International Journal of Hybrid Information Technology
Vol. 9, No.8 (2016), pp. 23-32
http://dx.doi.org/10.14257/ijhit.2016.9.8.03

Bayesian Genome-wide Mapping Interacting QTL for Survival Time

Xiaojing Zhou[1] and Dan Zhao[2],

[1]*Department of Mathematics, Heilongjiang Bayi Agriculture University, Daqing, 163319, People's Republic of China*
[2]*Department of Mathematics, Anshan Normal University, Anshan,114007,, People's Republic of China*

Abstract

In survival analysis, the accelerated failure time (AFT) model, as it is a linear function of the logarithms of survival times on covariates with easily interpretable parameters, is considered to be a useful alternative to the proportional hazards model (PHM). To dissect complex genetic architecture for survival traits, which have a skewed distribution and are often subject to censoring, we construct a multiple interacting QTL model based on the parametric AFT model with the baseline distribution of log-t distribution. Bayesian model selection is proposed to estimate the main and epistatic effects of QTLs in a computationally efficient manner, in which, the prior distribution of the scaled parameter in the AFT model is specified as the inverted chi-square distribution rather than a constant. Simulation experiments showed that our proposed method was superior to Bayesian method for normal phenotypes in terms of both the statistical powers of QTL detection and the precision of QTL parameter estimation. Three new pairs of epistatic QTLs were found in analyzing a real dataset for flowering time in rice.

Keywords*: survival trait, accelerated failure time (AFT) model, Bayesian model selection, interacting loci*

1. Introduction

Survival traits, which are broadly defined as the length of time between two events, have started to draw the attention of some researchers for exploiting the approaches to mapping the traits. Except for the characteristic of skewed distribution, there traits are also difficult to follow up. Many methodologies in survival analysis, including a cure-rate model[1], distribution-free non-parametric[2-5], Cox parametric [6,7] and Cox semi-parametric models[8-10] are sequentially applied into the interval mapping[11] of survival trait loci. In outbred populations, the variance component based on the methods of Epstein *et al.* [12] or Pankratz *et al*. [13] are appropriate for mapping QTL of survival traits. However, all those mapping approaches just estimate and test one locus at a time. Subsequently, Bayesian mapping method, which is able to simultaneously identify multiple QTLs[14-19], has been introduced in QTL detection. The number of QTL is determined either by the Bayes factor[20-21] or by reversible-jump MCMC[14]. Although such Bayesian mapping approach improves statistical power of QTL detection, it has been noted that drawing number of QTLs with a reversible-jump MCMC procedure may result in lower convergence efficiency and poor mixing. Moreover, the effects of deviation from the assumption on normal distributions have not been fully addressed, nor the censoring mechanism of survival traits has been taken into account. On the basis of the Bayesian shrinkage mapping, Wang *et al.* have developed a robust mapping strategy for analyzing continuous non-normal quantitative traits[22], by replacing the normal distribution for residuals in multiple QTL model with a Student-t distribution[23-25].

ISSN: 1738-9968 IJHIT

Their method only focused on the estimates for main-effect QTL, whereas ignored epistatic effects of the detected QTL. The genetic architecture of quantitative trait includes not only the number and locations of QTL, but also their main and epistatic effects. In particular, the unknown number of QTL and possible huge epistatic effects make the dissection for genetic architecture of quantitative trait extremely complex.

In survival analysis, besides the Cox proportional hazard model (PHM), the accelerated failure time (AFT) model is also the natural choice for formulating the association of phenotypes with QTL effects on the survival times[26]. The AFT model has an intuitive physical interpretation for real-life examples, as it directly expresses the failure time, rather than the probability as in the PHM, and therefore would be an important alternative to the PHM[27-28]. The AFT model makes modeling simple as it relates the logarithm of the failure time linearly to the covariates [29-30]. It also reduces the potential error amplifications from linking models with different structures. However, few studies have used AFT model to map QTL. Cheng and Tzeng[26] have proposed parametric and semi-parametric methods that based on AFT models for interval mapping, but the fact that using the likelihood derived by Diao *et al*.[6] to estimate model parameter greatly increases the computational complexity. Furthermore, the mapping approach does not take any account of the interactions among QTLs.

In this study, for dissecting complex genetic architecture for survival traits, we firstly construct a multiple interacting QTL model based on the parametric AFT model, in which the log-*t* distribution is considered as baseline function, and then develop a Bayesian model selection strategy for estimating the QTL parameters, finally, demonstrate the flexibility and utility of the method by conducting simulation experiments. Then apply the method to a real dataset for flowering time in rice. The comparing results between the method proposed here with the traditional Bayesian mapping under the normal distribution on both the simulation and the real data analysis indicate that our method has an improved power in mapping QTL with normal and non-normal phenotypes.

2. Theory and Methods

1) Genetic model

For simplicity, we consider a backcross population derived from two inbred lines to describe the Bayesian mapping model for quantitative traits. Certainly, the method can be easily applied to other experimental designs, such as F_2 design and four-way crosses. We measure genotypes of a set of co-dominant molecular markers with a known genetic linkage map as well as phenotypes for the trait of interest on *n* individuals. Supposing there are *m* quantitative trait loci controlling the trait of interest, we formulate the multiplicative effects of multiple interacting QTL on the survival time T_i by using the following AFT model:

$$\log(T_i) = \mu + \sum_{j=1}^{m} \gamma_j x_{ij} \alpha_j + \sum_{k>j}^{m} \gamma_{jk} z_{ijk} \delta_{jk} + \frac{e_i}{\sqrt{w_i}} \quad (1)$$

Where, as usual, μ is the population mean; α_j for $j = 1, \cdots, m$ is the additive effect of the *j*th QTL; δ_{jk} is the epistatic effect between *j*th QTL and *k*th QTL for $j = 1, 2, \cdots, m; k = j, j+1, \cdots, m$. Variable x_{ij} is a genotype indicator variable for individual *i* at locus *j* and defined as +1 for genotype Qq and -1 for genotype qq, $z_{ijk} = x_{ij} x_{ik}$; $\gamma_\bullet$ is a binary variable for each genetic effect, indicating that the corresponding effect is included ($\gamma_\bullet$ =1) or excluded ($\gamma_\bullet$ = 0) from model (1). Through inferring the $\gamma_\bullet$, we shall adopt Bayesian model selection to Markov Chain Monte Carlo (MCMC) sampling in an optimal model space; e_i is a random environmental error,

distributed as $N(0, \sigma^2)$ with σ^2 being residual variance; and w_i is a positive random variable which subjects to $Gamma(df/2, df/2)$ distribution with df being a scalar parameter.

2) Bayesian model selection

After organizing all genetic effects into β and all indicator variables or dummy variables into x_i, we can simplify the multiple interacting QTL model as the following linear model:

$$y_i = \mu + x_i\beta + \frac{e_i}{\sqrt{w_i}} \quad \text{with} \quad y_i = \log(T_i) \tag{2}$$

This, in fact, is not a common linear model, because the number of independent variables in the model and the associated design matrix are all unknown due to the unknown number of QTLs. Moreover, the residuals subject to student-*t* distribution rather than normal distribution. We approximate positions for all possible QTLs using a partition of the entire genome into evenly spaced loci, including all observed markers and additional loci between flanking markers, and then calculate the expected values for elements in relative design matrix to each locus based on the conditional probabilities of the loci genotypes on two flanking markers. As a result, a huge number of genetic effects for main-effect and epistatic QTLs will be required to estimate. It has been hypothesized that the genetic variation of most quantitative traits is actually controlled by a few loci with large effects and a large number of loci with small effects[31]. This suggests that among those estimated genetic effects, only a few are large or significant and most of them are small or neglectable. Therefore, Bayesian model selection based on a composite space representation[17-19, 32] provides a simple and efficient way to identify a small number of large or significant genetic effects in multiple interacting QTL model.

The Bayesian mapping approach starts with presetting the upper bound of the number of QTLs included in the model[18], which is greater than the number of detectable QTL in a given data set. Given the upper bound of the number of QTL, these QTL will be drawn from densely spaced loci across the genome. Even with a moderate number of the upper bound, there are many genetic effects being estimated in the model (2). To make inference of the existence of these effects, we introduce a random binary variables γ to indicate which genetic effects are included in ($\gamma = 1$) or excluded from ($\gamma = 0$) the model[33-37]. Let $\Omega = \mathrm{diag}(\gamma)$, model (2) becomes

$$y_i = \mu + x_i\Omega\beta + \frac{e_i}{\sqrt{w_i}} \tag{3}$$

Within the framework of Bayesian model selection, Bayesian sampling for unknown parameters including μ, Ω, β and w_i in model (3), is implemented with MCMC algorithm. In the course of Bayesian sampling, the realized sampling value for γ in the matrix Ω at this round determines which genetic effect and position of QTL will be drawn or estimated at next round. As a rule, the large genetic effects are included in the model with higher probabilities than the small ones. This will greatly save the sampling time, as just a few large main and epistatic effects are drawn in each round.

We implement MCMC sampling by following the simplified and computationally efficient procedures:

(1) Compute the expected values for the associated design matrix with all spaced loci over the genome:

$$E(x) = \pi_{QQ} - \pi_{Qq}$$

with π_{QQ} and π_{Qq} being the conditional probabilities of genotypes QQ and Qq on two flanking markers, respectively.

(2) Initialize all variables with some legal values or values sampled from their prior

distributions. Herein, the upper bound for the number of QTL *L* is estimated by $L = l_0 + 3\sqrt{l_0}$, with l_0 being the prior expected number of QTL (including main and epistatic effect QTLs) that is determined according to initial investigations with traditional methods; the binary indicator γ is assigned to be independence prior $p(\gamma) = \prod w^{\gamma}(1-w)^{(1-\gamma)}$, where, the prior inclusion probability for main effect $w_m = 1 - \left[1 - \frac{l_m}{L}\right]^{\frac{1}{3}}$ with l_m being the expected number of main effect QTL and $w_e = 1 - \left[\frac{1 - l_0 / L}{(1-w_m)^3}\right]^{1/9(L-1)}$ for epistatic effect.

(3) Update population mean μ by sampling from a normal distribution with mean $\left(\sum_{i=1}^{n} w_i\right)^{-1} \sum_{i=1}^{n} w_i (y_i - x_i\beta)$ and variance $\left(\sum_{i=1}^{n} w_i\right)^{-1} \sigma^2$.

(4) Update the binary indicators γ by adopting an efficient Metropolis-Hastings algorithm [19, 37] with the probability of acceptance min (1, ρ), where $\rho = \left(\frac{wR}{1-w}\right)^{1-2\gamma}$ with $R = \sqrt{\frac{c}{c+1}} \exp\left(-\frac{\hat{\beta}_j^2}{2\hat{\sigma}_j^2}\right)$.

(5) Update the QTL effects corresponding to $\gamma = 1$ by drawing from the normal distribution with mean $\hat{\beta}_j = \frac{c}{c+1}(\sum_{i=1}^{n} w_i x_{ij}^2)^{-1} \sum_{i=1}^{n} w_i x_{ij}(y_i - x_i\beta + x_{ij}\beta_j)$ and variance $\hat{\sigma}_j^2 = \frac{c}{c+1}(\sum_{i=1}^{n} w_i x_{ij}^2)^{-1}\sigma^2$, where *j* is the numbering of corresponding genetic effect, *c* takes *n*. Note that if $\gamma = 0$, then the corresponding β_j is taken to be zero.

(6) Update the residual variance σ^2 by sampling from an inverted Chi-square distribution with parameters $v_e + n$ and $(v_e + n)S_e + \sum_{i=1}^{n} w_i (y_i - \mu - x_i\beta)^2$, where v_e and S_e are prior hyperparameters.

(7) Update w by drawing from a Gamma distribution with parameters $\frac{1+df}{2}$ and

$$2\left[df + \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \mu - \sum_{j=1}^{q} x_{ij}b_j)^2\right]^{-1}.$$

(8) Update *df* using the Metropolis–Hastings algorithm[19, 36], based on the conditional posterior density of *df*:

(9) Update the QTL position by drawing from all spaced loci over the genome, corresponding $\gamma = 1$. Note that the existence of QTL depends on whether the $\gamma = 1$ for either main or epistatic effect. Each locus is sampled from a variable interval whose boundaries are the positions of adjoining QTLs. Metropolis-Hastings algorithm is used to determine whether each proposed (new) position should be accepted or not [38].

(10) Repeat steps (3) – (9) until the Markov chain reaches a desirable length.

Post MCMC analysis includes the monitor of the mixing behavior and convergence rates of MCMC algorithms and the assessment of characteristics of genomic architecture. The former can adopt visually inspecting trace plots of the sample values of scalar quantities of interest or formal diagnostic methods provided in the package R/coda, and the latter use model averaging accounts for model uncertainty[39-41] that averages over possible models weighted by their posterior probabilities. We can use various methods to

graphically and numerically summarize and interpret the posterior samples. The posterior inclusion probability for each locus is estimated as its frequency in the posterior samples. Bayes factor (BF) is used to show evidence for inclusion against exclusion of each QTL locus or effect[18, 20]. Generally, a threshold of BF is empirically determined as 3, or $2\ln BF = 2.1$, for declaring statistical significance for each QTL effect[20].

3. Simulation Studies

A single large chromosome segment with lengh of 500 cM was simulated for a BC population with sample sizes of 150 and 300, on which sixty-one co-dominant markers were spaced evenly. Four main-effect QTLs were put along the chromosome, two pairs of which interact. We simulated these QTLs and marker genotypes with Bayesian mapping based on AFT model and the traditional Bayesian mapping procedure with normal residuals, respectively. The population mean was taken to be $\mu = 5.0$. The degree of freedom (*df*) was assigned to be 3, so that the residual variance was 3.0. Given these parameters, the phenotype of quantitative trait is randomly generated on each individual according to model (1).

Before Bayesian sampling with MCMC algorithm, we assign the prior number of main-effect QTL at $l_m = 3$ and the prior expected number of epistatic QTL at 3, so that the upper bound of the number of QTL is $L = 6 + 3\sqrt{6} \approx 13$. The actual values for the hyper parameters are taken to be $v_e = 0$ and $s_e = 1$. The initial values of all variables are sampled from their prior distributions. In the MCMC sampling, the burn-in is determined as 6000 cycles by visually inspecting the plots of some samples across the rounds. MCMC is additionally run for 160,000 cycles after the burn-in period (deleted). To reduce serial correlation, we save one observation in every 40 cycles and therefore obtain independent posterior sample of 4,000 observations for the post-MCMC analysis. Considering each simulation is more time consuming, the simulations are repeated 50 times to evaluate statistical power of QTL detection.

The mapping results with both methods are listed in Table 1. The standard deviations with the method proposed here are smaller than the ones with the traditional Bayesian mapping method, especially for the interacting QTLs. Apparently, Bayesian genome-wide mapping based on the AFT model performs significantly higher statistical power of QTL detection than traditional Bayesian mapping if the residual error subjects to heavy-tailed distribution. Our method can detect the QTL which cannot be detected by the traditional Bayesian mapping method. The results indicate that Bayesian mapping interacting loci based on AFT model is able to better estimate the effects and positions of detected QTLs. The estimating precision of parameters and the statistical power of QTL detection, as expected, increase as sample size increased. Table 1 also displays the mean estimates of the genetic effects for simulated QTL obtained with AFT mapping model. It is shown clearly that the estimates of QTL effects are fairly close to the true parameter values except a pair of interacting QTLs.

Table 1. Mean Estimates and Sds (In Parentheses) of QTL Positions and Effects, Statistical Power of QTL Detection Obtained with Bayesian Analysis Based on the AFT Model and Normal Model

Sample size	Methods	QTL parameters	QTL No					
			1	2	3	4	5	6
150		True position	56	148	267	359	56×267	148×359
		True effect	0.45	0.70	0.30	0.55	0.30	0.20
	AFT	Position	54.3(5.2)	147.4(3.1)	267.0(6.0)	358.7(3.5)	56.7(9.8)×270.2(10.7)	151.4(6.2)×357.8(9.1)
		effect	0.44(0.09)	0.71(0.15)	0.41(0.09)	0.58(0.12)	0.25(0.15)	0.05(0.17)
		Power	65.0%	90.0%	40.0%	70.0%	50.0%	35.0%
	Normal	Position	—	148.5(1.9)	260.0(0.0)	360.4(1.7)	46.6(11.8)×264.6(14.3)	145.3(14.1)×359.5(11.9)
		Power	0.0%	20.0%	5.0%	25.0%	5.0%	20.0%
300	AFT	Position	55.4(3.3)	148.2(2.7)	266.2(3.3)	359.2(2.9)	57.7(6.3)×268.2(6.2)	151.9(5.3)×359.2(3.4)
		effect	0.44(0.08)	0.71(0.10)	0.30(0.08)	0.56(0.06)	0.29(0.12)	0.10(0.10)
		Power	80.0%	100.0%	80.0%	95.0%	85.0%	55.0%
	Normal	Position	53.0(4.2)	148.0(4.9)	263.3(0.0)	359.7(4.3)	46.5(10.7)×259(13.3)	149.3(10.5)×358.6(10.7)
		Power	10.0%	25.0%	5.0%	30.0%	10.0%	30.0%

4. Case Analysis

HelveticaIn Bayesian epistatic analysis, except for model (1), we also assume the residual in model (1) to be a normal distribution and compare mapping results from the two models. When the residual is normal distribution, the $w=1$ in model (1) and Bayesian epistatic analysis for survival time is not required to sample *w*. According to the results from the interval nonepistatic mapping[11] and two-dimensional genome scan, the prior number of main-effect QTL was set at $l_m=3$ according to the interval mapping results and the prior expected number of all QTL (l_0) was taken to be l_m+5. The upper bounds of the number of QTL, *L*, were then 16. The hyper parameters v_e and s_e were assigned to be 0 and 1, respectively. The initial values of all variables were drawn from their prior distributions. The MCMC was run for 200,000 cycles after the burn-in period for 10000 cycles.

The estimates for main-effects QTL parameters and the estimates for epistatic-effect QTL parameters obtained with Bayesian mapping based on AFT model and traditional Bayesian mapping procedure with normal residuals were listed in Table 2 and 3, respectively. For the estimates of main-effects QTL parameters, the two models behave almost equally well. While, for the detecting results of interacting QTLs, the AFT model detected a total of six pairs interacting QTLs, covering all the ones identified with the traditional Bayesian mapping method. The detecting results sufficiently validated the effectiveness and flexibility of the Bayesian mapping approach based on the parametric

AFT model.

Table 2. Estimates for Main-Effect QTL Parameters Obtained With Bayesian QTL Analysis for Flowering Time in Rice Based On the AFT Model and Normal Model

Methods	QTL parameters	QTL no		
		1	2	3
AFT	LG-position	6-30.18	12-1.32	14-15.06
	Heritability (%)	**6.51**	14.20	5.25
	Additive effect	**0.50**	0.76	**0.38**
	2lnBF	4.37	8.26	3.34
Normal	LG-position	6-30.18	12-1.32	14-15.06
	Heritability (%)	**6.46**	13.81	**5.07**
	Additive effect	0.50	0.75	0.37
	2lnBF	4.54	8.24	2.96

Table 3. Estimates for Epistatic-Effect QTL Parameters Obtained With Bayesian Analysis for Flowering Time in Rice Based on the AFT Model and Normal Model

QTL no.	AFT				Normal			
	LG-positions	Heritability (%)	Additive effects	2lnBF	LG-position	Heritability (%)	Additive effects	2lnBF
1	(3-29.91) × (4- 74.22)	11.79	1.49	6.07	—	—	—	—
2	(3-40.96) × (9-0.00)	12.15	1.40	7.94	—	—	—	—
3	(5-8.25) × (14-1.04)	10.23	0.97	8.76	(5-9.29) × (14-1.04)	10.03	0.83	8.95
4	(6-16.48) × (8-9.55)	12.58	1.42	9.77	(6-17.55) × (8-8.26)	11.59	1.45	9.44
5	(10-4.52) × (14-0.00)	10.73	1.14	9.24	(10-6.72) × (14-1.04)	15.56	0.75	9.02
6	(10-3.39) × (15-26.43)	11.97	0.29	7.39	—	—	—	—

6. Conclusion

Survival traits, which have a skewed distribution and are usually subject to censoring due to random loss of follow-up or limited duration of the experiment, have been widely observed in nature, for instance, flowering time in plants, failure time or survival time in animals. In the context of the AFT model based on the log-*t* distribution, we develop a genome-wide mapping strategy for detecting interacting QTLs with the aid of Bayesian model selection, realizing the dissection of complex genetic architecture for survival traits. Our proposed approach performs highly computational efficiency because it allows us to carry out MCMC sampling for QTL parameters in the reduced model space. We conduct two simulation experiments to validate the flexibility of the Bayesian mapping strategy based on the AFT model and the superiority compared with traditional Bayesian model analysis assuming a normal distribution. A real data analysis for flowering time also indicates the flexibility of the method by comparing the results obtained with traditional Bayesian model analysis.

The AFT model we used herein is established under the assumption of log-*t* distribution for survival time. Other skewed distributions, such as exponential, Weibull, log-normal, log-logistic and Gamma distributions *etc.* can be also used to fit survival data and form the likelihood function of accelerated failure time model. With those distributions, the posteriors for genetic effects of QTL cannot be analytically integrated out, thus requiring that these parameters be drawn with Metropolis-Hastings algorithm at each MCMC iteration. Complex MCMC procedures increase computational burden on sampling all model parameters and tend to have poor mixing. In addition, the mapping strategy proposed here can be further extended to more complex experimental population, such as multiple line crosses and outbred population and more complex QTL models including epistatic effects between imprinted QTLs.

Acknowledgements

Fund: The preparation of this manuscript is partially supported by the National Natural Science Foundation of China (No. C170102) and program of Heilongjiang Province Education Administration (No.12531473).
zhouxiaojing7924@126.com
zhaodan_718@126.com

References

[1] K.W. Broman, "Mapping quantitative trait loci in the case of a spike in the phenotype distribution",Genetics, vol. 163, no. 3, **(2003)**, pp. 1169-1175.
[2] W. Coppieters, A. Kvasz and F. Farnir, "A rank-based nonparametric method for mapping quantitative trait loci in outbred half-sib pedigrees", application to milk production in a granddaughter design, Genetics, vol.149, no. 3, **(1998)**, pp. 1547-1555.
[3] J.P. Fine, F. Zou and B.S. Yandell, "Nonparametric estimation of the effects of quantitative trait loci", Biostatistics, vol. 5, no. 4, **(2004)**, pp. 501-513.
[4] L. Kruglyak and E.S. Lander, "A nonparametric approach for mapping quantitative trait loci", Genetics, vol.139, no. 3, **(1995)**, pp. 1421-1428.
[5] F. Zou, B.S. Yandell and J.P. Fine, "Rank-based statistical methodologies for quantitative trait locus mapping", Genetics, vol. 165, no. 3, **(2003)**, pp.1599-1605.
[6] G. Diao, D.Y. Lin and F. Zou, "Mapping quantitative trait loci with censored observations", Genetics, vol. 168, no. 3, **(2004)**, pp. 1689-1698.
[7] C. Symons, M.J. Daly and J. Fridlyand, "Multiple genetic loci modify susceptibility to plasmacytoma-related morbidity in E (mu)-v-abl transgenic mice", Proc Natl Acad Sci USA, vol. 99, no. 17, **(2002)**, pp.11299-11304.
[8] G. Diao and D.Y. Lin, "Semiparametric methods for mapping quantitative trait loci with censored data", Biometrics, vol. 61, no. 3, **(2005)**, pp. 789-798.
[9] Y. Fang, "A note on QTL detecting for censored traits", Genet Sel Evol, vol. 38, no. 2, **(2006)**, pp. 221-229.
[10] C.R. Moreno, J.M. Elsen, P. Le Roy and V. Ducrocq, "Interval mapping methods for detecting QTL affecting survival and time-to-event phenotypes", Genet Res, vol. 85, no. 2, **(2005)**, pp. 139-149.
[11] E.S. Lander and D. Botstein, "Mapping mendelian factors underlying quantitative traits using RFLP linkage maps", Genetics, vol. 121, no. 1, **(1989)**, pp. 185-199.
[12] M.P. Epstein, X. Lin and M. Boehnke, "A tobit variance-component method for linkage analysis of censored trait data", Am J Hum Genet, vol. 72, no. 3, **(2003)**, pp. 611-620.
[13] V.S. Pankratz, M. de Andrade and T.M. Therneau, "Random-effects Cox proportional hazards model, general variance components methods for time-to-event data", Genet Epidemiol, vol. 28, no. 2, **(2005)**, pp. 97-109.
[14] M.J. Sillanpää and E. Arjas, "Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data", Genetics, vol. 148, no. 3, **(1998)**, pp. 1373-1388.
[15] M.J. Sillanpää and E. Arjas, "Bayesian mapping of multiple quantitative trait loci from incomplete outbred offspring data", Genetics, vol. 151, no. 4, **(1999)**, pp. 1605-1619.
[16] H. Wang, Y.M. Zhang and X. Li, "Bayesian shrinkage estimation of quantitative trait loci parameters", Genetics, vol. 170, no. 1, **(2005)**, pp. 465-480.
[17] N. Yi and S. Xu, "Bayesian mapping of quantitative trait loci for complex binary traits", Genetics, vol. 155, no. 3, **(2000)**, pp. 1391-1403.
[18] N. Yi, B.S. Yandell, G.A. Churchill and D.B. Allison, "Bayesian model selection for genome-wide epistatic quantitative trait loci analysis", Genetics, vol. 170, no. 3, **(2005)**, pp. 1333-1344.

[19] N. Yi, S. Banerjee, D. Pomp and B.S. Yandell, "Bayesian mapping of genomewide interacting quantitative trait loci for ordinal traits", Genetics, vol. 176, no. 3, **(2007)**, pp. 1855-1864.
[20] R.E. Kass and A.E. Raftery, "Bayes Factors", J, Am, Stat, Assoc, vol. 90, no. 430, **(1995)**, pp. 773-795.
[21] J.M. Satagopan, B.S. Yandell, M.A. Newton and T.C. Osborn, "A bayesian approach to detect quantitative trait loci using Markov chain Monte Carlo", Genetics, vol. 144, no. 2, **(1996)**, pp. 805-816.
[22] X. Wang, Z.Z. Piao, B. Wang, R.Q. Yang and Z.X. Luo, "Robust Bayesian mapping of quantitative trait loci using Student-t distribution for residual", Theor Appl Genet, vol. 118, **(2009)**, pp. 609-617.
[23] W.H. Rogers and J.W. Tukey, "Understanding some long-tailed distributions", Stat Neerl, vol. 26, **(1972)**, pp. 211-226.
[24] A.P. Dempster, N.M. Laird and D.B. Rubin, "Iteratively reweighted least squares for linear regression when errors are normal/independent distributed", In, Krishnaiah, P, R, (Ed,), Multivariate Analysis, North-Holland, Amsterdam, **(1980)**.
[25] K.L. Lange, R.J.A. Little and J.M.G. Taylor, "Robust statistical modelling using the t-distribution", Journal of the American Statistical Association, vol. 84, **(1989)**,pp. 881–896.
[26] J.Y. Cheng and S. Tzeng, "Parametric and semiparametric methods for mapping quantitative trait loci", Comput Statist Data Anal, vol. 53, no. 5, **(2009)**, pp. 1843-1849.
[27] Z. Jin, D.Y. Lin and L.J. Wei, "Rank-based inference for the accelerated failure time model", Biometrika, vol. 90, **(2003)**, pp. 341-353.
[28] Z.S. Ma and E.J. Bechinski, "Accelerated failure time (AFT) modeling for the development and survival of Russian wheat aphid", Diuraphis noxia (Mordvilko), Popul Ecol, vol. 51, **(2009)**, pp. 543-548.
[29] D. Cox and D. Oakes, "Analysis of Survival Data", Chapman and Hall, London, **(1984)**.
[30] J.D. Kalbfleisch and R.L. Prentice, "The statistical analysis of failure time data", 2nd ed,Wiley, New York, **(2002)**.
[31] M. Lynch and B. Walsh, "Genetic and Analysis of Quantitative Traits", 1st edn, ed,Sinauer Associates, Sunderland, **(1998)**.
[32] B.P. Carlin and S. Chib, "Bayesian model choice via Markov chain Monte Carlo", J Am Stat Assoc, vol. 88, **(1995)**, pp. 881-889.
[33] N. Yi, "A unified Markov chain Monte Carlo framework for mapping multiple quantitative trait loci", Genetics, vol. 167, no. 2, **(2004)**, pp. 967-975.
[34] H. Chipman, E.I. Edwards and R.E. Mcculloch, "The practical implementation of Bayesian model selection", vol. 63, no. 4, **(2001)**, pp. 122-127.
[35] E.I. George and R.E. Mccullochr, "Approaches for Bayesian variable selection", Stat, Sin, vol. 7, no. 2, **(1997)**, pp. 339-373.
[36] R. Kohn, M. Smith and D. Chan, "Nonparametric regression using linear combinations of basis functions", Stat, Comput, vol. 11, **(2001)**, pp. 313-322.
[37] L. Kuo and B. Mallick, "Variable selection for regression models", Sankhyā Ser,B, vol. 60, no. 1), **(1998)**, pp. 65-81.
[38] Y.M. Zhang and S. Xu, "Advanced statistical methods for detecting multiple quantitative trait loci", Recent Res, Devel, Genet, Breeding, vol. 2, **(2005)**, pp. 1–23.
[39] R.D. Ball, "Bayesian methods for quantitative trait loci mapping based on model selection", approximate analysis using the Bayesian information criterion, Genetics, vol. 159, no. 3, **(2001)**, pp. 1351-1364.
[40] A.E. Raftery, D. Madigan and J.A. Hoteing, "Bayesian Model Averaging for Linear Regression Models J", Am, Stat, Assoc, vol. 92, no. 437, **(1997)**, pp. 179-191.
[41] M.J. Sillanpää and J. Corander, "Model choice in gene mapping, what and why", Trends Genet, vol. 18, no. 6, **(2002)**, pp. 301-307.

Authors

Xiaojing Zhou, She comes from Department of Mathematics, Heilongjiang Bayi Agriculture University Daqing, China. Her research interests include mathematical biology, biostatistics and Genome-wide association study.

Dan Zhao, She comes from Department of Mathematics, Anshan Normal University, Anshan, China. Her research interests include mathematical statistics and statistical analysis.

