File S1. Markov Chain Monte Carlo Implementation Strategy for Ante-BayesA and Ante-BayesB and Supplementary Tables and Figures.

Joint Posterior Density of All Unknown Parameters: In order to conduct MCMC, it is necessary to first specify the joint posterior density of all unknown parameters (SORENSEN and GIANOLA 2002). To do this, we interchangeably reparameterize the joint density of the data $\mathbf{y}$ and the random SNP effects, using $\mathbf{g}$ for ante-BayesA and $\boldsymbol{\delta}$ for ante-BayesB in order to exploit algorithmic efficiencies that are unique to either model. For instance with ante-BayesA, we write

$$p\left(\mathbf{y},\mathbf{g} \mid \boldsymbol{\beta},\mathbf{u},\boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t},\sigma_e^2\right) = p\left(\mathbf{y} \mid \boldsymbol{\beta},\mathbf{u},\mathbf{g},\sigma_e^2\right) p\left(\mathbf{g} \mid \boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t}\right) \qquad [A1]$$

Note the component $p\left(\mathbf{y} \mid \boldsymbol{\beta},\mathbf{u},\mathbf{g},\sigma_e^2\right) = N\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{g}+\mathbf{W}\mathbf{u},\mathbf{I}\sigma_e^2\right)$ is based on Equation [1] whereas $\mathbf{g} \sim p\left(\mathbf{g} \mid \boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t}\right) = N\left(\mathbf{0},\mathbf{G}\right)$ with $\mathbf{G} = \left(\mathbf{I}-\mathbf{T}\right)^{-1}\boldsymbol{\Delta}\left(\mathbf{I}-\mathbf{T}\right)^{-1\prime}$ are defined by elements in $\boldsymbol{\sigma}_{\boldsymbol{\delta}} = \left[\sigma_{\delta_1}^2 \quad \sigma_{\delta_2}^2 \quad \sigma_{\delta_3}^2 \quad \ldots \quad \sigma_{\delta_m}^2\right]$ specified along the diagonal of $\boldsymbol{\Delta}$, and by $\mathbf{t} = \left[t_{2,1},t_{3,2},...,t_{m,m-1}\right]'$ specified just below the diagonal elements in $\mathbf{T}$ as previously indicated. For ante-BayesB, we reparameterize [A1] differently:

$$p\left(\mathbf{y},\boldsymbol{\delta} \mid \boldsymbol{\beta},\mathbf{u},\boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t},\sigma_e^2\right) = p\left(\mathbf{y} \mid \boldsymbol{\beta},\boldsymbol{\delta},\mathbf{t},\sigma_e^2\right) p\left(\boldsymbol{\delta} \mid \boldsymbol{\sigma}_{\boldsymbol{\delta}}\right) \qquad [A2]$$

recognizing that $\boldsymbol{\delta} = (\mathbf{I}-\mathbf{T})\mathbf{g}$ such that [A2] represents a linear transformation of [A1]. That is, the first component of [A2] is based on $p\left(\mathbf{y} \mid \boldsymbol{\beta},\boldsymbol{\delta},\mathbf{t},\sigma_e^2\right) = N\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\left(\mathbf{I}-\mathbf{T}\right)^{-1}\boldsymbol{\delta}+\mathbf{W}\mathbf{u},\mathbf{I}\sigma_e^2\right)$ whereas $p\left(\boldsymbol{\delta} \mid \boldsymbol{\sigma}_{\boldsymbol{\delta}}\right) = \prod_{j=1}^{m} N\left(0,\sigma_{\delta_j}^2\right)$. We'll subsequently represent [A1] and [A2] together as

$p\left(\mathbf{y},\mathbf{g}(\boldsymbol{\delta})\,|\,\boldsymbol{\beta},\mathbf{u},\boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t},\sigma_e^2,\mu_t,\sigma_t^2\right)$ to recognize the interchangeability between **g** and **δ** when conditioning on **t**. The joint posterior density of all unknown parameters can be written as products of specifications provided previously:

$$\begin{aligned}&p\left(\boldsymbol{\beta},\mathbf{g},\mathbf{u},\sigma_e^2,\boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t},\mu_t,\sigma_t^2,\nu_\delta,s_\delta^2\,|\,\mathbf{y}\right)\propto\\&p\left(\mathbf{y},\mathbf{g}(\boldsymbol{\delta})\,|\,\boldsymbol{\beta},\mathbf{u},\boldsymbol{\sigma}_{\boldsymbol{\delta}},\mathbf{t},\sigma_e^2,\mu_t,\sigma_t^2\right)p(\boldsymbol{\beta})\left(\prod_{j=2}^{m}p\left(t_{j,j-1}\,|\,\mu_t,\sigma_t^2\right)\right)\\&\left(\prod_{j=1}^{m}p\left(\sigma_{\delta_j}^2\,|\,\nu_\delta,s_\delta^2,\pi_\delta\right)\right)p\left(\sigma_u^2\,|\,\nu_u,s_u^2\right),p\left(\sigma_e^2\,|\,\nu_e,S_e\right)p(\nu_\delta)\\&p\left(s_\delta^2\,|\,\alpha_s,\beta_s\right)p\left(\mu_t\,|\,\mu_{t0},s_{t0}^2\right)p\left(\sigma_t^2\,|\,\nu_t,s_t^2\right)p\left(\pi_\delta\,|\,\alpha_\pi,\beta_\pi\right)\end{aligned} \quad \text{[A3]}$$

From the paper, $p(\boldsymbol{\beta})=N\left(\boldsymbol{\beta},\mathbf{V}_{\boldsymbol{\beta}}\right)$, $p\left(t_{j,j-1}\,|\,\mu_t,\sigma_t^2\right)=N\left(\mu_t,\sigma_t^2\right)$, $p\left(\sigma_u^2\,|\,\nu_u,s_u^2\right)=\chi^{-2}\left(\nu_u,\nu_u s_u^2\right)$, $p\left(\sigma_e^2\,|\,\nu_e,s_e^2\right)=\chi^{-2}\left(\nu_e,\nu_e s_e^2\right)$, $p\left(s_\delta^2\,|\,\alpha_s,\beta_s\right)=Gamma\left(\alpha_s,\beta_s\right)$, $p\left(\mu_t\,|\,\mu_{t0},s_{t0}^2\right)=N\left(\mu_{t0},s_{t0}^2\right)$, $p\left(\sigma_t^2\,|\,\nu_t,s_t^2\right)=\chi^{-2}\left(\nu_t,\nu_t s_t^2\right)$, and $p\left(\pi_\delta\,|\,\alpha_\pi,\beta_\pi\right)=Beta\left(\alpha_\pi,\beta_\pi\right)$. Furthermore, $p\left(\sigma_{\delta_j}^2\,|\,\nu_\delta,s_\delta^2,\pi_\delta\right)$ is a mixture analogous to Equation [2] for ante-BayesB whereas $p\left(\sigma_{\delta_j}^2\,|\,\nu_\delta,s_\delta^2,\pi_\delta=1\right)=\chi^{-2}\left(\nu_\delta,\nu_\delta s_\delta^2\right)$ for ante-BayesA as described in the paper. For some parameters, we subsequently derive and present FCD separately for ante-BayesA (π_δ = 1) from ante-BayesB (π_δ < 1) as some MCMC sampling strategies appear to be simpler or more computationally efficient for one or the other model.

Now MCMC requires random draws from the full conditional densities of each unknown parameter (or blocks thereof) conditional on all other parameters and the data(SORENSEN and GIANOLA 2002). These full conditional densities are provided below for various classes of these unknown parameters.

Sampling All Fixed and Random Effects in Ante-BayesA: Write $\boldsymbol{\theta}=\left[\boldsymbol{\beta}' \quad \mathbf{g}' \quad \mathbf{u}'\right]'$ as the (p+m+q) vector of fixed and random effects, $\mathbf{Q}=\left[\mathbf{X} \vdots \quad \mathbf{Z} \vdots \quad \mathbf{W}\right]$ as the n x (p+m+q) overall model incidence matrix with $\boldsymbol{\Sigma}^{-}=diag\left(\mathbf{V}_{\boldsymbol{\beta}}^{-1} \quad \mathbf{G}^{-1} \quad \mathbf{A}^{-1}\sigma_u^{-2}\right)$ as a block diagonal matrix with the corresponding listed components as the various blocks. It can be readily demonstrated (SORENSEN and GIANOLA 2002) that the FCD of $\boldsymbol{\theta}$ is

$$\boldsymbol{\theta} \mid \mathbf{y}, ELSE \sim N\left(\hat{\boldsymbol{\theta}}, \mathbf{C}\right) \qquad \text{[A4]}$$

where $ELSE$ denotes all other parameters in [A3] other than $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}=\mathbf{CQ}'\mathbf{y}+\left[\boldsymbol{\beta}_0{}'\mathbf{V}_{\boldsymbol{\beta}}^{-1} \quad \mathbf{0}_{1x(m+q)}\right]'$ for $\mathbf{C}=\left(\mathbf{Q}'\mathbf{Q}+\Sigma^{-}\right)^{-1}\sigma_e^2$. Note that with a typical "flat" prior for $\boldsymbol{\beta}$ is defined by $\mathbf{V}_{\boldsymbol{\beta}}^{-1}=\mathbf{0}$ such that $\hat{\boldsymbol{\theta}}=\mathbf{CQ}'\mathbf{y}$. Also, note that univariate or multivariate block FCD subsets of $\boldsymbol{\theta}$ could also be partitioned and sampled using [A4] based on results from Wang and Gianola(1994) The structure of $\mathbf{G}^{-1}=\left\{G^{jj'}\right\}$ contained within $\boldsymbol{\Sigma}^{-}$ is a simple tridiagonal matrix: using Zimmerman and Nunez-Anton (2010 pg 52), the diagonal elements are $G^{jj}=\sigma_{\delta_j}^{-2}+t_{j+1,j}^2\sigma_{\delta_{j+1}}^{-2}$ for j = 1,2,….,m-1 with $G^{mm}=\sigma_{\delta_m}^{-2}$ whereas the elements adjacent to the diagonal are $G^{j,j+1}=G^{j+1,j}=-t_{j+1,j}\sigma_{\delta_{j+1}}^{-2}$.

Sampling marker-specific variances in AnteBayesA: Consider now the FCD for $\sigma_{\delta_j}^2$, j=1,2,…,m:

$$p\left(\sigma_{\delta_j}^2 \mid \mathbf{y}, ELSE\right) \propto p\left(\mathbf{g} \mid t_{21}, t_{32}, \ldots, t_{m,m-1}, \sigma_{\delta_1}^2, \sigma_{\delta_2}^2, \sigma_{\delta_3}^2, \ldots, \sigma_{\delta_m}^2\right) p\left(\sigma_{\delta_j}^2 \mid \nu_\delta, s_\delta^2\right) \qquad \text{[A5]}$$

We use Chan and Jeliazkov (2009, pg 461) to simplify the first component of [A5] as follows:

$$p\left(\mathbf{g}\,|\,t_{21},t_{32},\ldots,t_{G,G-1},\boldsymbol{\sigma}^2_{\delta_1},\boldsymbol{\sigma}^2_{\delta_2},\boldsymbol{\sigma}^2_{\delta_3},\ldots,\boldsymbol{\sigma}^2_{\delta_G}\right)$$
$$\propto\left|\mathbf{G}^{-1}\right|^{1/2}\exp\left(-\frac{1}{2}\mathbf{g}'\mathbf{G}^{-1}\mathbf{g}\right)=\left|\left(\mathbf{I}-\mathbf{T}\right)'\boldsymbol{\Delta}^{-1}\left(\mathbf{I}-\mathbf{T}\right)\right|^{1/2}\exp\left(-\frac{1}{2}\mathbf{g}'\left(\mathbf{I}-\mathbf{T}\right)'\boldsymbol{\Delta}^{-1}\left(\mathbf{I}-\mathbf{T}\right)\mathbf{g}\right)\text{[A6]}$$
$$\propto\left|\boldsymbol{\Delta}^{-1}\right|^{1/2}\exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{\Delta}^{-1}\boldsymbol{\delta}\right)\propto\prod_{j=1}^{G}\left(\boldsymbol{\sigma}^2_{\delta_j}\right)^{-1/2}\exp\left(-\frac{1}{2}\frac{\boldsymbol{\delta}_j^2}{\boldsymbol{\sigma}^2_{\delta_j}}\right)$$

Using the component in [A6] pertaining to $\sigma^2_{\delta_j}$ in [A5] and $p\left(\sigma^2_{\delta_j}\,|\,\nu_\delta,s^2_\delta\right)\propto\sigma^2_{\delta_j}{}^{-\left(\frac{\nu_\delta}{2}+1\right)}e^{-\frac{\nu_\delta s^2_\delta}{2\sigma^2_{\delta_j}}}$ then

$$p\left(\sigma^2_{\delta_j}\,|\,\mathbf{y},ELSE\right)\propto\propto\left(\sigma^2_{\delta_j}\right)^{-1/2}\exp\left(-\frac{1}{2}\frac{\delta_j^2}{\sigma^2_{\delta_j}}\right)\frac{\left(\frac{\nu_\delta s^2_\delta}{2}\right)^{\frac{\nu_\delta}{2}}}{\Gamma\left(\frac{\nu_\delta}{2}\right)}\sigma^2_{\delta_j}{}^{-\left(\frac{\nu_\delta}{2}+1\right)}e^{-\frac{\nu_\delta s^2_\delta}{2\sigma^2_{\delta_j}}} \quad \text{[A7]}$$
$$\propto\left(\sigma^2_{\delta_j}\right)^{-\left(\frac{\nu_\delta+1}{2}+1\right)}\exp\left(-\frac{1}{2}\frac{\left(\delta_j^2+\nu_\delta s^2_\delta\right)}{\sigma^2_{\delta_j}}\right)$$

i.e. $p\left(\sigma^2_{\delta_j}\,|\,\mathbf{y},ELSE\right)=\chi^{-2}\left(\nu_\delta+1,\delta_j^2+\nu_\delta s^2_\delta\right)$. As a sidenote, elements of $\boldsymbol{\delta}$ can be recursively derived from **g**:

$$\boldsymbol{\delta}=\left(\mathbf{I}-\mathbf{T}\right)\mathbf{g}=\begin{bmatrix}1&0&\cdots&0&0\\-t_{21}&1&\cdots&0&0\\0&-t_{32}&&0&0\\\vdots&&\ddots&1&0\\0&0&&-t_{m,m-1}&1\end{bmatrix}\begin{bmatrix}g_1\\g_2\\g_3\\\\g_m\end{bmatrix}=\begin{bmatrix}g_1\\g_2-t_{21}g_1\\g_3-t_{32}g_2\\\\g_m-t_{m,m-1}g_{m-1}\end{bmatrix} \quad \text{[A8]}$$

Sampling fixed and random effects other than SNP effects in Ante-BayesB. Here we deem it computationally tractable to sample the rest of the location parameters separately from **g**. We again use Equation [A4] except that now we define $\boldsymbol{\theta}=\left[\boldsymbol{\beta}'\;\;\mathbf{u}'\right]'$ as a (*p*+ *q*)x1 vector of fixed and random polygenic effects with $\mathbf{Q}=\left[\mathbf{X}\vdots\;\;\mathbf{W}\right]$ being the corresponding *n* x (*p*+ *q*) submodel incidence matrix and $\boldsymbol{\Sigma}^{-}=diag\left(\mathbf{0}_{pxp}\;\;\mathbf{A}^{-1}\sigma_u^{-2}\right)$ being the corresponding block diagonal matrix. We then sample using Equation [A4] and $\hat{\boldsymbol{\theta}}=\mathbf{CQ}'\left(\mathbf{y}\text{-}\mathbf{Zg}\right)+\left[\boldsymbol{\beta}_0'\mathbf{V}_{\boldsymbol{\beta}}^{-1}\;\;\mathbf{0}_{1xq}\right]'$ for $\mathbf{C}=\left(\mathbf{Q}'\mathbf{Q}+\Sigma^{-}\right)^{-1}\sigma_e^2$.

Sampling random SNP effects and variances in Ante-BayesB: We consider the collapsed sampling strategy (Liu 1994) for jointly sampling $\sigma^2_{\delta_j}$ and δ_j as previously adapted for Bayes B in Meuwissen et al (2001). Consider the previously described mixture prior on the conditional variances

$$p\left(\boldsymbol{\sigma}^2_{\delta_j} \mid v_\delta, s^2_\delta, \boldsymbol{\pi}_\delta\right) = \begin{cases} 0 \textit{ with probability } \boldsymbol{\pi}_\delta \\ \boldsymbol{\chi}^{-2}\left(v_\delta, v_\delta s^2_\delta\right) \textit{ with probability } \mathbf{1\text{-}}\boldsymbol{\pi}_\delta \end{cases} \quad \text{[A9]}$$

We jointly sample $\sigma^2_{\delta_j}$ and δ_j from $p\left(\sigma^2_{\delta_j}, \delta_j \mid ELSE, \mathbf{y}\right)$, by sampling first from $p\left(\sigma^2_{\delta_j} \mid \mathbf{y}, ELSE \text{ except } \delta_j\right)$ and then from $p\left(\delta_j \mid ELSE, \mathbf{y}\right)$. The first component of [A2] implies the following linear model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{H}\boldsymbol{\delta} + \mathbf{W}\mathbf{u} + \mathbf{e} \quad \text{[A10]}$$

where $\mathbf{H} = \mathbf{Z}\left(\mathbf{I} - \mathbf{T}\right)^{-1}$. Let's further partition $\mathbf{H}$ into the jth column, $\mathbf{h}_j$, and other remaining columns $\mathbf{H}_{-j}$; similarly, we represent $\boldsymbol{\delta}_{-j}$ as all elements of $\boldsymbol{\delta}$ other than δ_j. Then we further rewrite [A10] as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{H}_{-j}\boldsymbol{\delta}_{-j} + \mathbf{h}_j\delta_j + \mathbf{W}\mathbf{u} + \mathbf{e} \quad \text{[A11]}$$

It can be readily demonstrated, following similar developments for BayesB provided by Meuwissen et al. (2001), that:

$$\begin{aligned} p\left(\sigma^2_{\delta_j} \mid \mathbf{y}, ELSE \text{ except } \delta_j\right) &= \int_{\delta_j} p\left(\sigma^2_{\delta_j}, \delta_j \mid \mathbf{y}, ELSE\right) \\ &\propto \int_{\delta_j} p\left(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\delta}, \mathbf{u}, \mathbf{t}, \sigma^2_e\right) p\left(\delta_j \mid \sigma^2_{\delta_j}\right) p\left(\sigma^2_{\delta_j} \mid v_\delta, s^2_\delta, \pi\right) d\delta_j \\ &\propto p\left(\sigma^2_{\delta_j} \mid v_\delta, s^2_\delta, \pi\right) \int_{\delta_j} \exp\left(-\frac{1}{2\sigma^2_e}\left(\mathbf{y}^*_j - \mathbf{h}_j\delta_j\right)'\left(\mathbf{y}^*_j - \mathbf{h}_j\delta_j\right)\right) \exp\left(-\frac{1}{2\sigma^2_{\delta_j}}\delta^2_j\right) d\delta_j \\ &\propto p\left(\sigma^2_{\delta_j} \mid v_\delta, s^2_\delta, \pi\right) \left|\mathbf{V}_j\right|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{y}^{*\prime}_j \mathbf{V}^{-1}_j \mathbf{y}^*_j\right) \end{aligned} \quad \text{[A12]}$$

where $\mathbf{y}^*_j = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{H}_{-j}\boldsymbol{\delta}_j - \mathbf{W}\mathbf{u}$ and $\mathbf{V}_j = \mathbf{h}_j\mathbf{h}'_j\sigma^2_{\delta_j} + \mathbf{I}\sigma^2_e$.

Since [A12] is not a recognizable distributional form, a Metropolis Hastings step is required. We adapt the independence chains implementation (CHIB and GREENBERG 1995) as also adapted by Meuwissen et al. (2001) using the prior $p\left(\sigma^2_{\delta_j} \mid v_\delta, s^2_\delta, \pi_\delta\right)$ as the candidate density. That is, at MCMC cycle [*k*], one samples a candidate, say, $\sigma^{2[*]}_{\delta_j}$, from $p\left(\sigma^2_{\delta_j} \mid \nu_\delta, s^2_\delta, \pi_\delta\right)$ conditioned upon the updated values for ν_δ, s^2_δ and π_δ. One accepts $\sigma^2_{\delta_{j[k]}} = \sigma^2_{\delta_{j*}}$ as the value for in cycle [*k*] with probability based on the Metropolis-Hastings acceptance ratio $q\left(\sigma^2_{\delta_{j[k-1]}} \rightarrow \sigma^2_{\delta_{j*}}\right)$:

$$q\left(\sigma^2_{\delta_{j[k-1]}} \rightarrow \sigma^2_{\delta_{j*}}\right) = \begin{cases} \min\left(\dfrac{p\left(\sigma^2_{\delta_{j*}} \mid \mathbf{y}, ELSE \text{ except } \delta_j\right) p\left(\sigma^2_{\delta_{j[k-1]}} \mid \nu_\delta, s^2_\delta, \pi\right)}{p\left(\sigma^2_{\delta_{j[k-1]}} \mid \mathbf{y}, ELSE \text{ except } \delta_j\right) p\left(\sigma^2_{\delta_{j*}} \mid \nu_\delta, s^2_\delta, \pi\right)}, 1\right); \\ 1, otherwise \end{cases} \quad \text{[A13]}$$

If the proposal $\sigma^2_{\delta_{j*}}$ is rejected, then set $\sigma^2_{\delta_{j[k]}} = \sigma^2_{\delta_{j[k-1]}}$; i.e., the value of $\sigma^2_{\delta_j}$ in the previous MCMC cycle. It can be demonstrated that using Meuwissen et al. (2001) that [A13] is further equal to:

$$q\left(\sigma^2_{\delta_{j[k-1]}} \rightarrow \sigma^2_{\delta_{j*}}\right) = \begin{cases} \min\left(\dfrac{\left|\mathbf{V}_j^*\right|^{-1/2} \exp\left(-\dfrac{1}{2}\mathbf{y}^{*\prime}\mathbf{V}_j^{*-1}\mathbf{y}^*\right)}{\left|\mathbf{V}_j^{[t-1]}\right|^{-1/2} \exp\left(-\dfrac{1}{2}\mathbf{y}^{*\prime}\left(\mathbf{V}_j^{[t-1]}\right)^{-1}\mathbf{y}^*\right)}, 1\right) \\ 1, otherwise \end{cases} \quad \text{[A14]}$$

Note that neither the determinant $\left|\mathbf{V}_j\right|$ nor the inverse $\mathbf{V}_j^{-1}$ are trivial computations since *m* is typically large. Adapting a development from Rohan Fernando (personal communication) for BayesB, it can be readily shown that [A14] further simplifies:

$$q\left(\sigma^2_{\delta_{j[k-1]}} \rightarrow \sigma^2_{\delta_{j*}}\right) = \begin{cases} \min\left(\dfrac{\left(v_j^*\right)^{-1/2} \exp\left(-\dfrac{1}{2}\dfrac{\left(\mathbf{h}_j'\mathbf{y}_j^*\right)^2}{v_j^*}\right)}{\left(v_j^{[t+1]}\right)^{-1/2} \exp\left(-\dfrac{1}{2}\dfrac{\left(\mathbf{h}_j'\mathbf{y}_j^*\right)^2}{v_j^{[t+1]}}\right)}, 1\right) \\ 1, otherwise \end{cases}$$

[A15]

where $v_j^* = \left(\mathbf{h}_j'\mathbf{h}_j\right)^2 \sigma^2_{\delta_{j}*} + \left(\mathbf{h}_j'\mathbf{h}_j\right)\sigma^2_e$ and $v_j^{[k-1]} = \left(\mathbf{h}_j'\mathbf{h}_j\right)^2 \sigma^2_{\delta_{j[k-1]}} + \left(\mathbf{h}_j'\mathbf{h}_j\right)\sigma^2_e$. Once $\sigma^2_{\delta_j}$ is sampled, one could immediately draw δ_j from $p\left(\delta_j \mid ELSE, \mathbf{y}\right)$ readily seen to be

$$p\left(\delta_j \mid ELSE,\mathbf{y}\right)=N\left(\frac{\mathbf{h}_j^{'}\mathbf{y}_j^{*}}{\mathbf{h}_j^{'}\mathbf{h}_j+\sigma_{\delta_j}^{-2}},\frac{\sigma_e^2}{\mathbf{h}_j^{'}\mathbf{h}_j+\sigma_{\delta_j}^{-2}}\right) \qquad [A16]$$

in order to complete the joint collapsed sampler draw from $p\left(\sigma_{\delta_j}^2,\delta_j \mid ELSE,\mathbf{y}\right)$. One could demonstrate the following backward recursive relationship $\mathbf{h}_{j-1}=t_{j,j-1}\mathbf{h}_j+\mathbf{z}_{j-1}$, $j = m, m -2,\ldots,2$ with $\mathbf{z}_j$ denoting column j of **Z** and $\mathbf{h}_m = \mathbf{z}_m$. Hence for computational tractability, one could use this relationship in sampling pairs from $p\left(\sigma_{\delta_j}^2,\delta_j \mid ELSE,\mathbf{y}\right)$ starting with $j = m$ and working recursively backwards to j=1.

Sampling proportion of SNP markers associated with zero-effects in Ante-BayesB: The FCD of π_δ is based on the following:

$$p\left(\boldsymbol{\pi}_\delta \mid \mathbf{y},ELSE\right)\propto\prod_{j=1}^{m}p\left(\boldsymbol{\sigma}_{\delta_j}^2 \mid \boldsymbol{\nu}_\delta,s_\delta^2,\boldsymbol{\pi}_\delta\right)p\left(\boldsymbol{\pi}_\delta \mid \boldsymbol{\alpha}_\delta,\boldsymbol{\beta}_\delta\right) \qquad [A17]$$

where $p\left(\boldsymbol{\pi}_\delta \mid \boldsymbol{\alpha}_\delta,\boldsymbol{\beta}_\delta\right)=Beta\left(\boldsymbol{\alpha}_\delta,\boldsymbol{\beta}_\delta\right)$. Let $m_1=\sum_{j=1}^{m}I\left(\boldsymbol{\sigma}_{\delta_j}^2=0\right)$ denote the number of zero-valued elements sampled in $\boldsymbol{\sigma}_{\boldsymbol{\delta}}$ for a particular MCMC cycle where $I(.)$ denotes the indicator function. Then it can be readily demonstrated that Equation[A17] is simply
$p\left(\boldsymbol{\pi}_\delta \mid \mathbf{y},ELSE\right)=Beta\left(\boldsymbol{\alpha}_\delta+m_1,\boldsymbol{\beta}_\delta+m-m_1\right)$.

Sampling antedependence parameters and their corresponding hyperparameters: Consider now deriving the joint FCD of $\mathbf{t}=\left[t_{2,1},t_{3,2},\ldots,t_{m,m-1}\right]'$:

$$\begin{aligned}&p\left(\mathbf{t} \mid \mathbf{y},ELSE\right)\\ &\propto p\left(\mathbf{g} \mid t_{2,1},t_{3,2},\ldots,t_{m,m-1},\sigma_{\delta_1}^2,\sigma_{\delta_2}^2,\sigma_{\delta_3}^2,\ldots,\sigma_{\delta_G}^2\right)\left(\prod_{j=2}^{m}p\left(t_{j,j-1} \mid \mu_t,\sigma_t^2\right)\right)\end{aligned} \qquad [A18]$$

Borrowing developments, again from Chan and Jeliakov (2009, pg 462), the first component of [A18] can be rewritten as:

$$
\begin{aligned}
&p\left(\mathbf{g} \mid t_{21}, t_{32}, \ldots, t_{m,m-1}, \boldsymbol{\sigma}^2_{\delta_1}, \boldsymbol{\sigma}^2_{\delta_2}, \boldsymbol{\sigma}^2_{\delta_3}, \ldots, \boldsymbol{\sigma}^2_{\delta_m}\right) \\
&\propto \left|(\mathbf{I}-\mathbf{T})'\boldsymbol{\Delta}^{-1}(\mathbf{I}-\mathbf{T})\right|^{1/2} \exp\left(-\frac{1}{2}\mathbf{g}'(\mathbf{I}-\mathbf{T})'\boldsymbol{\Delta}^{-1}(\mathbf{I}-\mathbf{T})\mathbf{g}\right) \\
&\propto \mathrm{xp}\left(-\frac{1}{2}\frac{\left(g_2 - t_{21}g_1\right)^2}{\boldsymbol{\sigma}^2_{\delta_2}}\right)\exp\left(-\frac{1}{2}\frac{\left(g_3 - t_{32}g_2\right)^2}{\boldsymbol{\sigma}^2_{\delta_3}}\right)\ldots\exp\left(-\frac{1}{2}\frac{\left(g_m - t_{m,m-1}g_{m-1}\right)^2}{\boldsymbol{\sigma}^2_{\delta_m}}\right) \\
&\propto \exp\left(-\frac{1}{2}\left(\mathbf{g}_{(-1)} - \boldsymbol{\Psi}\mathbf{t}\right)'\boldsymbol{\Delta}^{-1}_{(-1)}\left(\mathbf{g}_{(-1)} - \boldsymbol{\Psi}\mathbf{t}\right)\right)
\end{aligned}
\qquad \text{[A19]}
$$

saving only terms that are functions of $\mathbf{t}$ with $\boldsymbol{\Psi} = diag\left(g_1, g_{2,} \ldots, g_{m-1}\right)$ being a diagonal *m*-1 x *m*-1 matrix with the listed elements, $\mathbf{g}_{(-1)} = \left[g_2 \quad g_3 \quad \ldots \quad g_m\right]'$, and $\boldsymbol{\Delta}_{(-1)} = diag\left(\boldsymbol{\sigma}^2_{\delta_2}, \boldsymbol{\sigma}^2_{\delta_3} \ldots, \boldsymbol{\sigma}^2_{\delta_m}\right)$ being a diagonal *m*-1 x *m*-1 matrix with the listed elements. Hence, [A18] can be rewritten as follows:

$$
\begin{aligned}
&p\left(\mathbf{t} \mid \mathbf{y}, ELSE\right) \propto p\left(\mathbf{g}_{(-1)} \mid \mathbf{t}, \boldsymbol{\Delta}_{(-1)}\right) p\left(\mathbf{t} \mid \mathbf{1}\mu_t, \mathbf{I}\sigma^2_t\right) \\
&\propto \left(\exp\left(-\frac{1}{2}\left(\mathbf{g}_{(-1)} - \Psi\mathbf{t}\right)'\boldsymbol{\Delta}^{-1}_{(-1)}\left(\mathbf{g}_{(-1)} - \Psi\mathbf{t}\right)\right)\right)\exp\left(-\frac{1}{2\sigma^2_t}(\mathbf{t} - \mathbf{1}\mu_t)'(\mathbf{t} - \mathbf{1}\mu_t)\right) \\
&\propto \left(\exp\left(-\frac{1}{2}\left(\mathbf{t} - \hat{\mathbf{t}}\right)'\boldsymbol{\Sigma}^{-1}_t\left(\mathbf{t} - \hat{\mathbf{t}}\right)\right)\right)
\end{aligned}
\qquad \text{[A20]}
$$

where

$$
\hat{\boldsymbol{\Sigma}}_t = \left(\Psi'\boldsymbol{\Delta}^{-1}_{(-1)}\Psi + \mathbf{I}\sigma^{-2}_t\right)^{-1} \qquad \text{[A21]}
$$

and

$$
\hat{\mathbf{t}} = \left(\Psi'\boldsymbol{\Delta}^{-1}_{(-1)}\Psi + \mathbf{I}\sigma^{-2}_t\right)^{-1}\left(\Psi'\boldsymbol{\Delta}^{-1}_{(-1)}\mathbf{g}_{(-1)} + \mathbf{1}\sigma^{-2}_t\mu_t\right) \qquad \text{[A22]}
$$

Note that $\Psi'\boldsymbol{\Delta}^{-1}_{(-1)}\Psi + \mathbf{I}\sigma^{-2}_t$ is diagonal with elements $\left(g_j\right)^2 \sigma^{-2}_{\delta_{j+1}} + \sigma^{-2}_t$, j = 1,2,…,*m*-1, whereas element j of $\Psi'\boldsymbol{\Delta}^{-1}_{(-1)}\mathbf{g}_{(-1)} + \mathbf{1}\sigma^{-2}_t\mu_t$ is $g_j g_{j+1}\sigma^{-2}_{\delta_{j+1}} + \sigma^{-2}_t\mu_t$, j = 1,2,…,*m*-1. In other words, the FCD of $t_{j+1,j}$ is $t_{j+1,j} \mid ELSE, \mathbf{y} \sim N\left(\hat{t}_{j+1,j}, \hat{\sigma}^{\;2}_{t(j+1,j)}\right)$ where

$$\hat{t}_{j+1,j} = \frac{g_j g_{j+1} \sigma^{-2}_{\delta_{j+1}} + \sigma_t^{-2} \mu_t}{\left(g_j\right)^2 \sigma^{-2}_{\delta_{j+1}} + \sigma_t^{-2}} \quad \text{[A23]}$$

and

$$\hat{\sigma}_{t(j+1,j)}^{\;2} = \left(\left(g_j\right)^2 \sigma^{-2}_{\delta_{j+1}} + \sigma_t^{-2}\right)^{-1} \quad \text{[A24]}$$

Note further that $\hat{t}_{j+1,j}$ can be written as a weighted average:

$$\hat{t}_{j+1,j} = \frac{g_j g_{j+1} \sigma^{-2}_{\delta_{j+1}} + \sigma_t^{-2} \mu_t}{\sigma^{-2}_{\delta_{j+1}} \left(g_j\right)^2 + \sigma_t^{-2}} = \frac{\sigma^{-2}_{\delta_{j+1}} \left(g_j\right)^2}{\sigma^{-2}_{\delta_{j+1}} \left(g_j\right)^2 + \sigma_t^{-2}} \frac{g_{j+1}}{g_j} + \frac{\sigma_t^{-2}}{\sigma^{-2}_{\delta_{j+1}} \left(g_j\right)^2 + \sigma_t^{-2}} \mu_t \quad \text{[A25]}$$

Now with $g_j = 0$, as one might anticipate occasionally with ante-BayesB with markers defined at the beginning of a linkage group, $\hat{t}_{j+1,j} = \mu_t$ and $\hat{\sigma}_{t(j+1,j)}^{\;2} = \sigma_t^2$ such that one draws $t_{j+1,j}$ from its prior density based on updated values of μ_t and σ_t^2. For the much more common situation in ante-BayesB (assuming large $\boldsymbol{\pi}_{\boldsymbol{\delta}}$) where $g_j \neq 0$ but $\sigma^2_{\delta_{j+1}} = 0$, the FCD of $t_{j+1,j}$ can be shown to be a point mass on $\frac{g_{j+1}}{g_j}$.

With $p\left(\mu_t\right)$ specified to be normal with prior mean μ_{t0} and prior variance s^2_{t0} then Gibbs sampling can be used for the corresponding parameters.

$$p\left(\mu_t, | \, \mathbf{y,ELSE}\right) = N\left(\tilde{\mu}_t, \tilde{\sigma}_t^2\right) \quad \text{[A26]}$$

where

$$\tilde{\mu}_t = \frac{\frac{m-1}{\sigma_t^2}\bar{t} + \frac{1}{s_{t0}^2}\mu_{t0}}{\frac{1}{s_{t0}^2} + \frac{m-1}{\sigma_t^2}} \quad \text{[A27]}$$

for $\bar{t} = \frac{\sum_{2}^{m} t_{j,j-1}}{m-1}$ and

$$\tilde{\sigma}_t^2 = \left(\frac{1}{s_{t0}^2} + \frac{m-1}{\sigma_t^2} \right)^{-1} \quad [A28]$$

The FCD of σ_t^2 given that the prior $p\left(\sigma_t^2 \mid \nu_t, s_t^2\right)$ is scaled inverted chi-square with known hyperparameters ν_t and s_t^2 can be derived as follows:

$$\begin{aligned} & p\left(\sigma_t^2, \mid \mathbf{y}, ELSE\right) \\ & \propto \left(\prod_{j=2}^{m} p\left(t_{j,j-1} \mid \mu_t, \sigma_t^2\right) \right) p\left(\sigma_t^2 \mid \nu_t, s_t^2\right) \\ & \propto \left(\left(2\pi\sigma_t^2\right)^{-(m-1)/2} \prod_{j=2}^{m} \exp\left(-\frac{1}{2\sigma_t^2} \sum_{j=2}^{m} \left(t_{j,j-1} - \mu_t\right)^2 \right) \right) \sigma_t^{2^{-\left(\frac{\nu_t}{2}+1\right)}} e^{-\frac{\nu_t s_t^2}{2\sigma_t^2}} \\ & \propto \left(\left(\sigma_t^2\right)^{-\left(\frac{\nu_t+m-1}{2}+1\right)} \prod_{j=1}^{G} \exp\left(-\frac{1}{2\sigma_t^2} \left(\sum_{j=2}^{m} \left(t_{j,j-1} - \mu_t\right)^2 + \nu_t s_t^2 \right) \right) \right) \end{aligned} \quad [A29]$$

That is, $p\left(\sigma_t^2, \mid \mathbf{y}, ELSE\right) = \chi^{-2}\left(m + \nu_t, \sum_{j=2}^{m} \left(t_{j,j-1} - \mu_t\right)^2 + \nu_t s_t^2 \right)$. Note that we advocate the non-informative specifications $\nu_t = -1$ and $s_t^2 = 0$ in the paper.

Sampling the scale parameter for the random SNP effects: Borrowing results from Yi and Xu (2008), the FCD for s_δ^2 based on the specification of a conjugate prior $p\left(s_\delta^2 \mid \alpha_s, \beta_s\right)$ = Gamma $\left(\alpha_s, \beta_s\right)$ can be written as follows:

$$\begin{aligned} & p\left(s_\delta^2 \mid \mathbf{y}, ELSE\right) \propto \left(\prod_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right) p\left(\sigma_{\delta_j}^2 \mid \nu_\delta, s_\delta^2\right) \right) p\left(s_\delta^2 \mid \alpha_s, \beta_s\right) \\ & = \left(\prod_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right) \frac{\left(\frac{s_\delta^2 \nu_\delta}{2}\right)^{\frac{\nu_\delta}{2}}}{\Gamma\left(\frac{\nu_\delta}{2}\right)} \sigma_{\delta_j}^{2^{-\left(\frac{\nu_\delta}{2}+1\right)}} e^{-\frac{s_\delta^2 \nu_\delta}{2\sigma_{\delta_j}^2}} \right) \frac{\left(\beta_s\right)^{\alpha_s}}{\Gamma\left(\alpha_s\right)} \left(s_\delta^2\right)^{\alpha_s - 1} e^{-\beta_s s_\delta^2} \\ & \propto \left(s_\delta^2\right)^{\alpha_s + \frac{m_1 \nu_\delta + 1}{2} - 1} \exp\left(-s_\delta^2 \left(\frac{\nu_\delta}{2} \sum_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right) \sigma_{\delta_j}^{-2} + \beta_s \right) \right) \end{aligned} \quad [A30]$$

i.e., a Gamma distribution with parameters $\alpha_s + \frac{m_1\nu_\delta+1}{2}$ and $\frac{\nu_\delta}{2}\sum_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right)\sigma_{\delta_j}^{-2} + \beta_s$.

Sampling the degrees of freedom parameter for the random SNP effects: Simple Metropolis updates could be used for sampling ν_δ For an arbitrary prior $p(\nu_\delta)$, the corresponding FCD is as follows:

$$p(\nu_\delta \mid ELSE) \propto \left(\prod_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right) p\left(\sigma_{\delta_j}^2 \mid \nu_\delta, s_\delta^2\right)\right) p(\nu_\delta)$$
$$= \left(\prod_{j=1}^{m} I\left(\sigma_{\delta_j}^2 > 0\right) \frac{\left(\frac{\nu_\delta s_\delta^2}{2}\right)^{\frac{\nu_\delta}{2}}}{\Gamma\left(\frac{\nu_\delta}{2}\right)} \sigma_{g_j}^{2^{-\left(\frac{\nu_\delta}{2}+1\right)}} e^{-\frac{\nu_\delta s_\delta^2}{2\sigma_{g_j}^2}}\right) p(\nu_\delta) \qquad \text{[A31]}$$

Details on how to ν_δ can be based on a random walk Metropolis Hastings step; we have provided details on this in other non-genomic applications involving the sampling of degrees of freedom parameters (BELLO *et al.* 2010; KIZILKAYA and TEMPELMAN 2005).

Sampling the residual variance: Given a specified scaled inverted chi-square prior $p\left(\sigma_e^2, \mid \alpha_e, s_e^2\right) = \chi^{-2}\left(\alpha_e, \alpha_e s_e^2\right)$, the corresponding FCD of σ_e^2 can be written as follows:

$$p\left(\sigma_e^2 \mid \mathbf{y}, ELSE\right)$$
$$\propto \left(\left(2\pi\sigma_e^2\right)^{-n/2}\right)\exp\left(-\frac{1}{2\sigma_e^2}(\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})'(\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})\right)\sigma_e^{2^{-\left(\frac{\nu_e}{2}+1\right)}} e^{-\frac{\nu_e s_e^2}{2\sigma_e^2}}$$
$$\propto \sigma_e^{2^{-\left(\frac{\nu_e+n}{2}+1\right)}} \exp\left(-\frac{1}{2\sigma_e^2}\left((\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})'(\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})+\nu_e s_e^2\right)\right) \qquad \text{[A32]}$$

In other words, [A32] is $\chi^{-2}\left(\nu_e + n, (\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})'(\mathbf{y}\text{-}\mathbf{X}\boldsymbol{\beta}-\mathbf{Zg}-\mathbf{Wu})+\nu_e s_e^2\right)$. Note that we advocate the non-informative specifications $\nu_e = -1$ and $s_e^2 = 0$ in the paper.

Sampling the polygenic variance: Given a conjugate scaled inverted-chi square prior $p\left(\sigma_u^2, \mid \alpha_u, s_u^2\right) = \chi^{-2}\left(\alpha_u, \alpha_u s_u^2\right)$ The FCD of σ_e^2 is classically given as follows:

$$p\left(\sigma_u^2 \mid \mathbf{y},\mathbf{ELSE}\right) \propto \left(\sigma_u^2\right)^{-\left(\frac{\nu_u+q}{2}+1\right)} \exp\left(-\frac{1}{2}\frac{\left(\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}+\nu_u s_u^2\right)}{\sigma_u^2}\right) \quad \text{[A33]}$$

In other words, [A33] is $\chi^{-2}\left(\nu_u+q,\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}+\nu_u s_u^2\right)$. Note that we advocate the non-informative specifications $\nu_u=-1$ and $s_u^2=0$ in the paper.

References for File S1

BELLO, N. M., J. P. STEIBEL and R. J. TEMPELMAN, 2010 Hierarchial Bayesian modeling of random and residual variance-covariance matrices in mixed effects models. Biometrical Journal **52:** 297-313.

CHAN, J. C. C., and I. JELIAZKOV, 2009 MCMC Estimation of Restricted Covariance Matrices. Journal of Computational and Graphical Statistics **18:** 457-480.

CHIB, S., and E. GREENBERG, 1995 Understanding the Metropolis-Hastings Algorithm. American Statistician **49:** 327-335.

KIZILKAYA, K., and R. J. TEMPELMAN, 2005 A general approach to mixed effects modeling of residual variances in generalized linear mixed models. Genetics Selection Evolution **37:** 31-56.

LIU, J. S., 1994 The Collapsed Gibbs Sampler in Bayesian Computations with Applications to a Gene-Regulation Problem. Journal of the American Statistical Association **89:** 958-966.

MEUWISSEN, T. H. E., B. J. HAYES and M. E. GODDARD, 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics **157:** 1819-1829.

SORENSEN, D., and D. GIANOLA, 2002 *Likelihood, Bayesian, and MCMC methods in quantitative genetics*. Springer-Verlag, New York.

WANG, C. S., J. J. RUTLEDGE and D. GIANOLA, 1994 Bayesian-Analysis of Mixed Linear-Models Via Gibbs Sampling with an Application to Litter Size in Iberian Pigs. Genetics Selection Evolution **26:** 91-115.

YI, N., and S. XU, 2008 Bayesian Lasso for quantitative trait loci mapping. Genetics **179:** 1045 - 1055.

ZIMMERMAN, D. L., and V. A. NUNOZ-ANTON, 2010 *Antedependence Models for Longitudinal Data*. Chapman and Hall/CRC.

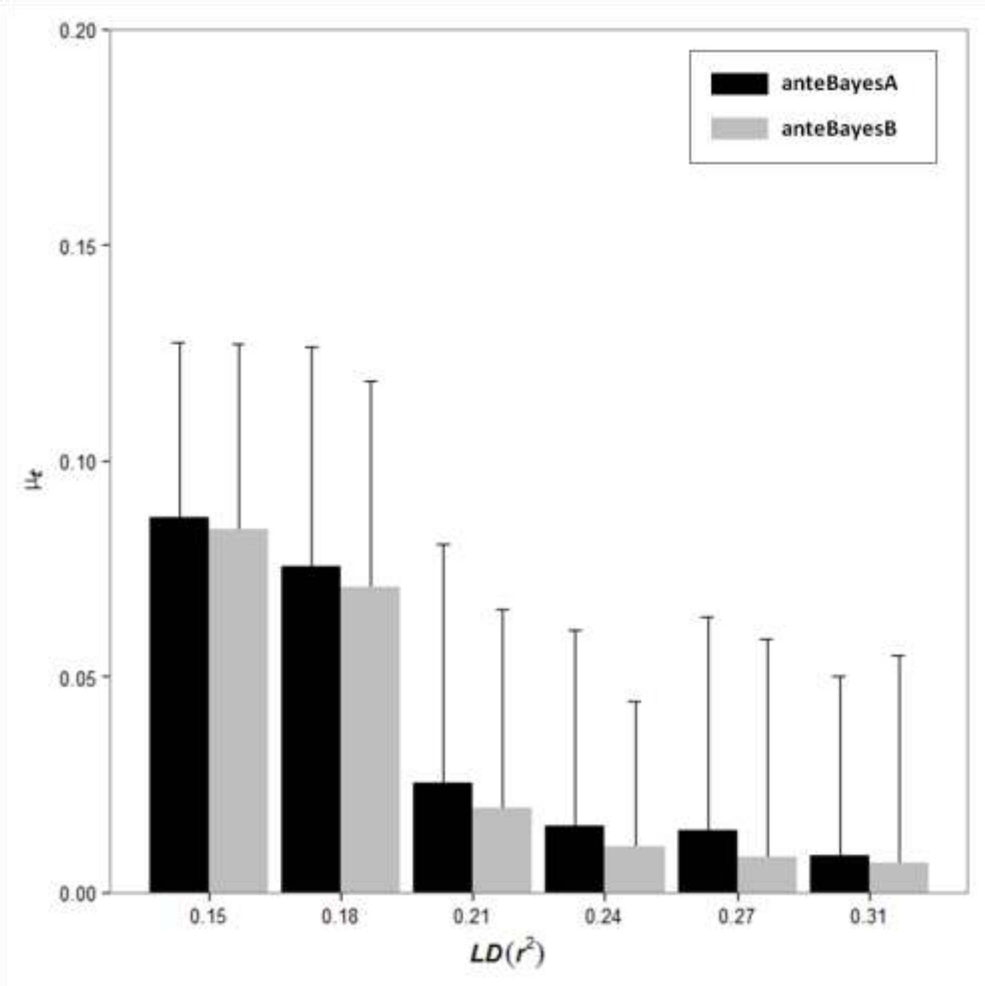


FIGURE S1. Average posterior means of μ_t and empirical standard errors across 20 replicates for each of six different LD levels using ante-BayesA and ante-BayesB. No significant differences (P>.01) were determined between the competing procedures with each other or from zero at each LD level.

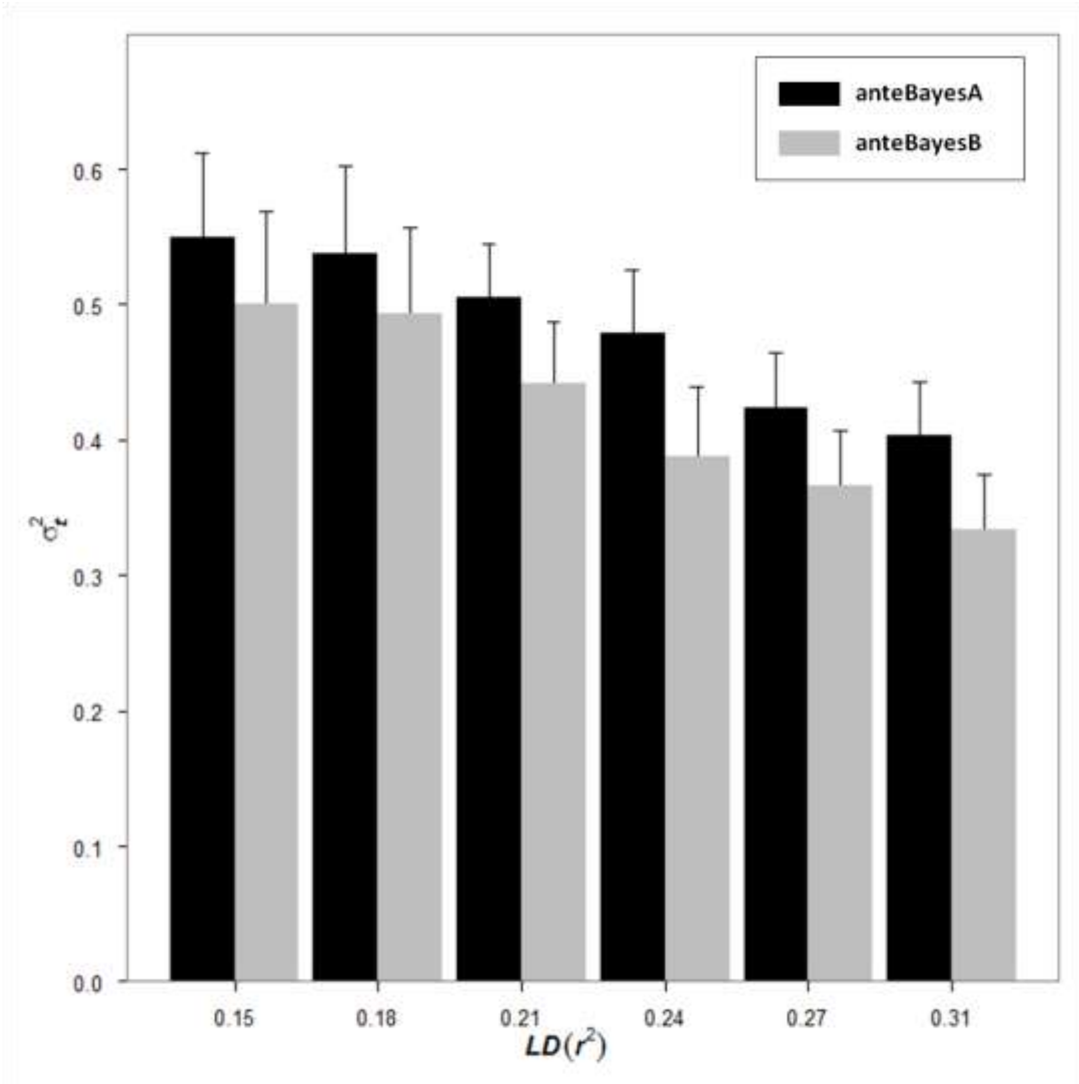


FIGURE S2. Average posterior means of σ_t^2 and empirical standard errors across 20 replicates for each of six different LD levels using ante-BayesA and ante-BayesB. No significant differences (P>.01) were determined between the two sets of competing procedures at each LD level.

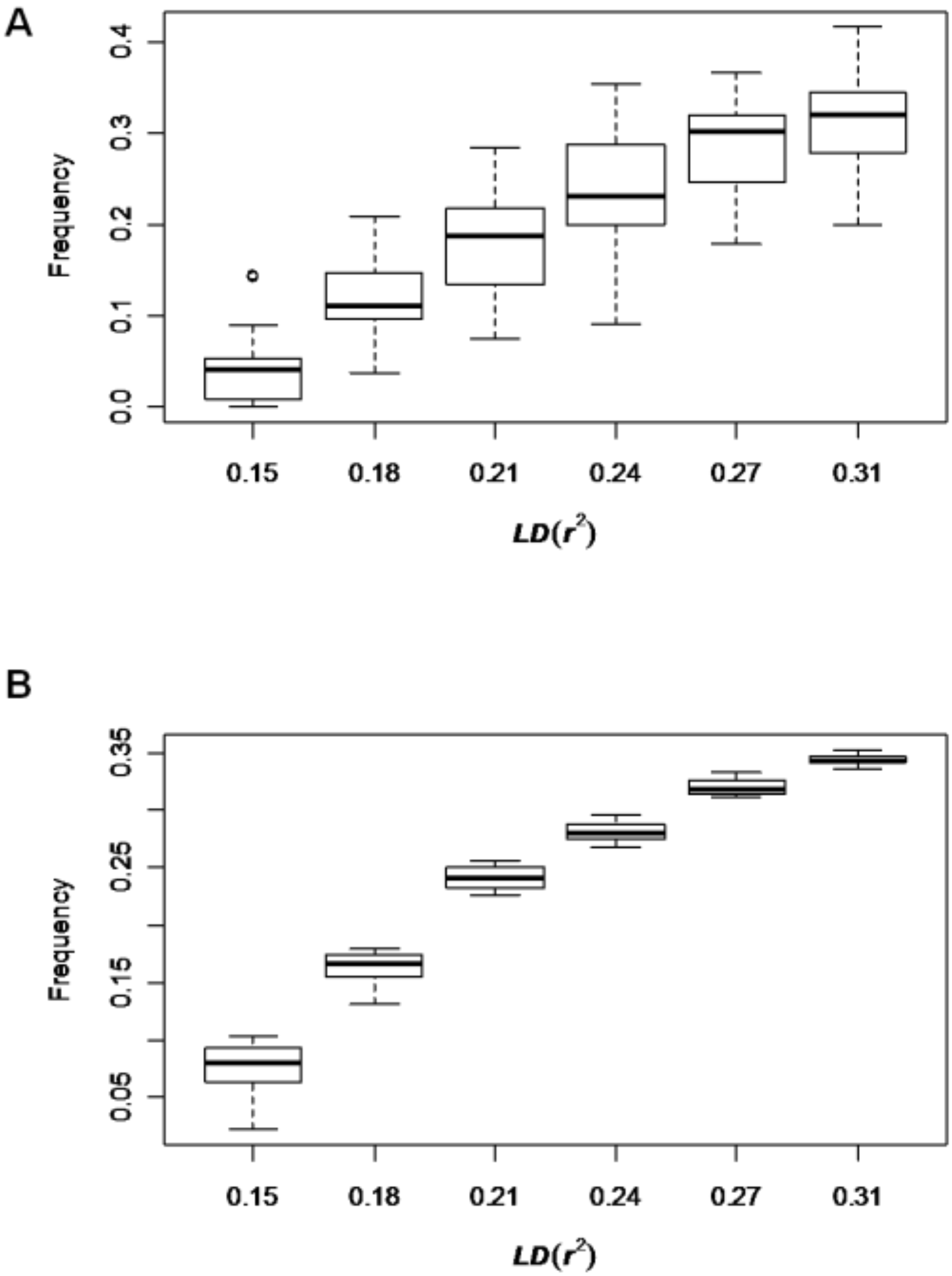


FIGURE S3: Box-plot of proportions of the absolute posterior means of elements of $\left\{t_{j,j-1}\right\}_{j=2}^{m}$ divided by their respective posterior standard deviations that exceeded 2 across all 20 replicates for each of six different levels of LD using ante-BayesA (A) and ante-BayesB (B).

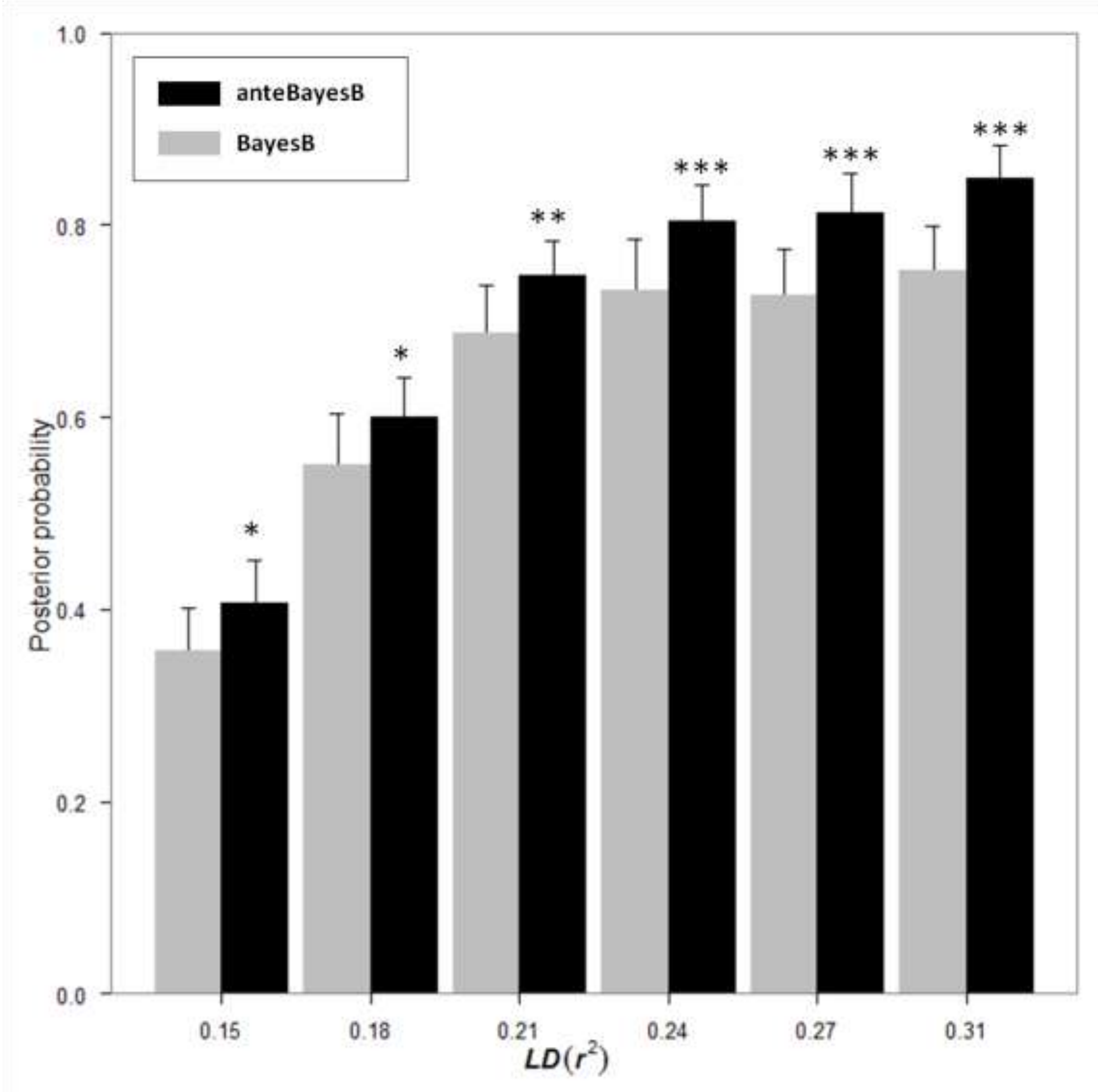
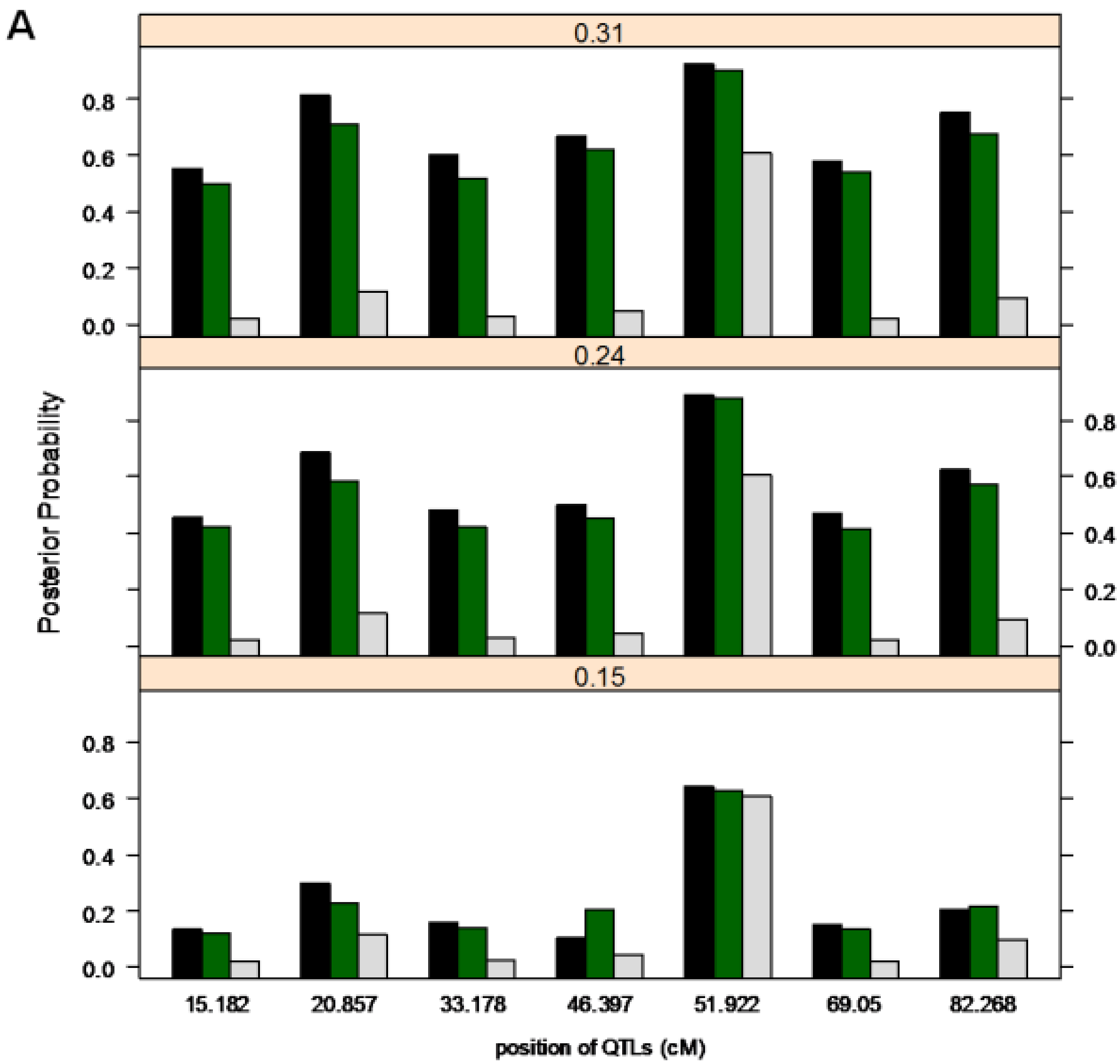
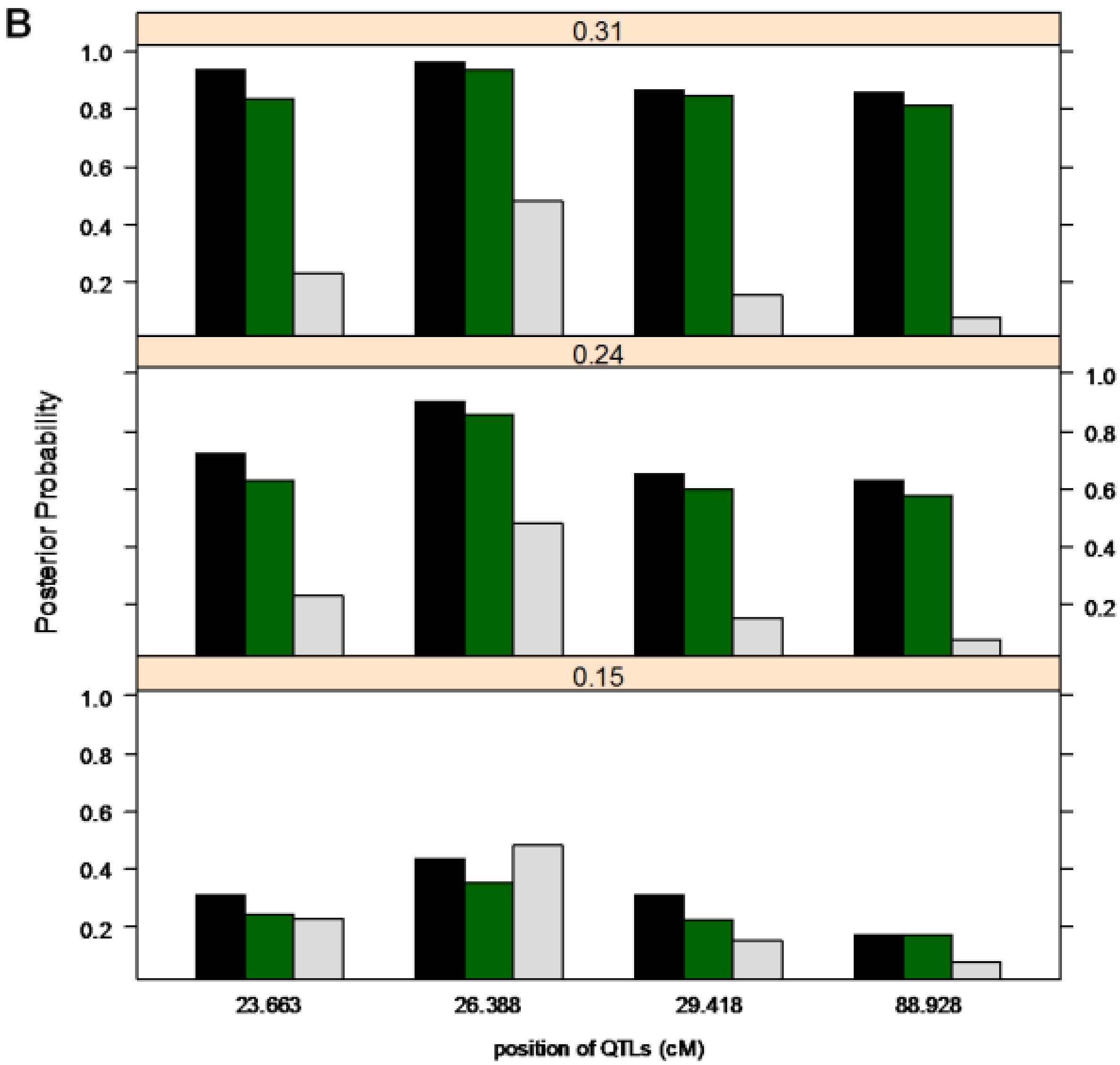


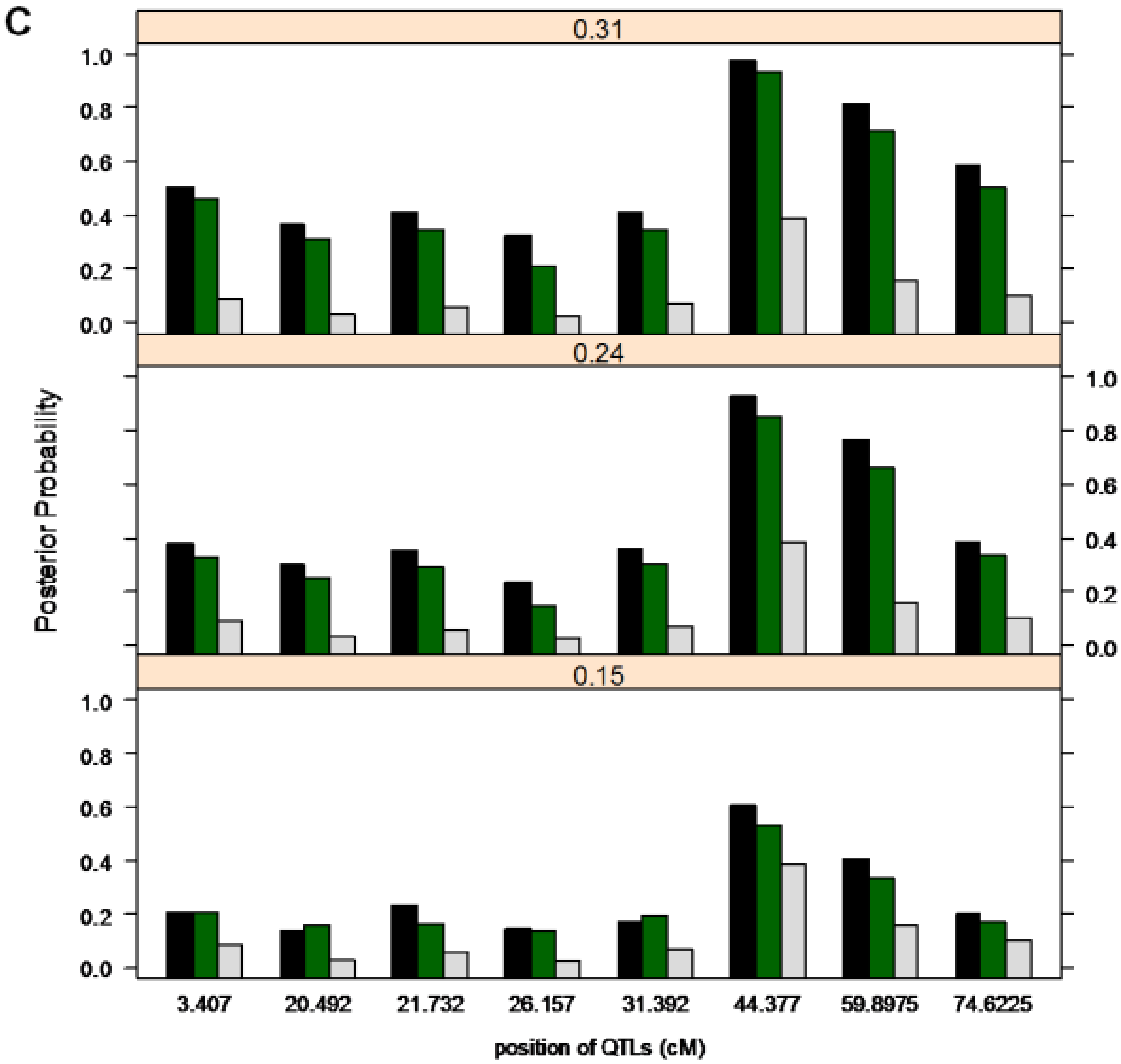
FIGURE S4. Average posterior probabilities of association for the top QTL within each of 20 replicates using BayesB and ante-BayesB for each of six different LD levels. LD-specific differences between the two methods declared significant by *(P<0.01), **(P <0.001), or ***(P<0.0001).


A
0.31
0.24
0.15
0.8
0.6
0.4
0.2
0.0
Posterior Probability
15.182
20.857
33.178
46.397
51.922
69.05
82.268
position of QTLs (cM)

1


B
0.31
0.24
0.15
1.0
0.8
0.6
0.4
0.2
Posterior Probability
23.663
26.388
29.418
88.928
position of QTLs (cM)

1


C
0.31
0.24
0.15
1.0
0.8
0.6
0.4
0.2
0.0
Posterior Probability
3.407
20.492
21.732
26.157
31.392
44.377
59.8975
74.6225
position of QTLs (cM)

1

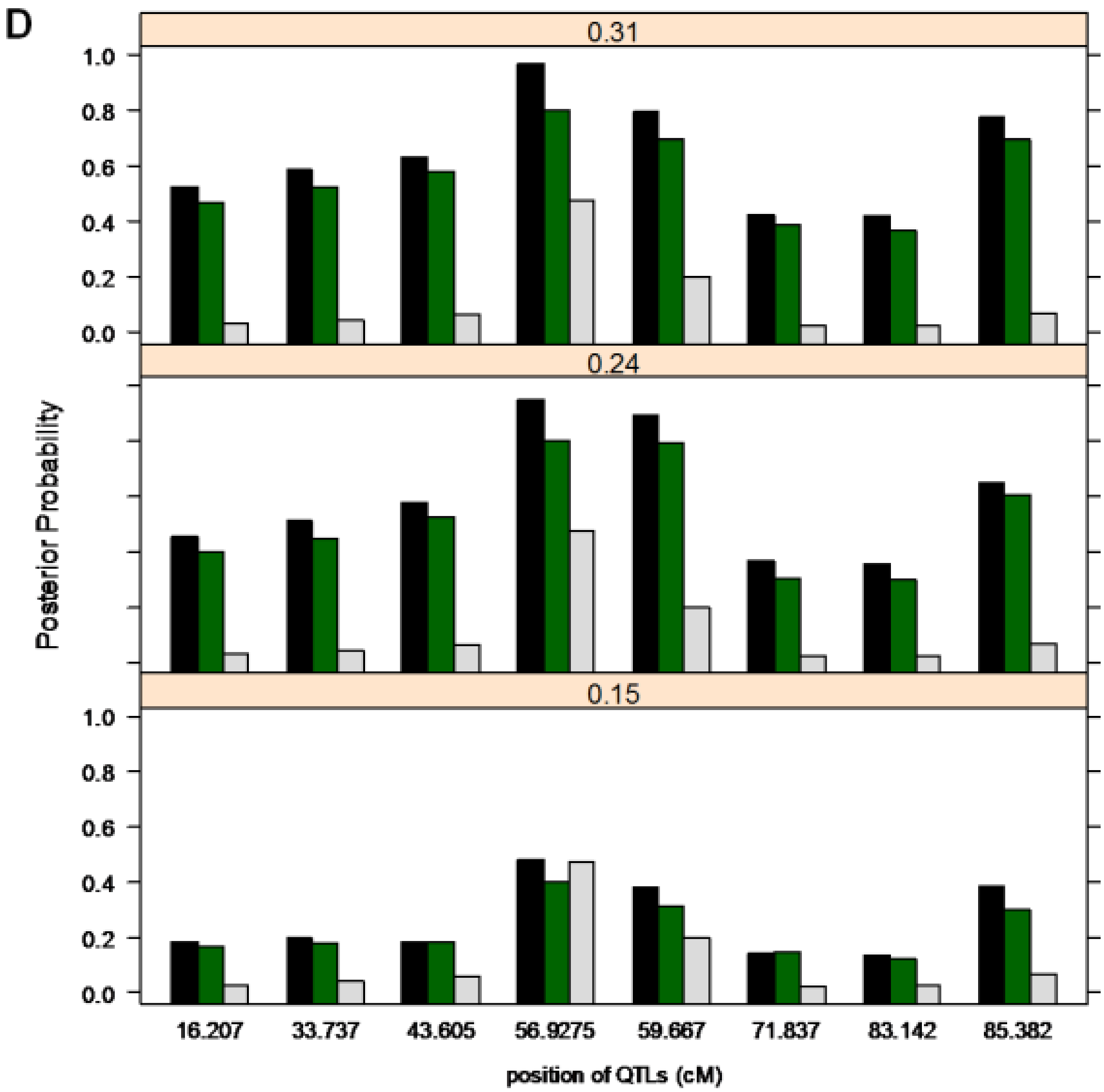


FIGURE S5. Bar plots of posterior probabilities of association of either or both of two bracketing SNP to each of the six largest QTL effects within each of the first four replicates (A,B,C,D) at the highest (r^2=0.31), medium (r^2=0.24) and lowest (r^2=0.15) average LD levels.. Posterior probabilities using BayesB and ante-BayesB are represented by green and black bars, respectively, whereas gray bars represent the proportion of the genetic variance accounted for by the corresponding QTL. QTL location is labeled on x-axis for each replicate.

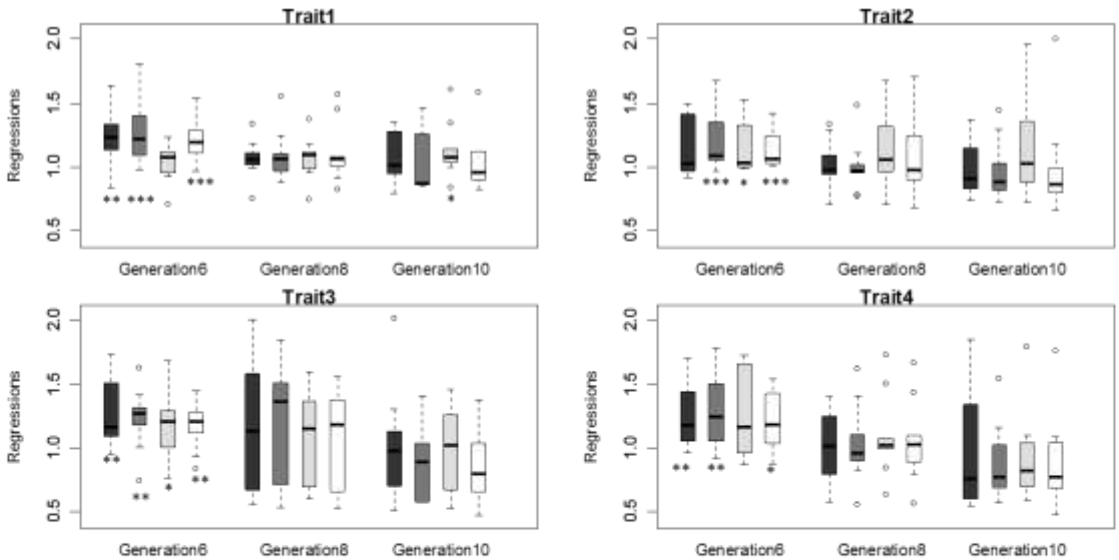


FIGURE S6: Boxplots of estimated slopes for within-replicate regressions of true breeding values on estimated breeding values across 9 replicates for four traits in Generations 6, 8 and 10 from benchmark data of Hickey and Gorjanc (2011) using ante-BayesB (black), BayesB (dark gray), anteBayesA (light gray) and BayesA (white). Differences from unity indicated as significant by *(0.05<*P*<.10), **(0.01<*P*<.05) or ***(*P*<.01).

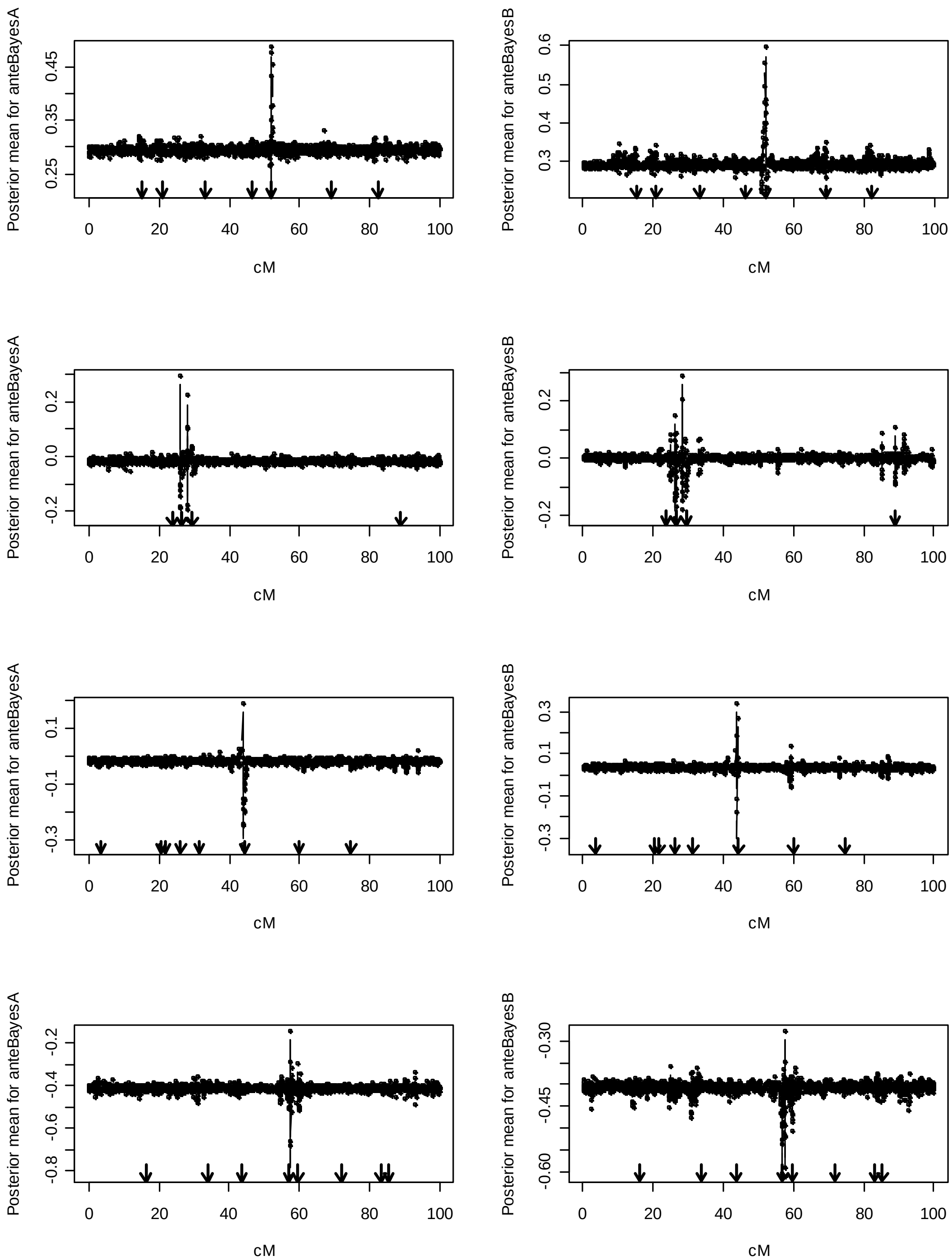


FIGURE S7. Posterior means of antedependence parameters $\left\{t_{j,j-1}\right\}_{j=2}^{m}$ versus corresponding SNP bracket location based on anteBayesA (left column) and anteBayesB (right column) for each of the first four replicates (rows 1 through 4) based on analyses using highest average marker density ($r^2 = 0.31$) Arrows denote the position for any QTL that accounted for greater than 2% of the total variance in each replicate.

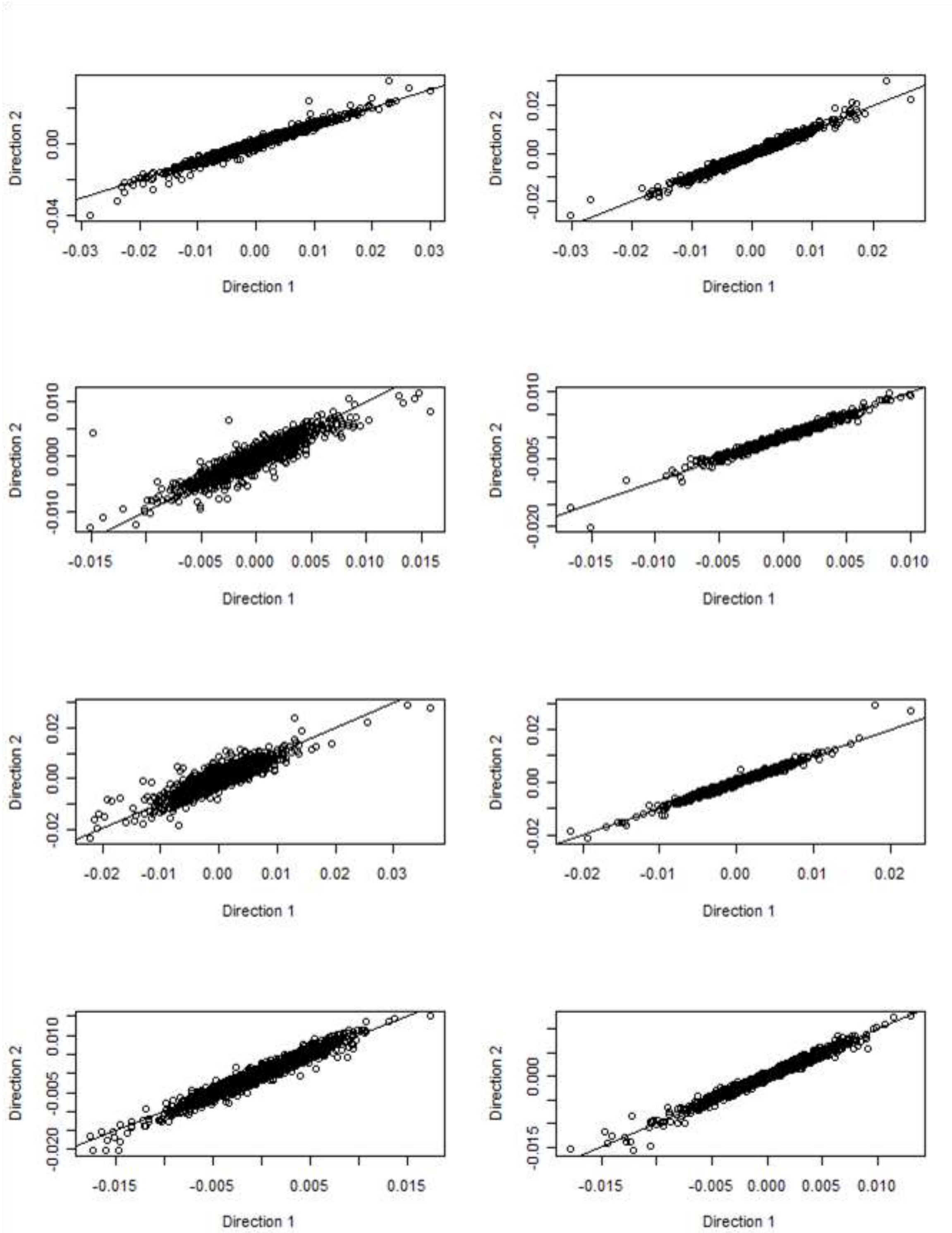


FIGURE S8. Posterior means of **g** using ante-BayesA (left-column) and ante-BayesB (right-column) based on specifying antedependence in one direction along the chromosome against corresponding posterior means based on the same analyses but specifying antedependence in the opposite direction for each of the first four replicates (rows 1 through 4) and the highest average marker density ($r^2 = 0.31$). Reference lines of intercept 0 and slope 1 are superimposed.

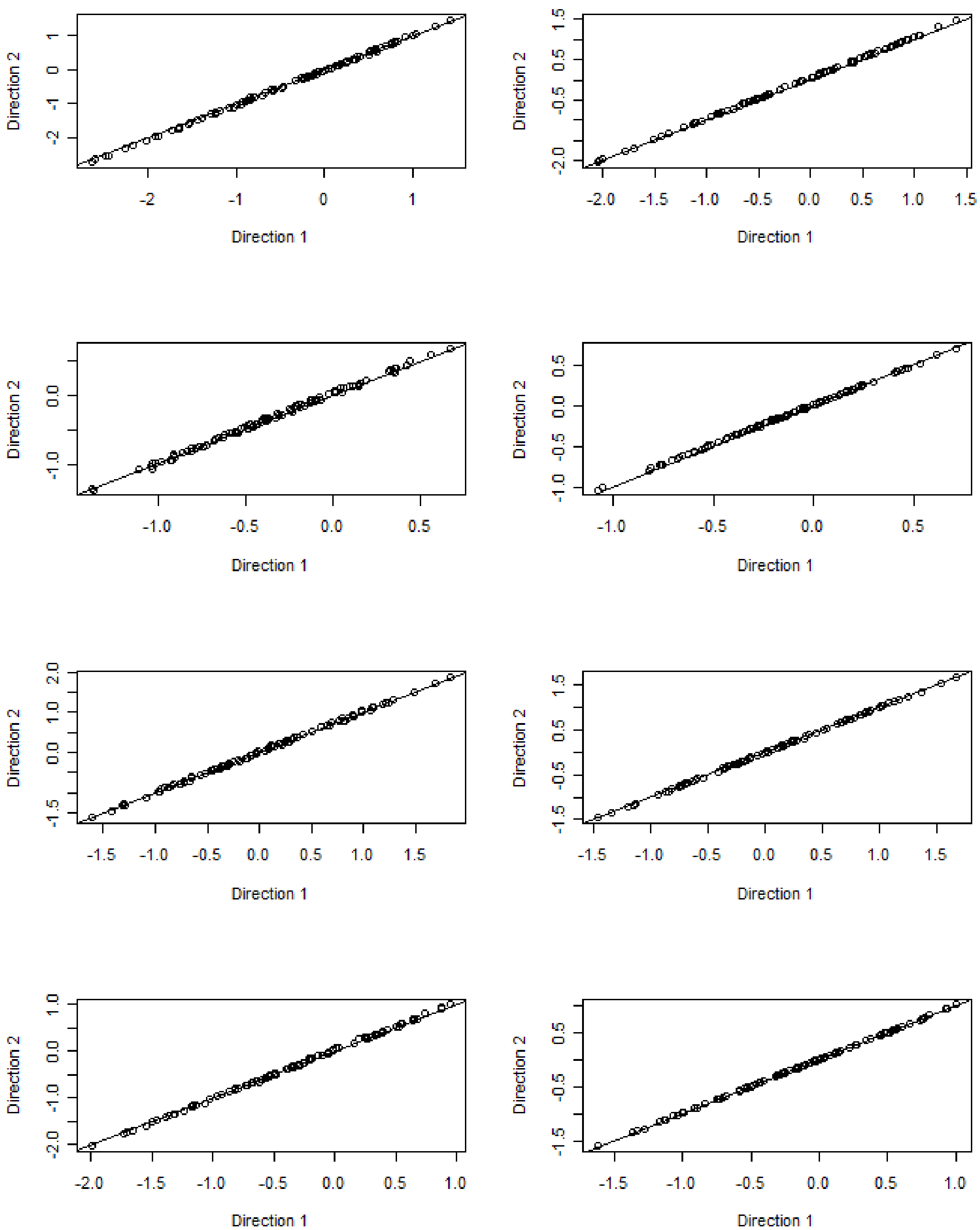


FIGURE S9. Posterior means of EBV using ante-BayesA (left-column) and ante-BayesB (right-column) based on specifying antedependence in one direction along the chromosome against corresponding posterior means based on the same analyses but specifying antedependence in the opposite direction for each of the first four replicates (rows 1 through 4) and the highest average marker density ($r^2 = 0.31$). Reference lines of intercept 0 and slope 1 are superimposed.

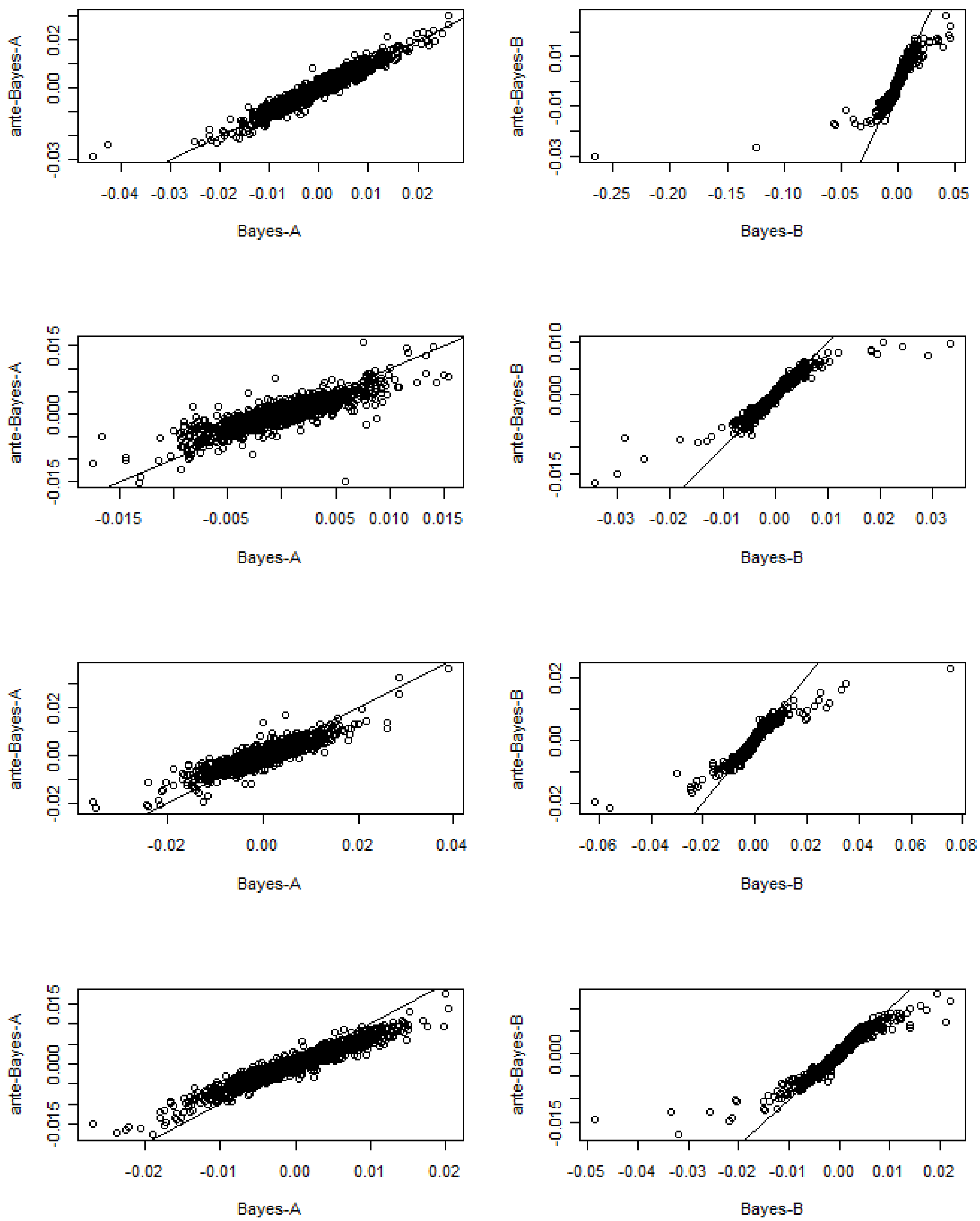


FIGURE S10. Posterior means of **g** based on ante-BayesA versus BayesA (left-column) and ante-BayesB versus BayesB (right-column) for each of the first four replicates (rows 1 through 4) and the highest marker density ($r^2 = 0.31$). Reference lines of intercept 0 and slope 1 are superimposed.

TABLE S1

Average posterior means (PMEAN), posterior standard deviations (PSD), posterior medians (PMED), and effective sample size (ESS) for residual variance (σ_e^2), cage variance (σ_c^2), polygenic variance (σ_u^2) and key hyperparameters (ν_g, s_g^2, and π_g) based on BayesA and BayesB analyses of training data subsets derived from 10 different partitions of the heterogeneous stock mice dataset. Empirical standard deviations across the 10 partitions are provided in parentheses.

Parameter	PMEAN	PSD	PMED	ESS
BayesA				
σ_e^2	0.32(0.05)	0.15(0.003)	0.31(0.05)	6126(283)
σ_c^2	2.08(0.13)	0.31(0.014)	2.05(0.12)	6048(247)
σ_u^2	3.21(0.15)	0.49(0.018)	3.18(0.13)	5728(213)
ν_g	16.12(1.83)	23.11(2.72)	7.62(0.56)	285(20)
s_g^2	0.002(0.00014)	0.0007(0.0001)	0.002(0.00014)	266(17)
BayesB				
σ_e^2	0.34(0.04)	0.13(0.002)	0.34(0.04)	12745(1213)
σ_c^2	2.03(0.12)	0.35(0.011)	2.04(0.13)	11642(1086)
σ_u^2	3.25(0.17)	0.53(0.021)	3.25(0.16)	9892(924)
ν_g	19.31(1.03)	24.14(1.09)	9.47(0.68)	536(65)
s_g^2	0.02(0.003)	0.0006(0.003)	0.02(0.0004)	493(48)
π_g	0.81(0.03)	0.01(0.004)	0.81(0.03)	517(61)

TABLE S2

Average posterior means (PMEAN), posterior standard deviations (PSD), posterior medians (PMED), and effective sample size (ESS) for residual variance (σ_e^2), cage variance (σ_c^2), polygenic variance (σ_u^2) and key hyperparameters (ν_δ , s_δ^2 , π_g , μ_t and σ_t^2) based on ante-BayesA and ante-BayesB analyses of training data subsets derived from 10 different partitions of the heterogeneous stock mice dataset. Empirical standard deviations across the 10 partitions are provided in parentheses.

Parameter	PMEAN	PSD	PMED	ESS
Ante-BayesA				
σ_e^2	0.32(0.04)	0.14(0.002)	0.32(0.03)	2517(134)
σ_c^2	2.01(0.13)	0.34(0.012)	2.02(0.13)	2483(128)
σ_u^2	3.22(0.14)	0.53(0.03)	3.21(0.14)	2357(122)
ν_δ	15.54(1.32)	21.23(1.54)	7.27(0.46)	185(26)
s_δ^2	0.001(0.0004)	0.0006(0.00005)	0.001(0.0004)	144(20)
μ_t	0.030(0.002)	0.0139(0.004)	0.030(0.001)	2561(52)
σ_t^2	0.037(0.004)	0.010(0.0003)	0.034(0.004)	986(43)
Ante-BayesB				
σ_e^2	0.34(0.04)	0.15(0.004)	0.33(0.04)	9512(765)
σ_c^2	2.02(0.12)	0.39(0.017)	2.01(0.10)	9134(726)
σ_u^2	3.23(0.15)	0.52(0.03)	3.22(0.14)	9038(689)
ν_δ	18.21(1.04)	22.46(1.22)	9.26(0.74)	376(33)
s_δ^2	0.14(0.05)	0.32(0.18)	0.07(0.15)	259(21)
π_δ	0.80(0.04)	0.01(0.004)	0.79(0.03)	343(32)
μ_t	0.02(0.003)	0.011(0.002)	0.02(0.004)	1480(49)
σ_t^2	0.032(3e-3)	0.05(2e-4)	0.031(3e-3)	638(41)

TABLE S3

Average posterior means for residual variance (σ_e^2), polygenic variance (σ_u^2), key hyperparameters (ν_g , s_g^2 , and π_g) based on BayesA and BayesB analyses and key hyperparameters (ν_δ , s_δ^2 , π_g , μ_t and σ_t^2) based on ante-BayesB and ante-BayesA analyses for 4 different traits using Hickey and Gorjanc (2011) benchmark data. Empirical standard deviations across 9 replicates are provided in parentheses.

	Trait 1	Trait 2	Trait 3	Trait 4
anteBayesB				
σ_e^2	0.834(0.019)	1.348(0.125)	0.593(0.015)	0.939(0.038)
σ_u^2	0.131(0.017)	0.161(0.033)	0.101(0.010)	0.144(0.020)
ν_δ	28.043(1.841)	15.707(2.430)	22.941(2.158)	17.792(2.953)
s_δ^2	7.13e-4(4.63e-5)	1.34e-3(1.74e-4)	4.40e-4(5.30e-5)	9.14e-4(8.21e-5)
π_δ	0.789(0.016)	0.834(0.015)	0.771(0.017)	0.826(0.019)
μ_t	0.032(0.011)	0.033(0.009)	0.011(0.015)	0.013(0.016)
σ_t^2	0.038(0.002)	0.034(0.001)	0.031(0.006)	0.046(0.009)
BayesB				
σ_e^2	0.833(0.020)	1.339(0.123)	0.591(0.015)	0.937(0.040)
σ_u^2	0.108(0.017)	0.141(0.034)	0.085(0.010)	0.114(0.012)
ν_g	31.643(1.623)	23.475(3.750)	33.536(2.501)	27.203(3.674)
s_g^2	8.23e-4(5.60e-5)	1.97e-3(3.21e-4)	5.31e-4(6.75e-5)	4.91e-3(3.71e-3)
π_g	0.826(0.011)	0.861(0.015)	0.823(0.017)	0.871(0.020)
anteBayesA				
σ_e^2	0.827(0.020)	1.339(0.124)	0.590(0.015)	0.935(0.040)
σ_u^2	0.157(0.030)	0.208(0.051)	0.143(0.018)	0.169(0.030)
ν_δ	22.556(1.192)	12.501(2.840)	20.909(1.194)	14.941(3.003)
s_δ^2	6.40e-5(1.45e-5)	6.84e-5(1.91e-5)	2.22e-5(1.06e-5)	4.85e-5(1.81e-5)
μ_t	0.021(0.013)	0.007(0.020)	0.002(0.016)	0.021(0.025)
σ_t^2	0.035(0.003)	0.022(0.005)	0.032(0.002)	0.027(0.004)
BayesA				
σ_e^2	0.829(0.019)	1.338(0.124)	0.592(0.015)	0.939(0.040)
σ_u^2	0.099(0.015)	0.123(0.034)	0.071(0.009)	0.097(0.012)
ν_g	23.871(2.033)	14.639(3.147)	23.032(2.832)	15.882(3.812)
s_g^2	1.05e-4(8.70e-6)	1.22e-4(1.68e-5)	6.33e-5(7.23e-6)	8.30e-5(1.61e-5)

TABLE S4

Average effective sample size for residual variance (σ_e^2), polygenic variance (σ_u^2), key hyperparameters (ν_g, s_g^2, and π_g) based on BayesA and BayesB analyses and key hyperparameters (ν_δ, s_δ^2, π_g, μ_t and σ_t^2) based on ante-BayesB and ante-BayesA analyses for 4 different traits using Hickey and Gorjanc (2011) benchmark data. Empirical standard deviations across 9 replicates are provided in parentheses.

	Trait 1	Trait 2	Trait 3	Trait 4
anteBayesB				
σ_e^2	4910(879)	4883(757)	6139(1065)	4163(599)
σ_u^2	1423(317)	905(131)	1164(174)	961(177)
ν_δ	136(12)	157(25)	125(10)	146(24)
s_δ^2	108(7)	111(9)	113(12)	121(19)
π_δ	216(19)	236(24)	210(20)	249(25)
μ_t	168(29)	184(35)	150(26)	132(27)
σ_t^2	241(46)	302(63)	207(45)	225(53)
BayesB				
σ_e^2	3058(541)	4814(1181)	3462(555)	3354(344)
σ_u^2	1056(238)	968(284)	993(154)	836(85)
ν_g	137(6)	162(18)	127(5)	128(13)
s_g^2	138(4)	122(28)	119(16)	126(16)
π_g	264(36)	347(41)	297(26)	275(39)
anteBayesA				
σ_e^2	4887(1131)	6352(1900)	9318(1823)	5667(1638)
σ_u^2	426(64)	1429(1026)	782(225)	810(321)
ν_δ	105(4)	164(48)	101(8)	122(25)
s_δ^2	114(9)	120(34)	107(20)	118(24)
μ_t	137(10)	110(4)	101(18)	107(10)
σ_t^2	191(25)	273(66)	201(25)	259(51)
BayesA				
σ_e^2	2794(342)	4212(1053)	2986(402)	2541(200)
σ_u^2	677(93)	566(111)	642(77)	494(38)
ν_g	114(6)	211(69)	123(4)	229(65)
s_g^2	150(14)	187(58)	145(11)	183(42)