

## RESEARCH

## Open Access

# A comparison of methods for whole-genome QTL mapping using dense markers in four livestock species

Andres Legarra<sup>1\*</sup>, Pascal Croiseau<sup>2</sup>, Marie Pierre Sanchez<sup>2</sup>, Simon Teyssède<sup>1,3</sup>, Guillaume Sallé<sup>4,5</sup>, Sophie Allais<sup>6,7,8</sup>, Sébastien Fritz<sup>9</sup>, Carole Rénée Moreno<sup>1</sup>, Anne Ricard<sup>2,10</sup> and Jean-Michel Elsen<sup>1</sup>

## Abstract

**Background:** With dense genotyping, many choices exist for methods to detect quantitative trait loci (QTL) in livestock populations. However, no across-species study has been conducted on the performance of different methods using real data. We compared three methods that correct for relatedness either implicitly or explicitly: linkage and linkage disequilibrium haplotype-based analysis (LDLA), efficient mixed-model association (EMMA) analysis, and Bayesian whole-genome regression (BayesC). We analyzed one chromosome in each of five datasets (dairy cattle, beef cattle, sheep, horses, and pigs) using real genotypes based on dense single nucleotide polymorphisms and phenotypes. The *P* values corrected for multiple testing or Bayes factors greater than 150 were considered to be significant. To complete the real data study, we also simulated quantitative trait loci (QTL) for the same datasets based on the real genotypes. Several scenarios were chosen, with different QTL effects and linkage disequilibrium patterns. A pseudo-null statistical distribution was chosen to make the significance thresholds comparable across methods.

**Results:** For the real data, the three methods generally agreed within 1 or 2 cM for the locations of QTL regions and disagreed when no signals were significant (e.g. in pigs). For certain datasets, LDLA had more significant signals than EMMA or BayesC, but they were concentrated around the same peaks. Therefore, the three methods detected approximately the same number of QTL regions. For the simulated data, LDLA was slightly less powerful and accurate than either EMMA or BayesC but this depended strongly on how thresholds were set in the simulations.

**Conclusions:** All three methods performed similarly for real and simulated data. No method was clearly superior across all datasets or for any particular dataset. For computational efficiency and ease of interpretation, EMMA is recommended, but using more than one method is suggested.

## Background

Many methods to detect and localize quantitative trait loci (QTL) in humans and animals are reported in the literature. Current methods work on the basis of identity, either by descent or by state, at either single nucleotide polymorphisms (SNPs) or a short series of SNPs (i.e., a haplotype). Current state-of-the-art methods, especially in human genetics, use a consecutive series of single-marker tests (e.g. association analysis), often corrected by stratification or coancestry [1-4]. The high density of molecular markers ensures that common QTL variants of medium effect will

be captured by random linkage disequilibrium of close markers. However, for certain cases, such as causal QTL that are rare variants, simultaneous consideration of consecutive loci (e.g. haplotype association or linkage analysis) can be more powerful or accurate than locus-by-locus analysis (e.g. [5]). In addition, recent research has proposed fitting all markers simultaneously to better model the genetic background and thus improve the analysis [6-8]. Sahana et al. [9] used simulation to compare methods, but comparison of QTL mapping methods using real data has not been reported in animal genetics. The purpose of this study was to compare methods for QTL detection and localization based primarily on real datasets for five traits and four species and complemented with some simulations that mimicked the real data.

\* Correspondence: [andres.legarra@toulouse.inra.fr](mailto:andres.legarra@toulouse.inra.fr)

<sup>1</sup>INRA, UMR 1388 GenPhySE, BP52627, 31326 Castanet Tolosan, France  
Full list of author information is available at the end of the article

Three state-of-the-art methods were compared: LDLA (linkage disequilibrium and linkage analysis), a method that considers haplotypes clustered by approximate identical-by-descent probabilities [10]; EMMA (efficient mixed-model association), a regular single-marker association analysis method for a genome-wide association study (GWAS) but with correction for relatedness in the population [3,11,12]; and BayesC, a Bayesian method that fits effects of all SNPs simultaneously [13]. We examined the following features: number of positive signals, agreement of location of positive signals between the methods, shape of signals, and redundancy of test statistics between neighbouring positions. Data-based simulation was used to characterize method properties (power, accuracy, and false discovery rate) under a pseudo-null hypothesis that considered presence of a QTL effect.

## Methods

Several research projects in the INRA (French National Institute for Agriculture Research) Animal Genetics division have generated a number of datasets that include SNPs and phenotypes. For this study, we considered a variety of distinct experimental designs and species. First, we chose and analyzed chromosomes on which at least one QTL (but not a very large one) had already been detected using the datasets. Then, we considered populations with different family structures: very related (cattle), minimally related but with structure (horses), structured in breeds (pigs), and breed crosses (sheep).

### Description of data

Five datasets, which included four species separately, were examined. One chromosome was analyzed for each dataset.

Cattle data in this study came from on-farm recording. For the horses, the procedures involving animals consisted of radiography and collection of blood samples, which are routinely performed in veterinary practice and are considered noninvasive. All horses were sedated in accordance with clinical guidelines. For the sheep study, all animals were kept indoors, handled with care, and managed as a commercial flock following the INRA ethics policy. At the end of the experimental infection, animals were slaughtered at the INRA-Nouzilly abattoir following the EU rules. For the pig study, animals involved were reared and slaughtered in compliance with national regulations applicable to animal research and commercial slaughtering.

#### Dairy cattle (*Bos taurus*)

Milk yield (305-day) was examined for a population of 1221 Montbéliarde bulls with a complex pedigree structure that overlapped several generations. Pseudo-performance of bulls was measured by daughter yield deviation [14], which is the average performance of a bull's daughters corrected

for other effects such as herd and genetic merit of dams. The bulls had been genotyped with the Illumina Bovine SNP50 BeadChip. Markers were discarded based on low call rate (<95%), lack of known position on the genome, or very high Mendelian inconsistency (more than 20% parent-offspring discordance), resulting in 43 582 SNPs that were used. Chromosome 1 with 2854 markers was analyzed.

#### Beef cattle (*Bos taurus*)

Meat tenderness [15] was examined for a population of 936 Blonde d'Aquitaine bulls that had been genotyped with the Illumina Bovine SNP50 BeadChip (43 582 SNPs), filtered as for dairy cattle. Chromosome 7 with 1889 markers was analyzed.

#### Sheep (*Ovis aries*)

Resistance to nematode infestation measured as faecal egg count after a first infection at 90 days postpartum (FEC12t) [16] was examined for a population of 1067 meat sheep that resulted from a BlackBelly × Romane backcross [17]. Three generations were included: F1, backcross, and backcross × backcross. Animals had been genotyped with the Illumina OvineSNP50 BeadChip (42 469 SNPs, filtered as in [17]). Chromosome 12 with 1424 markers was analyzed.

#### Horse (*Equus caballus*)

Incidence of hock osteochondrosis [18] was examined for a minimally related but family structured population of 627 French Trotter horses; 102 stallions sired the 525 horses that had been scored for hock osteochondrosis. Animals had been genotyped with the Illumina EquineSNP50 BeadChip (41 249 SNPs, filtered as in [18]). Chromosome 3 with 2267 markers was analyzed.

#### Pig (*Sus scrofa*)

Length of carcass pre-corrected for breed and other environmental factors [19] was examined for a population of 764 pigs (495 Large White, 129 Landrace, and 140 Piétrain) from 327 sires; 656 pigs had been measured for length of carcass. The entire population had been genotyped with the Illumina PorcineSNP60 DNA Analysis kit (46 865 SNPs, edited as in [19]). Chromosome 17 with 1672 markers was studied.

A brief description of the data is in Table 1.

### Data analysis

We used three methods to analyse the data: LDLA, EMMA, and BayesC.

#### LDLA

This method was originally developed by Meuwissen et al. [20] and our implementation was as in [10]. Genotypes were first phased based on family and linkage

**Table 1 Basic data description by species**

Descriptor	Dairy cattle	Beef cattle	Sheep	Horses	Pigs
Nb of animals	1221	936	1067	627	764
Trait	305-day milk yield (DYD)	meat tenderness	fecal egg count	hock osteochondrosis score	length of carcass
Population structure	large half-sib families, complex pedigree	small half-sib families, complex pedigree	F1, backcross, backcross × backcross	many small families	three breeds
Chr studied	1	7	12	3	17
Nb of chr markers	2854	1889	1424	2267	1672
Length (cM)	161	112	79	119	60
Heritability	0.90	0.20	0.45	0.40	0.30

Nb = number; chr = chromosome; DYD = daughter yield deviation.

disequilibrium information [21]. Then, haplotypes were defined based on four consecutive (polymorphic) SNPs. Transmission of these haplotypes was traced along the pedigree; if no recombination occurred, haplotypes transmitted from parents to offspring were considered to be identical and carriers of the same allele at the QTL. Founder haplotypes, which cannot have identity-by-descent ascertained, were clustered based on their resemblance, which was based on likeness in state according to a simple coalescence model. Finally, a variance component model was fitted to the phenotypes for each locus on the chromosome of interest and maximized by restricted maximum likelihood, either with (alternative hypothesis) or without haplotype effects (null hypothesis). The model included a polygenic effect structured according to a pedigree-based relationship matrix:  $\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{T}\mathbf{h} + \mathbf{e}$ , where  $\mathbf{y}$  is the vector of phenotypes,  $\mathbf{u}$  and  $\mathbf{h}$  are vectors of polygenic and haplotype effects (with respective incidence matrices  $\mathbf{Z}$  and  $\mathbf{T}$ ), respectively, and  $\mathbf{e}$  is a vector of residuals;  $Var(\mathbf{h}) = \mathbf{H}\sigma_h^2$  and  $Var(\mathbf{u}) = \mathbf{A}\sigma_u^2$  are haplotypic and polygenic covariance matrices, with respective variance components  $\sigma_h^2$  and  $\sigma_u^2$ , with matrix  $\mathbf{H}$  set up according to [10]. A likelihood ratio test comparing the null and alternative hypothesis above was then calculated, and  $P$  values were computed according to [22]. The process was repeated for each locus on the chromosome of interest; for each locus, new incidence and covariance matrices  $\mathbf{T}$  and  $\mathbf{H}$  were created.

#### EMMA

We used a simple extension of regular regression-based association analysis. The extension fitted a polygenic background via a relationship matrix to deal with structure, as regularly practiced in animal genetics (e.g. [23]) and recently applied in human genetics (e.g. [3]). In our implementation, we used a whole-genome SNP-based relationship matrix ( $\mathbf{G}$ ) [24] instead of a pedigree-based relationship matrix. At each marker locus on the chromosome of interest, we fitted the model  $\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{w}s + \mathbf{e}$ , where  $\mathbf{u}$  is a vector of polygenic effects as in

the LDLA model but with  $Var(\mathbf{u}) = \mathbf{G}\sigma_u^2$ ,  $\mathbf{w}$  is a vector of covariates coded as (0,1,2) for each genotype at the analysed SNP, and  $s$  is the substitution effect of the marker. Effects were estimated using blup90 [25]. A  $t$ -statistic was constructed from the estimate of the substitution effect  $s$  as  $t = \hat{s}/s.e.(\hat{s})$  where  $s.e.(\hat{s})$  is the standard error of the estimate  $\hat{s}$ .

#### BayesC

This method [13] is one of a family of Bayesian methods that were originally conceived for prediction of genetic merit and phenotypes and later extended to map gene locations [6,26]. It has the potential to analyze the whole genome simultaneously. The BayesC model included a set of variable indicators,  $\mathbf{d} = \{d_1, \dots, d_m\}$ , which indicates if a SNP is ( $d_i = 1$ ) or is not ( $d_i = 0$ ) in the model. The phenotypes were thus modelled as  $\mathbf{y} = \mathbf{1}\mu + \sum \mathbf{w}_i d_i s_i + \mathbf{e}$ . Solutions of the model were obtained by Gibbs sampling, which provides marginal *a posteriori* inference on  $\mathbf{d}$  in the form of Bayes factors (BF) [27,28]. A key parameter of the model is the *a priori* number of SNPs, which was fixed at  $Pr(d_i = 1) = 1/1000$ . Experimentation with 1/100 or 1/10000 did not qualitatively change the results. Gibbs-sampling chains were run for 100 000 iterations using GS3 [29]. All SNPs in the genome were included in the model, although only the SNP estimates for the chromosome of interest were further analyzed. The statistic used was the BF, which corresponds to the increase from prior to posterior probabilities of the SNP being “in” the model [27,30], which in our case can be written as:

$$BF_i = \frac{Pr(d_i = 0|\mathbf{y})/Pr(d_i = 1|\mathbf{y})}{Pr(d_i = 0)/Pr(d_i = 1)} = 999 \frac{\hat{d}_i}{1 - \hat{d}_i},$$

where  $\hat{d}_i$  is the posterior probability of a locus being in the model. For simple models, the BF is numerically identical to the likelihood ratio if *a priori* probabilities are identical [27,30]. Numerically, the scales of the BF and the likelihood ratio are fairly similar [27]. An advantage of BF is that it has a clear interpretation, contrary to the *ad hoc* thresholds used for the likelihood

ratio (e.g. [31]), which are not comparable from one study to another.

### QTL detection

We used chromosome-wise Bonferroni correction to infer rejection thresholds for  $P$  values of LDLA and EMMA, and a BF of 150 for BayesC, because this value was suggested as “very strong evidence” [27] and later used for QTL mapping [32].

Potentially, different methods can reveal different QTL in the same dataset. For instance, for two closely positioned QTL on the same chromosome, classical linkage-based analysis will show a large QTL signal around one location, whereas linkage disequilibrium methods may indicate two separate narrow signals. To determine approximately the ability of methods to disentangle closely positioned QTL, we estimated a measure of redundancy for each method along the chromosome, which was based on the equivalent number of independent tests using Geyer’s [4] effective sample size. This method estimates the number of independent points in a Markov chain process, which in our case consisted of QTL signals along the genome produced by a method.

### Simulations

We simulated data for five scenarios. One hundred replicates were run for each scenario. For each real dataset, we picked at random one or two SNPs as QTL (the same for each replicate within each scenario but not across scenarios) and assigned an effect (the same for each replicate; the effect could be additive or dominant). The effect type, minor allele frequency, and linkage disequilibrium (measured as the correlation between a QTL and the neighbouring marker in strongest linkage disequilibrium with QTL) are in Table 2 for each scenario. The SNP acting as a QTL was then masked. We also simulated a polygenic component based on pedigree relationships, as well as residual deviations, which were added to the QTL genotype effects to simulate phenotypes. Each replicate was analyzed with the three methods (LDLA, EMMA, and BayesC) and for each replicate, the most significant QTL was considered and other significant signals were

disregarded. We evaluated three indicators of the quality of QTL mapping: mean squared error in localization (MSE), false discovery rate (FDR), and power. The MSE and power were based on the declared QTL falling within a region around the simulated QTL, and FDR was the number of QTL declared outside the QTL region. Three sizes of QTL regions were examined:  $\pm 0.5$ ,  $\pm 1$ , and  $\pm 2$  cM around the true QTL. The QTL explained 5% of total genetic variance in all scenarios, and the broad-sense heritability of the trait was 0.30.

Statistics for significance decisions (acceptance or rejection of the null hypothesis) generally were not comparable across methods. For example, BF does not have defined type-I and type-II error rates. Therefore, we used the simulations to establish the appropriate rejection thresholds for each statistic (likelihood ratio test for LDLA,  $t$ -statistic for EMMA, and BF for BayesC) within each scenario. We examined three thresholds. First, a 5% empirical threshold was computed from the distribution of statistics across all non-QTL positions (those that did not fall in the intervals mentioned above). Because a QTL was included in the simulated data but only non-QTL positions were included in the computation of the statistic, we refer to this as a pseudo-null distribution. Based on 1000 non-QTL positions and 20 simulations, the distribution of the statistic under the pseudo-null hypothesis was described by 20 000 samples. Two more stringent thresholds were also defined. For each of the 100 simulations, the most significant  $P$  value in the non-QTL regions was retained; then, the 10% and 20% quantiles (computed following [33]) were set as the two thresholds.

After running all simulations, we had five datasets, five scenarios, 100 replicates for each scenario, three methods, three alternative thresholds, and three alternative QTL region sizes to consider as true positives. The average (across 100 replicates) of each QTL quality indicator (MSE, FDR, or power) was calculated, which resulted in 675 values, one for each combination of the above factors.

## Results

### Data analysis

Manhattan plots of QTL signals for the three detection methods are shown for each data set in Figures 1, 2, 3, 4 and 5.

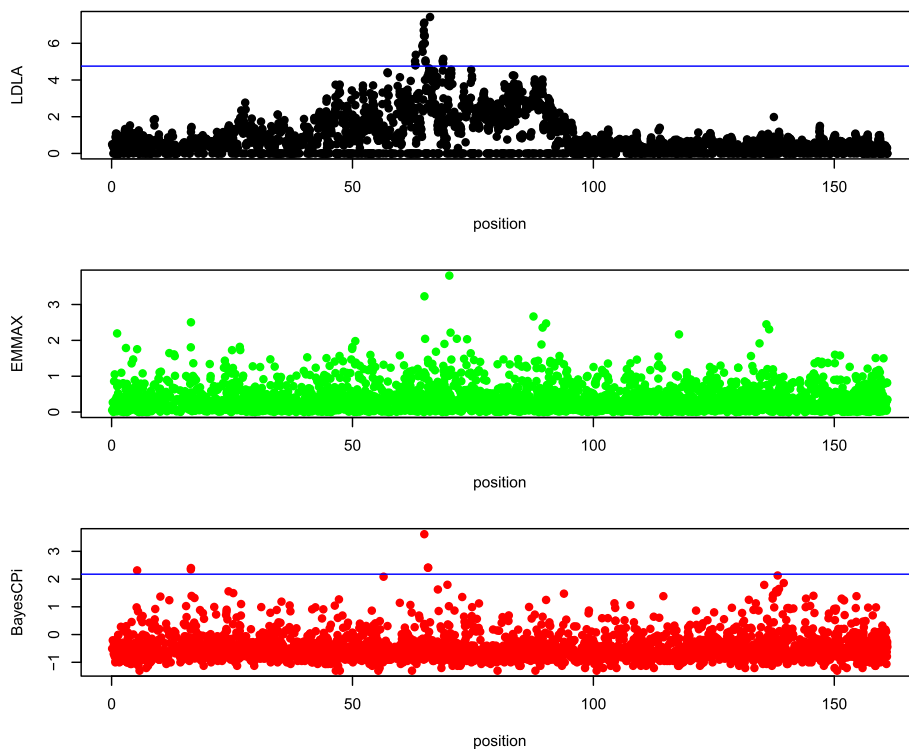
### Numbers of positive signals

Table 3 shows the numbers of significant QTL positions detected. These numbers should be regarded with caution because the actual numbers of QTL are not known. Of the three methods, LDLA had the largest number of signals for dairy cattle, horses, and sheep. For beef cattle and pigs, no significant association was found using any of the three detection methods.

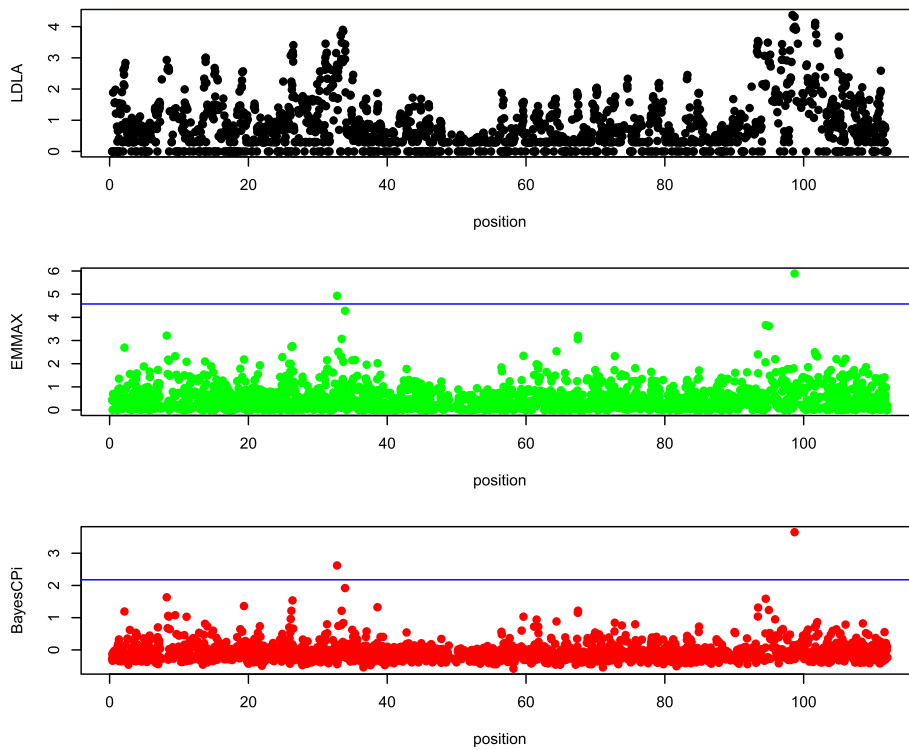
**Table 2 Description of simulated scenarios**

Scenario	Nb of QTL	Type of effect	MAF	Linkage disequilibrium*
1	1	Additive	0.3	0.8
2	1	Additive	0.3	0.4
3	1	Additive	0.1	0.8
4	1	Dominant	0.3	0.8
5	2	Additive	0.3	0.8

Nb = number; QTL = quantitative trait loci, MAF = minor allele frequency  
\*correlation between QTL and a close marker.

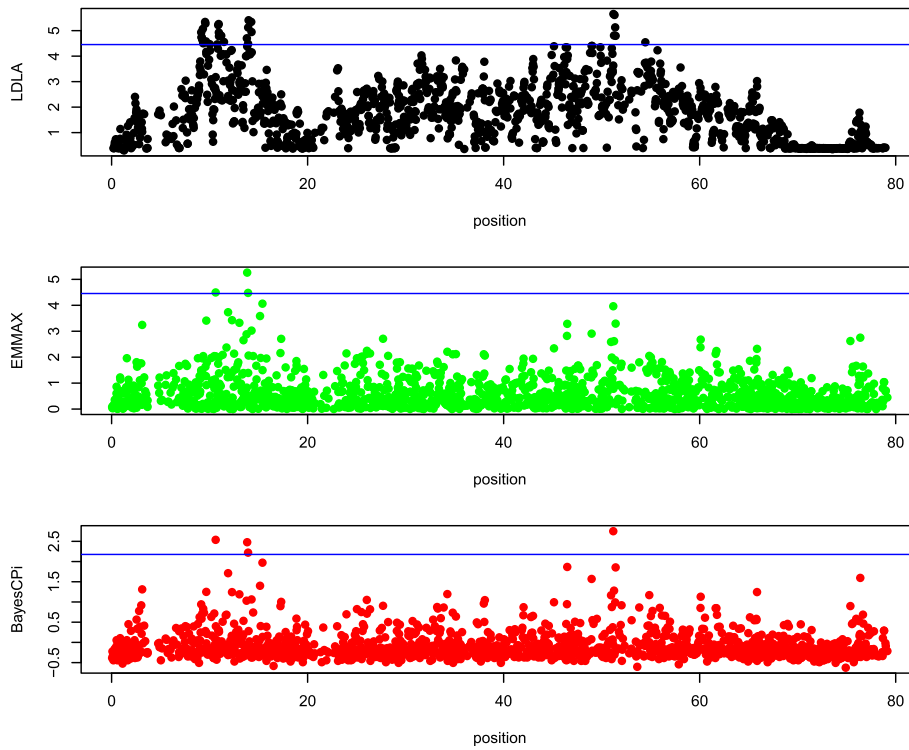


**Figure 1** Manhattan plot of chromosome 1 in dairy cattle. The y-axis is  $\log_{10}(1/P)$  value for LDLA and EMMA and  $\log_{10}(\text{Bayes factor})$  for BayesC; the x-axis is the position along the chromosome in cM; the blue line (if any) is the rejection threshold.

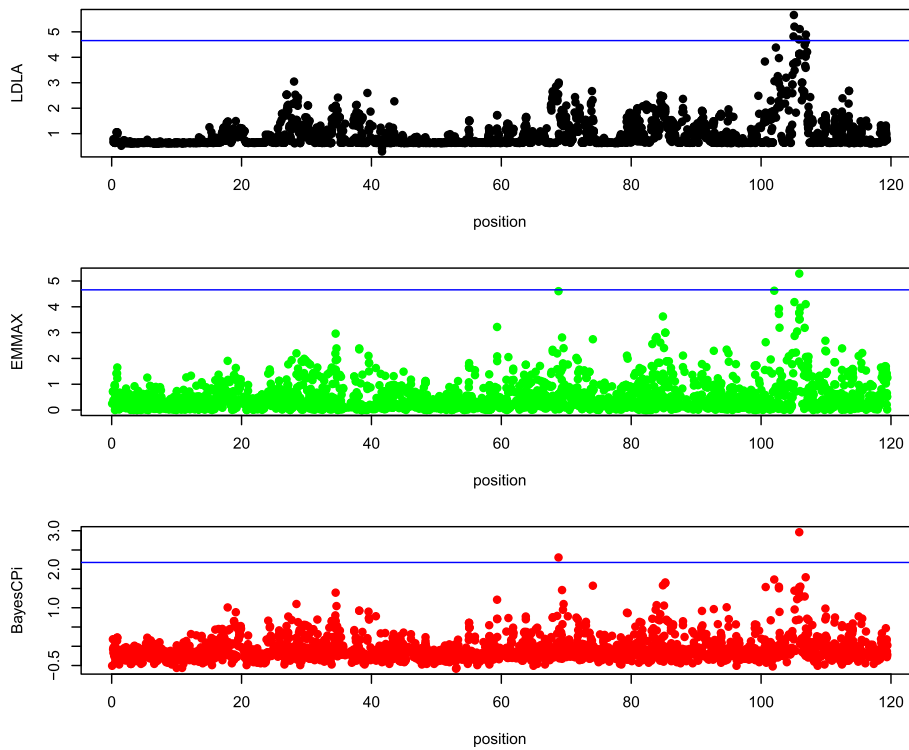


**Figure 2** Manhattan plot of chromosome 7 in beef cattle. The y-axis is  $\log_{10}(1/P)$  value for LDLA and EMMA and  $\log_{10}(\text{Bayes Factor})$  for BayesC; the x-axis is the position along the chromosome in cM; the blue line (if any) is the rejection threshold.

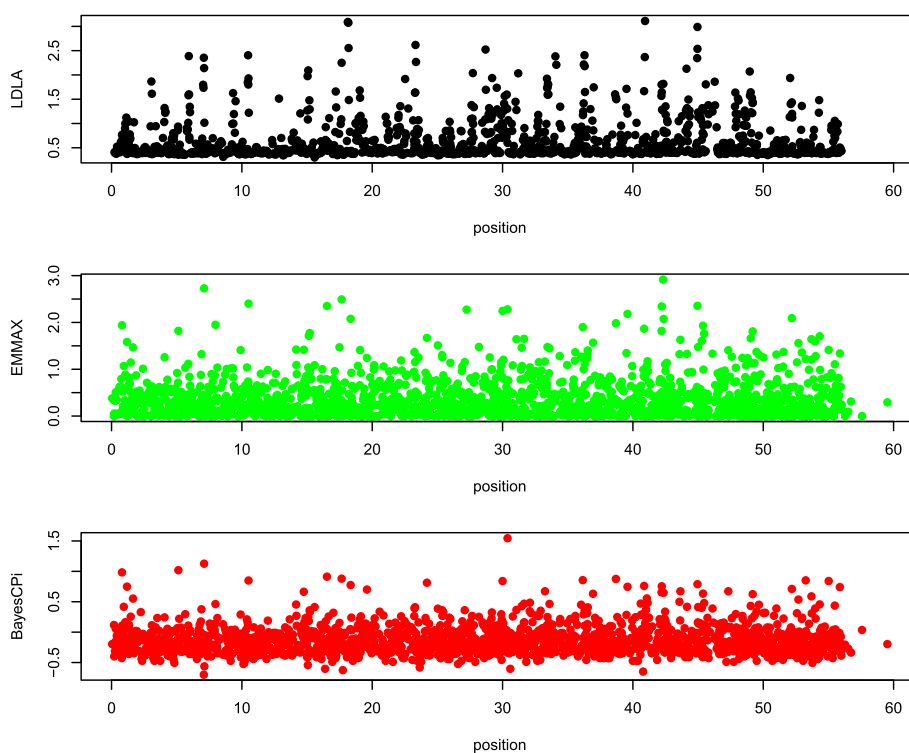




**Figure 3** Manhattan plot of chromosome 12 in sheep. The y-axis is  $\log_{10}(1/P)$  value for LDLA and EMMA and  $\log_{10}(\text{Bayes Factor})$  for BayesC; the x-axis is the position along the chromosome in cM; the blue line (if any) is the rejection threshold.



**Figure 4** Manhattan plot of chromosome 3 in horses. The y-axis is  $\log_{10}(1/P)$  value for LDLA and EMMA and  $\log_{10}(\text{Bayes Factor})$  for BayesC; the x-axis is the position along the chromosome in cM; the blue line (if any) is the rejection threshold.



**Figure 5** Manhattan plot of chromosome 17 in pigs. The y-axis is  $\log_{10}(1/P)$  value for LDLA and EMMA and  $\log_{10}(\text{Bayes Factor})$  for BayesC; the x-axis is the position along the chromosome in cM; the blue line (if any) is the rejection threshold.

**Agreement**

Because the objective of QTL mapping is not only to detect QTL but to provide locations for further investigation by molecular geneticists, agreement of the location of positive signals between the three detection methods was evaluated. Informal visual agreement between methods is evident from Figures 1, 2, 3, 4 and 5. Additional Tables S1, S2, S3, S4 and S5 [See Additional file 1: Tables S1, S2, S3, S4 and S5] show exact locations and statistics for the two major signals found in each analysis. For example, the dairy cattle data showed a QTL across methods in a window of 2 cM around 65 cM and beef cattle data show agreement between methods of a QTL around 32 and 98 cM, etc. For sheep and beef cattle, the figures show more agreement than the additional tables, because the peaks in agreement may be first or second highest signals in one method but third or fourth highest signals in

another. Only the pig dataset did not show clear agreement for detected QTL, but no significant positive signals were found for pigs. We believe that the true signals that were present in these datasets were found by all methods (within a bracket of 1 to 5 cM) and that false signals were random noise and, therefore, without agreement.

**Redundancy**

Estimates of the numbers of independent tests are in Table 4. Redundancy increased with family structure (*i.e.* number of half- and full-sibs within family); *e.g.* in horses, the large number of small families resulted in low redundancy. In general, LDLA tended to be the most redundant (*i.e.* a small number of independent tests). Redundancy was less for EMMA and BayesC than for LDLA because (1) haplotypes (or haplotype structure) are highly correlated from one position to the next (but SNP effects are

**Table 3** Numbers of significant positions detected for quantitative trait loci by species and detection method

Method	Dairy cattle	Beef cattle	Sheep	Horses	Pigs
LDLA*	20	0	29	6	0
EMMA**	0	2	3	1	0
BayesC***	6	2	4	2	0

\*statistic = likelihood ratio test corrected by Bonferroni; \*\*statistic = t-statistic corrected by Bonferroni; \*\*\*statistic = Bayes factor.

**Table 4** Numbers of independent tests by species and detection method

Method	Dairy	Beef	Sheep	Horses	Pigs
LDLA*	237	183	153	978	380
EMMA**	2323	1082	770	968	1672
BayesC***	2390	1889	1108	1303	1509

\*statistic = likelihood ratio test corrected by Bonferroni; \*\*statistic = t-statistic corrected by Bonferroni; \*\*\*statistic = Bayes factor.

**Table 5 Estimated mean squared errors (MSE), power, and false discovery rates (FDR) of detection methods across simulation factors**

Method	MSE	Power	FDR
BayesC*	0.22	0.36	0.007
EMMA**	0.21	0.36	0.006
LDLA***	0.39	0.26	0.017

\*statistic = likelihood ratio test corrected by Bonferroni; \*\*statistic = t-statistic corrected by Bonferroni; \*\*\*statistic = Bayes factor; standard errors = 0.02 for MSE, 0.005 for power, and 0.0005 for FDR.

not) and (2) LDLA is partly family (linkage) based (but EMMA and BayesC are not). In addition, BayesC estimators are marginalized over all other loci.

### Simulations

Quality indicators for QTL detection methods are in Table 5 as least-squares means (estimated marginal means over a balanced design across all other simulation factors [34]). The LDLA method had not only lower power but also higher MSE and FDR. The EMMA and BayesC methods were clearly superior to LDLA in the simulations, regardless of scenario or simulation factor (data not shown).

### Discussion

Overall, the results were similar with the three QTL detection methods. The Manhattan plots had similar profiles and indicated the same QTL locations for significant (or close to significant) points. None of the methods appeared to produce spurious peaks because of population stratification. All three methods accounted for relationships either through pedigree or genome-wide markers.

Table 3 suggests that LDLA is more powerful than EMMA or BayesC. However, in Figures 1, 2, 3, 4 and 5, some multiple signals obtained with LDLA corresponded to a single peak obtained with EMMA and BayesC, which suggests that many LDLA signals echo a single QTL. This makes sense because the covariance structure for haplotypes changes little from one position to another (as evident in Table 4); i.e. LDLA uses linkage information. All QTL regions that were detected by LDLA were also detected by EMMA and BayesC; thus, LDLA was not more powerful than EMMA or BayesC.

The performance of the methods differed slightly between the real and the simulated data. For example, LDLA produced more positive signals with real data than EMMA and BayesC. In contrast, for simulated data (Table 5), LDLA was the least powerful method. Several explanations for this contradiction are possible. One is that only one QTL was simulated, whereas many potential QTL exist in the real data. Another explanation is that LDLA produces several correlated signals for a single QTL, as discussed previously. Still another

explanation is that the nominal threshold used for the real data was too liberal and, therefore, produced false positives. The threshold to be used for the real data is also debatable (e.g. [35]). Although a reasonable consensus exists on how to calculate a threshold under a strict null hypothesis (no QTL on the chromosome), such a hypothesis is nonsensical because biologically there must be many, possibly small QTL on each chromosome. Therefore, a pseudo-null hypothesis was used for our simulations. As an example of the problems caused by the use of rejection thresholds, the two highest (although non-significant) peaks of EMMA in Figure 1 agree with the significant peaks obtained with LDLA and BayesC. In Figure 2, (non-significant) results from LDLA agree with results from the other methods. Therefore, no one method is clearly the most powerful for a given dataset because the ideal rejection threshold depends on the true genetic architecture of a dataset. Thus, a common sense approach is to use several different methods for QTL detection.

For QTL location, the methods were in agreement for the real data. This was as expected because the chromosomes were suspected to contain QTL. For the pig data, no agreement for QTL location was found among the methods, but no signals were significant; using several methods and obtaining no significant signals indicates the absence of detectable QTL for this dataset. However, for the beef cattle data, no significant QTL were found with LDLA but significant QTL were evident when using EMMA and BayesC.

The results from the BayesC and EMMA methods agreed to a very large extent for the real and simulated data because they both model each SNP as having an effect. The complexity of BayesC and lack of a consensual rejection threshold for BayesC might make EMMA preferable. If BayesC is used, Bayesian testing (i.e. using BF) is recommended, since it produced results that are comparable within and across datasets.

Finally, LDLA results in a more correlated structure of signals across the genome. This was evident as less power and accuracy in QTL mapping with simulated data but was not clear in the real data. Therefore, the larger number of independent tests found for EMMA and BayesC does not necessarily imply greater accuracy.

Some discussion about the scope of this work is also needed. Many studies have compared QTL detection methods based on simulations. In those simulations, the QTL location and action are precisely described. However, very little is known about QTL action in real data and, thus, such simulations tend to provide incomplete results because they do not account well for the range of possible gene actions. Phenomena such as self-regulation, epistasis, and pleiotropy are particularly difficult to simulate. Our study, which was based on real data,



attempts to address the question of whether analysis of a dataset with another method will provide a different result. We have given some empirical answers to this question.

## Conclusions

All three methods (haplotype-based, association analysis with relationship matrices, and Bayesian analysis) were suitable for QTL detection. Using several methods for the same dataset is recommended. EMMA was easier to use and had slightly higher accuracy and power for both real and simulated data. For QTL with large signals, the methods agreed well. Establishment of significance thresholds is difficult because it is unclear what the null hypothesis should be [35]. The use of nominal  $P$  values corrected for multiple-testing for the frequentist methods and of BF for the Bayesian method resulted in acceptable statistical properties and led to similar conclusions across methods and datasets.

## Additional file

**Additional file 1: Table S1.** Dairy cattle: positions of the first and second largest signals for each method. Positions of two largest QTL detection signals. **Table S2.** Beef cattle: positions of the first and second largest signals for each method. Positions of two largest QTL detection signals. **Table S3.** Horses: positions of the first and second largest signals for each method. Positions of two largest QTL detection signals. **Table S4.** Title: Sheep: positions of the first and second largest signals for each method. Positions of two largest QTL detection signals. **Table S5.** Pigs: positions of the first and second largest signals for each method. Positions of two largest QTL detection signals.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

SA, SF, CM, GS, MPS, and AR prepared the real datasets. The datasets were analyzed by SA, ST, GS, PC, MPS, and AL. AL, JME, AR, PC, and MPS discussed, prepared, and analyzed the simulations and prepared a first draft that was completed by the rest of the authors. All authors read and approved the final manuscript.

## Acknowledgements

Most of the research was funded by the French National Research Agency (ANR, Paris, France) program Rules & Tools. All data were acquired with funding by APISGENE (Paris, France) and the following grants. Dairy cattle data were funded by ANR grant CARTOFINE. Beef cattle data were funded by the QUALVIGENE program supported by ANR (Contracts n°05-GANI-017-04, ANR-GANI-017, ANR-06-GANI-001), the Office de l'Élevage (Montreuil sous Bois, France), and the Fonds National de l'Élevage (Paris, France). Sheep data were funded by ANR program SHEEPSNPQTL and European grant 3SR. Horse data were funded by ANR program GENEQUIN, the Fonds Eperon, the French Horse and Riding Institute (IFCE), and the Basse-Normandie Regional Council. Pig data were funded by ANR programs Delisus (ANR-07-GANI-001) and Immopig (ANR-06-GANI-008) as well as BIOPORC (ADN, Nucleus and Gene + breeding organizations). We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrénées for providing computing and storage resources. We thank Suzanne Hubbard for technical editing of the manuscript. Editors and reviewers are acknowledged for useful comments and suggestions.

## Author details

<sup>1</sup>INRA, UMR 1388 GenPhySE, BP52627, 31326 Castanet Tolosan, France. <sup>2</sup>INRA, UMR 1313 GABI, Domaine de Vilvert, 78352, Jouy-en-Josas, France. <sup>3</sup>Current

address: RAGT-R2n, Le bourg, 12510 Druelle, France. <sup>4</sup>INRA, UMR1282 Infectiologie et Santé Publique, F-37380 Nouzilly, France. <sup>5</sup>Université François Rabelais de Tours, UMR1282 Infectiologie et Santé Publique, 37000 Tours, France. <sup>6</sup>Agrocampus Ouest, UMR1348 Pegase, F-35000 Rennes, France. <sup>7</sup>INRA, UMR1348 Pegase, F-35590 Saint-Gilles, France. <sup>8</sup>Université Européenne de Bretagne, Rennes, France. <sup>9</sup>UNCEIA, Genetics Team, 75595 Paris, France. <sup>10</sup>Recherche et Innovation, IFCE, 61310 Exmes, Paris, France.

Received: 10 January 2014 Accepted: 6 January 2015

Published online: 12 February 2015

## References

1. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet.* 2006;38:203–8.
2. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38:904–9.
3. Kang HM, Sul JH, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet.* 2010;42:348–54.
4. Geyer CJ. Practical Markov Chain Monte Carlo. *Stat Sci.* 1992;7:473–83.
5. Browning SR, Thompson EA. Detecting rare variant associations by identity-by-descent mapping in case-control studies. *Genetics.* 2012;190:1521–31.
6. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* 2001;157:1819–29.
7. Xu S. Estimating polygenic effects using markers of the entire genome. *Genetics.* 2003;163:789–801.
8. Yi N, Xu S. Bayesian LASSO for quantitative trait loci mapping. *Genetics.* 2008;179:1045–55.
9. Sahana G, Gulbrandsen B, Janss L, Lund MS. Comparison of association mapping methods in a complex pedigree population. *Genet Epidemiol.* 2010;34:455–62.
10. Druet T, Fritz S, Boussaha M, Ben-Jemaa S, Guillaume F, Derbala D, et al. Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. *Genetics.* 2008;178:2227–35.
11. Teysnière S, Elsen J-M, Ricard A. Statistical distributions of test statistics used for quantitative trait association mapping in structured populations. *Genet Sel Evol.* 2012;44:32.
12. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet.* 2012;44:821–4.
13. Habier D, Fernando RL, Kizilkaya G, Garrick DJ. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics.* 2011;12:186.
14. VanRaden PM, Wiggins GR. Derivation, calculation, and use of national animal model information. *J Dairy Sci.* 1991;74:2737–46.
15. Allais S, Levéziel H, Payet-Duprat N, Hocquette J, Lepetit J, Rousset S, et al. The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. *J Anim Sci.* 2010;88:446–54.
16. Bishop SC, Morris CA. Genetics of disease resistance in sheep and goats. *Small Ruminant Res.* 2007;70:48–59.
17. Salle G, Jacquet P, Gruner L, Cortet J, Sauve C, Prevot F, et al. A genome scan for QTL affecting resistance to *Haemonchus contortus* in sheep. *J Anim Sci.* 2012;90:4690–705.
18. Teysnière S, Dupuis MC, Guérin G, Schibler L, Denoix JM, Elsen JM, et al. Genome-wide association studies for osteochondrosis in French Trotter horses. *J Anim Sci.* 2012;90:45–53.
19. Sanchez MP, Tribut T, Iannuccelli N, Bouffaud M, Servin B, Tenghe A, Dehais P, Muller D, Del Schneider MP, Mercat MJ, Rogel-Gaillard C, Milan D, Bidanel JP, Gilbert H. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. *Genet Sel Evol* 2014, 46:12.
20. Meuwissen THE, Karlsten A, Lien S, Olsaker I, Goddard ME. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics.* 2002;161:373–9.
21. Druet T, Georges M. A hidden Markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. *Genetics.* 2010;184:789–98.
22. Visscher PM. A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Res Hum Genet.* 2006;9:490–5.

23. Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, et al. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res.* 2002;12:222–31.
24. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci.* 2008;91:4414–23.
25. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. BLUPF90 and related programs (BGF90). In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production: 19–23 August 2002. Montpellier. CD-ROM Communication N° 28–07; 2002.*
26. Hoggart CJ, Whittaker JC, De Iorio M, Balding DJ. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS Genet.* 2008;4:e1000130.
27. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc.* 1995;90:773–95.
28. Wakefield J. Bayes factors for genome-wide association studies: comparison with P-values. *Genet Epidemiol.* 2009;33:79–86.
29. Legarra A, Ricardi A, Filangi O. GS3: Genomic selection, Gibbs sampling, Gauss-Seidel (and BayesCpi): [<http://genotoul.toulouse.inra.fr/~alegarra>] 2011.
30. Sorensen D, Gianola D. Likelihood, bayesian and MCMC methods in quantitative genetics. New York: Springer-Verlag; 2002.
31. Boddicker N, Waide EH, Rowland RR, Lunney JK, Garrick DJ, Reecy JM, et al. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. *J Anim Sci.* 2012;90:1733–46.
32. Vidal O, Noguera JL, Amills M, Varona L, Gil M, Jiménez N. Identification of carcass and meat quality quantitative trait loci in a Landrace pig population selected for growth and leanness. *J Anim Sci.* 2005;83:293–300.
33. Harrell FE, Davis C. A new distribution-free quantile estimator. *Biometrika.* 1982;69:635–40.
34. Institute S. SAS/STAT 9.3 user's guide. Cary, NC: SAS Institute; 2011.
35. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. *Nat Rev Genet.* 2014;15:335–46.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

