Mapping Environment Specific Quantitative Trait Loci

Xin Chen\textsuperscript{1}, Fuping Zhao\textsuperscript{2} and Shizhong Xu\textsuperscript{2}

\textsuperscript{1}Department of Statistics, University of California, Riverside, CA 92521
\textsuperscript{2}Department of Botany and Plant Sciences, University of California, Riverside, CA 92521

CORRESPONDING AUTHOR

Shizhong Xu, Ph.D
Department of Botany and Plant Sciences
University of California
Riverside, CA 92521
E-mail: shizhong.xu@ucr.edu
Phone: (951)-827-5898

KEYWORDS

Empirical Bayes, Factor Analytical Model, Mixed Model, $Q \times E$ Interaction, Shrinkage Estimation
ABSTRACT

Environment specific quantitative trait loci (QTL) refer to QTL that express differently in different environments, a phenomenon called QTL by environment (Q×E) interaction. Q×E interaction is a difficult problem extended from traditional QTL mapping. The mixture model maximum likelihood method is commonly adopted for interval mapping of QTL, but the method is not optimal in handling QTL interacting with environments. We partitioned QTL effects into main and interaction effects. The main effects are represented by the means of QTL effects in all environments and the interaction effects are represented by the variances of the QTL effects across environments. We used the MCMC implemented Bayesian method to estimate both the main and interaction effects. The residual error covariance matrix was modeled using the factor analytic covariance structure. A simulation study showed that the factor analytic structure is robust and can handle other structures as special cases. The method was also applied to Q×E interaction mapping for the yield trait of barley. Eight markers showed significant main effects and 18 markers showed significant Q×E interaction. The 18 interacting markers were distributed across all seven chromosomes of the entire genome. Only one marker had both the main and the Q×E interaction effects. Each of the other markers had either a main effect or a Q×E interaction effect but not both.
INTRODUCTION

Genotype-by-environment (G×E) interaction is a very important phenomenon in quantitative genetics. With the advanced molecular technology and statistical methods for QTL mapping (JANSEN 1993; LANDER and BOTSTEIN 1989; ZENG 1994), G×E interaction analysis has shifted to QTL by environment (Q×E) interaction. In the early stage of QTL mapping, almost all statistical methods were developed in a single environment (PATERSON et al. 1991; STUBER et al. 1992). Data from different environments were analyzed separately and the conclusions were drawn from the separate analyses of QTL across environments. These methods do not consider the correlation of data under different environments and thus may not extract maximum information from the data. Composite interval mapping for multiple traits can be used for Q×E interaction if different traits are treated as the same trait measured in different environments (JIANG and ZENG 1995). This multivariate composite interval mapping approach makes good use of all data simultaneously and increases statistical power of QTL detection and accuracy of the estimated QTL positions. However, the number of parameters of this method increases dramatically as the number of environments increases. Therefore, the method may not be applied when the number of environments is large. Several other models have been proposed to solve the problem of large number of environments (BEAVIS and KEIM 1996; JANSEN et al. 1995; ROMAGOSA et al. 1996). These methods were based on some special situations and assumptions. One typical assumption was independent errors or constant variances across environments. These assumptions are often violated in real QTL mapping.
experiments.

Earlier investigators have realized the problem and adopted the mixed model methodology to solve the problem (Boer et al. 2007; Piepho 2000). Under the mixed model framework, people can choose which model effects are random and which are fixed. The mixed model methodology is very flexible, leading to an easy way to model genetic and environmental correlation between environments using a suitable error structures. Piepho (2000) proposed a mixed model to detect QTL main effect across environments. Similar to the composite interval mapping analysis, his model incorporated one putative QTL and a few cofactors. The Q×E effects in the model were assumed to be random, which has greatly reduced the number of estimated parameters. However, the fact that only one QTL is included in the model means that Piepho's (2000) model remains a single QTL model rather than a multivariate model. Boer (2007) proposed a step-by-step mixed model approach to detecting QTL main effects, Q×E interaction effects, and QTL responses to specific environmental covariates. In the final step, Boer (2007) rewrote the model to include all QTL in a multiple-QTL model and re-estimated their effects.

In this study, we extended the Bayesian shrinkage method (Xu 2003) to map Q×E interaction effects of QTL. In the original study (Xu 2003), we treated each marker as a putative QTL and used the shrinkage method to simultaneously estimate marker effects of the entire genome. In the multiple environment case, we can still use this approach to simultaneously evaluating marker effects under multiple environments but we can further partition the marker effects into main and Q×E
interaction effects. For any particular marker, the mean of the marker effects represents the main effect and the variance of the marker effects represents the Q×E interaction effect for that marker. Under the Bayesian framework, we assigned a normal prior with zero mean and an unknown variance to each marker main effect and a multivariate normal prior with zero vector mean and homogeneous diagonal variance-covariance matrix to the Q×E interaction effects of each maker. In multiple environments, the structure of the error terms might be very complicated since we need to consider the correlation of the same genotype under different environments. In our analysis, we used different variance-covariance structures to model the error terms. The simplest case was homogeneous diagonal matrix, and the most complex choice was an unstructured matrix. We also used a heterogeneous diagonal matrix whose parameters are somewhere between the two models. Finally, we considered several factor analytic models. The reason to use the factor analytic structure is that it can separate genetic effects into common effect and environment specific effects. In additional, the factor analytic structure is parsimonious and thus can substantially reduce the computational burden of the mixed model analyses.

THEORY AND METHOD

Hierarchical model

Let \( y_j = [y_{j1}, y_{j2}, \ldots, y_{jm}]^T \) be an \( m \times 1 \) column vector for the observed phenotypic values of individual \( j \) measured from \( m \) environments for \( j = 1, \ldots, n \), where \( n \) is the sample size. Let \( q \) be the number of QTL included in the model. The
The multivariate linear model is

\[
y_j = \beta + \sum_{k=1}^{q} Z_{jk} \gamma_k + \xi_j
\]  

(1)

In the above model, \( \beta = [\beta_1, \beta_2, \ldots, \beta_m]^T \) is an \( m \times 1 \) vector for the intercepts.

The dependent variable \( Z_{jk} \) is a genotype indicator variable for individual \( j \) at marker \( k \) and it is defined as \( Z_{jk} = \{-1, 1\} \) for the two genotypes of a backcross (BC) individual or \( Z_{jk} = \{-1, 0, 1\} \) for the three genotypes of an \( F_2 \) individual. The regression coefficient \( \gamma_k = [\gamma_{k1}, \gamma_{k2}, \ldots, \gamma_{km}]^T \) is an \( m \times 1 \) vector of QTL effects for the \( m \) environments. Finally, \( \xi_j = [\xi_{j1}, \xi_{j2}, \ldots, \xi_{jm}]^T \) is an \( m \times 1 \) vector for the residual errors. To model the Q×E interaction, we assume that \( \gamma_k \) follow a multivariate normal distribution,

\[
p(\gamma_k) = N(\gamma_k | \alpha_k, \sigma_k^2)
\]  

(2)

where \( \alpha_k \) is a unity vector with dimension \( m \), \( \sigma_k^2 \) is the variance of \( \gamma_k \) representing the Q×E interaction. This type of model with further modeling on \( \gamma_k \) is called a hierarchical model. In the hierarchical model, the first moment parameter \( \alpha_k \) is the main effect and the second moment parameter \( \sigma_k^2 \) represents the degree of Q×E interaction. The residual error vector \( \xi_j \) is assumed to be multivariate normal with density

\[
p(\xi_j) = N(\xi_j | 0, \Theta)
\]  

(3)

where \( \Theta \) is an \( m \times m \) variance-covariance matrix, which can be chosen from a class of available forms (to be discussed later). We have now defined the data and parameters. The next step of the Bayesian analysis is to choose prior distribution and
infer the posterior distribution for each parameter.

**Prior distribution**

We often have enough information from the data to estimate $\beta$ and thus a flat (uninformative) prior was chosen for $\beta$, i.e., $p(\beta) = 1$. The main effect for the $k$th QTL was assigned the following normal prior,

$$p(\alpha_k) = N(\alpha_k | 0, \phi_k^2)$$  \hspace{1cm} (4)

where $\phi_k^2$ is the prior variance. A scaled inverse chi-square distribution was assigned to $\phi_k^2$, which is

$$p(\phi_k^2) = \text{Inv} - \chi^2(\phi_k^2 | \tau, \omega)$$  \hspace{1cm} (5)

A special case of this prior is $\tau = \omega = 0$, leading to $p(\phi_k^2) = 1 / \phi_k^2$, called the Jeffreys' prior. However, as mentioned by TER BRAAK (2005) this prior is improper and leads to an improper posterior distribution. The revised prior is proposed by them and claimed to lead to a proper posterior distribution. The revised prior is

$$p(\phi_k^2) = \text{Inv} - \chi^2(\phi_k^2 | -2\delta, 0)$$  \hspace{1cm} (6)

where $0 < \delta \leq 0.5$. In this study we used the proper prior to avoid any potential problems caused by the improper posterior distribution. The same scaled inverse chi-square distribution was also assigned to $\sigma_k^2$,

$$p(\sigma_k^2) = \text{Inv} - \chi^2(\sigma_k^2 | -2\delta, 0)$$  \hspace{1cm} (7)

Finally, we assumed $\Theta = I_r \sigma^2$, where $\sigma^2$ is a common residual variance and was assigned the same scaled inverse prior

$$p(\sigma^2) = \text{Inv} - \chi^2(\sigma^2 | -2\delta, 0)$$  \hspace{1cm} (8)

Other structures of $\Theta$ will be considered and described in a later section.
Posterior distribution

The Markov chain Monte Carlo (MCMC) algorithm was used to implement the Bayesian shrinkage analysis. In the MCMC sampling, we only need to derive the fully conditional posterior distribution for each parameter. For example, the fully conditional posterior distribution for $\beta$ is denoted by $p(\beta \mid \cdots)$, where the dots after the vertical bar represents the data and all other parameters. Except for the prior of $\beta$, all other priors we chose in the previous section are conjugate. Therefore, the fully conditional posterior has the same form as the prior distribution. Derivation of the posterior distribution was not given and we simply provided the parameters of the fully conditional posterior distribution for each variable.

The posterior distribution for $\beta$ is multivariate normal

$$p(\beta \mid \cdots) = N(\beta \mid \mu_\beta, \Sigma_\beta)$$ (9)

where

$$\mu_\beta = \frac{1}{n} \sum_{j=1}^{n} \left( y_j - \sum_{k=1}^{Q} Z_{jk} \gamma_k \right)$$ (10)

and

$$\Sigma_\beta = \frac{1}{n} \Theta$$ (11)

The posterior for $\gamma_k$ is also multivariate normal,

$$p(\gamma_k \mid \cdots) = N(\gamma_k \mid \mu_k, \Sigma_k)$$ (12)

where

$$\mu_k = \left[ \frac{1}{\sigma_k^2} I_{m \times m} + \sum_{j=1}^{n} Z_{jk}^T \Theta^{-1} Z_{jk} \right]^{-1} \left[ \frac{1}{\sigma_k^2} I_{m \times m} \alpha_k + \sum_{j=1}^{n} Z_{jk}^T \Theta^{-1} \left( y_j - \beta - \sum_{k'=1}^{Q} Z_{jk'} \gamma_{k'} \right) \right]$$ (13)
\[ \Sigma_k = \left[ \frac{1}{\sigma_k^2} I_{m \times m} + \sum_{j=1}^{n} Z_j^T \Theta^{-1} Z_j \right]^{-1} \quad (14) \]

The posterior distribution for \( \alpha_k \) is normal,

\[ p(\alpha_k | \cdots) = N(\alpha_k | \zeta_k, \nu_k) \quad (15) \]

where

\[ \zeta_k = \left( \frac{1}{\varphi_k^2} + \frac{m}{\sigma_k^2} \right)^{-1} \left( \frac{1}{\sigma_k^2} \sum_{i=1}^{m} \gamma_{ki} \right) \quad (16) \]

and

\[ \nu_k = \left( \frac{1}{\varphi_k^2} + \frac{m}{\sigma_k^2} \right)^{-1} \quad (17) \]

We now discuss the posterior distributions for all the variance components, \( \sigma^2 \), \( \sigma_k^2 \) and \( \varphi_k^2 \) for \( k = 1, \cdots, q \). All of them are scaled inverse chi-squares as given below,

\[ p(\sigma_k^2 | \cdots) = \text{Inv-}\chi^2 \left[ \sigma_k^2 | \tau + m, \omega + (\gamma_k - 1_m \alpha_k)^T (\gamma_k - 1_m \alpha_k) \right] \quad (18) \]

\[ p(\varphi_k^2 | \cdots) = \text{Inv-}\chi^2 (\varphi_k^2 | \tau + 1, \omega + \alpha_k^2) \quad (19) \]

and

\[ p(\sigma^2 | \cdots) = \text{Inv-}\chi^2 (\sigma^2 | \tau + nm, \omega + SS) \quad (20) \]

where

\[ SS = \sum_{j=1}^{n} \left( y_j - \beta - \sum_{k=1}^{q} Z_j \gamma_k \right)^T \left( y_j - \beta - \sum_{k=1}^{q} Z_j \gamma_k \right) \]

\[ \sum_{j=1}^{n} \left( y_j - \beta - \sum_{k=1}^{q} Z_j \gamma_k \right)^T \left( y_j - \beta - \sum_{k=1}^{q} Z_j \gamma_k \right) \quad (21) \]

**MCMC sampling**

Since all the fully conditional posterior distributions have closed form distributions, either a normal or a scaled inverse chi-square, Gibbs sampler was used for sampling all the variables, which is summarized below.
(1) Initialize all variables by sampling the values from their prior distributions.

(2) Sample the parameters sequentially from their corresponding posterior distributions.

(3) Repeat the sampling cycle until the chain reaches a desired length.

The posterior sample contains all the observations after burn-in deletion and chain thinning. Post MCMC analysis was performed on the posterior sample. We often ran multiple chains and took the average posterior statistics across the chains as the Bayesian estimates of the parameters.

**Covariance structure**

We now introduce several alternative covariance structures for the residual errors.

1. **Identity matrix**

   The simple structure described earlier \( \Theta = I_m \sigma^2 \) is called the scaled identity matrix structure. This assumption is over simplified and should be relaxed in real data analysis.

2. **Diagonal matrix**

   The covariance matrix is defined as
   \[
   \Theta = D = \text{diag}[d_1, d_2, \ldots, d_m]
   \]  

   which represents uncorrelated residual errors but has taken into account non-homogenous residual variances for different environments. This assumption may hold in most situations. Each \( d \) was assigned a scaled inverse chi-square distribution and the fully conditional posterior distribution for \( d_i \) is

\[
10
\]
\[ p(d, \cdots) = \text{Inv} - \chi^2(d, \tau + n, \omega + SS) \] (23)

where

\[ SS_i = \sum_{j=1}^{n} \left( y_{ji} - \beta_i - \sum_{k=1}^{q} Z_{jk} \gamma_k \right)^2 \] (24)

3. Unstructured matrix

The unstructured covariance matrix has been used before by JIANG and Zeng (1995) for multivariate QTL mapping. The only restriction for matrix \( \Theta \) is positive definite. We assigned an inverse Wishart prior distribution to \( \Theta \). This prior is the multivariate version of the scaled inverse chi-square distribution,

\[ p(\Theta) = \text{Inv-Wishart}(\Theta | \tau, \omega) \] (25)

where \( \tau > m - 1 \) is the prior degree of belief and \( \omega > 0 \) is a positive definite matrix with the same dimension as matrix \( \Theta \). The posterior distribution remains inverse Wishart and thus

\[ p(\Theta \mid \cdots) = \text{Inv-Wishart}(\Theta | \tau + n, \omega + SS) \] (26)

where

\[ SS = \sum_{j=1}^{q} \left( y_j - \beta - \sum_{k=1}^{q} Z_{jk} \gamma_k \right) \left( y_j - \beta - \sum_{k=1}^{q} Z_{jk} \gamma_k \right)^T \] (27)

is an \( m \times m \) sum of squares matrix, different from the \( SS \) defined in equation (20).

4. Factor analytic structured matrix

The covariance matrix has the following structure,

\[ \Theta = BB^T + D \] (28)

It is called factor analytic structure because this structure has been used in factor analysis. This factor analytic structure was derived based on the following latent
variable linear model for the residual errors,

\[ \xi_j = Bu_j + e_j \]  

(29)

where \( u_j \) is an \( r \times 1 \) latent factors \( (r < m) \) with a

\[ p(u_j) = N(u_j \mid 0, I_r) \]  

(30)
distribution, \( B \) is an \( m \times r \) matrix called factor loading and \( e_j \sim N(0, D) \) is a vector of independent errors and \( D = \text{diag}[d_1, d_2, \ldots, d_m] \) is a diagonal matrix for the independent error variances.

Under the factor analytic structure, the MCMC algorithm requires sampling \( B \) and \( u_j \) for \( j = 1, \ldots, n \), in addition to other parameters. We now describe the prior and posterior of these new variables. The prior for \( u_j \) is standardized multivariate normal given in equation (30). The fully conditional posterior distribution remains multivariate normal,

\[ p(u_j \mid \cdots) = N(u_j \mid \mu_j, \Sigma_j) \]  

(31)

where

\[ \mu_j = \left[ I_r + (B^T D^{-1} B)^{-1} \right]^{-1} B^T D^{-1} \left( y_j - \beta - \sum_{k=1}^q Z_{jk} \gamma_k \right) \]  

(32)

and

\[ \Sigma_j = \left[ I_r + (B^T D^{-1} B)^{-1} \right]^{-1} \]  

(33)

The factor loadings are represented by an \( m \times r \) matrix \( B \). Let \( B_l = [B_{l1}, \ldots, B_{lm}]^T \) be the \( l \)th column of matrix \( B \) for \( l = 1, \ldots, r \). We now rewrite equation (29) as

\[ \xi_j = \sum_{l=1}^r B_l u_{jl} + e_j \]  

(34)

Given \( u_j \) and knowing that
\[ \xi_j = y_j - \beta - \sum_{k=1}^{q} Z_{jk} \gamma_k \]  

(35)

equation (34) is a typical multivariate regression problem. The fully conditional posterior distribution of \( B_j \) is multivariate normal,

\[ p(B_j | \cdots) = N(B_j | \mu_i, \Sigma_i) \]  

(36)

where

\[
\mu_i = \left[ \sum_{j=1}^{n} u_{ji}^2 D^{-1} \right]^{-1} \left[ \sum_{j=1}^{n} u_{ji} D^{-1} (\xi_j - \sum_{j' \neq i} B_{j'} u_{ji'}) \right] = \left[ \sum_{j=1}^{n} u_{ji}^2 \right]^{-1} \left[ \sum_{j=1}^{n} u_{ji} (\xi_j - \sum_{j' \neq i} B_{j'} u_{ji'}) \right]  
\]

(37)

and

\[
\Sigma_i = \left[ \sum_{j=1}^{n} u_{ji}^2 D^{-1} \right]^{-1} = \left[ \sum_{j=1}^{n} u_{ji}^2 \right]^{-1} D  
\]

(38)

Having provided the fully conditional posterior distribution for every variable, we are now ready to conduct the Gibbs sampler to infer the empirical posterior distribution for each variable.

**APPLICATIONS**

**Barley data analysis**

We used a barley data obtained from the North American Barley Genome Mapping Project (TINKER et al. 1996) to demonstrate the application of the new method. In the barley QTL mapping project, there were 127 mapped markers covering 1500 cM of the barley genome. Seven traits were investigated in the project. In this study, we used the yield trait analysis for the demonstration. The doubled haploid (DH) population was initiated from the cross between Harrington and TR306. The DH
population consisted of 145 lines, each grown in 28 different environments. The dataset was updated after it was first published in 1996, but the difference between the original and the updated data was minor so that we can still compare the current result with that from the original study.

We used six different covariance structures to analyze the data, which were (1) the homogeneous (constant) variance $\Theta = I_{28}\sigma^2$, (2) the heterogeneous variances $\Theta = D$, (3) unstructured matrix $\Theta$ (positive definite), (4) the first order factor analytic structure $\Theta = B_{28 \times 1}B^T_{1 \times 28} + D$, (5) the second order factor analytic structure $\Theta = B_{28 \times 2}B^T_{2 \times 28} + D$ and (6) the third order factor analytic structure $\Theta = B_{28 \times 3}B^T_{3 \times 28} + D$.

The length of the Markov chain consisted of 200,000 sweeps. The first 100,000 sweeps were deleted as burn-in and thereafter the chain was thinned by keeping one observation in every 100 sweeps, producing 1,000 observations in the collected posterior sample for post MCMC analysis.

In order to test the significance of the QTL effects, we conducted a permutation test to generate the null distribution of each main effect and each Q×E interaction effect. In the permutation analysis, we repeated the MCMC sampling method as described before but reshuffled the phenotypic values. The permutation analysis was proposed by CHE AND XU (2010) who called it permutation inside the Markov chain. In the permutation analysis, the length of the Markov chain was 200,000 sweeps. The first 100,000 sweeps were deleted as burn-in and the chain thinning rate was one over 25. The posterior sample contained 4,000 observations. From the null distribution, we drew confidence interval for each estimated effect. An
effect was claimed to be significant if the estimated value fell outside of the 99% confidence interval of the null distribution.

Among the six covariance structures, the second structure $\Theta = D$ (a diagonal matrix) detected the maximum number of QTL (main and Q×E interaction effects). Although more QTL does not mean better, it is hard to use cross validation to evaluate different structures under the MCMC implemented Bayesian analysis. Bayes factors are often used to evaluate different models. However, the complexity of our proposed model makes the calculation of the Bayes factors difficult. Therefore, we used the Bayesian information criteria (BIC) to evaluate the performance of the six different models. The BIC score was calculated using

$$BIC = -2\log(L) + p\log(n)$$

(39)

where $L$ is the likelihood function evaluated at the estimated parameters, $p$ is the number of parameters, and $n$ is the sample size. The Bayesian estimates of the parameters in $L$ are the posterior means of $\beta$, $\gamma$ and $\Theta$. The BIC scores are shown in Table 1, which indicates that the second (heterogeneous residual variance) model performed better than all other models. The first order factor analytic model was the second best model with a BIC score slightly larger than that of the best model. The result of the best model is depicted in Figure 1, where the posterior means of the main effects and the Q×E interaction effects are plotted against the genome locations of the markers. Figure 1 also gives the 99% confidence intervals for the main and Q×E interaction effects. Eight markers showed significant main effects and 18 markers showed significant Q×E interaction. The 18 interacting markers were
distributed across all seven chromosomes of the entire genome. Only one marker had both the main and the Q×E interaction effects. Each of the other markers had either a main effect or a Q×E interaction effect but not both. The estimated main and Q×E interaction effects for the markers are given in Table 2.

We also performed an individual marker analysis to compare the result with that of the Bayesian analysis. For the individual marker analysis, QTL mapping was conducted separately for each environment. The average estimated effect for each marker across the 28 environments represented the main effect while the variance of the estimated effects across the environments represented the Q×E interaction effect. The estimated main and Q×E interaction effects of the single marker analysis are shown in Figure 2. We can see that this figure is quite similar to Figure 1 of the Bayesian analysis. The main difference between the two figures is the different sharpness of the marker effects. The Bayesian analysis generated very clean (sharp) signals of the plots.

Arabidopsis data analysis

The barley data contains many environments which is hard to find in most studies. So we also applied our model to a recombinant inbred line data of Arabidopsis data (LOUDET et al. 2002), where two parents initiating the line cross were Bay-0 and Shahdara with Bay-0 as the female parent. Flowering time was recorded for each line in two environments: long day (16-hour photoperiod) and short day (8-hour photoperiod). The population contained 420 lines. A total of 38 microsatellite markers were used for QTL mapping. We inserted a pseudo marker in every 2 cM of the
genome and had a total of 200 markers (38 true markers plus 162 pseudo markers) in our analysis.

The variance of Q×E interaction $\sigma_k^2$ may not be estimated accurately due to small environments. So in small environments the variance would then simply serve as a tool to shrink the environment specific QTL effects. The bias of $\sigma_k^2$ would lead to biased estimation of main effect as well. Although the MCMC algorithm remains the same as before, we need to revise our post-MCMC procedure. We use vector $\gamma_k$ to estimate Q×E interaction effects of kth marker. The differences between vector $\gamma_k$ and its mean represent Q×E interaction effects. In two environments case, we can just use the differences between the two components of $\gamma_k$ as the Q×E interaction effects because vector $\gamma_k$ is a 2×1 vector. Since there are only two environments, we did not use factor analytic model to analyze the data. The BIC scores for the three models (homogeneous, heterogeneous and unstructured covariance matrices) are 3798.49, 3775.64, and 3645.55. Figure 3 shows the main and Q×E interaction effects of the Arabidopsis data under the unstructured covariance model. The 95% confidence intervals for the main and Q×E interaction effects are also given. Four markers showed significant main effects and six markers showed significant Q×E interaction effects.

**Simulation Study**

The barley data analysis did not show the advantage of fitting appropriate covariance structures over the simple diagonal covariance matrix because the 28 different environments did not seem to be correlated. Therefore, we conducted two
simulation experiments (Simulations 1 and 2) in this section to demonstrate the importance of covariance structure to the Bayesian analysis of Q×E interaction. We also did a two-environment simulation (Simulation 3) to demonstrate the fitness of our model for small environments.

In simulation 1, we used the real marker information from the North American Barley Genome Mapping Project (TINKER et al. 1996) to simulate the genome. We simulated 127 markers from seven chromosomes with marker distances exactly the same as the real data. We simulated 145 DH lines in 28 environments. The intercept \( \beta \) was given values ranging from 200 to 605 for the 28 different environments. We assumed that 10 out of the 127 markers had main effects and also Q×E interaction effects in the 28 environments. In the simulation experiment, we chose the factor analytic covariance structure \( \Theta = BB^T + D \) with \( B \) defined as a \( 28 \times 3 \) matrix, indicating that correlations had occurred between different environments. The true values of \( \beta \), the QTL effects, the \( B \) matrix and the \( D \) matrix are given in Table 4 and 5. The simulated data were analyzed using the six different covariance structures described early in the barley data analysis. We expected that the first three structures (homogeneous variance, heterogeneous variance and unstructured matrices) would perform poorly but the last three structures (first order, second order and third order factor analytic structures) perform better, especially the third order factor analytic structure.

In the MCMC implemented Bayesian analysis, the length of the Markov chain was 50,000 sweeps. The first 25,000 sweeps (burn-in period) were deleted. The chain
thinning rate was one in 50. The empirical posterior sample contained 500 observations for the post MCMC analysis. The MCMC experiment with the same simulated data was repeated a few times to make sure that the chain had converged to the stationary distribution.

The results of the simulation studies are depicted in Figure 4 for the average main effects of 20 replicates and Figure 5 for the average Q×E interaction effects of 20 replicates. These two figures show that the estimated QTL effects agreed well with the true effects. Figures 4 and 5 also show that the first four covariance structures (homogeneous residual, heterogeneous residual, unstructured covariance and first order factor analytic structure) have some noticeable background noise, indicating some false positive had occurred. However, the last two factor analytic structures have very little background noise. From the two figures, the background noise of the first four covariance structures may not be very clear. So we calculated standard deviations of each marker’s main effect among the 20 replicates and plotted them in figure 6 from which we can see the difference among the six models. The performance of the last two factor analytic models is very stable for majority of the markers without main effects. While the first four structures, especially the first two, cannot not achieve such a nice performance, which means that among the 20 replicates these models generated some false main effects. Table 3 shows the average BIC scores for the six different covariance structures. We see that the factor analytic structure outperformed the other three models. This is consistent with our expectation. The lowest BIC occurred in the second order factor analytic structure. However, the
third order factor analytic structure (the true model) was just slightly higher than the second order structure. The log likelihood value of the third order factor was higher than the second-order. Table 4 gives the average estimated main and Q×E interaction effects obtained from the 20 replicates based on the second order factor analytic model. When we compared the estimated main effects and the true effects, we noticed that large main effects were estimated quite accurately but small effects were shrunken to zero. The Q×E interaction effects were always detectable regardless the sizes of the effects. Table 5 gives the true intercepts and the residual error variances (the D matrix) along with their estimated values. The true B matrix is also given in Table 5. The estimated B was not given here because the columns of B are independent and thus are exchangeable. This does not affect the estimate of the covariance structure. We have checked the average estimated variance-covariance matrix and did observe three separate environment groups.

Although the stability test and BIC scores showed the advantage of factor analytic model, the differences of marker main effects for the six models are not very obvious in figure 4. So performed the second simulation experiment (simulation 2) to further demonstrate the advantage of factor analytic model. We mainly focused on comparison of heterogeneous diagonal structure and three factor analytic models. In this simulation, 100 DH lines in 8 environments were generated with 30 markers. The distance between two nearby markers is 30 cM. The intercept $\beta$ was given values ranging from 200 to 305 for the eight environments. We assumed that three out of the 30 markers had main effects and also Q×E interaction effects in the eight
environments. We also chose the factor analytic covariance structure \( \Theta = BB^T + D \) with \( B \) defined as a \( 8 \times 2 \) matrix. The first column of \( B \) had values of 20, 10, 10, 5, 0, 0, 0, 0. The second column had values of 0, 0, 0, 0, 15, 15, 10, 2. Matrix \( D \) was an identity matrix. Figure 7 shows the estimated main effects of the four models. We can clearly see many false positive main effects in the heterogeneous diagonal structure. The BIC scores for the four models are 4264, 3362, 2495, and 2558, which also are in favor of the factor analytic models.

In simulation 3, we also generated 100 DH lines using the same marker information given by simulation 2, but this time we only simulated two environments. The intercept were 200 and 215 for the two environments. The factor analytic covariance structure was used with \( B \) defined as a \( 2 \times 2 \) matrix. The first column of \( B \) had values of 1 and 2. The second column had values of 2 and 1. Matrix \( D \) was an identity matrix. The MCMC and post-MCMC analysis of this data used the setup as Arabidopsis data analysis. Figure 8 gave the comparison of the true and the estimated main and Q×E interaction effects. From the figure, the true and the estimated marker effects are very close to each other for all the three models. The promising results also demonstrate that our proposed method is a good choice to handle data with small environments.

**DISCUSSION**

The importance of this study is reflected by two major contributions to Q×E study, the multiple QTL model and the factor analytic covariance structure. The
multiple QTL model for Q×E is an extension of the Bayesian shrinkage analysis for mapping QTL in a single environment (Xu 2003). The factor analytic covariance structure is available in the literature but has never been applied to QTL mapping. Other covariance structures may be considered in future studies, e.g., the AR(1) and CS (compound symmetry) covariance structures. These alternative structures can be used to fit models when the environments represent temporal or spatial variation. The 28 environments in the barley experiments represent 28 different locations (spatial variation). However, the information about the location was not available to us. We believe that the factor analytic structure is robust and can be fit to a wide variety of covariance structures, ranging from the simplest diagonal matrix to the most complicated unstructured matrix by choosing different orders of the factors. This has been demonstrated by the similarity of the diagonal matrix and the first order factor analytic model in our data analyses. The factor analytic model is also easy to fit under the general linear model framework. Both the factor loadings and the factors themselves have normal posterior distributions and can be sampled using the Gibbs sampler approach.

The most significant contribution of this study was to use the variance of QTL effects across environments to measure the size of the Q×E interaction for a particular QTL. This has significantly simplified the Q×E study. If the number of environments were small, however, the variance would not be accurately estimated. In this case, one should use some kind of linear contrasts of the environment specific effects as a measure of the Q×E interaction. Arabidopsis data and simulation 3 are two examples
of such a treatment. The variance would then simply serve as a tool to shrink the
environment specific QTL effects. The MCMC sampling procedure remains the same,
but the post MCMC analysis needs to be modified.

The method developed in the current study only applies to plants where the
same genotype can be replicated in multiple environments. In animals where the same
genotype cannot be replicated (except identical twins), some modification is required.
For example, if an F2 family is raised in three environments, each animal may have a
different genotype from other animals. This argument also applies to QTL by sex
interaction where the same individual cannot be split into male and female. The
modification is not trivial and thus deserves a further study.

Although the environment specific QTL effects, denoted by vector $\gamma_k$ for the
$k$th marker, are only used to draw the posterior distributions for the main and Q×E
interaction effects, they may be interesting parameters in their own rights. The
posterior mean of each $\gamma_k$ can be used to predict the molecular breeding value of
each line in a particular environment. This information may facilitate marker assisted
selection (using a few markers) or genome selection (using all markers of the entire
genome). Genome selection has been an important strategy for animal (MEUWISSEN et
al. 2001) and plant breeding (Xu 2003).

The Bayesian method presented here only applies to multiple marker analysis,
i.e., each marker is treated as a putative QTL. If the markers are not evenly placed in
the genome, one may insert some pseudo markers in regions not well covered by
markers. In the regions with saturated markers, one may only use a few selected
markers to avoid a potential multicollinearity problem. With the current molecular
technology, genomes of most species of agricultural importance may be saturated very
soon with high density markers. Pseudo marker insertion will no longer be necessary,
but marker selection will become important. One strategy for marker selection is to
include one marker in every $d$ centiMorgan to the Bayesian model. The optimal
strategy may be the moving interval approach proposed by Wang et al. (2005), in
which a fixed number of putative QTL was included in the model for each
chromosome and the position of the putative QTL can move (jump) among a few
neighboring markers. This approach may be adopted in the second stage of mapping,
i.e., fine mapping after the important QTL regions have been identified.

One drawback of the MCMC implemented Bayesian method is the slow
computation process due to the large number of environments and the high
dimensionality of the model. A quick method may be the posterior mode estimation in
which only the conditional posterior modes are presented as the Bayesian estimates
for the interested parameters. Although the estimates are no longer Bayesian estimates,
the results may be comparable. This quick posterior mode estimation may provide
preliminary results to be used for further analysis using the fully Bayesian analysis.

Finally, the entire data analyses were conducted using a program developed in
R. Interested readers may visit our website (www.statgen.ucr.edu) to download the
program and the sample data to test the method and analyze their own data.
ACKNOWLEDGEMENTS

This project was supported by the National Research Initiative (NRI) Plant Genome of the USDA Cooperative State Research, Education and Extension Service (CSREES) 2007-02784 to SX.
REFERENCES


JANSEN, R. C., 1993 Interval mapping of multiple quantitative trait loci. Genetics 135: 205-211.


MEUWISSEN, T. H., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic

Mendelian factors underlying quantitative traits in tomato: comparison across

PIEPOHO, H. P., 2000 A mixed-model approach to mapping quantitative trait loci in
barley on the basis of multiple environment data. Genetics 156: 2043-2050.

ROMAGOSA, I., S. E. ULLRICH, F. HAN and P. M. HAYES, 1996 Use of the additive main
effects and multiplicative interaction model in QTL mapping for adaptation in
barley. TAG Theoretical and Applied Genetics 93: 30-37.

STUBER, C. W., S. E. LINCOLN, D. W. WOLFF, T. HELENTJARIS and E. S. LANDER, 1992
Identification of genetic factors contributing to heterosis in a hybrid from two

TER BRAAK, C. J. F., M. P. BOER and M. BINK, 2005 Extending Xu's Bayesian model
for estimating polygenic effects using markers of the entire genome. Genetics
170: 1435.

TINKER, N. A., D. E. MATHER, B. G. ROSSNAGEL, K. J. KASHA, A. KLEINHOFs et al.,
1996 Regions of the genome that affect agronomic performance in two-row

WANG, H., Y. M. ZHANG, X. LI, G. L. MASINDE, S. MOHAN et al., 2005 Bayesian
shrinkage estimation of quantitative trait loci parameters. Genetics 170:
465-480.

XIAOHONG, C., and X. SHIZHONG, 2010 Significance Test and Genome Selection in

Xu, S., 2003 Estimating polygenic effects using markers of the entire genome.

Genetics 163: 789-801.

Zeng, Z. B., 1994 Precision mapping of quantitative trait loci. Genetics 136:

1457-1468.
FIGURE LEGENDS

Figure 1. The estimated main and Q×E interaction effects for markers of the entire barley genome. (a) main effects under the heterogeneous residual variance model; (b) Q×E interaction effects under the heterogeneous residual variance model. Chromosomes are separated by the dotted vertical reference lines. The dashed curves represent the 99% confidence intervals generated from the permutation analysis.

Figure 2. Estimated main and Q×E interaction effects using single marker analysis. (a) main effects; (b) Q×E interaction effects; Chromosomes are separated by the dotted vertical reference lines.

Figure 3. Estimated main and Q×E interaction effects for markers of the entire Arabidopsis genome. (a) main effects under the unstructured model; (b) Q×E interaction effects (expressed as differences between the QTL effects in the two environments) under the unstructured model. Chromosomes are separated by the dotted vertical reference lines. The marker positions are represented by the ticks on the horizontal line. The dashed curves represent the 95% confidence intervals generated from the permutation analysis.

Figure 4. The average main effects across 20 replicated simulation experiments for the entire genome in the first simulation experiment. (a) homogeneous residual variance (scaled identity matrix); (b) heterogeneous residual variances (diagonal matrix); (c) unstructured covariance matrix; (d) first order factor analytic structure; (e) second order factor analytic structure; (f) third order
factor analytic structure. The dotted vertical reference lines divide the genome into seven chromosomes. The black solid needles are the average marker main effects. The true effects of the markers are indicated by the grey needles.

**Figure 5.** The average Q×E interaction effects across 20 replicated simulation experiments for the entire genome in the first simulation experiment. Chromosomes are separated by the dotted vertical reference lines. The black solid needles are the average QxE interaction effects. The true effects of QxE interaction are indicated by the grey needles.

**Figure 6.** The standard deviations (stabilities) of estimated main effects across 20 replicated simulation experiments for the entire genome in the first simulation experiment. The positions of markers with (non-zero) simulated effects are indicated by the ticks on the horizontal lines at value of -2.

**Figure 7.** The marker main effects in the second simulation experiment. (a) heterogeneous residual variances (diagonal matrix); (b) first order factor analytic structure; (c) second order factor analytic structure; (d) third order factor analytic structure. The black solid needles are the estimated marker main effects. The grey needles are the estimated main effects.

**Figure 8.** The marker main effects and Q×E interaction effects in the third simulation experiment. (a) and (b): homogeneous residual variance (scaled identity matrix); (c) and (d): heterogeneous residual variances (diagonal matrix); (e) and (f): unstructured covariance matrix. The black needles are the estimated effects (main and interaction). The grey needles are the true values (main and interaction).
interaction).
Table 1. BIC scores of the six variance-covariance structures for the barley data analysis. The number of parameters is denoted by $p$.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Log likelihood</th>
<th>$p$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>-18055.18</td>
<td>1</td>
<td>36118.67</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>-17239.85</td>
<td>28</td>
<td>34712.35</td>
</tr>
<tr>
<td>Unstructured</td>
<td>-17234.20</td>
<td>406</td>
<td>37841.83</td>
</tr>
<tr>
<td>First-order factor</td>
<td>-17171.71</td>
<td>56</td>
<td>34808.72</td>
</tr>
<tr>
<td>Second-order factor</td>
<td>-17155.33</td>
<td>84</td>
<td>35008.60</td>
</tr>
<tr>
<td>Third-order factor</td>
<td>-17143.79</td>
<td>118</td>
<td>35218.19</td>
</tr>
</tbody>
</table>
Table 2. Estimated main and Q×E interaction effects and their 99% confidence intervals of the null distributions for the barley yield data. All the estimated effects are outside of the 99% confidence intervals.

<table>
<thead>
<tr>
<th>Marker Number (Position, Chromosome)</th>
<th>Main effect</th>
<th>Q×E interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (3.516, 1)</td>
<td>--</td>
<td>10.240 (0, 7.719)</td>
</tr>
<tr>
<td>3 (7.892, 1)</td>
<td>--</td>
<td>15.805 (0, 8.066)</td>
</tr>
<tr>
<td>13 (111.273, 1)</td>
<td>--</td>
<td>42.046 (0, 6.797)</td>
</tr>
<tr>
<td>18 (182.978, 1)</td>
<td>--</td>
<td>7.070 (0, 6.356)</td>
</tr>
<tr>
<td>23 (24.59, 2)</td>
<td>4.379 (-0.904, 1.063)</td>
<td>--</td>
</tr>
<tr>
<td>32 (145.863, 2)</td>
<td>--</td>
<td>8.974 (0, 5.420)</td>
</tr>
<tr>
<td>38 (247.586, 2)</td>
<td>-4.447 (-1.868, 1.340)</td>
<td>--</td>
</tr>
<tr>
<td>50 (168.297, 3)</td>
<td>-7.927 (-0.739, 0.725)</td>
<td>--</td>
</tr>
<tr>
<td>51 (172.134, 3)</td>
<td>--</td>
<td>9.226 (0, 6.239)</td>
</tr>
<tr>
<td>52 (178.493, 3)</td>
<td>--</td>
<td>20.160 (0, 4.541)</td>
</tr>
<tr>
<td>53 (182.222, 3)</td>
<td>--</td>
<td>4.557 (0, 4.214)</td>
</tr>
<tr>
<td>54 (185.101, 3)</td>
<td>3.875 (-0.975, 1.018)</td>
<td>--</td>
</tr>
<tr>
<td>59 (63.732, 4)</td>
<td>--</td>
<td>9.724 (0, 5.383)</td>
</tr>
<tr>
<td>64 (145.469, 4)</td>
<td>4.711 (-1.230, 1.541)</td>
<td>19.317 (0, 5.378)</td>
</tr>
<tr>
<td>65 (163.074, 4)</td>
<td>--</td>
<td>22.422 (0, 5.337)</td>
</tr>
<tr>
<td>77 (178.118, 5)</td>
<td>-4.542 (-1.256, 0.836)</td>
<td>--</td>
</tr>
<tr>
<td>86 (41.568, 6)</td>
<td>--</td>
<td>6.603 (0, 4.198)</td>
</tr>
<tr>
<td>89 (85.438, 6)</td>
<td>--</td>
<td>7.548 (0, 7.476)</td>
</tr>
<tr>
<td>95 (122.839, 6)</td>
<td>--</td>
<td>21.108 (0, 7.386)</td>
</tr>
<tr>
<td>96 (131.476, 6)</td>
<td>-5.819 (-1.539, 0.884)</td>
<td>--</td>
</tr>
<tr>
<td>102 (7.247, 7)</td>
<td>--</td>
<td>77.311 (0, 5.606)</td>
</tr>
<tr>
<td>105 (43.111, 7)</td>
<td>5.605 (-0.704, 0.958)</td>
<td>--</td>
</tr>
<tr>
<td>106 (47.487, 7)</td>
<td>--</td>
<td>16.614 (0, 6.140)</td>
</tr>
<tr>
<td>116 (164.789, 7)</td>
<td>--</td>
<td>9.899 (0, 5.959)</td>
</tr>
<tr>
<td>118 (173.542, 7)</td>
<td>--</td>
<td>33.210 (0, 3.860)</td>
</tr>
</tbody>
</table>
Table 3. Average BIC score for six different variance-covariance structures in the simulation study. The number of parameters is denoted by $p$.

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>Log likelihood</th>
<th>$p$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>-18096.83</td>
<td>1</td>
<td>36201.97</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>-17564.80</td>
<td>28</td>
<td>35362.26</td>
</tr>
<tr>
<td>Unstructured</td>
<td>-16815.18</td>
<td>406</td>
<td>37003.79</td>
</tr>
<tr>
<td>First-order factor</td>
<td>-17176.53</td>
<td>56</td>
<td>34818.36</td>
</tr>
<tr>
<td>Second-order factor</td>
<td>-16887.74</td>
<td>84</td>
<td>34473.44</td>
</tr>
<tr>
<td>Third-order factor</td>
<td>-16858.56</td>
<td>118</td>
<td>34647.73</td>
</tr>
</tbody>
</table>
**Table 4.** True and estimated QTL main effects and Q×E interaction effects from 20 replicated simulation experiments under the second order factor analytic covariance structure.

<table>
<thead>
<tr>
<th>Marker Number (Position, Chromosome)</th>
<th>Main Effect</th>
<th>Q×E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Estimated</td>
</tr>
<tr>
<td>1 (0, 1)</td>
<td>5</td>
<td>1.14</td>
</tr>
<tr>
<td>14 (127.208, 1)</td>
<td>-7</td>
<td>-6.51</td>
</tr>
<tr>
<td>32 (145.863, 2)</td>
<td>-1</td>
<td>-0.01</td>
</tr>
<tr>
<td>48 (153.931, 3)</td>
<td>10</td>
<td>8.08</td>
</tr>
<tr>
<td>52 (178.493, 3)</td>
<td>-4</td>
<td>-0.75</td>
</tr>
<tr>
<td>65 (163.074, 4)</td>
<td>-8</td>
<td>-8.14</td>
</tr>
<tr>
<td>71 (109.389, 5)</td>
<td>15</td>
<td>15.52</td>
</tr>
<tr>
<td>84 (10.533, 6)</td>
<td>-1</td>
<td>0.00</td>
</tr>
<tr>
<td>96 (131.476, 6)</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>110 (122.584, 7)</td>
<td>3</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table 5. The true and estimated intercepts, the $B$ and $D$ matrices used in simulation 1.

The estimated parameters are from the second order factor analytic model. The estimated $B$ is not given (see details in the simulation study). The symbol E stands for environment.

<table>
<thead>
<tr>
<th>E</th>
<th>Intercept</th>
<th></th>
<th>$B$ matrix</th>
<th></th>
<th>$D$ matrix</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Estimated</td>
<td>True</td>
<td>True</td>
<td>Estimated</td>
<td>Estimated</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>199.59</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>7569</td>
</tr>
<tr>
<td>2</td>
<td>215</td>
<td>215.11</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>6561</td>
</tr>
<tr>
<td>3</td>
<td>230</td>
<td>230.28</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>1849</td>
</tr>
<tr>
<td>4</td>
<td>245</td>
<td>243.15</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>2116</td>
</tr>
<tr>
<td>5</td>
<td>260</td>
<td>259.01</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>784</td>
</tr>
<tr>
<td>6</td>
<td>275</td>
<td>276.03</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>729</td>
</tr>
<tr>
<td>7</td>
<td>290</td>
<td>290.02</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>1296</td>
</tr>
<tr>
<td>8</td>
<td>305</td>
<td>305.01</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>1296</td>
</tr>
<tr>
<td>9</td>
<td>320</td>
<td>320.02</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>5184</td>
</tr>
<tr>
<td>10</td>
<td>335</td>
<td>335.68</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>4225</td>
</tr>
<tr>
<td>11</td>
<td>350</td>
<td>350.78</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>121</td>
</tr>
<tr>
<td>12</td>
<td>365</td>
<td>365.58</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>1444</td>
</tr>
<tr>
<td>13</td>
<td>380</td>
<td>381.38</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>529</td>
</tr>
<tr>
<td>14</td>
<td>395</td>
<td>397.56</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>1296</td>
</tr>
<tr>
<td>15</td>
<td>410</td>
<td>411.56</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>1521</td>
</tr>
<tr>
<td>16</td>
<td>425</td>
<td>425.82</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>529</td>
</tr>
<tr>
<td>17</td>
<td>440</td>
<td>441.16</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>2809</td>
</tr>
<tr>
<td>18</td>
<td>455</td>
<td>456.33</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>529</td>
</tr>
<tr>
<td>19</td>
<td>470</td>
<td>469.16</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>2116</td>
</tr>
<tr>
<td>20</td>
<td>485</td>
<td>484.17</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>2304</td>
</tr>
<tr>
<td>21</td>
<td>500</td>
<td>498.4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1156</td>
</tr>
<tr>
<td>22</td>
<td>515</td>
<td>515.16</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>729</td>
</tr>
<tr>
<td>23</td>
<td>530</td>
<td>530.17</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1369</td>
</tr>
<tr>
<td>24</td>
<td>545</td>
<td>544.75</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>400</td>
</tr>
<tr>
<td>25</td>
<td>560</td>
<td>559.78</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1369</td>
</tr>
<tr>
<td>26</td>
<td>575</td>
<td>575.4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>625</td>
</tr>
<tr>
<td>27</td>
<td>590</td>
<td>589.9</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>2401</td>
</tr>
<tr>
<td>28</td>
<td>605</td>
<td>604.43</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>841</td>
</tr>
</tbody>
</table>
Figure 1
Figure 2
Figure 4
Figure 5
Figure 6
Figure 7
Figure 8