In the quantitative trait case, because the disease prevalence is not defined it is difficult to solve both parameters α and θ using only the trait heritability. To remove the parameter α while maintaining the interaction pattern in each epistasis model, each genotype value in Table_S2_1 was divided by the quantity of α(1 - θ) (Table_S2_3). For simplicity the θ value for each epistatic scenario in Table_S2_2 was applied to calculate the genotype values for the counterpart epistatic scenario for quantitative traits. Trait heritability was also set to 0.03 for Model 1 and 0.02 for Models 2 to 4 used in simulating phenotypes (see below). Therefore, the epistatic scenarios for quantitative traits may not be directly comparable against the counterparts for binary traits.

Table_S2_3. Genotype values of the simulated epistatic SNP pair in four epistasis models used in quantitative trait simulation scenarios*

	SNP_2		
SNP_1	BB	Bb	bb
Model 1			
AA	$1/(1 + \theta)$	$1/(1 + \theta)$	$1/(1 + \theta)$
Aa	$1/(1 + \theta)$	1	$(1 + \theta)$
aa	$1/(1 + \theta)$	$(1 + \theta)$	$(1 + \theta)^2$
Model 2			
AA	$1/(1 + \theta)$	1	1
Aa	1	$1/(1 + \theta)$	$1/(1 + \theta)$
aa	1	$1/(1 + \theta)$	$1/(1 + \theta)$
Model 3			
AA	$1/(1 + \theta)$	$1/(1 + \theta)$	1
Aa	$1/(1 + \theta)$	1	$1/(1 + \theta)$
aa	1	1	$1/(1 + \theta)$
Model 4			
AA	$1/(1 + \theta)$	1	$1/(1 + \theta)$
Aa	1	$1/(1 + \theta)$	1
aa	$1/(1 + \theta)$	1	$1/(1 + \theta)$

*: Parameter θ takes the values defined in Table_S2_2 for each epistatic scenario

For each epistatic scenario, genotypes of 1,000 SNPs were generated using the program gs (version 2.0) (Li and Chen, 2008) by randomly sampling from SNPs located on chromosome 11 of the CEU HapMap (phase 3 release 2) phased data under the assumption of Hardy-Weinberg equilibrium and selecting SNPs with the desired MAF to use as disease SNPs. Binary phenotypes were also generated by gs according to the genotype values calculated for the scenario as showed in Table_S2_1 and Table_S2_2. Quantitative phenotypes were simulated by an R script (Table_S2_4) based on the genotype values calculated for the scenario as showed in Table_S2_2 and Table_S2_3 and trait heritability with random noise sampled from a normal distribution with a mean of zero and variance of 1 minus heritability. The simulated quantitative phenotypes were standardized before the association analysis.

NULL scenarios

Following the simulation design in the BOOST study (Wan, *et al.*, 2010), two NULL scenarios were simulated each using 1,000 samples and 500 replicates: (1) the program genomeSIMLA (Dudek, *et al.*, 2006) was used to generate SNP genotype data based on the Affymetrix 500k SNP array to accommodate linkage disequilibrium in real GWAS. We used only information of SNPs located on chromosome 1 of the Affymetrix 500k SNP array to generate 38,836 SNPs as the BOOST study; (2) 1,000 SNPs were randomly generated with MAFs uniformly from [0.05, 0.5]. In both cases, samples were generated by sampling from a Bernoulli distribution for binary traits and a Gaussian distribution (mean=0, variance=1) for quantitative traits.

Table_S2_4. The R script for generating quantitative phenotypes

```
epiSim <- function(m1, m2, pattern, v){
  # Take two SNPs and simulate epistatic phenotype
  # with genotype-phenotype as defined in pattern
  # Args:
  #
  # m1, m2: SNP markers, numeric vectors {0, 1, 2}
  # pattern: genotype phenotype map, 3x3 matrix
  # v: variance explained by gp map
  # return array of length of m1 representing standardized phenotype

  N <- length(m1)
  if(v == 0) return(rnorm(N))
  m12 <- m1 + m2*3
  phen <- array(0,N)
  for(i in 0:8)
  {
    phen[m12 == i] <- c(pattern)[i+1]
  }
  if(v == 1) return(phen)
  var_orig <- var(phen)
  sd_req <- sqrt((var_orig / v) - var_orig)
  phen <- phen + rnorm(N,mean=0,sd=sd_req)
  return((phen - mean(phen)) / sd(phen))
}
```

References

Dudek, S.M., *et al.* (2006) Data simulation software for whole-genome association and other studies in human genetics, *Pac Symp Biocomput*, 499-510.

Levy, J. and Nagylaki, T. (1972) A model for the genetics of handedness, *Genetics*, **72**, 117-128.

Li, J. and Chen, Y. (2008) Generating samples for association studies based on HapMap data, *BMC Bioinformatics*, **9**, 44.

Li, W. and Reich, J. (2000) A complete enumeration and classification of two-locus disease models, *Hum Hered*, **50**, 334-349.

Marchini, J., *et al.* (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases, *Nat Genet*, **37**, 413-417.

Moore, J.H. and Williams, S.M. (2009) Epistasis and its implications for personal genetics, *Am J Hum Genet*, **85**, 309-320.

Neuman, R.J. and Rice, J.P. (1992) Two-locus models of disease, *Genet Epidemiol*, **9**, 347-365.

Purcell, S., *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses, *Am J Hum Genet*, **81**, 559-575.

Wan, X., *et al.* (2010) BOOST: A Fast Approach to Detecting Gene-Gene Interactions in Genome-wide Case-Control Studies, *American journal of human genetics*, **87**, 325-340.