9. Curtis, J. T. and McIntosh, R. P., The interrelationship of certain analytic and synthetic phytosociological characters. *Ecology*, 1950, **31**, 4343–4445.
10. Gomez, K. A. and Gomez, A. A., *Statistical Procedures for Agriculture Research*, John Wiley, New York, USA, 1984, pp. 304–308.
11. Bisht, A. S. and Bhat, A. B., Vegetation structure and plant diversity relation in a sub-alpine region of Garhwal Himalaya, Uttarakhand, India. *Afr. J. Plant Sci.*, 2013, **7**, 401–406.
12. Man, V., Verma, R. K., Chauhan, N. S. and Kapoor, K. S., Phytosociological attributes of Porang valley in Lippa–Asrang wildlife sanctuary of District Kinnaur, Himachal Pradesh. *Ann. For.*, 2012, **20**(1), 1–16.
13. Shameem, S. A. and Kangroo, I. N., Comparative assessment of edaphic features and phytodiversity in Lower Dachigam National Park, Kashmir Himalaya, India. *Afr. J. Environ. Sci. Technol.*, 2011, **5**(11), 972–984.
14. Molla, A. and Kewessa, G., Woody species diversity in traditional agroforestry practices of Dellomenna District, Southeastern Ethiopia: implication for maintaining native woody species. *Int. J. Biodiver.*, 2015, 1–13.
15. Amjad, M. S., Ethnobotanical profiling and floristic diversity of Bana Valley, Kotli (Azad Jammu and Kashmir), Pakistan. *Asian Pac. J. Trop. Biomed.*, 2015, **5**(4), 292–299.
16. Baig, B. A., Ramamoorthy, D. and Bhat, T. A., Threatened medicinal plants of Menwarsar Pahalgam, Kashmir Himalayas: distribution pattern and current conservation status. *Proc. Int. Acad. Ecol. Environ. Sci.*, 2013, **3**(1), 25–35.
17. Lebret, M., Nys, C. and Forgeard, N., Litter production in an Atlantic beech (*Fagus sylvatica* L.) time sequence. *Ann. For. Sci.*, 2001, **58**, 755–768.
18. Sharma, P., Rana, J. C., Devi, U., Randhawa, S. S. and Kumar, R., Floristic diversity and distribution pattern of plant communities along altitudinal gradient Northwest Himalaya. *Sci. World J.*, 2014, 1–11.
19. Pappoe, A. N. M., Armah, F. A., Quaye, E. C., Kwakye, P. K. and Buxton, G. N. T., Composition and stand structure of a tropical moist semi-deciduous forest in Ghana. *Int. Res. J. Plant Sci.*, 2010, **1**(4), 95–106.
20. Behera, M. D., Kushwaha, S. P. S. and Roy, P. S., Geo-spatial modeling for rapid biodiversity assessment in Eastern Himalayan region. *For. Ecol. Manage.*, 2005, **207**, 363–384.
21. Bijalwan, A., Swamy, S. L., Sharma, C. M., Umrao, R. and Paliwal, H. B., Structure, composition and diversity of tree vegetation in Sal mixed dry tropical forest in Chhattisgarh plains of India. *Indian For.*, 2011, 453–462.
22. Nabi, S., Qaisar, K. N., Rather, S. A., Khan, P. A. and Nabi, B., Plant diversity in horti-agricultural system of district Budgam, Kashmir valley: seasonal variation in structure, composition and vegetation indices. *Ecoscan*, 2016, **10**(3&4), 543–551.

ACKNOWLEDGEMENT. The first author thanks the Department of Science and Technology, Government of India for financial support (INSPIRE Fellowship) during this study.

Received 26 April 2021; revised accepted 31 March 2022

doi: 10.18520/cs/v122/i10/1209-1214

Identification of *cis*- and *trans*-expression quantitative trait loci using Bayesian framework

Himadri Shekhar Roy*, Amrit Kumar Paul and Ranjit Kumar Paul

ICAR-Indian Agricultural Statistics Research Institute,
New Delhi 110 012, India

The detection and identification of expression quantitative trait loci (eQTLs) for biological characteristics like gene expression is an important focus of genomics. The existence of *cis*- and *trans*-eQTLs is crucial for establishing their cumulative significance to the desired traits. A crucial aspect of genomics is identifying the *cis*- and *trans*-eQTLs that capture substantial changes in the expression of distant genes. The goal of the present study was to use an integrated hierarchical Bayesian model to identify the *cis*- and *trans*-eQTLS. Molecular approaches are utilized to categorize just the candidate genes when quantitative trait loci or eQTLs are identified. Variations inside or near the gene are hypothesized to determine the genetic variances that reflect transcript levels. The identification of eQTLs has helped us better understand gene regulation and complex trait analysis. The present study focused on barley crops, and only *cis*-eQTLs were identified; no additional eQTL hotspots were determined. Mouse gene expressions were used to study *trans*-eQTLs and substantial *cis*- and *trans*-eQTLs, as well as four eQTL hotspots were identified.

Keywords: Barley, gene expression, hotspots, integrated hierarchical model, quantitative trait loci.

A quantitative characteristic is described in biological research as a quantifiable phenotype, such as yield, disease resistance, etc. that differs numerically and is dependent on the collective activities of numerous genes and the environment. The initial step in quantitative characteristic loci (QTL) mapping is to identify potential genes for detecting chromosomal regions linked with a certain quantitative trait[1]. Recently, expression quantitative trait loci (eQTL) mapping has become a popular and successful approach for detecting regulatory areas for genes using transcriptome and genotyping data[2]. These are the genomic regions that control mRNA or protein expression. The most significant molecular phenotypes are gene expressions, which operate as quantitative qualities that link genetic diversity to phenotypic variance. These differences in gene expression are expected to be important in phenotypic differences and species evolution. As a result, identifying eQTL has become a significant element of biological research[3]. Traditional QTL mapping often finds broad areas

*For correspondence. (e-mail: himadriiasri@gmail.com)

of application. Following the discovery of QTLs and eQTLs, molecular methods are used to categorize just the candidate genes. The genetic differences that reflect transcript levels are considered to be determined by variations inside or close to the gene. The discovery of eQTLs has expanded our understanding of gene regulation and complicated trait analysis[4,5]. It is now possible to map eQTLs for virtually all genes in the genome using DNA microarrays. Expression QTLs are objectively classified into two types: *cis* and *trans*. In the case of *cis*-eQTL, the observed position of the eQTL is the same as the gene site (Figure 1 *a*), but in the case of *trans*-eQTL, the observed location of the eQTL is not the same as the location of the gene (Figure 1 *b*). For example, a *trans*-eQTL might reflect the position of a transcription factor that regulates the expression of the target alone or, perhaps, the linked expression of multiple functionally related genes.

The discovery of eQTLs can help us better understand genetics and gene expression control systems in many species[6]. The identification of *trans*-eQTL is substantially more challenging than *cis*-eQTL identification due to the enormous number of single-nucleotide polymorphism (SNP)–gene combinations that must be examined for *trans*-association mapping. The *trans*-eQTL analysis requires a greater sample size and/or effect than the *cis*-eQTL analysis to attain the same power. However, *trans*-eQTLs exhibit lesser effects than *cis*-eQTLs[7]. Several approaches for determining the real number of *trans*-eQTLs have recently been developed.

Several researchers have found a substantial clustering of *trans*-eQTLs (many genes linked with the same locus) into so-called eQTL hotspots, indicating that these genomic areas contain variants that alter the dynamic and global character of transcriptional regulation[8]. Certain genomic areas known as regulatory hotspots can alter the expression levels of multiple genes, according to eQTL mapping studies.

Since the main difference between eQTL and regular QTL studies is the number of phenotypes, it is not unexpected that traditional QTL approaches have been employed to find eQTLs one gene at a time. This 'one gene at a time' strategy ignores the numerous essential combinatorial effects and gene interactions. Using an empirical Bayes method, Kendziorski *et al.*[9] suggested a mixture over marker (MOM) modelling technique to promote information-sharing across both markers and transcripts.

Wen[10] developed a simple method for simultaneous testing of multiple testing hypotheses that is robust in terms of false discovery rate (FDR) control. In applications of single and multi-tissue eQTL mapping, the suggested technique is computationally efficient. Recently, Bayesian models have been widely used to tackle the excessive multiplicity problem in eQTL studies. Several approaches based on sparse Bayesian regression (SBR) modelling have been developed for QTL research, such as the R-QTLBIM (QTL-Bayesian interval mapping) package[11]. Later, Banerjee *et al.*[12] improved this approach to handle many characteristics (genes) at once. Due to computational challenges, they have restricted the number of traits to five. In the present study, an integrated hierarchical Bayesian model is used to analyse a large number of SNP–gene pairs[13]. All of the Bayesian frameworks discussed earlier involve a common prior probability for including a marker in the sparse regression model, which increases the risk of a high number of false-positive hotspots across all genes.

The following model (eq. (1)) may be used to represent gene expression measurements across individuals

$$y_{ig} = \mu_g + \sum_{j=1}^{S} x_{ij}\gamma_{jg}\beta_{jg} + \varepsilon_{ig}, \quad (1)$$

where $g = 1, ..., G$ indicates a particular gene or feature, $i = 1, ..., n$ denotes a specific individual, $j = 1, ..., S$ denotes a specific SNP, y_{ig} the gene expression level for gene g for the ith individual, μ_g the overall mean of gene expression of gene g (across all individuals), x_{ij} the genotype at locus j for the ith individual under an additive, dominant or recessive genetic model and β_{jg} is the effect size of SNP j on gene g.

Since only a few markers directly affect gene expression, many of the βs are nearly zero, and binary indicator variables are added to the model as γ_{jg} to specify which markers should be included, i.e. $\gamma_{jg} = 1$, if SNP j should be included in the model for gene g and $\gamma_{jg} = 0$ otherwise. ε_{ig} is a Gaussian error term with gene-specific variance.

Gene/marker-specific QTL probability as $w_{jg} = P(\gamma_{jg} = 1)$ is considered a priori, with strength borrowed across genes to estimate this probability using flexible genome-wide prior distribution representation. The hierarchical structure of the model aids in the detection of eQTLs linked to several genes. It is also assumed that the random error term, i.e. ε_{ig} is independent and identically distributed, so that genes are conditionally independent given all model parameters. However, dependence between genes is introduced using an exchangeable prior on the γ_{jg}s, which is computationally easy to understand.

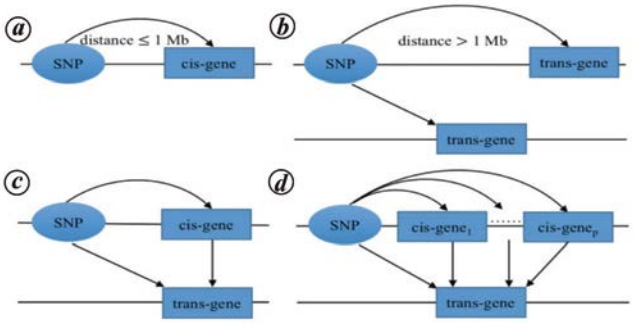

Figure 1. Graphical depiction of expression quantitative trait loci (eQTLs). ***a***, *cis*-eQTLs; ***b***, *trans*-eQTLs; ***c***, *trans*-eQTLs mediated by a single *cis*-mediator; ***d***, *trans*-eQTLs regulated by numerous *cis*-mediators[25].

Thousands of gene expression levels are employed as quantitative phenotypes in eQTL studies. Univariate QTL analysis for each gene expression profile is often required to analyse such type of data. The model outlined above was applied to the most essential elements – (1) majority of eQTLs have multiple effects on gene expression and (2) genes in the same pathway are more likely to be linked. As a result, there is a possibility to pool information from hundreds or thousands of gene expression features in order to derive more insightful conclusions.

The different prior distributions of the model are described as follows. The parameter set-up is given by $\theta = (\mu_g, \sigma_g^2, \gamma_{jg}, \beta_{jg}, w_{jg}, p_j, a_j, b_j)$. The set-up of the priors is defined as follows. The prior probability for the inclusion of SNP j in the model for gene g is given by $\gamma_{jg} \sim$ Bernoulli (w_{jg}), where $P(\gamma_{jg} = 1) = w_{jg}$ is an unknown parameter.

As only a small number of SNPs act as a determinant of gene expression, the inclusion probability parameter w_{jg} takes the value 0 a priori most of the time. When w_{jg} is not 0, it is assumed to come from a beta distribution, beta(a_j, b_j). This can be expressed as a Dirac mass at 0 and a beta distribution with weights p_j and $1 - p_j$ as follows

$$w_{jg} \sim p_j \delta_0 (w_{jg}) + (1 - p_j) \text{ beta}(a_j, b_j) (w_{jg}). \quad (2)$$

The parameter p_j (the probability that w_{jg} is 0) is equal for all genes, which helps in detecting an SNP that is weakly associated with many gene expressions[14]. Now, for p_j with hyperparameters a_0 and b_0, a common conjugate beta prior is used: $p_j \sim \text{beta}(a_0, b_0)$. a_j and b_j are assumed to follow exponential distributions with hyperparameters λ_a and λ_b: $a_j \sim \exp(\lambda_a) b_j \sim \exp(\lambda_b)$. Now, $\mu_g \sim N(m_g, \tau_g^2)$, where m_g and τ_g are the empirical mean and variance of gene expression g respectively.

$\beta_{jg} = 0$ if $\gamma_{jg} = 0$ and $\beta_{jg} \sim N(0, v_g^2)$, if $\gamma_{jg} = 1$, with $v_{jg}^2 = c(x_j^T x_j)^{-1} \sigma_g^2$, where c is the scaling factor parameter and

$$(x_j^T x_j)^{-1} = \left(\sum_{i=1}^{n} x_{ij}^2 \right)^{-1}$$

mimics the regressor variance, which leads to the well-known g-prior of Zellner[15].

Here the approach of Yi and Shriner[16] is followed, and c is considered to be a constant equal to S, the number of SNPs. Bottolo and Richardson[17] considered an inverse-gamma prior $c \sim$ inverse Gamma $(\frac{1}{2}, \frac{n}{2})$ based on the Zellner and Siow prior. Petretto *et al.*[18] considered a common c for all genes with the prior of Liang *et al.*[19]. $c \sim \frac{1}{1+c}$ in the interval $(0, M)$, where the end-point $M = \max(n, S^2)$. The term σ_g^2, the overall variance of v_{jg}^2, ensures that σ_g^2 is a nuisance parameter in the model and can be integrated out; thus, $\sigma_g^2 \sim$ inverse gamma $(\frac{1}{2}, \frac{1}{2})$.

The posterior distribution was generated using Markov Chain Monte Carlo (MCMC) methods[20]. Gibbs sampling was used for all updates, and there were no closed forms for full conditionals for a_j and b_j. Adaptive rejection sampling was used to update these two parameters[21]. The number of iterations was determined using the approach of Raftery and Lewis[22]. To estimate standard posterior values for the two datasets utilized here, 100,000 iterations with 50,000 burn-in iterations were employed. Finally, the model was based on four hyperparameters, viz. a_0, b_0, λ_a, λ_b, that were predetermined. The expected number of e-QTLs $(E(\eta_\gamma))$ and their dispersion $(V(\eta_\gamma))$ were utilized to calculate these values a priori. In this method, $a_0 = \lambda_a = 10$ and $b_0 = \gamma_b = 0.1$ is chosen, which favours models with fewer eQTLs. The final goal is to find gene/SNP relationships, which may be done with the help of parameter estimations from the proposed model. If the posterior probability of association for gene g at SNP j is more than 0.80, an eQTL is identified.

In this study, two real datasets were analysed to identify the type of eQTLs and also the hotspots.

Dataset 1: An experiment was conducted to characterize quantitative resistance to the barley leaf rust pathogen

Table 1. Detected expression quantitative trait loci (eQTLs) with posterior probability of association (PPA) > 0.80

Gene	SNP	PPA	Type
Contig12563_s_at	Contig20996_10	0.81	cis
Contig10533_at	Contig12729_5	1.00	cis
Contig1031_at	Contig5754_9	0.80	cis

SNP, Singe-nucleotide polymorphism.

Table 2. PPA with different genes and SNP beyond the cut-off value of PPA > 0.8 and with false discovery rate (FDR) level 10%

Gene	SNP	PPA
1422462_at	rs6263067	1
1438426_at	rs6263067	0.97
1438852_x_at	rs6263067	0.93
1450813_a_at	rs6263067	0.88
1416647_at	rs4222763	0.87
1428844_a_at	rs4222763	0.88
1436955_at	rs4222763	0.9
1441568_at	rs4222763	0.88
1448604_at	rs4222763	0.97
1448958_at	rs4222763	0.98
1427711_a_at	rs13476267	1
1444320_at	mCV23574676	1
1439075_at	rs6331493	1
1424811_at	rs3142215	0.92
1415673_at	rs13479058	0.92
1415677_at	rs13479058	0.89
1415703_at	rs13479058	0.9
1415713_a_at	rs13479058	0.88
1415715_at	rs13479058	0.9
1450003_at	rs3693435	0.93

Table 3. Some detected *cis*- and *trans*-eQTLs presented with PPA

Gene	SNP	PPA	Gene chrm	Gene start	Gene end	Marker chrm	Marker position	Type
1448958_at	rs4222763	0.98	1	1.58E+08	1.58E+08	1	168084791	trans
1427711_a_at	rs13476267	1	7	26246721	26262644	1	183253943	trans
1444320_at	mCV23574676	1	8	26835796	26865068	2	26699898	trans
1439075_at	rs6331493	1	2	1.44E+08	1.44E+08	2	144676277	cis
1424811_at	rs3142215	0.92	6	85767215	85770966	6	89630672	trans
1415673_at	rs13479058	0.92	5	1.3E+08	1.3E+08	6	137525036	trans
1415677_at	rs13479058	0.89	14	56357857	56364527	6	137525036	trans
1415703_at	rs13479058	0.9	21	1.48E+08	1.48E+08	6	137525036	trans
1415713_a_at	rs13479058	0.88	12	1.05E+08	1.05E+08	6	137525036	trans
1415715_at	rs13479058	0.9	5	33995865	34000626	6	137525036	trans
1415734_at	rs13479058	0.86	6	87949100	87995264	6	137525036	trans
1415746_at	rs13479058	0.9	7	26052723	26079178	6	137525036	trans
1415749_a_at	rs13479058	0.96	4	1.24E+08	1.24E+08	6	137525036	trans
1415754_at	rs13479058	0.89	15	78971797	78982197	6	137525036	trans
1415758_at	rs13479058	0.87	5	73411430	73647857	6	137525036	trans
1415771_at	rs13479058	0.9	1	88241294	88256030	6	137525036	trans
1415772_at	rs13479058	0.88	1	88241294	88256030	6	137525036	trans
1415790_at	rs13479058	0.9	11	44268073	44284000	6	137525036	trans
1415816_at	rs13479058	0.86	6	85402067	85418466	6	137525036	trans
1415830_at	rs13479058	0.87	5	21992303	22056247	6	137525036	trans

chrm, Chromosome.

Table 4. A sample of eQTLs for the hotspot of marker chromosome 6

Gene	SNP	PPA	Gene chrm	Gene start	Gene end	Marker chrm	Marker position	Type
1415673_at	rs13479058	0.92	5	1.3E+08	1.3E+08	6	1.38E+08	trans
1415677_at	rs13479058	0.89	14	56357857	56364527	6	1.38E+08	trans
1415703_at	rs13479058	0.9	21	1.48E+08	1.48E+08	6	1.38E+08	trans
1415713_a_at	rs13479058	0.88	12	1.05E+08	1.05E+08	6	1.38E+08	trans
1415715_at	rs13479058	0.9	5	33995865	34000626	6	1.38E+08	trans
1415734_at	rs13479058	0.86	6	87949100	87995264	6	1.38E+08	trans
1415746_at	rs13479058	0.9	7	26052723	26079178	6	1.38E+08	trans
1415749_a_at	rs13479058	0.96	4	1.24E+08	1.24E+08	6	1.38E+08	trans
1415754_at	rs13479058	0.89	15	78971797	78982197	6	1.38E+08	trans
1415758_at	rs13479058	0.87	5	73411430	73647857	6	1.38E+08	trans
1415771_at	rs13479058	0.9	1	88241294	88256030	6	1.38E+08	trans
1415772_at	rs13479058	0.88	1	88241294	88256030	6	1.38E+08	trans
1415790_at	rs13479058	0.9	11	44268073	44284000	6	1.38E+08	trans
1415816_at	rs13479058	0.86	6	85402067	85418466	6	1.38E+08	trans

Table 5. A sample **o**f eQTLs for the hotspot of marker chromosome 16

Gene	SNP	PPA	Gene chrm	Gene start	Gene end	Marker chrm	Marker position	Type
1415771_at	rs3693435	0.9	1	88241294	88256030	16	63508672	trans
1415830_at	rs3693435	0.87	5	21992303	22056247	16	63508672	trans
1415927_at	rs3693435	0.88	2	1.14E+08	1.14E+08	16	63508672	trans
1416066_at	rs3693435	0.88	6	1.25E+08	1.25E+08	16	63508672	trans
1416237_at	rs3693435	0.88	9	44850508	44862098	16	63508672	trans
1416270_at	rs3693435	0.87	19	8867619	8873047	16	63508672	trans
1416575_at	rs3693435	0.86	16	18780540	18812080	16	63508672	trans
1416687_at	rs3693435	0.86	9	92437061	92503266	16	63508672	trans
1416979_at	rs3693435	0.88	5	1.49E+08	1.49E+08	16	63508672	trans
1417160_s_at	rs3693435	0.93	11	83522517	83524850	16	63508672	trans
1417294_at	rs3693435	0.87	4	1.39E+08	1.39E+08	16	63508672	trans

Puccinia hordei in the St/Mx population. Agilent barley custom microarray was used to assess transcript abundance in 139 DH lines of the St/Mx population challenged with *P. hordei*. The datasets were generated previously by Rostoks *et al.*[23] and Chen *et al.*[24] respectively, having 4286 SNPs and 595,754 expressions.

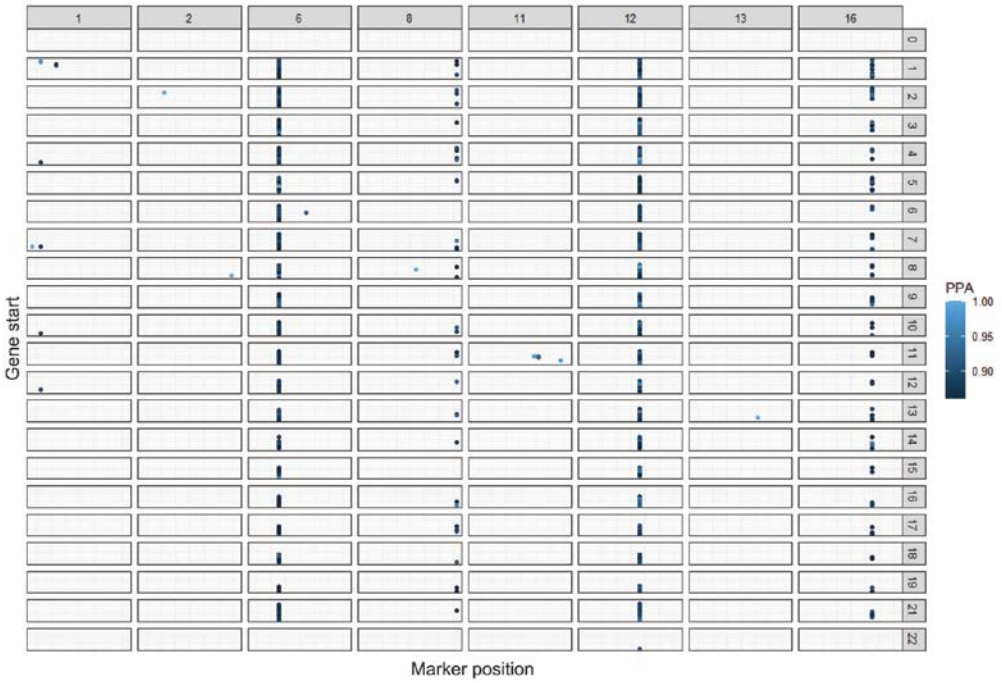

Figure 2. The *x*-axis represents the position of each eQTL along the genome, while the *y*-axis represents the position of the probe set target itself. The chromosomal borders are shown by grey lines. A diagonal line is formed by *cis*-eQTLs. Transcript groups related to a single *trans*-eQTL are shown by vertical bands[26].

Table 6. A sample of eQTLs for the hotspot of marker chromosome 12

Gene	SNP	PPA	Gene chrm	Gene start	Gene end	Marker chrm	Marker position	Type
1415683_at	rs6170344	0.91	11	1.03E+08	1.03E+08	12	83729204	trans
1415684_at	rs6170344	0.94	10	43988164	44084097	12	83729204	trans
1415690_at	rs6170344	0.94	11	94515081	94521403	12	83729204	trans
1415693_at	rs6170344	0.95	15	57701057	57723973	12	83729204	trans
1415695_at	rs6170344	0.95	7	1.21E+08	1.21E+08	12	83729204	trans
1415697_at	rs6170344	0.92	5	92481172	92512761	12	83729204	trans
1415703_at	rs6170344	0.92	21	1.48E+08	1.48E+08	12	83729204	trans
1415718_at	rs6170344	0.86	11	57615139	57623719	12	83729204	trans
1415735_at	rs6170344	0.86	19	10680115	10704312	12	83729204	trans
1415742_at	rs6170344	0.92	6	83004647	83007674	12	83729204	trans

Dataset 2: We used the FDR approach to analyse publicly available eQTL data. The data are available at the QTL Archive, now part of the Mouse Phenome Database, at http://phenome.jax.org/db/q?rtn=projects/projdet&reqprojid=532. Custom Agilent two-colour ink-jet microarrays were used to determine gene expression. The Affymetrix 5K Gene Chip was used to genotype mice. Finally, there were 28 samples, 918 SNPs and 23,238 expression values in the dataset. The number of identified eQTLS in the dataset was 4243. The mechanism of detection in this case was based on the posterior likelihood of connection (PPA). Significant SNP–gene connections were evaluated if PPA was greater than 0.80, which is the cut-off value of this method. Table 2 shows all the PPA values as well as the gene–SNP relationships. Table 3 lists the detected *cis*- and *trans*-eQTLs.

Based on PPA between eQTL and the gene, four hotspots were identified (Figure 2). Figure 2 shows that a large number of *trans*-eQTLs have a similar position with marker chromosomes 6, 8, 12 and 16. Table 4 shows that SNP rs13479058 is linked to a significant number of genes and is found in common places on chromosome 6 (marker chromosome 6). As a result, these genes are known as hotspots. Similarly, using marker chromosome 16, SNP rs3693435 was found to be associated with a significant number of genes and had similar positions (Table 5). Table 6 shows that SNP rs6170344 is linked to a significant number of genes and has a common position on chromosome 12 (marker chromosome 12). SNP rs6343961 was linked to a significant number of genes and shared sites on chromosome 8 using marker chromosome 8 (Table 7).

One of the important aims of genomics is to identify the location of eQTLs for molecular traits like gene expression. Finding the *cis*- and *trans*-eQTLs is critical for determining the cumulative importance of eQTLs to the targeted characteristics. Identifying the *cis*- and *trans*-eQTLs that

Table 7. A sample of eQTLs for the hotspot of marker chromosome 8

Gene	SNP	PPA	Gene chrm	Gene start	Gene end	Marker chrm	Marker position	Type
1418669_at	rs6343961	0.86	4	1.37E+08	1.37E+08	8	18316422	trans
1419009_at	rs6343961	0.86	4	56756285	56757797	8	18316422	trans
1420351_at	rs6343961	0.9	4	1.55E+08	1.55E+08	8	18316422	trans
1420716_at	rs6343961	0.86	1	1.68E+08	1.68E+08	8	18316422	trans
1422628_at	rs6343961	0.9	19	12620230	12664258	8	18316422	trans
1423419_at	rs6343961	0.87	11	82594610	82616364	8	18316422	trans
1423774_a_at	rs6343961	0.95	7	87439403	87461145	8	18316422	trans
1424187_at	rs6343961	0.86	16	45093515	45127778	8	18316422	trans
1424923_at	rs6343961	0.92	12	1.05E+08	1.05E+08	8	18316422	trans
1426190_at	rs6343961	0.9	5	1.14E+08	1.14E+08	8	18316422	trans
1426638_at	rs6343961	0.87	17	86012948	86031153	8	18316422	trans
1427324_at	rs6343961	0.89	5	1.24E+08	1.24E+08	8	18316422	trans
1427786_at	rs6343961	0.94	11	96874708	96886048	8	18316422	trans

capture large changes in the expression of distant genes is a key concept in genomics. In the present study, major focus was given to identifying these *cis*- and *trans*-eQTLs using integrated hierarchical Bayesian model for eQTLs. This study was done mainly for barley crop and only *cis*-eQTLs were detected; no other eQTL hotspots were found. To explore the trans-acting eQTLs, mouse gene expressions were considered. Both *cis*- and *trans*-eQTLs were identified as well as four eQTL hotspots.

1. Druka, A., Potokina, E., Luo, Z., Jiang, N., Chen, X., Kearsey, M. and Waugh, R., Expression quantitative trait loci analysis in plants. *Plant Biotechnol. J.*, 2010, **8**(1), 10–27.
2. Gelfond, J. A., Ibrahim, J. G. and Zou, F., Proximity model for expression quantitative trait loci (eQTL) detection. *Biometrics*, 2007, **63**(4), 1108–1116.
3. Potokina, E., Druka, A., Luo, Z., Wise, R., Waugh, R. and Kearsey, M., Gene expression quantitative trait locus analysis of 16,000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. *Plant J.*, 2008, **53**(1), 90–101.
4. Brem, R. B., Yvert, G., Clinton, R. and Kruglyak, L., Genetic dissection of transcriptional regulation in budding yeast. *Science*, 2002, **296**(5568), 752–755.
5. Göring, H. H. *et al.*, Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nature Genet.*, 2007, **39**(10), 1208–1216.
6. Flutre, T., Wen, X., Pritchard, J. and Stephens, M., A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genet.*, 2013, **9**(5), e1003486.
7. Pierce, B. L. *et al.*, Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: a genome-wide analysis among 1,800 South Asians. *PLoS Genet.*, 2014, **10**(12), e1004818.
8. Zhu, Z. *et al.*, Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genet.*, 2016, **48**(5), 481–487.
9. Kendziorski, C. M., Chen, M., Yuan, M., Lan, H. and Attie, A. D., Statistical methods for expression quantitative trait loci (eQTL) mapping. *Biometrics*, 2006, **62**(1), 19–27.
10. Wen, X., Molecular QTL discovery incorporating genomic annotations using Bayesian false discovery rate control. *Ann. Appl. Stat.*, 2016, **10**(3), 1619–1638.
11. Yandell, B. S. *et al.*, R/qtlbim: QTL with Bayesian interval mapping in experimental crosses. *Bioinformatics*, 2007, **23**(5), 641–643.
12. Banerjee, S., Yandell, B. S. and Yi, N., Bayesian quantitative trait loci mapping for multiple traits. *Genetics*, 2008, **179**(4), 2275–2289.
13. Scott-Boyer, M. P., Imholte, G. C., Tayeb, A., Labbe, A., Deschepper, C. F. and Gottardo, R., An integrated hierarchical Bayesian model for multivariate eQTL mapping. *Stat. Appl. Genet. Mol. Biol.*, 2012, **11**(4).
14. Lucas, J., Carvalho, C., Wang, Q., Bild, A. N., Nevins, J. R., Michael, A. J. and West, M., Sparse statistical modelling in gene expression genomics. *Bayes. Infer. Gene Expr. Proteom.*, 2006, **1**(1), 155–176.
15. Zellner, A., On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, 1986.
16. Yi, N. and Shriner, D., Advances in Bayesian multiple quantitative trait loci mapping in experimental crosses. *Heredity*, 2008, **100**(3), 240–252.
17. Bottolo, L. and Richardson, S., Evolutionary stochastic search for Bayesian model exploration. *Bayesian Anal.*, 2010, **5**(3), 583–618.
18. Petretto, E. *et al.*, New insights into the genetic control of gene expression using a Bayesian multi-tissue approach. *PLoS Comput. Biol.*, 2010, **6**(4), e1000737.
19. Liang, F., Paulo, R., Molina, G., Clyde, M. A. and Berger, J. O., Mixtures of g-priors for Bayesian variable selection. *J. Am. Stat. Assoc.*, 2008, **103**(481), 410–423.
20. Gelfand, A. E. and Smith, A. F., Samping-based approaches to calculating marginal densities. *J. Am. Stat. Assoc.*, 1990, **85**, 398–409.
21. Gilks, W. R. and Wild, P., Adaptive rejection sampling for Gibbs sampling. *J. Royal Stat. Soc.: Series C (App. Stat.)*, 1992, **41**(2), 337–348.
22. Raftery, A. E. and Lewis, S. M., Implementing mcmc. In Markov chain Monte Carlo in Practice, 1996, pp. 115–130.
23. Rostoks, N. *et al.*, Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol. Genet. Genomics*, 2005, **274**(5), 515–527.
24. Chen, X. *et al.*, An eQTL analysis of partial resistance to *Puccinia hordei* in barley. *PLoS ONE*, 2010, **5**(1), e8598.
25. Shan, N., Wang, Z. and Hou, L., Identification of trans-eQTLs using mediation analysis with multiple mediators. *BMC Bioinformatics*, 2019, **20**(3), 87–97.
26. Imholte, G. C., Scott-Boyer, M. P., Labbe, A., Deschepper, C. F. and Gottardo, R., iBMQ: a R/Bioconductor package for integrated Bayesian modeling of eQTL data. *Bioinformatics*, 2013, **29**(21), 2797–2798.

Received 10 February 2022; accepted 23 March 2022

doi: 10.18520/cs/v122/i10/1214-1219