Flexible multivariate linear mixed models for structured multiple traits

Hyeonju Kim[1], Gregory Farage[1], John T. Lovell[2], John K. Mckay[3], Thomas E. Juenger[4], & Śaunak Sen[1]

[1]*Division of Biostatistics, Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA.*

[2]*Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA.*

[3]*Department of Bioagricultural Sciences and Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA.*

[4]*Department of Integrative Biology, University of Texas, Austin, Texas, USA.*

Many genetic studies collect structured multivariate traits containing rich information across the traits. We present a flexible multivariate linear mixed model for Quantitative Trait Loci analysis (FlxQTL) incorporating trait covariates and a trait kernel to test genetic associations across multiple correlated traits while capturing genetic relatedness. FlxQTL effectively handles a trait across multi-environments, multi-time points, etc. enhancing computation speed.

Multivariate traits embrace spatial, temporal [1,2] structure or more complex structures because they are measured in multiple environments or treatments, possibly over time (sometimes repeated mearsures), with diverse methods in breeding designs. For instance, a trait of body weights might be measured weekly, crop yield (or biomass) may be in a particular location/year combination, i.e.

multi-environment trials (MET), and gene expression may be in different brain regions of heterogeneous rats. Such traits have been explored in the context of linear mixed models (LMMs) [3–6] by controlling confounding due to genetic relatedness among individuals to identify genetic loci contributing to quantitative traits of interest (QTL) in genomewide association studies (GWAS) [7–13]. Multivariate LMMs (MLMMs) enhance statistical power over univariate LMMs [14] because they can accumulate signals across traits [8,15] with a common genetic locus.

For these models, parameter estimation is a multidimensional optimization problem; there is a risk of reaching suboptimal solutions, and it is computationally expensive when fitting multiple traits [9,14]. Existing algorithms using an expectation-maximization (EM) method followed by second order schemes such as Newton-Rapson (NR), average information (AI), etc. provide stable, fast convergence [5,14] (GCTA [16,17], ASReml [18], WOMBAT [19], GEMMA [14]). The computational complexity is $O(n^3m^3)$ for EM, NR and $O(n^3m^7)$ for AI (m traits, n individuals), which indicates GCTA, WOMBAT, and ASReml are unfeasible for GWAS with a great deal of SNPs and a moderate number of individuals [14]. GCTA can fit only up to 2 traits, ASReml and WOMBAT have flexibility in choosing some covariance structures and including fixed effect covariates regarding individuals while restricting the number of traits as GEMMA, but they are much slower than GEMMA. Further, all those algorithms cannot incorporate the multivariate correlated structure of traits in the model, which benefits computational efficacy, increase in power and flexibility in data analysis.

Here, we introduce a FlxQTL, which can test associations between genetic markers and

multiple correlated traits (**Supplementary Software** and https://github.com/hkim89/FlxQTL.jl). Our algorithm, FlxQTL, builds on a MLMM modeled by the sum of two Kronecker products: the first component consists of a genetic kinship matrix and a trait kernel, i.e. a climatic relatedness matrix for MET, yielding multiplicative random GxE effects (**Supplementary Note**); the second is unstructured residual noise.

$$Y = XBZ' + R + E. \tag{1}$$

Let $Y_{n\times m}$ be a matrix for a trait measured in n individuals across m environments (or over m time points), or phenotypes of n individuals for m traits, $X_{n\times p}$ be p genotype probabilities (or genotype including the intercept) of a marker to be tested without (or with) covariates in regard to individuals. $Z_{m\times q}$ be q trait covariates [1] such as contrasts, (orthonomal) basis functions, etc. with B effect sizes, R be random effects, and E be residual errors. The likelihood ratio test (LRT) is expressed as a LOD score ($-\log_{10}(LRT)$) to test the association between a marker and a multivariate trait of interest. Cutoffs are obtained from the null distribution of maximum LOD scores by permutation test (Online Methods).

The algorithm, FlxQTL, is marked by simplifying parameter estimation by its model structure (**Fig. 1**) and the choice of numerical methods. The trait covariates reduce the matrix of fixed effect sizes from $p \times m$, which needs to be estimated in existing MLMMs [8],[14], to $p \times q$, thereby playing a substantial role when m gets large. The trait kernel reduces a $m \times m$ random effects covariance matrix to a scalar parameter. With an unstructured environment error (or model error) matrix, only $pq+1+\frac{m(m+1)}{2}$ parameters in total are estimated rather than $pm+m(m+1)$ ($q << m$). FlxQTL has been tested with up to a 36-dimensional quantitative trait. Unlike NR, AI, a Speed

restarting Nesterov's accelerated gradient method requires no second derivative providing stable linear convergence [20]. It only needs the first order approximation, here an expectation-conditional maximization (ECM) [21], followed by the acceleration step using a particularly tuned momentum coefficient, which behaves like the second derivative, to update parameters (**Supplementary Note**), thereby offering simpler time complexity of the two methods (Online Methods).

The algorithm, FlxQTL, is flexible in handling varied trait structures, as well as genotype (probability) data of multi-parental crosses. Not only can the trait covariates be trait contrasts to capture GXE interactions but they can be orthonormal basis functions, such as B-splines, wavelet, Fourier, etc. to account for a trait trend across the environmental gradient or a trend by elapsed time. One can efficiently analyze high-dimensional structured traits such as a trait across multi-environments, multi-traits, and time-valued traits for plant, mouse, and human data, etc. For genotype data, one can also use genotype probability data for multi-parental crosses, so that the information loss could be minimal in QTL analysis.

To assess the performance of FlxQTL, we compared with GEMMA on Mouse HS1940 data available in the GEMMA package. The trait kernel and trait covariates set to be identities in FlxQTL since they are not supported by GEMMA. We proceeded the three-trait data analysis followed by measuring computation times when increasing the number of simulated traits and SNPs between the two algorithms (Data processing and analysis in Online Methods). The difference in P values between the algorithms was negligible (**Fig. 2b** and **2c**). GEMMA was shown to be fast for only the small size traits (3 traits), whereas FlxQTL was 1.5-21 times faster than GEMMA with

larger traits (6-12 traits) when increasing the number of SNPs (**Fig. 2a**). These results show that FlxQTL exhibits strength in fast implementation for high-dimensional traits.

We also evaluated the statistical performance of FlxQTL simulating 6-dimensional quantitative trait datasets employing genotypes and annotated environmental information from *Arabidopsis thaliana* data [22] (Online Methods). We carried out two comprehensive simulation studies on contributions of a climatic relatedness matrix as a trait kernel and trait covariates, respectively. A comparison of power with and without the climatic relatedness was made when including the same trait contrasts (**Supplementary Table 1** and **Supplementary Figs. 3-5**). We next compared power with and without trait covariates, allowing to have main QTL effects and QTL and site interactions, keeping those climatic relatedness matrices the same (**Supplementary Fig. 6**). The thresholds for the type I error rate were calibrated by the distribution of maximum LOD scores under the null model of existing multivariate gene and environment random effects. The power was measured under the alternative model based on the thresholds. The overall result showed that including climatic relatedness yielded only subtle differences in power. In assessing the trait covariates, for each genetic variance in plasticity varying fixed effect sizes, the power was the highest when including trait covariates (**Supplementary Fig. 6**).

We tested FlxQTL with the empirical data from *A. thaliana*, where the trait was the mean number of fruits per seedling planted in two distinct regions, Sweden and Italy, from July, 2009 to June, 2012. Due to lack of statistical frameworks for testing the presence of QTL in such reciprocal transplant experiments, Ågren et al. opted to conduct six separate univariate QTL scans and dissect

effects post hoc. However, FlxQTL can directly test evolutionary hypotheses of local adaptation in replicated reciprocal transplants (The data processing and analysis is detailed in Online Methods, **Supplementary Tables 2** and **3** and **Supplementary Figs. 7** and **16**).

FlxQTL was further tested with more complex data of a four-way outbred mapping population in outcrossing switchgrass (*Panicum virgatum*), where the trait was flowering time measured across 10 latitudinal gradients for 4 years (2016-2019), considered as a 36-dimensional multivariate trait without 4 missing site by year combination. To augment benefits from trait covariates, we attempted environmental scan just as genome scan using environmental information such as monthly averaged photoperiod to obtain LOD scores conditional on each QTL after selecting three major QTL from the initial genome scan. We found that the photoperiod may influence a QTL@6.6012 in Chromosome 5N, in particular, compared with the effects on the other two QTL (**Supplementary Fig. 17** and Data processing and analysis in Online Methods).

To gauge the performance of the trait covariates on genomewide association studies (GWAS), we analyzed a trait of weekly measured body weights after birth to 16 weeks in 1212 F2 intercrosses between Gough Island mice and WSB/EiJ for the genetic basis of adaptation [2]. Using a 4 degrees of freedom cubic spline as trait covariates, FlxQTL captured the weight trend and corresponding weight growth rate simultaneously and yielded the result similar to R/qtl [23] with two exceptions, which was resolved by 2D-genome scan with several more detection (Online Methods and **Supplementary Figs. 18** and **22**).

Overall, our results demonstrate how the model and the corresponding algorithm can ef-

fectively deal with higher dimensional traits with spatial or temporal structure while running fast parameter estimation. On current simulations and real data analysis, FlxQTL using trait covariates increases power compared with existing algorithms while leveraging information from climatic relatedness generated by various annotated environmental information including an independent case turns out to be insensitive to the information, which suggests genome scan using an identity trait kernel offer quick data analysis to have an idea for further analysis.

Our algorithm is not free from limitations. The model assumes to follow a joint multivariate normal distribution, i.e. the sum of two Kronecker products of covariance matrices, which requires fully observed genotype and phenotype data for efficient inference [8,14]. If genotype, phenotype data contain missing values, one possibly takes advantage of any software package containing imputation methods for genotype data [23,24], phenotype data [25]. The statistical inference relies on LOD scores and permutation test, which are better than using $-\log_{10} P$ from an asymptotic χ^2-distribution when the pattern of linkage distribution is complex or extensive. However, the computational burden of permutations will be notable when matrix dimensions of data get considerably large even working with multi-nodes. It is possible to develop an efficient threshold estimation scheme [26] for such data. Like other existing MLMM based algorithms, FlxQTL cannot fit more complex trait structure directly, for instance, multiple traits across multiple sites measured over years. One way to fit such complex data could be to juxtapose one trait across sites per year and build several large trait matrices for analyses selecting contrasts or B-spline basis functions to account for sites. This would eventually require a great deal of physical memory and intensive computation. Possibly one can extend the model to a multi-linear model which performs multi-

Figure 1: FlxQTL model structure.

Figure 2: Comparison of computation time between FlxQTL and GEMMA (v.0.98.3) using Mouse_HS1940 data (n=1940 individuals, p=10783 SNPs) and simulated data. Version information: OS: Linux Debian 4.19.37-5, CPU: Intel(R) Xeon(R) CPU E5-2630 2.4GHz, Software: FlxQTL (Julia v.1.5.3, adding 32 cores), GEMMA (v.0.98.3, max 32 threads)

way array (tensor) decompositions for dimension reduction. With the growth of more complex and high-throughput phenotyping, we expect the use of our models to help shed light on the genetic control of structured multivariate traits.

METHODS

Methods and any associated references are available in the online version of the paper.

Acknowledgements Sen's funding Grant: R01 GM123489 (NIGMS), NIDA (P30DA044223). John Lovell's funding grant: The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No DE-AC02-05CH11231. John Mckay's funding grant: NSF Awards DEB 1022196 and 1556262 to JKM. We thank Jon Ågren and Grey Monroe for providing climate data from the *Arabidopsis thaliana* reciprocal transplant experiments. Juenger's funding grant: This research was supported by the U.S. Department of Energy (DOE) Award DESC0014156 and by National Science PGRP Awards IOS0922457 and IOS1444533 to T.E.J. *A.thaliana* fitness: https://datadryad.org/stash/downloads/file_stream/72691 F2 intercrosses between

Gough Island mice and WSB/EiJ: https://phenome.jax.org/projects/Payseur1

Competing Financial Interests The authors declare that they have no competing financial interests.

Author Contributions

ONLINE METHODS

Code availability. A Julia software implementation of FlxQTL is available at https://github.com/hkim89/FlxQTL.jl.

FlxQTL. Let $Y_{n\times m}$ be a trait for n individuals across m environments (or over m time points), or phenotypes of n individuals for m traits. $X_{n\times p}$ is p genotypes including the intercept or genotype probabilities of a marker to be tested without (or with) covariates, $Z_{m\times q}$ is q trait covariates such as treatments (or contrasts), (orthonormal) basis functions, etc. [1] FlxQTL is then constructed as follows:

$$Y = XBZ' + R + E. \tag{2}$$

One can observe that Z reduces the size of B to $p \times q$ compared with the standard model, where $Z = I_m$, $q = m$. This dimension reduction is easily discovered ($q << m$) when working with higher dimensional trait data, so as to resolve degrees-of-freedom problems for hypothesis testing. Two independent random effects follow multivariate normal distributions, such that

$$R \sim MVN(0, \tau^2 K_C, K_G), \quad E \sim MVN(0, \Sigma, I_n). \tag{3}$$

In the first variance component, K_G is a genetic relatedness (kinship) matrix between individuals computed by background genetic markers. K_C is a trait kernel generated by extra information associated with the traits; for METs, the trait kernel can be a climatic relatedness matrix using high-dimensional environment (site) information such as daily minimun, or maximum temperature, precipitation, etc. unless independence across sites (or environments) is assumed ($K_C = I$). The Kronecker product of the two relatedness matrices produces multivariate random GxE effects. τ^2 is a scaler by parameterization of a $m \times m$ covariance matrix proportional to K_C for dimension reduction. It can be particularly defined as a genetic variance in plasticity or GxE for METs. Σ is unstructured trait (or environment) errors, so that the second variance component assumes correlations among traits but iid noise between individuals. The two relatedness matrices are precomputed from corresponding datasets, and only τ^2 and Σ are estimated. The multi-level GxE random structure is verified in **Supplementary Note**. The MLMM in FlxQTL is then modeled by multivariate normal distribution whose covariance matrix is the sum of two Kronecker products for two independent random effects:

$$Y^v \sim N((X \otimes Z)B^v, \tau^2 K_C \otimes K_G + \Sigma \otimes I_n). \tag{4}$$

LOD scores ($-\log_{10}(LRT)$) are calculated by log-likelihood difference between the full model and null. At least 1000 permutations are implemented to obtain empirical LOD thresholds from the null distribution of maximal LOD scores in following steps: transforming the correlated trait to the iid by dividing corresponding individual-wise variance-covariance matrices, shuffling the iid trait by individuals, transforming back to the correlated trait, and then scanning genetic markers to pick up a maximal LOD score for each permutation. Note that LOD scores using trait covariates

(Z) are lower than those implemented by conventional MLMMs (i.e. $Z = I$), but the cutoff will be lower accordingly by the permutation test.

Computational costs. When it comes to algorithmic implementation, the eigen decomposition is applied to the two relatedness matrices only once for computational efficiency, which requires $O(n^3 + m^3) \approx O(n^3)$, $(m << n)$ as existing MLMMs. A LOCO (Leave One Chromosome Out) scheme [7,9,27] also works well because $O(dn^3) \approx O(n^3)$ $(d << n)$ for d chromosomes in genotype data, but in practice, genome scan with the LOCO scheme takes longer depending on types of data (about 30 % longer) than that without LOCO. For $t_0 >> t$, where t_0, t are the maximal number of iterations of ECM and ECM embedded in Speed restarting Nesterov's scheme, respectively, the computational complexity of two step implementation, which runs ECM for finding 'good' initial values followed by ECM embedded in Speed restarting Nesterov's scheme, is $O((c_1qn^2 + (m^3 + c_1m^2)n)(t_0 + t))$ per marker for n individuals, c_1 covariates. The whole time complexity including the eigen decomposition and rotation by corresponding eigen vectors is now $O(n^3 + (m + c_1)n^2 + s(c_1qn^2 + (m^3 + c_1m^2)n)(t_0 + t))$ with s markers.

Julia implementation. FlxQTL is implemented in Julia [28], a high-performance programming language, which enables users to add as many (even virtual) cores as possible for additional computational speed. The users take advantage of a high-performance computing (HPC) system to achieve a significant increase in speed or/and to do genome scan with millions of markers, so that practical high-dimensional data analysis is made possible. For faster genome scan for sufficiently numerous markers, one can add many cores in multi-nodes such as the HPC system when launching Julia. Julia then distributes each core to a marker to do genome scan independently. One can also em-

ploy MKL.jl to easily switch the BLAS library to the Intel MKL library for faster runtime on Intel hardware.

Data processing and analysis. *Mouse data.* For mouse_HS1940 data, SNPs were filtered out by minor allele frequency (MAF) less than 2% since we found GEMMA automatically processed the data with that criterion. For a fair comparison, we used a centered genetic relatedness matrix computed by GEMMA, but it needed some adjustment by adding very small value $(0.00001 * I)$ to make sure the matrix should be positive definite. The trait data consisted of 3 traits with missing values and was imputed using R-mice package with a 'pmm' option (predicted mean method). To compare the computation time between FlxQTL and GEMMA with regard to the number of traits and SNPs, we suffled the trait data individual-wise preserving the correlation among the three traits for non-singularity, placed the shuffled data next to the trait data continuously, and produced 6,9, and 12 simulated traits, respectively. We also generated 19 cumulative genotype datasets from the mouse data, that is, starting from a chromosome 1 SNP dataset (950 SNPs), building up a new dataset with 4-chromosome increments and reaching the whole SNP dataset (10783 SNPs). The median time of genome scan without LOCO for 64 replications was measured for each SNP dataset per simulated trait data including the actual 3 trait data.

For F2 intercrosses between Gough Island mice and WSB/EiJ, we computed genotype probabilities, excluding sex chromosomes, from R/qtl for minimal loss of information in QTL analysis adding pseudo markers in every 1cM between two markers. Traits were imputed using R-mice package with a 'pmm' option and then were standardized by overall mean and standard deviation because of their time-valued measures. We used a 4 degrees-of-freedom cubic spline from

R-splines package to produce trait covariates (Z) to capture the trend of weekly body weights and corresponding weight growth rate simultaneously. Here, we set the climatic relatedness to be an identity matrix since GxE interactions are not appropriate in this case and computed a linear genetic relatedness matrix from genotype probabilities due to its statistical interpretability. The thresholds at $\alpha = 0.1, 0.05$ were estimated by 1000 permutations. The result from 1D-multivariate genome scan by FlxQTL was overall consistent with the published one by the univariate linear model based R/qtl, except for Chromosome 6 and 7 perhaps due to polygenic effects. Since a SNP in Chromosome 7 was known to be crucial, we implemented 2D-genome scan for Chromosome 7, as well as Chromosome 8, 10 for comparison. FlxQTL detected several more QTL than R/qtl did despite the former requiring further analysis such as multiple-QTL analysis to finalize them.

Arabidopsis thaliana data. We employed R/qtl to add pseudo markers in every 1cM between two markers and dropped one of two consecutive markers when they are identical. Phenotypes of a trait were imputed using a 'pmm' option in R-mice package and were standardized by columns since *A. thaliana* is an annual species. The data comprises 699 markers (or SNPs) across 5 chromosomes by 400 lines (or individuals) for 6 traits. A matrix (Z) sets to be trait (orthonormal) contrasts, which consist of overall site (1's) and site difference (1's for Sweden and -1's for Italy), so that multiplying the site difference by QTL yields QTL$\times$site or GxE interactions. We computed both a genetic and a climatic relatedness matrices using Manhattan distance; thus, subtracting a marker-distance mean between two individuals from 1 yields a probability of genetic distance between two individuals. The climatic relatedness matrix was computed in a similar fashion, using daily range soil or air temperature, precipitation, etc. However, the difference between with and

without climatic relatedness matrix appears to be unnoticeable in this analysis as the simulation study showed (**Supplementary Fig. 8**). We performed 1-, 2-dimensional QTL analyses with a LOCO option followed by multiple QTL analysis, i.e. a stepwise model selection approach by forward selection and backward elimination adding or dropping one or two SNPs for each scan, respectively. The 95% of cutoff was estimated by permutation test and was used for a penalty term in the stepwise model selection since the false positive rate is maintained at the rate of α in the case of no QTL and with the search restricted to models with no more than one QTL [29]. Our result was by and large consistent and revealed one more significant candidate QTL in three chromosomes each but one less in two chromosomes each when comparing with the existing result [22].

Switchgrass (Panicum virgatum) data. Genotype probabilities were computed in every 1cM between two markers via R/qtl. Traits and climatic information data had no missing values, so that the whole dataset comprises 6118 genetic markers by 750 individuals for 36 quantitative traits of flowering time as a combination of 10 latitudinal sites and 4 years from 2016 to 2019, as well as a matrix of 12 by 36-monthly photoperiod. Three siginificant QTL (two in Chromosome 5N and one in Chromosome 4K) were selected by 1D-genome scan with a LOCO option followed by environmental scan conditional on each major QTL. The environmental scan was done by regressing each 36-monthly photoperiod under the alternative of existing an environmental factor that affects a QTL following the null scan and by computing LOD scores. For computational efficiency, traits and monthly photoperiod data were centered by their column means and scaled by overall standard deviations of the centered data. The genome scan took about 3 min 47.16 sec (Julia v.1.5.3. with 32 cores, OS: Linux Debian 4.19.37-5, CPU: Intel(R) Xeon(R) CPU E5-2630@2.4GHz), and the

environmental scan did about 12 sec (12 cores).

Simulation study. We generated simulated data sets of a quantitative trait for 3 years and 2 sites (or 6 quantitative traits) by simulating multivariate random GxE effects based on genotype data and annotated daily soil minimum and maximum data from the Arabidopsis data [22]. We rearranged soil daily range temperatures (difference between maximum and minimum temperatures) into a 3 year by 2 location combination, i.e. annotated information of 365 days by 6 environments, and standardized them by overall mean and standard deviation to keep their inherent information unchanged. To generate null trait data sets, we selected to standardize about 3.5% of genetic markers (about 25 markers of total 699 markers) and 8.2% of 365 soil daily range temperatures (about 30 days) to avoid sigularity, both of which constitute multivariate GxE random effects while τ^2 varies from small to large. The remaining errors indicating unmeasured environmental errors and iid noise were also added to the null trait data. Both random and error terms were respectively standardized. To add large effect QTL to the null data, we sampled one genetic marker apart from the selected many small effect QTL and generated large effect QTL multiplied by fixed effect sizes that vary from small to large and a (orthonormal) contrast matrix (Z) mentioned in the Data processing and analysis and **Supplementary Note**. Here we only considered one large QTL effect fixed when simulating data sets in order to eliminate complications that may arise for including an additional QTL. We carried out 1000 simulations under the null to attain a threshold at $\alpha = 0.05$ from the distribution of maximum LOD scores and measured power under the alternative for each instance given the climatic relatedness. The first simulation study to assess power for the effect of climatic relatedness varied with values of τ^2 for each fixed effect size matrix (B), which in

total resulted in 100,000 experiments. We then narrowed down to the feasible range of fixed effect sizes by excluding the effect sizes producing zeros and ones of power. Within that range, we ran simulations to establish a comparison between inclusion and exclusion of the climatic relatedness matrix (K_C) changing τ^2. Since corresponding results showed slight distinctions in power, we tackled unusual scenarios in genome scan, where K_C's are an autocorrelated matrix, an extreme case of positive definite matrix, etc. The results revealed analogous power to those shown in prior experiments. The second simulation study to assess the effect of trait covariates (Z) was implemented in the similar manner. In the given range of fixed effect sizes, we compared three cases: contrasts (FlxQTL), an identity matrix ($Z = I$ in FlxQTL), Julia-version MLMM. Here, the Julia-version MLMM was coded as an alternative of GEMMA to use as a reference for no trait covariate and kernel ($Z = K_C = I$) because both run by different algorithms under the same model while yielding very similar results. It is also easier to draw a comparison with the Julia-version MLMM than to do with GEMMA since GEMMA uses different data formats.

1. Xiong, H. *et al.* A flexible estimating equations approach for mapping function-valued traits. *Genetics* **189**, 305–316 (2011).

2. Gray, M. M. *et al.* Genetics of rapid and extreme size evolution in island mice. *Genetics* **201**, 213–228 (2015).

3. Garin, V., Malosetti, M. & van Eeuwijk, F. Multi-parent multi-environment qtl analysis: an illustration with the eu-nam flint population. *TAG. Theoretical and Applied genetics. Theoretische und Angewandte Genetik* (2020).

4. Smith, A., Cullis, B. & Thompson, R. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* **57**, 1138–1147 (2001).

5. Smith, A. B., Cullis, B. R. & Thompson, R. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science* **143**, 449–462 (2005).

6. Cullis, B. R., Smith, A. B., Beeck, C. P. & Cowling, W. A. Analysis of yield and oil from a series of canola breeding trials. part ii. exploring variety by environment interaction using factor analysis. *Genome* **53**, 1002–1016 (2010).

7. Lippert, C. *et al.* Fast linear mixed models for genome-wide association studies. *Nature methods* **8**, 833 (2011).

8. Casale, F. P., Rakitsch, B., Lippert, C. & Stegle, O. Efficient set tests for the genetic analysis of correlated traits. *Nature methods* **12**, 755–758 (2015).

9. Loh, P.-R. *et al.* Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics* **47**, 284 (2015).

10. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics* **11**, 459–463 (2010).

11. Kang, H. M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nature genetics* **42**, 348 (2010).

12. Yu, J. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics* **38**, 203–208 (2006).

13. Zhang, Z. *et al.* Mixed linear model approach adapted for genome-wide association studies. *Nature genetics* **42**, 355 (2010).

14. Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature methods* **11**, 407–409 (2014).

15. Korte, A. *et al.* A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature genetics* **44**, 1066 (2012).

16. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. Gcta: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics* **88**, 76–82 (2011).

17. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).

18. Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. & Thompson, R. Asreml user guide release 1.0 (2002).

19. Meyer, K. Wombata tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (reml). *Journal of Zhejiang University Science B* **8**, 815–821 (2007).

20. Su, W., Boyd, S. & Candes, E. A differential equation for modeling Nesterov's accelerated gradient method: Theory and Insights. In *Advances in Neural Information Processing Systems*, 2510–2518 (2014).

21. Meng, X. L. & Rubin, D. B. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **80**, 267–278 (1993).

22. Ågren, J., Oakley, C. G., McKay, J. K., Lovell, J. T. & Schemske, D. W. Genetic mapping of adaptation reveals fitness tradeoffs in arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **110**, 21077–21082 (2013).

23. Broman, K. W. & Sen, Ś. *A Guide to QTL Mapping with R/qtl*, vol. 46 (Springer, 2009).

24. Broman, K. W. *et al.* R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics* **211**, 495–502 (2019).

25. Buuren, S. v. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in r. *Journal of statistical software* 1–68 (2010).

26. Feingold, E., Brown, P. O. & Siegmund, D. Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. *American journal of human genetics* **53**, 234 (1993).

27. Listgarten, J. *et al.* Improved linear mixed models for genome-wide association studies. *Nature methods* **9**, 525 (2012).

28. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM review* **59**, 65–98 (2017).

29. Manichaikul, A., Moon, J.-Y., Sen, Ś., Yandell, B. S. & Broman, K. W. A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. *Genetics* **181**, 1077–1086 (2009).

30. Henderson, C. R. Applications of linear models in animal breeding. *University of Guelph Press, Guelph* **11**, 652–653 (1984).

31. Ledoit, O. & Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis* **88**, 365–411 (2004).

32. Schäfer, J. & Strimmer, K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical applications in genetics and molecular biology* **4** (2005).

1 Supplementary Note: FlxQTL for gxe detection

Condsider a multivariate linear mixed model [30],

$$Y_0 = X_0 B Z_0' + R_0 + E_0, \tag{5}$$

or its vectorized form,

$$Y_0^v = (Z_0 \otimes X_0) B^v + R_0^v + E_0^v,$$

where

$$R_0^v \sim MVN(0, \tau^2 K_C \otimes K_G), \quad E_0^v \sim MVN(0, \Sigma_0 \otimes I_n). \tag{6}$$

Y_0 is a matrix of $n \times m$ multivariate reponses (phenotypes or a trait in multiple environments), where n is the number of individuals(or lines), m is the number of environments(or sites). X_0 is $n \times p$ genotypes (or genotype probabilities) for a candidate marker including intercepts and/or covariates. Z_0 is $m \times q$ fixed column covariates in regard with environments (or sites). B is $p \times q$ fixed effects including main and interaction effects. R_0 is $n \times m$ gene by environment random effects. E_0 is $n \times m$ genetic variability among individuals that is not explained by R and residual errors in which I_n, Σ_0 are respective $n \times n$ indentity and $m \times m$ matrices. τ^2 is an overall genetic variance among environments. K_G is a $n \times n$ genetic background (or a genetic kinship) matrix, and K_C is a $m \times m$ environmental background (or weather profiles) matrix. Examples of C could be air, soil temperatures, water availiability (or draught) through growing season profile, or soil microbial composition. Note that we confine the relatedness matrices to full-rank and symmetric positive definite matrices. For a semi-positive definite relatedness matrix, one could take advantage of a Shrinkage approach [31,32] to force it to be positive definite without loss of information. Thus, $K_G = U_G \Lambda_G U_G'$ and $K_C = U_C \Lambda_C U_C'$. Multiplying $U_C' \otimes U_G'$ to the model yields

$$Y = XBZ' \quad + \quad R + E, \tag{7}$$

$$R^v \sim MVN(0, \tau^2 \Lambda_C \otimes \Lambda_G), \qquad E^v \sim MVN(0, \Sigma \otimes I_n), \tag{8}$$

where $Y_{n\times m} = [\mathbf{y}_1', \mathbf{y}_2', \cdots, \mathbf{y}_n']'$, $X_{n\times p} = [\mathbf{x}_1', \mathbf{x}_2', \cdots, \mathbf{x}_n']'$, $R_{n\times m} = [\mathbf{r}_1', \mathbf{r}_2', \cdots, \mathbf{r}_n']'$, and $E_{n\times m} = [\mathbf{e}_1', \mathbf{e}_2', \cdots, \mathbf{e}_n']'$.

The corresponding individual vector is now,

$$\mathbf{y}_i = ZB'\mathbf{x}_i + \mathbf{r}_i + \mathbf{e}_i, \;\; \mathbf{r}_i \sim N(0, \tau^2\lambda_i\Lambda_C), \;\; \mathbf{e}_i \sim N(0, \Sigma), \tag{9}$$

and its variance is $V_i = \tau^2\lambda_i\Lambda_C + \Sigma$ $(i = 1, \ldots, n)$.

Verification of infinitesimal gxe random terms Assume that a gene by environment random term has a form of $R_0 = GB_{rand}C'$, where $C_{m\times q_1}$ is a q_1 random climate (or weather) factor matrix, $G_{n\times p_1}$ is a p_1 background genotype matrix, and $B^v_{rand} \sim MVN(0, \tau^2 I_{q_1} \otimes I_{p_1})$. Then, the distribution of the gene by envrionment random effects in (6) is obtained in the following way.

$$\begin{aligned}
E(R_0^v R_0^{v\prime}) &= E((C \otimes G)B^v_{rand}{B^v_{rand}}'(C' \otimes G')) \\
&= (C \otimes G)E(B^v_{rand}{B^v_{rand}}')(C' \otimes G') \\
&= (C \otimes G)(\tau^2 I_{q_1} \otimes I_{p_1})(C' \otimes G') \\
&= \tau^2(CC' \otimes GG') \\
&= \tau^2 K_C \otimes K_G,
\end{aligned}$$

where C and G are scaled by $\sqrt{m}$, $\sqrt{n}$, respectively.

Expectation Conditional Maximization (ECM) algorithm In the ECM algorithm [21], the joint log-likelihood function, assuming the random term R to be missing, is

$$\begin{aligned}
& l(Y, R|B, \tau^2, \Sigma) \\
= \; & \sum_{i=1}^{n}\{-m\log(2\pi) - \frac{m}{2}\log\tau^2 - \frac{m}{2}\log(\lambda_i) - \frac{1}{2}\log|\Lambda_C| - \frac{1}{2}\log|\Sigma| \\
& -\frac{1}{2}\mathbf{e}_i'\Sigma^{-1}\mathbf{e}_i - \frac{1}{2}\mathbf{r}_i'(\tau^2\lambda_i\Lambda_C)^{-1}\mathbf{r}_i\}.
\end{aligned} \tag{10}$$

In the expectation step, the conditional distribution of $\mathbf{r}_i$ given $\mathbf{y}_i$ ($i = 1, \cdots, n$), $B^{(t)}$,$\tau^{2(t)}$, and $\Sigma^{(t)}$ at the current iteration t is [30]

$$\mathbf{r}_i|\mathbf{y}_i, B^{(t)}, \tau^{2(t)}, \Sigma^{(t)} \sim MVN(\hat{\mathbf{r}}_i^{(t)}, \hat{\Theta}_i^{(t)}), \tag{11}$$

where

$$\begin{aligned} V_i^{(t)} &= \tau^{2(t)}\lambda_i\Lambda_C + \Sigma^{(t)}, \quad \hat{\mathbf{r}}_i = \tau^{2(t)}\lambda_i\Lambda_C(V_i^{(t)})^{-1}(\mathbf{y}_i - ZB^{(t)\prime}\mathbf{x}_i), \\ \hat{\Theta}_i^{(t)} &= \tau^{2(t)}\lambda_i\Lambda_C - \tau^{2(t)}\lambda_i\Lambda_C(V_i^{(t)})^{-1}\tau^{2(t)}\lambda_i\Lambda_C. \end{aligned} \tag{12}$$

The expectation of the loglikelihood function in terms of R given Y, $B^{(t)}$,$\tau^{2(t)}$, and $\Sigma^{(t)}$ is obtained as follows.

$$\begin{aligned} & E\left(l(Y, R|B, \tau^2, \Sigma)|R|Y, B^{(t)}, \tau^{2(t)}, \Sigma^{(t)}\right) \\ = & \sum_{i=1}^{n}\{-m\log(2\pi) - \frac{m}{2}\log\tau^2 - \frac{m}{2}\log(\lambda_i) - \frac{1}{2}\log|\Lambda_C| - \frac{1}{2}\log|\Sigma| \\ & -\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) - \frac{1}{2}\hat{\mathbf{r}}_i^{(t)\prime}(\tau^{-2}\lambda_i^{-1}\Lambda_C^{-1} + \Sigma^{-1})\hat{\mathbf{r}}_i^{(t)} \\ & +\hat{\mathbf{r}}_i^{(t)\prime}\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) - \frac{1}{2}tr((\tau^{-2}\lambda_i^{-1}\Lambda_C^{-1} + \Sigma^{-1})\hat{\Theta}_i^{(t)})\}. \end{aligned} \tag{13}$$

In the conditional maximization step, one can update $B^{(t+1)}$ conditional on $\tau^{2(t)}$ and $\Sigma^{(t)}$, followed by updating $\{\tau^{2(t+1)}, \Sigma^{(t+1)}\}$ conditional on $B^{(t+1)}$, $\tau^{2(t)}$, and $\Sigma^{(t)}$ [14].

$$B^{(t+1)} = (X'X)^{-1}X'(Y - \hat{R}^{(t)})\Sigma^{-1}Z(Z'\Sigma^{-1}Z)^{-1}, \tag{14}$$

$$\tau^{2(t+1)} = \frac{1}{nm}\sum_{i=1}^{n} tr(\lambda_i^{-1}\Lambda_C^{-1}(\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)\prime} + \hat{\Theta}_i^{(t)})), \tag{15}$$

$$\Sigma^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mathbf{e}}_i^{(t)}\hat{\mathbf{e}}_i^{(t)\prime} + \hat{\Theta}_i^{(t)}), \tag{16}$$

$$\text{where} \quad \hat{\mathbf{e}}_i^{(t)} = \mathbf{y}_i - ZB^{(t+1)\prime}\mathbf{x}_i - \hat{\mathbf{r}}_i^{(t)}. \tag{17}$$

Log-likelihood function In the expectation step, the conditional distribution of $\mathbf{r}_i$ given $\mathbf{y}_i$ ($i = 1, \cdots, n$), $B^{(t)}$,$\tau^{2(t)}$, and $\Sigma^{(t)}$ at the current iteration t is [30]

$$
\begin{aligned}
& l(Y, R|B, \tau^2, \Sigma) = -nm\log(2\pi) \\
& -\frac{1}{2}\log|(U_C \otimes U_G)(\tau^2\Lambda_C \otimes \Lambda_G)(U_C' \otimes U_G')| - \frac{1}{2}\log|(U_C \otimes U_G)(\Sigma \otimes I_n)(U_C' \otimes U_G')| \\
& -\frac{1}{2}(Y_0 - X_0BZ_0' - R_0)^{v\prime}((U_C \otimes U_G)(\Sigma \otimes I_n)(U_C' \otimes U_G'))^{-1}(Y_0 - X_0BZ_0' - R_0)^v \\
& -\frac{1}{2}R_0^{v\prime}((U_C \otimes U_G)(\tau^2\Lambda_C \otimes \Lambda_G)(U_C' \otimes U_G'))^{-1}R_0^v \\
= \; & -nm\log(2\pi) - \frac{nm}{2}\log(\tau^2) - \frac{m}{2}\log|\Lambda_G| - \frac{n}{2}\log|\Lambda_C| - \frac{1}{2}\log|\Sigma \otimes I_n| \\
& +2\log|(U_C' \otimes U_G')| - \frac{1}{2}(Y - XBZ' - R)^{v\prime}(\Sigma \otimes I_n)^{-1}(Y - XBZ' - R)^v \\
& -\frac{1}{2}R^{v\prime}(\tau^2\Lambda_C \otimes \Lambda_G)^{-1}R^v \\
= \; & \sum_{i=1}^{n}\{-m\log(2\pi) - \frac{m}{2}\log\tau^2 - \frac{m}{2}\log(\lambda_i) - \frac{1}{2}\log|\Lambda_C| - \frac{1}{2}\log|\Sigma| \\
& -\frac{1}{2}\mathbf{e}_i'\Sigma^{-1}\mathbf{e}_i - \frac{1}{2}\mathbf{r}_i'(\tau^2\lambda_i\Lambda_C)^{-1}\mathbf{r}_i\}. \qquad (18)
\end{aligned}
$$

Derivation of conditional expected log-likelihood and its closed form solutions To obtain closed-form solutions of $B^{(t+1)}$, $\tau^{2(t+1)}$, and $\Sigma^{(t+1)}$ to the conditional expected loglikelihood in (13), one differentiates (13) with respect to each parameter after integrating out $\mathbf{r}_i$ ($i = 1, \ldots n$) in the joint log-likelihood function, (10).

In (10), taking expectation with respect to $\mathbf{r}_i$ ($i = 1, \cdots, n$) yields

$$
\begin{aligned}
El \; = \; & E\left(l(Y, R|B, \tau^2, \Sigma)|R|Y, B^{(t)}, \tau^{2(t)}, \Sigma^{(t)}\right) \\
= \; & \sum_{i=1}^{n}\{-m\log(2\pi) - \frac{m}{2}\log\tau^2 - \frac{m}{2}\log(\lambda_i) - \frac{1}{2}\log|\Lambda_C| - \frac{1}{2}\log|\Sigma| \\
& -\frac{1}{2}E\left((\mathbf{y}_i - ZB'\mathbf{x}_i - \mathbf{r}_i)'\Sigma^{-1}((\mathbf{y}_i - ZB'\mathbf{x}_i - \mathbf{r}_i)|R\right) - \frac{1}{2}E\left(\mathbf{r}_i'(\tau^2\lambda_i\Lambda_C)^{-1}\mathbf{r}_i|R\right)\}.
\end{aligned}
$$

For simplicity, let $E(\cdot|R|Y, B^{(t)}, \tau^{2(t)}, \Sigma^{(t)}) = E(\cdot|R)$. The last two terms then can be simplified as follows:

$$
\begin{aligned}
& \sum_{i=1}^{n}\{E(-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + \mathbf{r}_i'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) \\
& -\frac{1}{2}\mathbf{r}_i'\Sigma^{-1}\mathbf{r}_i - \frac{1}{2}\mathbf{r}_i'(\tau^2\lambda_i\Lambda_C)^{-1}\mathbf{r}_i|R)\} \\
= & \sum_{i=1}^{n}\{-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + E(\mathbf{r}_i'|R)\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) \\
& -\frac{1}{2}E(\mathbf{r}_i'(\Sigma^{-1} + (\tau^2\lambda_i\Lambda_C)^{-1})\mathbf{r}_i|R)\} \\
= & \sum_{i=1}^{n}\{-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + \hat{\mathbf{r}}_i^{(t)'}\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) \\
& -\frac{1}{2}tr((\Sigma^{-1} + (\tau^2\lambda_i\Lambda_C)^{-1})E(\mathbf{r}\mathbf{r}'|R))\}. \\
= & \sum_{i=1}^{n}\{-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + \hat{\mathbf{r}}_i^{(t)'}\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) \\
& -\frac{1}{2}tr((\Sigma^{-1} + (\tau^2\lambda_i\Lambda_C)^{-1})(\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)'} + \hat{\Theta}_i^{(t)}))\} \\
= & \sum_{i=1}^{n}\{-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + \hat{\mathbf{r}}_i^{(t)'}\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) \\
& -\frac{1}{2}tr((\Sigma^{-1} + (\tau^2\lambda_i\Lambda_C)^{-1})\hat{\Theta}_i^{(t)}) - \frac{1}{2}\hat{\mathbf{r}}_i^{(t)'}(\Sigma^{-1} + (\tau^2\lambda_i\Lambda_C)^{-1})(\hat{\mathbf{r}}_i^{(t)})\}.
\end{aligned}
$$

Now one can find closed-form solutions of B, τ^2, and Σ to the expected log-likelihood function using properties of the trace, calculus for matrices. For the solution for B,

$$
\begin{aligned}
El &= \sum_{i=1}^{n}\{-\frac{1}{2}(\mathbf{y}_i - ZB'\mathbf{x}_i)'\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i) + \hat{\mathbf{r}}_i^{(t)'}\Sigma^{-1}(\mathbf{y}_i - ZB'\mathbf{x}_i)\} + C, \\
&= \sum_{i=1}^{n}\{-\frac{1}{2}(\Sigma^{-\frac{1}{2}}\mathbf{y}_i - \Sigma^{-\frac{1}{2}}ZB'\mathbf{x}_i)'(\Sigma^{-\frac{1}{2}}\mathbf{y}_i - \Sigma^{-\frac{1}{2}}ZB'\mathbf{x}_i) \\
& +(\Sigma^{-\frac{1}{2}}\hat{\mathbf{r}}_i^{(t)})'(\Sigma^{-\frac{1}{2}}\mathbf{y}_i - \Sigma^{-\frac{1}{2}}ZB'\mathbf{x}_i)\} + C,
\end{aligned}
$$

where C is a constant.

$$\frac{\partial El}{\partial B'^v} = \sum_{i=1}^{n}((\mathbf{x}_i' \otimes \Sigma^{-\frac{1}{2}}Z)'(\Sigma^{-\frac{1}{2}}\mathbf{y}_i - (\mathbf{x}_i' \otimes \Sigma^{-\frac{1}{2}}Z)B'^v) - (\mathbf{x}_i' \otimes \Sigma^{-\frac{1}{2}}Z)'\Sigma^{-\frac{1}{2}}\hat{\mathbf{r}}_i^{(t)}) = 0$$

$$\sum_{i=1}^{n}(\mathbf{x}_i\mathbf{x}_i' \otimes Z'\Sigma^{-1}Z)B'^v = \sum_{i=1}^{n}(\mathbf{x}_i \otimes Z'\Sigma^{-\frac{1}{2}})(\Sigma^{-\frac{1}{2}}\mathbf{y}_i - \Sigma^{-\frac{1}{2}}\hat{\mathbf{r}}_i^{(t)})$$

$$B'^v = ((X'X)^{-1}X' \otimes (Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-\frac{1}{2}})(\Sigma^{-\frac{1}{2}}(Y' - \hat{R}^{(t)'}))^v$$

$$B' = (Z'\Sigma^{-1}Z)^{-1}Z'\Sigma^{-1}(Y' - \hat{R}^{(t)'})X(X'X)^{-1}$$

$$\therefore B = (X'X)^{-1}X'(Y - \hat{R}^{(t)})\Sigma^{-1}Z(Z'\Sigma^{-1}Z)^{-1}$$

$$\frac{\partial El}{\partial \tau^2} = \sum_{i=1}^{n}\{-\frac{m}{2}\tau^{-2} + \frac{1}{2}\hat{\mathbf{r}}_i^{(t)'}(\tau^{-4}\lambda_i^{-1}\Lambda_C^{-1})\hat{\mathbf{r}}_i^{(t)} + \frac{1}{2}tr(\tau^{-4}\lambda_i^{-1}\Lambda_C^{-1}\hat{\Theta}_i^{(t)})\} = 0 \qquad (19)$$

$$mn = \sum_{i=1}^{n} tr((\tau^{-2}\lambda_i^{-1}\Lambda_C^{-1})\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)'} + (\tau^{-2}\lambda_i^{-1}\Lambda_C^{-1})\hat{\Theta}_i^{(t)})$$

$$mn\tau^2 = tr(\sum_{i=1}^{n}\lambda_i^{-1}\Lambda_C^{-1}(\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)'} + \hat{\Theta}_i^{(t)}))$$

$$\therefore \tau^{2(t+1)} = \frac{1}{nm}\sum_{i=1}^{n} tr(\lambda_i^{-1}\Lambda_C^{-1}(\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)'} + \hat{\Theta}_i^{(t)}))$$

$$\begin{aligned}\frac{\partial El}{\partial \Sigma^v} &= \sum_{i=1}^{n}\{-\frac{1}{2}tr(\Sigma^{-1}\frac{\partial \Sigma}{\partial \Sigma^v}) + \frac{1}{2}(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i)'(\Sigma^{-1}\frac{\partial \Sigma}{\partial \Sigma^v}\Sigma^{-1})(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i) \\ &+ \hat{\mathbf{r}}_i^{(t)'}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \Sigma^v}\Sigma^{-1})\hat{\mathbf{r}}_i^{(t)} + \frac{1}{2}tr(\Sigma^{-1}\frac{\partial \Sigma}{\partial \Sigma^v}\Sigma^{-1}\hat{\Theta}_i^{(t)}) \\ &- \hat{\mathbf{r}}_i^{(t)'}(\Sigma^{-1}\frac{\partial \Sigma}{\partial \Sigma^v}\Sigma^{-1})(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i)\} = 0\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{n}\{tr(-I_m + \Sigma^{-1}(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i)(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i)' + \Sigma^{-1}\hat{\mathbf{r}}_i^{(t)}\hat{\mathbf{r}}_i^{(t)'} \\ +\Sigma^{-1}\hat{\Theta}_i^{(t)} - 2\Sigma^{-1}(\mathbf{y}_i - ZB^{(t+1)'}\mathbf{x}_i)\hat{\mathbf{r}}_i^{(t)'})\} = 0\end{aligned}$$

$$n \cdot tr(\Sigma) = \sum_{i=1}^{n} tr(\hat{\mathbf{e}}_i^{(t)} \hat{\mathbf{e}}_i^{(t)\prime} + \hat{\Theta}_i^{(t)})$$
$$\therefore \Sigma^{(t+1)} = \frac{1}{n} \sum_{i=1}^{n} (\hat{\mathbf{e}}_i^{(t)} \hat{\mathbf{e}}_i^{(t)\prime} + \hat{\Theta}_i^{(t)}),$$

where $\hat{\mathbf{e}}_i^{(t)} = \mathbf{y}_i - ZB^{(t+1)\prime}\mathbf{x}_i - \hat{\mathbf{r}}_i^{(t)}$.

ECM embeded in Nesterov's accelerated gradient with speed restart

Algorithm : ECM embeded in Speed Restarting Nesterov's Scheme [20]

input : $B^{(0)} \in \mathbb{R}^{q \times p}$, $\tau^{2(0)} \in \mathbb{R}$, $\Sigma^{(0)} \in \mathbb{R}^{m \times m}$, $B^{(-1)} = B^{(0)}$, $\tau^{2(-1)} = \tau^{2(0)}$, $\Sigma^{(-1)} = \Sigma^{(0)}$,

$t_{min} \in \mathbb{N}$

$j \leftarrow 1$

while $t = 1$, $\| B^{(t)} - B^{(t-1)} \| + \| \tau^{2(t)} - \tau^{2(t-1)} \| + \| \Sigma^{(t)} - \Sigma^{(t-1)} \| > \epsilon$

for small enough ϵ

E-step : Integrate out the unknown complete data loglikelihood by

$\mathbf{r}_i | \mathbf{y}_i, B^{(t-1)}, \tau^{2(t-1)}, \Sigma^{(t-1)} \sim MVN(\hat{\mathbf{r}}_i^{(t-1)}, \hat{\Theta}_i^{(t-1)})$ as in (11).

CM-step : Update each parameter by maximization over the full parameter space as in (14), (15), and (16).

Nesterov's Acceleration :

$$B^{(t)} \leftarrow B^{(t)} + \frac{j-1}{j+2}(B^{(t)} - B^{(t-1)})$$

$$\tau^{2(t)} \leftarrow \tau^{2(t)} + \frac{j-1}{j+2}(\tau^{2(t)} - \tau^{2(t-1)})$$
$$\Sigma^{(t)} \leftarrow \Sigma^{(t)} + \frac{j-1}{j+2}(\Sigma^{(t)} - \Sigma^{(t-1)})$$

if $\| B^{(t)} - B^{(t-1)} \| + \| \tau^{2(t)} - \tau^{2(t-1)} \| + \| \Sigma^{(t)} - \Sigma^{(t-1)} \| < \| B^{(t-1)} - B^{(t-2)} \| + \| \tau^{2(t-1)} - \tau^{2(t-2)} \| + \| \Sigma^{(t-1)} - \Sigma^{(t-2)} \|$ & $t \geq t_{min}$

then $j \leftarrow 1$

else $j \leftarrow j + 1$

end if

end while

2 Supplementary Figures

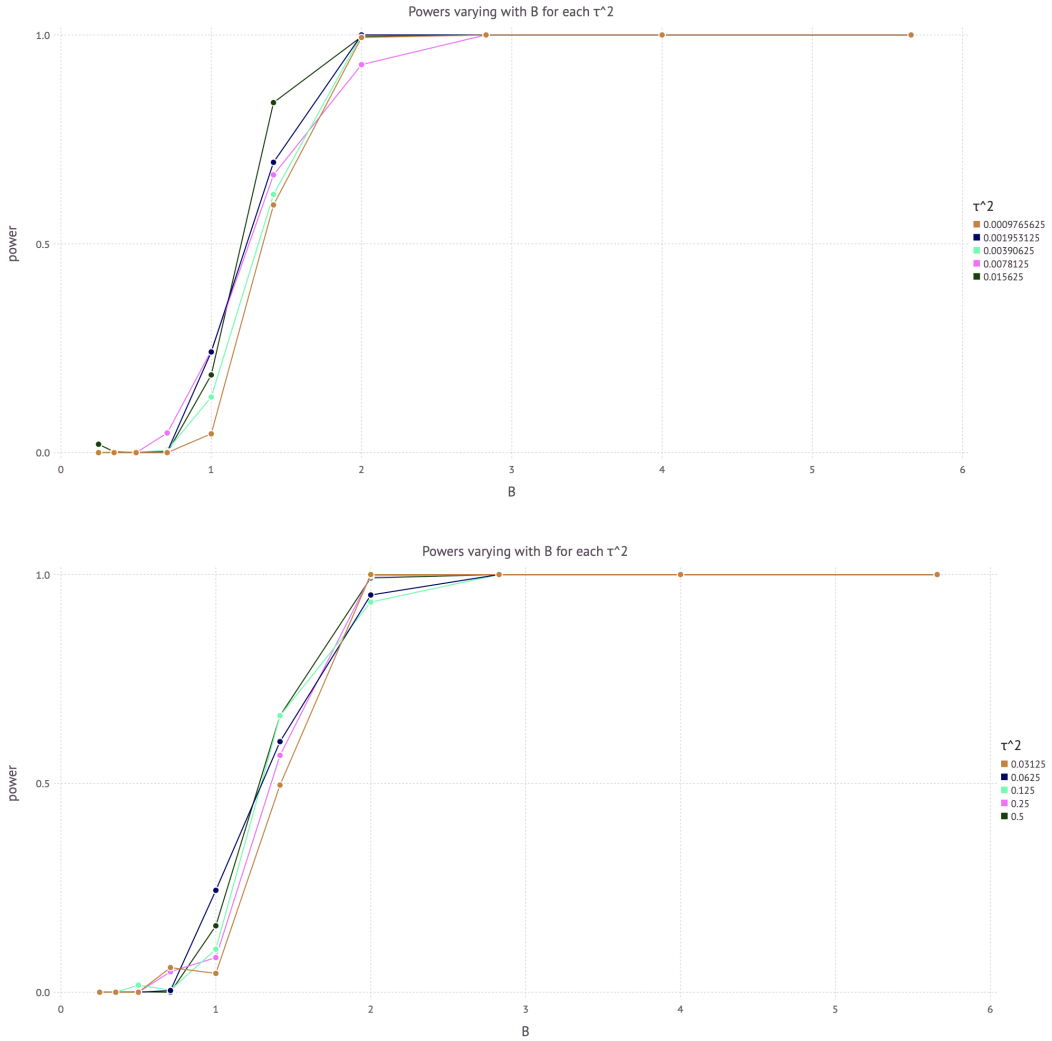


Figure 3: Power varied by τ^2 on simulated data using Arabidopsis thaliana and 3 year by 2 site soil daily range temperature data. Entries of B are randomly selected from $\{\pm 0.25\}$ multiplied by $\sqrt{2}$ raised to the power per τ^2.

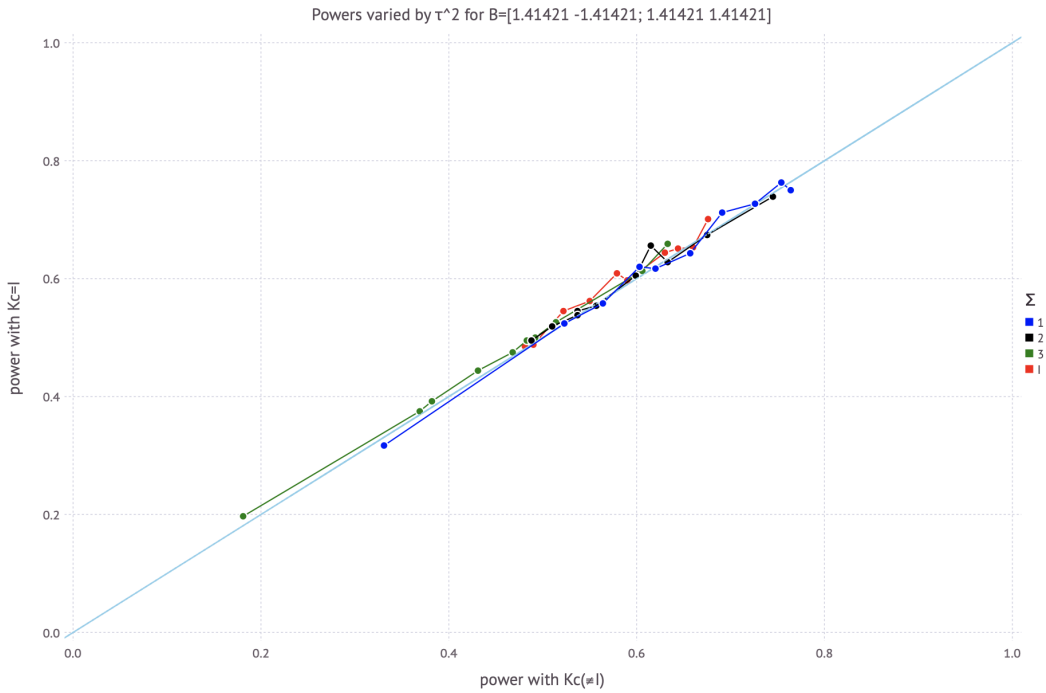


Figure 4: Power comparison of the existence of climatic relatedness (K_C) with its non-existence ($K_C = I$) on simulated data using Arabidopsis thaliana and 3 year by 2 site soil daily range temperature data, varying unstructured environment errors (Σ). Entries of B are randomly selected from $\{\pm 0.25\}$ multiplied by $\sqrt{2}$ raised to the power per τ^2.

3 Supplementary Tables

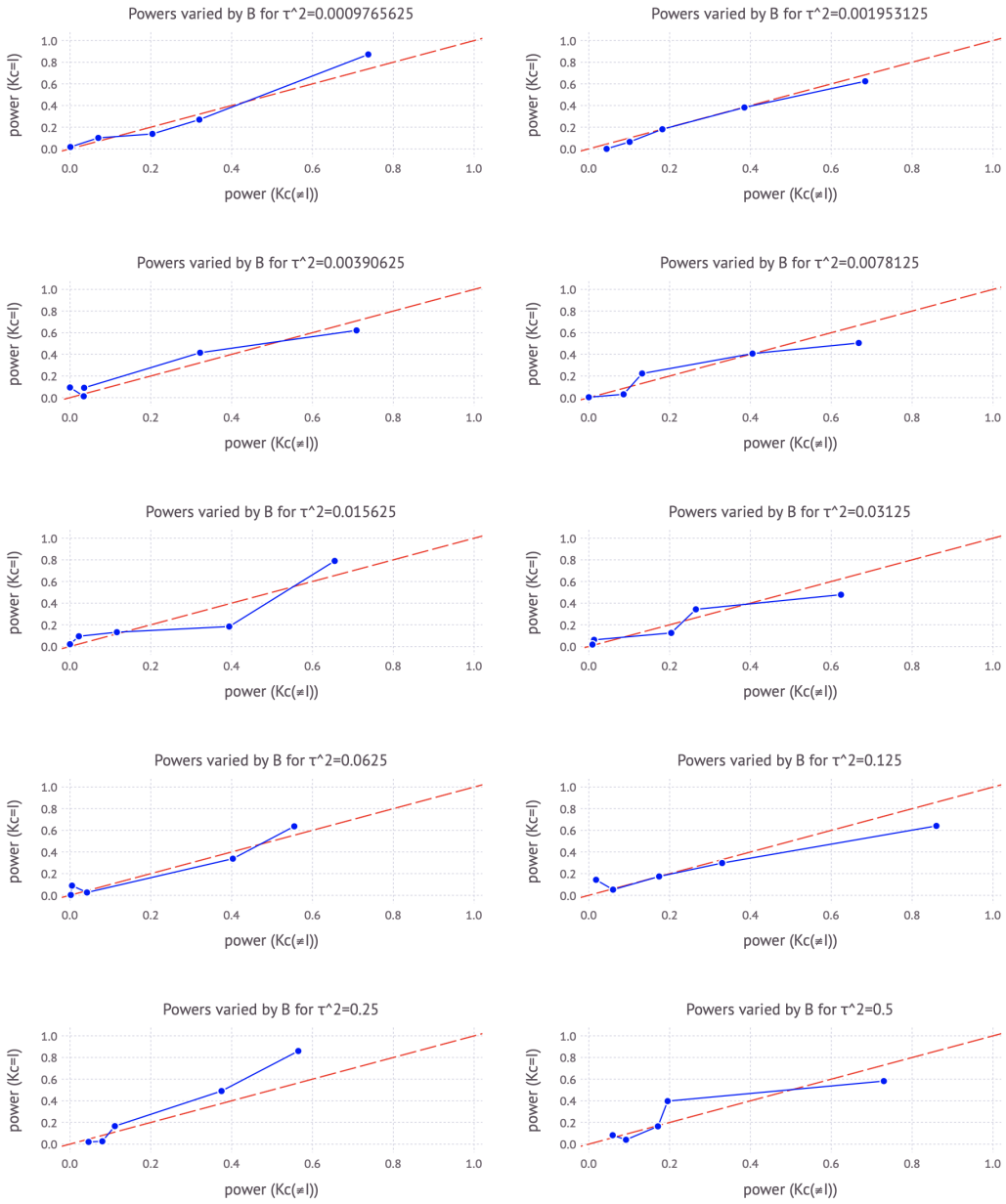


Figure 5: Contribution of climatic relatedness (K_C) to power on simulated data using Arabidopsis thaliana and 3 year by 2 site soil daily range temperature data. $K_C = I$ indicates not using the climatic relatedness. The range of true effect size matrices (B) narrows down based on power measured in Supplementary Figure 3.

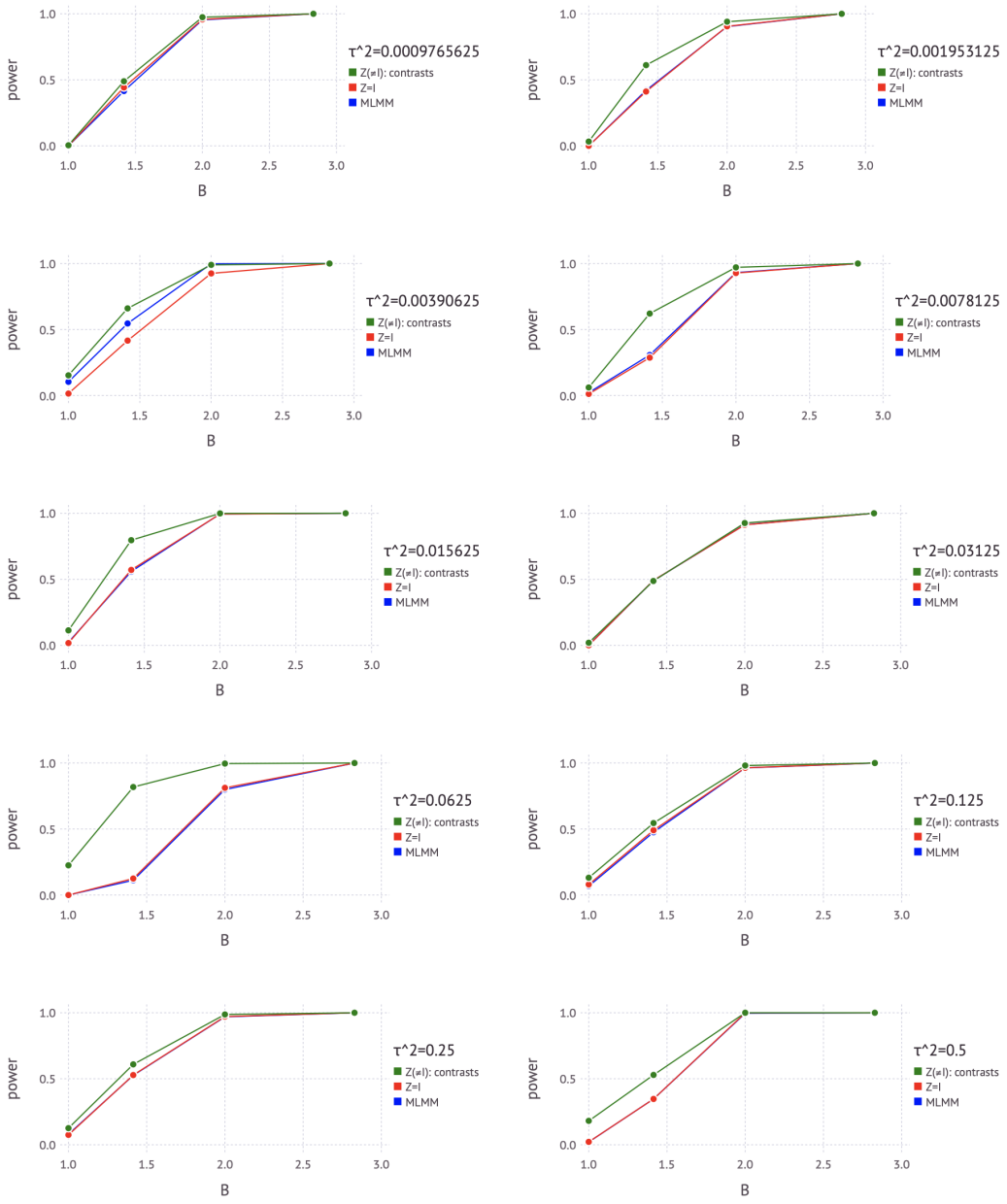


Figure 6: Contribution of trait covariates (Z) to power on simulated data using Arabidopsis thaliana and 3 year by 2 site soil daily range temperature data. Z: contrasts, $Z = I$ indicate employing the covariates (Z) or not in FlxQTL, and 'MLMM' indicates a Multivariate LMM implemented in Julia (Julia-version Multivariate LMM) as an alternative for GEMMA.

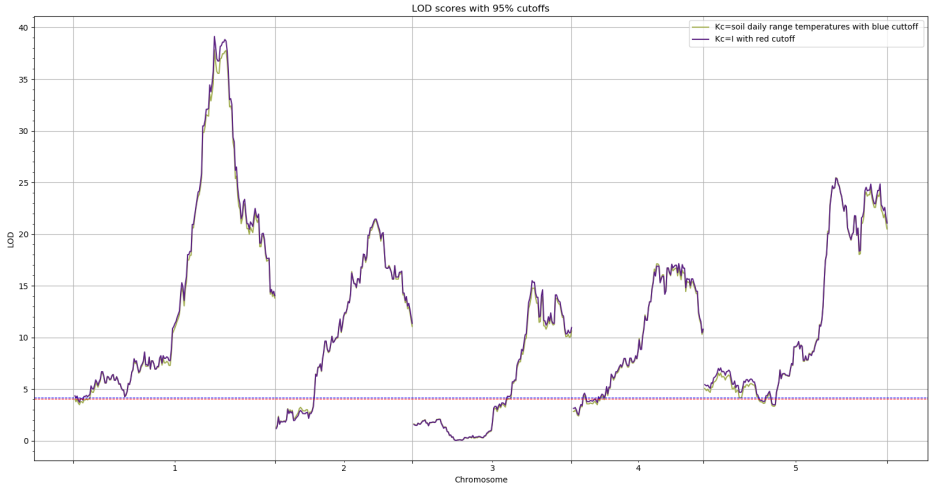


Figure 7: Multivariate genome scan for Arabidopsis thaliana: Comparison of LOD scores including a climatic relatedness (K_C) or not and corresponding 95% thresholds computed by 1000 permutations. Phenotypes of fitness (mean number of fruits per seedling planted) are considered as 6 environments, i.e. a 3 year by 2 site combination. The climatic relatedness matrix K_C is computed by using soil daily range temeperatures in two sites (Sweden and Italy) for 3 years. The trait covariates (Z) are contrasts containing 1's for overall mean for 6 environments and -1's (Italy), 1's (Sweden) for mean difference between sites to measure gene by environment interaction.

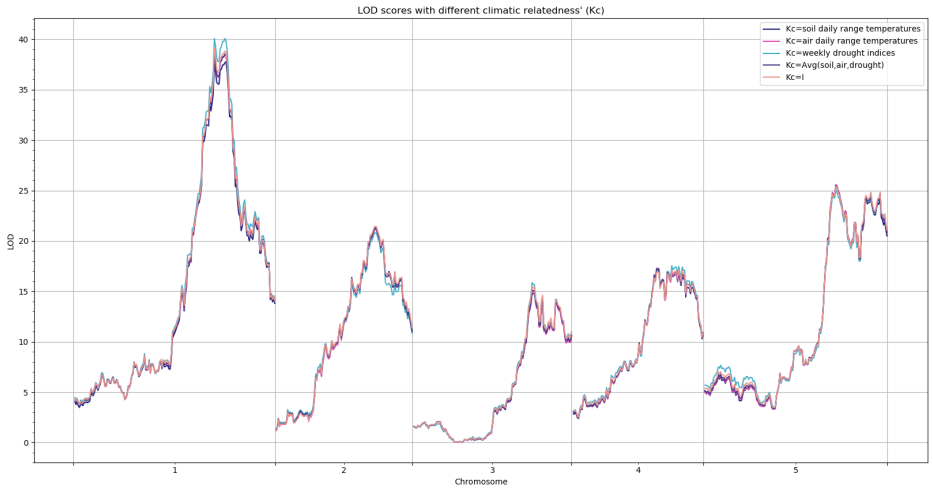


Figure 8: Multivariate genome scan for Arabidopsis thaliana: Comparison of LOD scores including various climatic relatedness matrices (K_C). Phenotypes of fitness (mean number of fruits per seedling planted) are considered as 6 environments, i.e. a 3 year by 2 site combination. The climatic relatedness matrices K_C are respectively computed by using soil/air daily range temperatures, weekly drought indices in two sites (Sweden and Italy) for 3 years. The trait covariates (Z) are contrasts containing 1's for overall mean for 6 environments and -1's (Italy), 1's (Sweden) for mean difference between sites to measure gene by environment interaction.

With Supplementary Figure 7, both point out that results are insensitive to types of climatic relatedness matrices (K_C's).

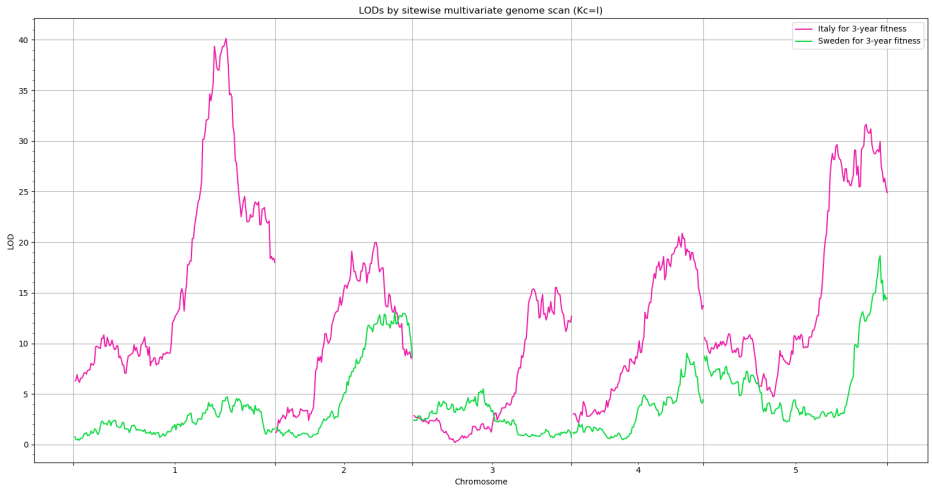


Figure 9: Site-wise multivariate genome scan for Arabidopsis thaliana: Comparison of LOD scores between Italy and Sweden, where K_C is not included according to Supplementary Figures 7, 8.

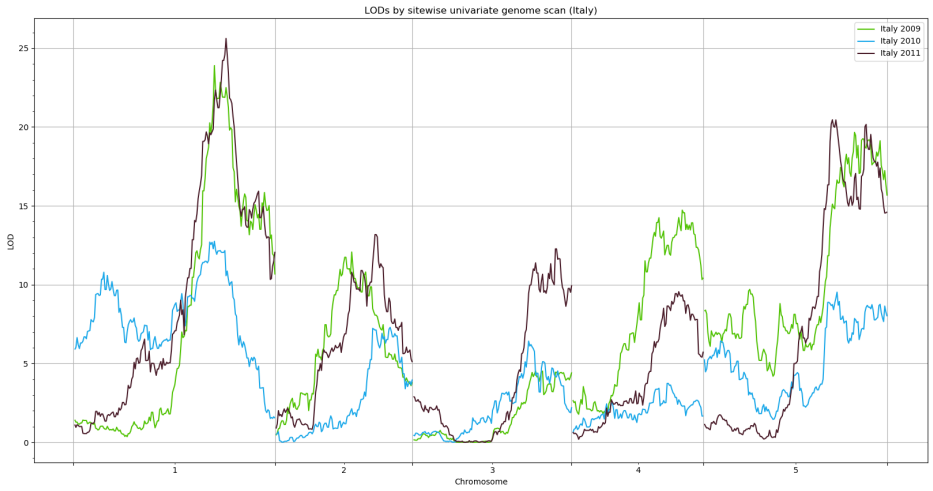


Figure 10: Univariate genome scans for Arabidopsis thaliana in Italy from 2009 to 2011.

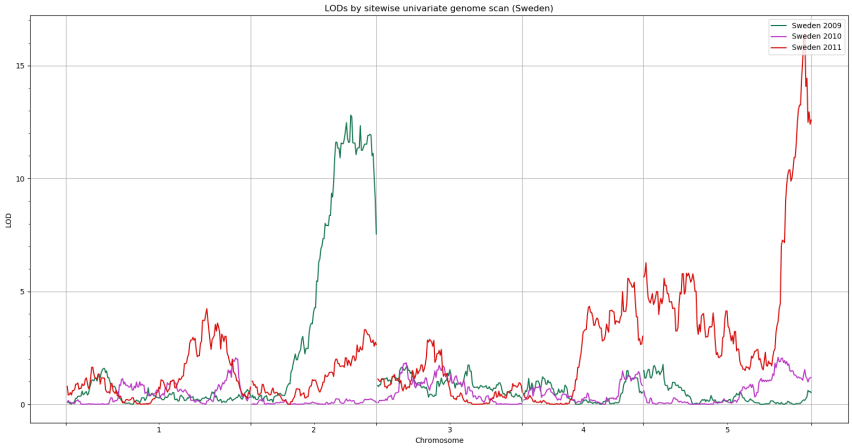


Figure 11: Univariate genome scans for Arabidopsis thaliana in Sweden from 2009 to 2011.

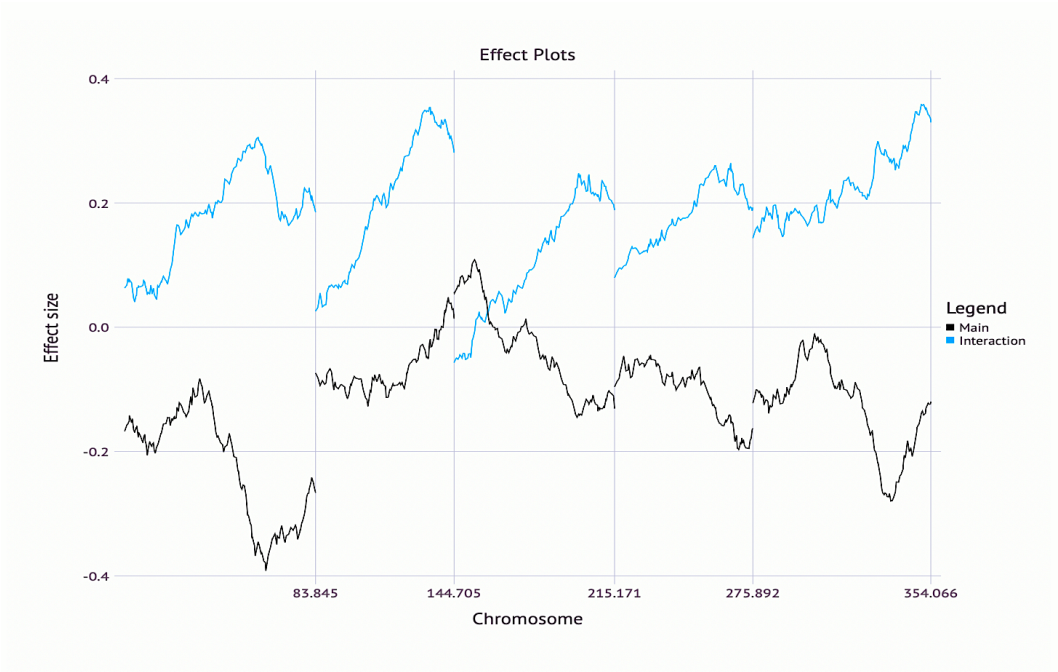


Figure 12: Effect plots for Arabidopsis thaliana with $K_C = I$ for 6 environments as described in Supplementary Figure 7.

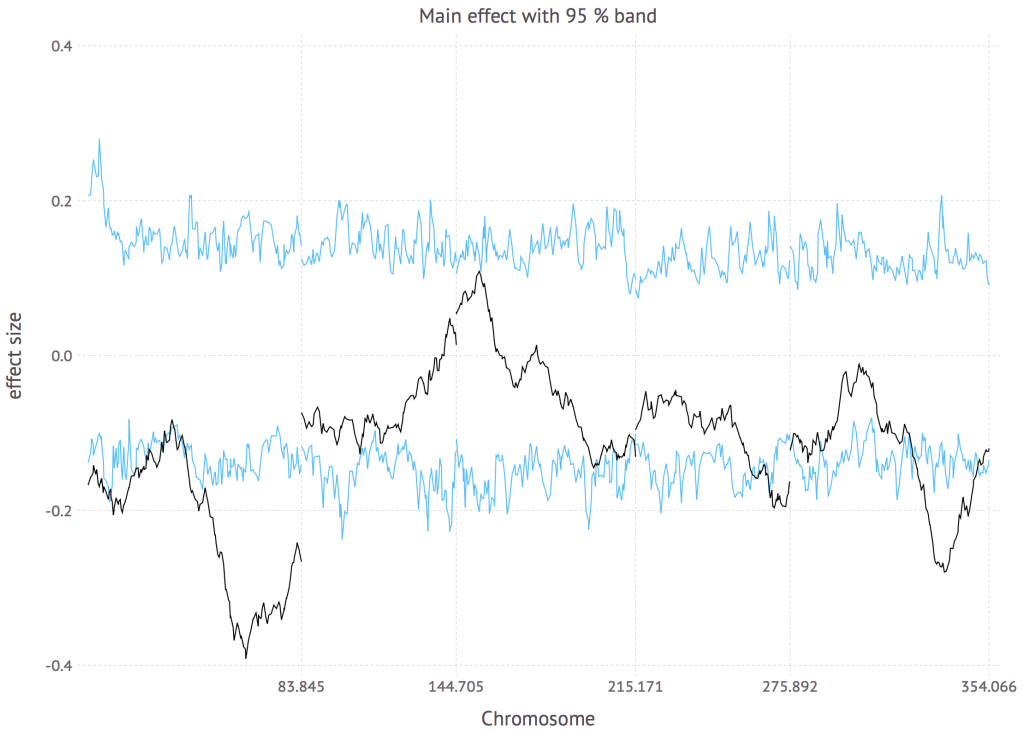


Figure 13: Main effects with 95% band for Arabidopsis thaliana. The band is obtained by 100 permutations

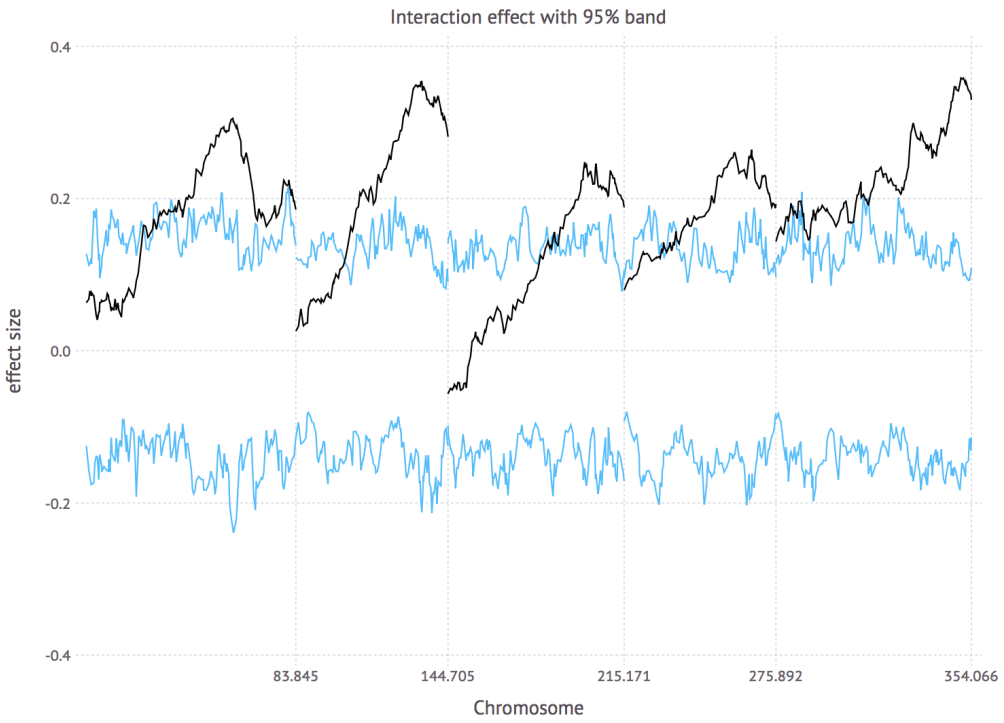


Figure 14: Interaction effects with 95% band for Arabidopsis thaliana. The band is obtained by 100 permutations

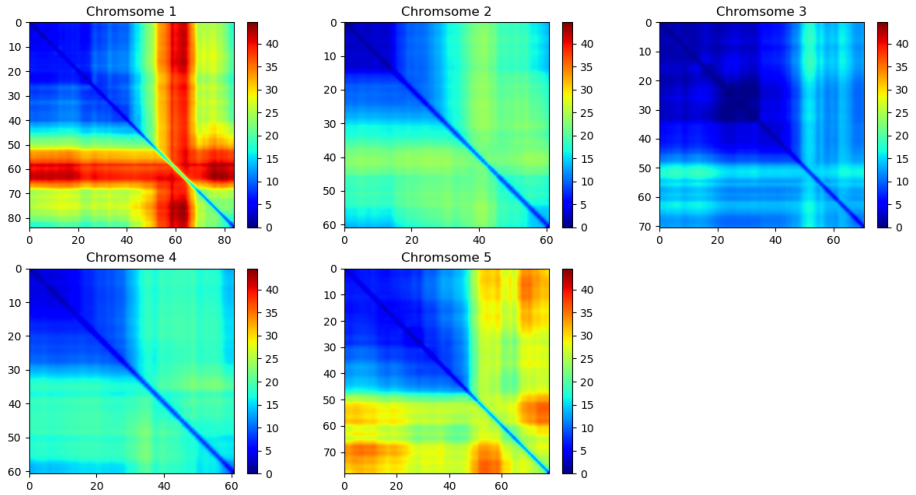
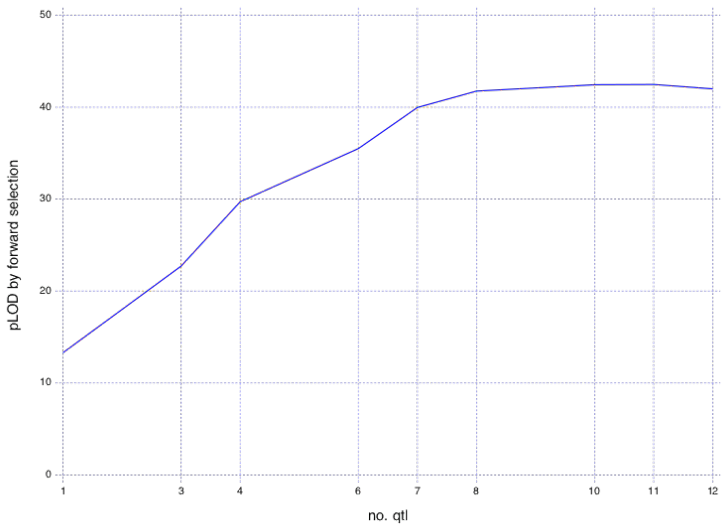
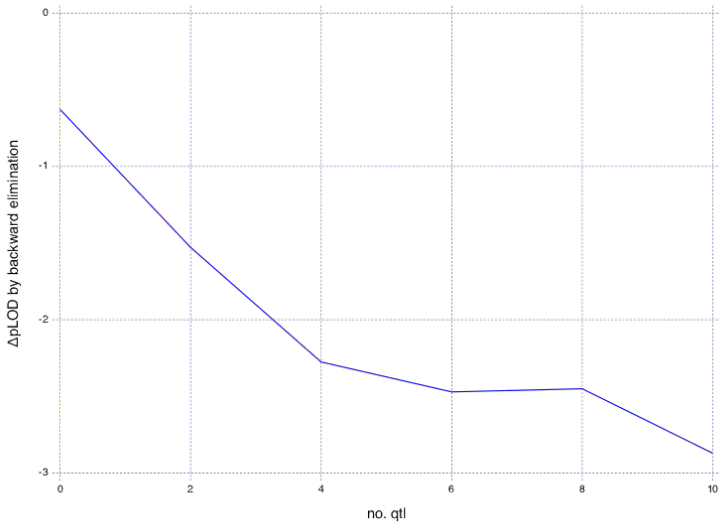


Figure 15: 2-d multivariate genome scan for Arabidopsis thaliana.



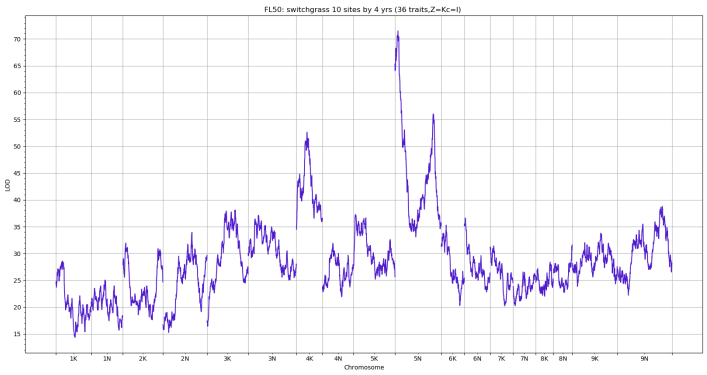
(a) Foward selection



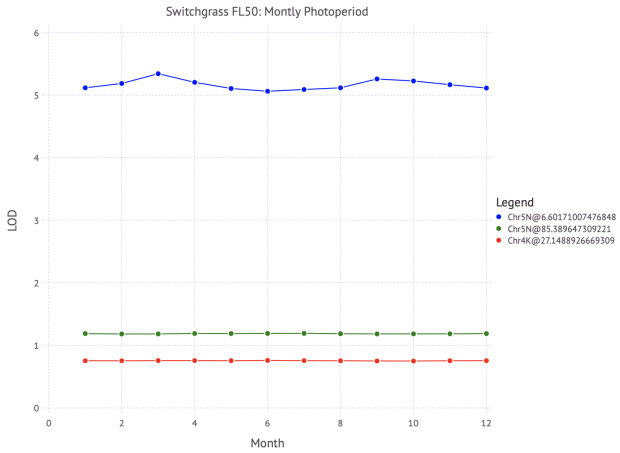
(b) Backward elimination

Figure 16: QTL selection for Arabidopsis thaliana (multiple QTL analysis).

A model selection approach by forward selection and backward elimination is performed to finialize significant QTL. A selection criterion is the penalized LOD score for additive models equivalent to BIC [29]. 1-d and 2-d genome scans are performed each step in both schemes based on QTL selected from 1-d and 2-d multivariate genome scans in Supplementary Figure 7, 15.



(a) 1D-genome scan



(b) Environmental scan for photoperiod

Figure 17: Switchgrass FL50

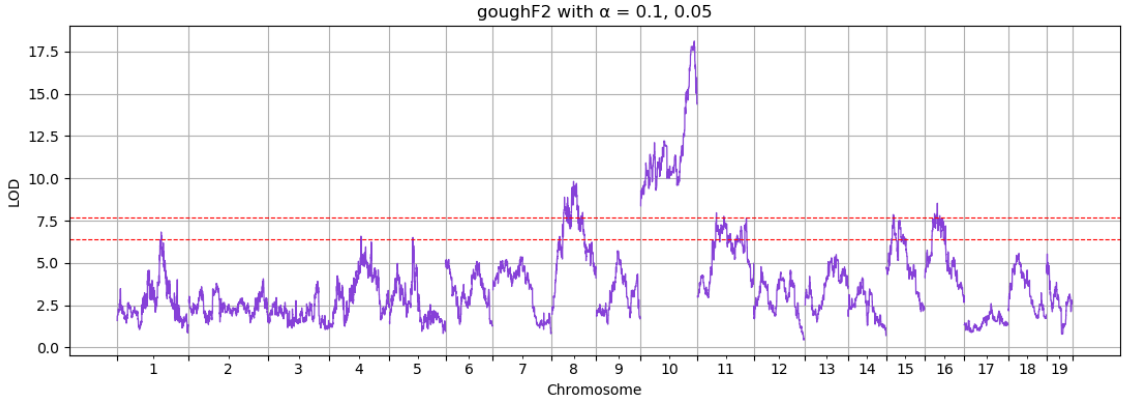


Figure 18: Application to functional data (F2 intercrosses between Gough Island mice and WSB/EiJ): LOD scores and thresholds at $\alpha = 0.1, 0.05$

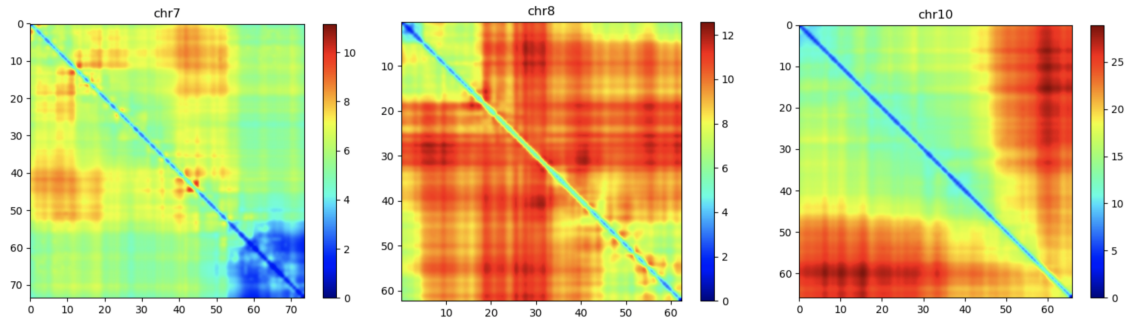


Figure 19: Application to functional data (F2 intercrosses between Gough Island mice and WSB/EiJ):LOD scores by 2D genome scan

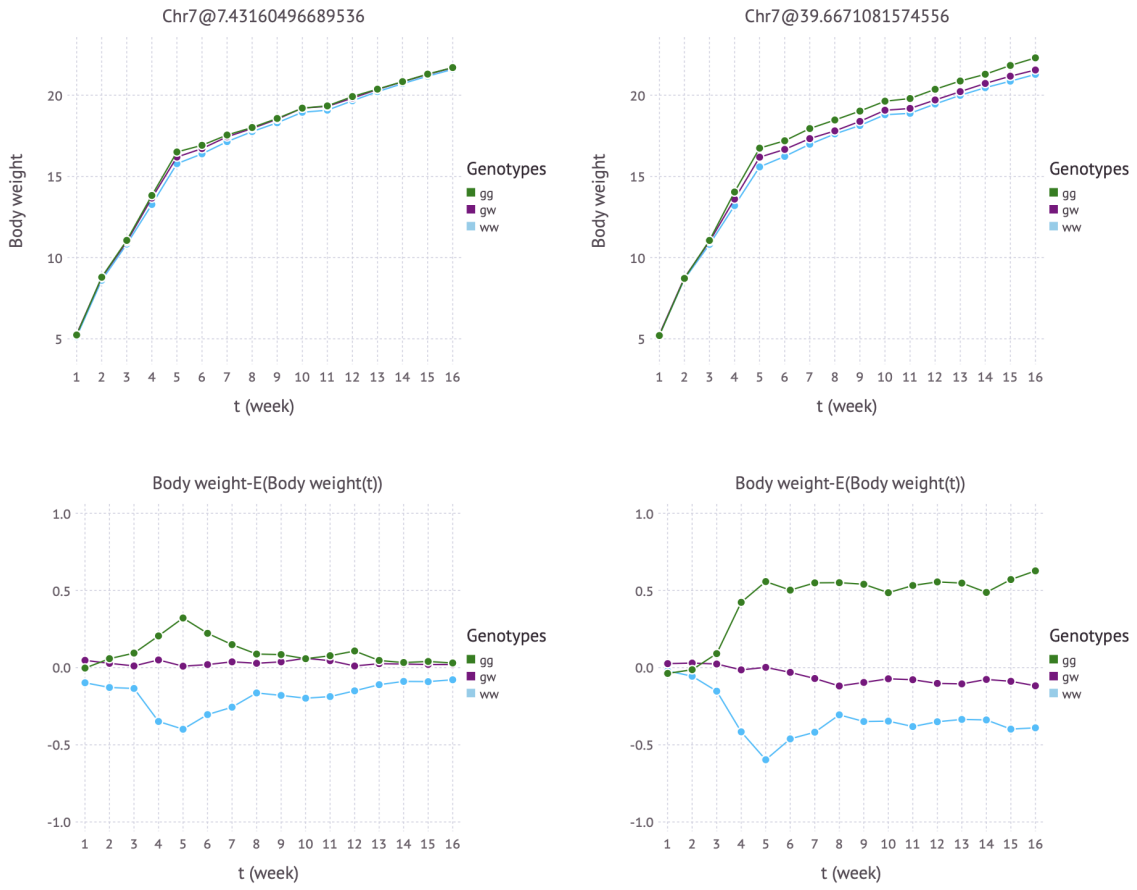


Figure 20: Application to functional data (F2 intercrosses between Gough Island mice and WSB/EiJ): Body weight effects by genotypes and corresponding deviations from weekly overall mean in Chromosome 7.

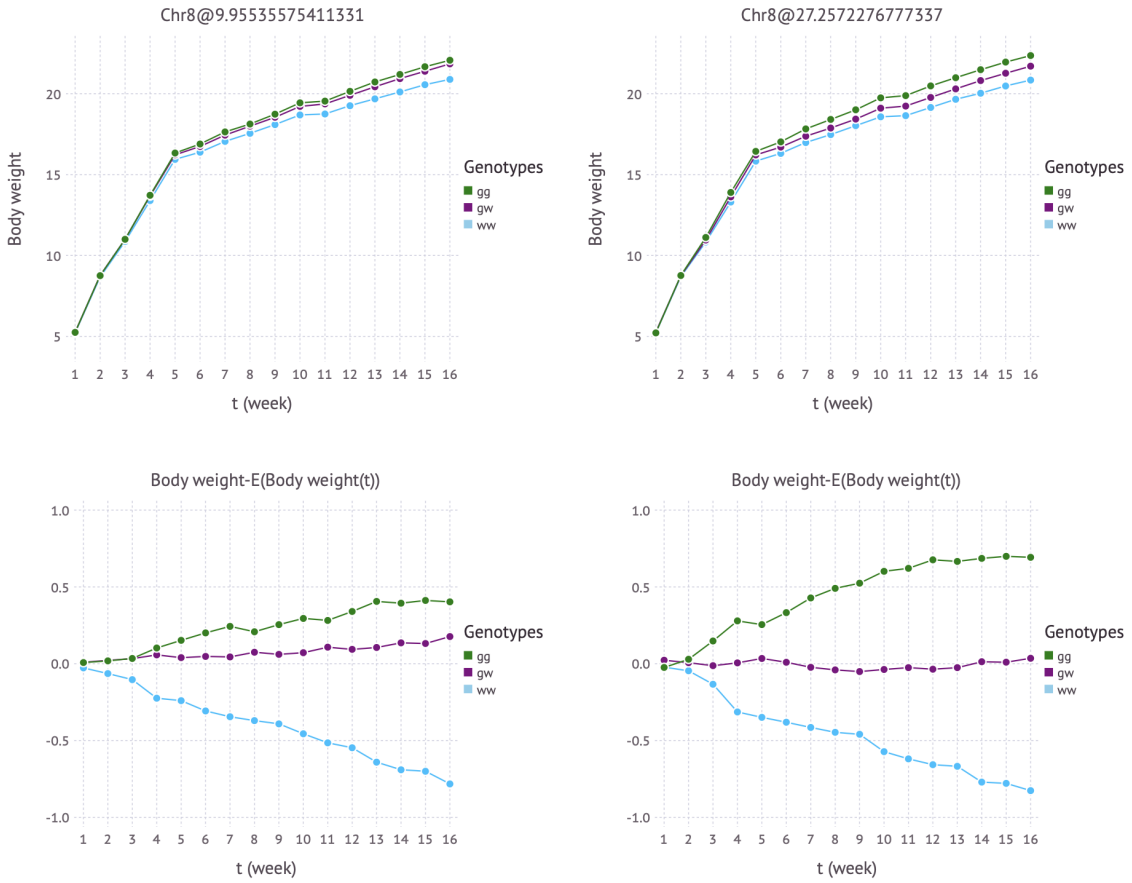


Figure 21: Application to functional data (F2 intercrosses between Gough Island mice and WSB/EiJ): Body weight effects by genotypes and corresponding deviations from weekly overall mean in Chromosome 8.

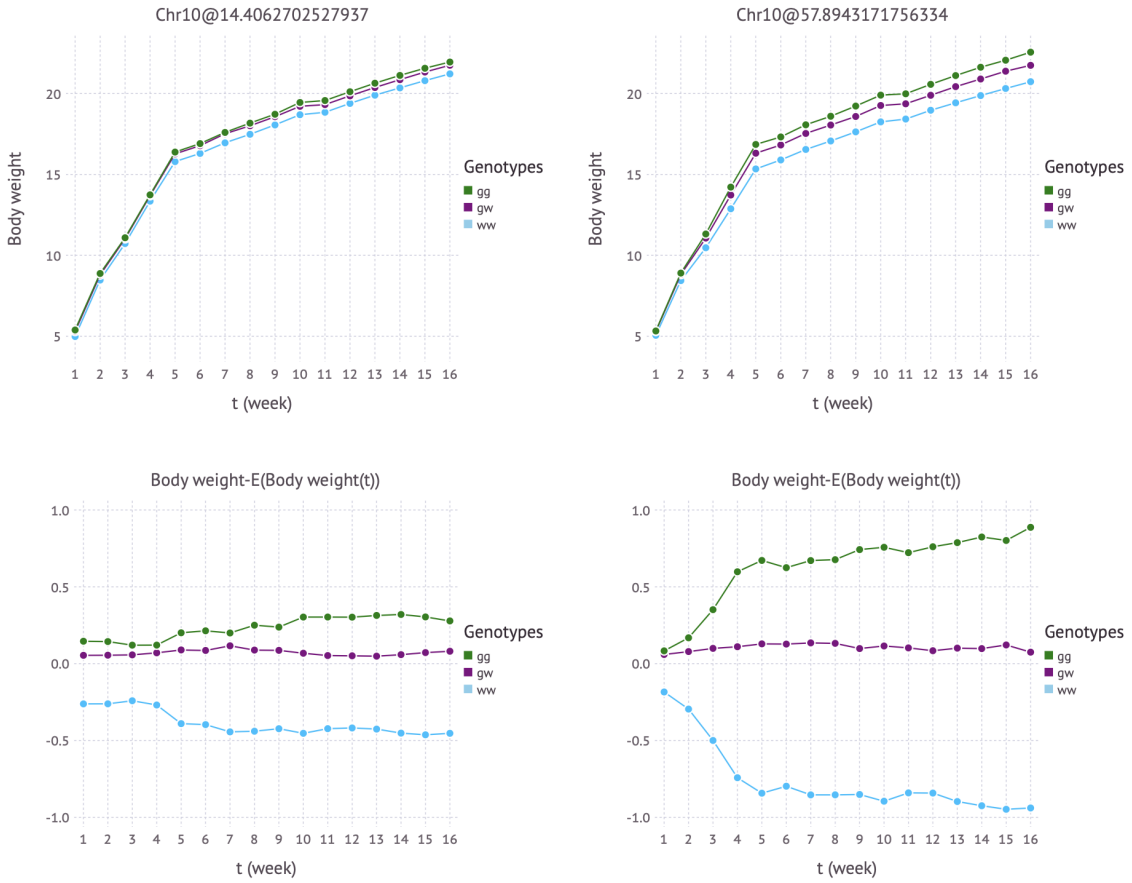


Figure 22: Application to functional data (F2 intercrosses between Gough Island mice and WSB/EiJ): Body weight effects by genotypes and corresponding deviations from weekly overall mean in Chromosome 10.

τ^2	Σ_1 ($K_C \neq I$)	Σ_1 ($K_C = I$)	Σ_2 ($K_C \neq I$)	Σ_2 ($K_C = I$)	Σ_3 ($K_C \neq I$)	Σ_3 ($K_C = I$)	$\Sigma = I$ ($K_C \neq I$)	$\Sigma, K_C = I$
0.00097656	0.726	0.727	0.675	0.674	0.369	0.375	0.481	0.485
0.00195313	0.62	0.617	0.537	0.538	0.492	0.5	0.676	0.701
0.00390625	0.691	0.712	0.633	0.628	0.431	0.444	0.59	0.597
0.0078125	0.564	0.558	0.488	0.495	0.181	0.197	0.522	0.545
0.015625	0.523	0.524	0.745	0.739	0.633	0.659	0.66	0.653
0.03125	0.603	0.62	0.51	0.519	0.468	0.475	0.644	0.651
0.0625	0.657	0.643	0.557	0.554	0.382	0.392	0.63	0.644
0.125	0.764	0.75	0.599	0.605	0.514	0.526	0.55	0.562
0.25	0.754	0.763	0.615	0.656	0.606	0.613	0.49	0.488
0.5	0.331	0.317	0.537	0.545	0.483	0.495	0.579	0.609

Table 1: Power comparison of with climatic relatedness (K_C) and without ($K_C = I$) in Supplementary Figure 4. 6×6 matrices of Σ_i for $i = 1, 2, 3$ are structured using following 3×3 sub-matrices: A, C (1's on the diagonal, a, c on off-diagonal entries, respectively), B (b in all entries). Σ_1 = [A B;B A], where $a = 0.4$, $b = 0.07$, Σ_2 =[A B;B A], where $a = 0.3$, $b = 0.07$, Σ_3=[A B;B C], where $a = 0.3$, $b = 0.0$, $c = 0.1$. I is an identity matrix.

Adding 1 or 2 QTL	Δ pLOD	pLOD	Deleting 1 or 2 QTL	Δ pLOD
10	0.692	42.442	10	-2.872
11	0.028	42.471	8	-2.451
12	-0.480	41.991	6	-2.471

Table 2: QTL selection for Arabidopsis thaliana (multiple QTL analysis) as in Supplementary Figure 16. 11 QTL can be finalized as significant QTL since the highest change (Δ pLOD = -2.872) in the backward elimination step when dropping one QTL from 11 QTL selected from the forward selection step yields.

Chromosome	no. of QTL	position (cM)
1	2	57, 77
2	2	133, 140
3	2	152, 199
4	2	259, 272
5	3	280,333,353

Table 3: Significant QTL for Arabidopsis thaliana as described in Supplementary Figure 16 and Supplementary Table 2.