Additional File 1
The derivation of the dynamic program algorithm for the step 2 in Algorithm 1

In the following, we will give a supplementary discussion about the step 2 in Algorithm 1 and the detailed derivations for the dynamic program algorithm. Firstly, we give some notes. Let \([q_i : q_j]\) denote a region from the locus \(q_i\) to the locus \(q_j\) and \(\ln(P(Z_i^{(c)}|\hat{\Psi}_c, m_c, w_c))\) be the maximum log-likelihood about the data from \([q_i : q_j]\) of the whole observes \(Z^{(c)}\), where \(q_i, q_j \in \{w_c\}, \ q_0 = 0 < q_i < q_j < q_{(\{w_c\}+1)} = L_c, \ 0 < i < j < \{w_c\}+1\). Then we define the dimension \(D_{(m_c, w_c)}\) as \((\{w_c\} + 1)D_{m_c}\), where \(D_{m_c}\) is the number of parameters only depending on \(m_c\). In our model, \(D_{m_c} = 3m_c + 2\) for chromosome \(c\). Thus, given \(m_c = i\) and \(w_c\), the objection function of the step 2 in algorithm 1 becomes:

\[
-\ln(P(Z_i^{(c)}|\hat{\Psi}_c, m_c = i, w_c)) + \lambda_c D_{(m_c = i, w_c)} = \sum_{k = 1}^{\{w_c\}+1} [-\ln(P(Z_i^{(c)}|\hat{\Psi}_c, m_c = i, w_c))] + \lambda_c D_{(m_c = i)}.
\]

Then given \(m_c = i\) and let \(w_c[i : jj]\) denote the change point set with the last region to be \([ii : jj]\), we define the score function for the change point set \(w_c[i : jj]\)

\[
\Delta_{(m_c = i)}[w_c[i : jj]] = \sum_{k = 1}^{\{w_c\}[ii : jj]} [-\ln(P(Z_i^{(c)}|\hat{\Psi}_c, m_c = i, w_c))] + \lambda_c D_{(m_c = i)}
\]

Next, we define \(\Delta^K_{(m_c = i)}[ii : jj]\) to be the minimum score of a change point set with the last region to be \([ii : jj]\) and the number of change points to be \(K_c\)

\[
\Delta^K_{(m_c = i)}[ii : jj] = \min_{w_c[i : jj], |w_c[i : jj]| = K_c} \{\Delta_{(m_c = i)}[w_c[i : jj]]\}.
\]

Thus, given \(m_c = i\) and \(K_c \leq K_{max}\), apply dynamic programming theory, we can get the following recursion

\[
\Delta^K_{(m_c = i)}[e + 1 : L_c] = \min_{w_c[1 : K_c]} \{\Delta^K_{(m_c = i)}[f + 1 : e] + \lambda_c D_{(m_c = i)}\} + \lambda_c D_{(m_c = i)}.
\]

where \(c, f \in w_c^1\), \(w_c^1 = \{0\} \cup w_c^0 = \{w_c^0, w_c^1, w_c^2, \ldots, w_c^i\}\) with the bound that if \(i < j\), \(w_c^i < w_c^j\) holds.

Given \(\lambda_c\) and \(m_c = i\), using the above recursion, we can design a dynamic programming algorithm to find the optimal change point set \(\hat{w}_{c,t,i}\) in step 2 of Algorithm 1.

The derivation of the RSPLIS procedure and proof of theorem 1
The derivation of the RSPLIS procedure

The derivation involves three steps:
(i) making connections between the multiple testing and weighted classification problems;
(ii) derive an oracle procedure for FDR control;
(iii) develop a data-driven procedure that mimics the oracle procedure.

Let $\beta$ be the relative cost of a false positive to a false negative. Consider a weighted classification problem with loss function $L(\theta, \delta) = 1/L = \sum_c \sum_r \sum_l \{\beta(1 - \theta_c rl)\delta_c rl + \theta_c rl (1 - \delta_c rl)\}$

where $L = \sum_c L_c$ is the total number of SNPs from all chromosomes. Under mild conditions, the multiple testing problem is equivalent to a weighted classification problem. Specifically, let $U_\alpha$ be the collection of all $\alpha$-level FDR procedures of the form $\delta = I(T < C)$. Suppose that the classification risk with the above loss function is minimized by $\delta^{(c)} = \{T, C^{(c)}(\beta)\}$, so that $T$ is optimal in the weighted classification problem. If $T \in T_\alpha$, then $T$ is also optimal in the multiple testing problem, in the sense that for each FDR level $\alpha$, there exists a unique $\beta^{(c)}(\alpha)$, and hence $C^{(c)}(\beta^{(c)}(\alpha)) = C(\alpha)$, such that $\delta^{(c)}(\alpha) = \{T, C(\alpha)\}$ controls the FDR at level $\alpha$ with the smallest FNR level among all testing rules in $U_\alpha$.

The optimal classification rule that minimizes $R(\beta) = E(L(\theta, \delta))$ is $\delta(\beta, 1/\beta) = \{\delta^{(c)}(\beta)\}$, where $\delta^{(c)}(\beta) = P_{\Psi_{cr}}(\theta^{(c)}(\beta) = 0|Z^{(c)}(\beta)) / P_{\Psi_{cr}}(\theta^{(c)}(\beta) = 1|Z^{(c)}(\beta))$ and $\delta^{(c)}(\beta) = I_{(\beta(Z^{(c)}(\beta)) < 1/\beta)}$. Note that $\beta^{(c)}(\beta)$ is strictly increasing in LIS$^{(c)}(\beta)$, the optimal testing procedure is of the form

$$\delta(\text{LIS}, C \mathbf{1}) = \{I_{\text{LIS}^{(c)}(\beta) < C} : c = 1, \ldots, C; r = 1, \ldots, R_c; l = 1, \ldots, L_{cr}\}.$$ 

Now the question is how to determine the optimal cutoff $C_{opt}$ for a given FDR level $\alpha$. Note that for a given threshold $C$, the FDR level of RSPLIS is

$$\text{FDR}(C) = E\left[\frac{\sum_c \sum_r \sum_l (1 - \theta^{(c)}(\beta)\delta^{(c)}(\beta))}{\sum_c \sum_r \sum_l \delta^{(c)}(\beta)} \lor 1\right]$$

$$= E[\frac{1}{\sum_c \sum_r \sum_l \delta^{(c)}(\beta)} \lor 1] \sum_c \sum_r \sum_l I_{\text{LIS}^{(c)}(\beta) < C} \text{LIS}^{(c)}(\beta).$$

From the above expression we can see that the group labels $c$ and $r$ are no longer needed and hence are dropped. Suppose the total number of rejections from all groups is $RN$, then according to the law of large numbers,

$$\text{FDR} = \frac{1}{RN} \sum_{i=1}^{RN} \text{LIS}(i) + o(1).$$

It is straightforward to see that we should choose the largest $RN$ such that

$$\frac{1}{RN} \sum_{i=1}^{RN} \text{LIS}(i) \leq \alpha.$$ 

Thus we have derived the RSPLIS procedure.
Proof of Theorem 1

(i) Validity. Let RN be the number of rejections by the RSPLIS procedure. Note that this is a pooled analysis, we neglect the group label c and r.

$$FDR_{RSPLIS} = E\left\{ \frac{\sum_{i=1,\ldots,L} \delta_i (1 - \theta_i)}{\left( \sum_{i=1,\ldots,L} \delta_i \right) \vee 1} \right\}$$

$$= E\left[ \frac{1}{\left( \sum_{i=1,\ldots,L} \delta_i \right) \vee 1} \sum_{i=1,\ldots,L} E\{\delta_i (1 - \theta_i) | Z\} \right]$$

$$= E\left[ \frac{1}{\left( \sum_{i=1,\ldots,L} \delta_i \right) \vee 1} \sum_{i=1,\ldots,L} \delta_i \text{LIS}_i \right]$$

$$= E\{ \frac{1}{RN \vee 1} \sum_{i=1,\ldots,RN} \text{LIS}_i \}.$$

The result follows by noting that for all realizations of Z, Our RSPLIS procedure guarantees that

$$\frac{1}{RN \vee 1} \sum_{i=1}^{RN} \text{LIS}_i \leq \alpha.$$

(ii) Asymptotic optimality. The asymptotic optimality can be shown without essential difficulty by generalizing the proof of Theorem 6 in Sun and Cai (2009) (for a single Markov chain). We refer to Sun and Cai (2009) for more technical details.