

MBB 291

Problem Set 4

Due Date: 28 October 2017, 11:59pm

General Instructions

- The answer sheet for this problem set should be submitted as a .R file containing the functions you used to perform the needed operations. Use $\langle last\ name \rangle$, $\langle first\ name \rangle.R$ as the name of the file. For example, *De la Cruz, Juan.R*.
- If you have consulted references (books, journal articles, online materials, other people), cite them as footnotes to the specific item where you used the resource/s as reference.
- Submission of the problem set answers should be done via e-mail. Attach the .pdf file, and write as the subject header of the e-mail: [MBB 291] < *Student Number* > – < *Last Name, First Name* > – Problem Set 4. For example: [MBB 291] 190800001 - De la Cruz, Juan - Problem Set 4. Send your answers to janmichaelyap@gmail.com.
- **You should receive a confirmation e-mail from me stating receipt of your deliverable within 24 hours upon your submission of the problem set.** If you have not received any, forward your previous submission using the same subject header once more.
- If you have any questions regarding an item (EXCEPT the answer and solution) in the problem set, do not hesitate to e-mail me to ask them. However, **questions regarding this problem set forwarded/received on or after 12:01am of 25 October 2017 will NOT be entertained.**

Specifications

Download the dataset from the website (in the QTL Mapping row of the right hand side of the web site) and name the file as *rice.csv*. The data set contains marker and phenotype data from a previous study on salt tolerance QTLs in rice¹. The task for the ME is to perform data preprocessing and QTL mapping on the dataset. To be included in the .R file are statements that were used to perform the following:

- Reading of the file and saving into a *cross* typed variable named *rice*
- Conversion of the *rice* into a *riself* variable
NOTE: There are at least two ways the first two tasks can be performed.
- Estimating the recombination fraction of the *rice*
- Creating a new re-estimated genetic map from *rice* and storing it in a variable named *newmap.rice*.
- Replacing the genetic map of *rice* with *newmap.rice*.
- Calculating the genotype error LOD scores of *rice* and inserting it into the variable itself.
- Perform single QTL mapping on *rice* using expectation maximization and storing the result in a variable named *out.em.rice*.
- Perform single QTL mapping on *rice* using Haley-Knott regression and storing the result in a variable named *out.hk.rice*.
- Perform single QTL mapping on *rice* using Haley-Knott regression with permutation testing storing the result in a variable named *out.hk.perm.rice*.
- Perform multiple QTL mapping on *rice* using Haley-Knott regression with permutation testing storing the result in a variable named *out.hk.mult.rice*. Set the maximum cofactors to be sought to 10.

¹Thomson MJ, de Ocampo M, Egdane J, Rahman MA, Sajise AG, Adorada DL, Tumimbang-Raiz E, Blumwald E, Seraj ZI, Singh RK, Gregorio GB, Ismail AM (2010). Characterizing the Saltol Quantitative Trait Locus for Salinity Tolerance in Rice. *Rice*; 3: 148-160.