

Homework 1

Computational approaches to statistical genetics

Important dates:

- Presentation date: 10 April 2017

Mark:

- 25 points

Instruction:

- Read Bessonov et al. (2015) paper and take Evans et al. (2013) as supporting document to address the following questions in your slides presentation.
- You can work in groups. One presentation per group is allowed (but everyone should present a part) and a score will be a group score.
- Presentation duration: 20min.
- Reference papers could be found from:

<http://www.montefiore.ulg.ac.be/~chaichoompu/CK/?Courses>

Questions

- 1) What is the study design:
 - a. Samples: unrelated individuals? Cohort? Case-control? Family? Would that make a difference for the subsequent analyses?
 - b. Markers: which markers are used? How many?
- 2) How were dependencies (LD) between markers treated? This relates to “pruning”. Why is this important?
- 3) How were markers coded in analytic models? Does this matter in terms of final epistasis results?
- 4) Which analytic framework / model was used? Regression? Parametric or non-parametric approach?
- 5) How was multiple testing dealt with?
- 6) What can be used as prior biological information? Is this information also relevant post analysis?
- 7) Which measures were taken to strengthen the findings?
 - a. Does agreement between different analytic tools imply validity of findings?
 - b. Replication?
 - c. Validation?

Address all of the questions and work out in more details, using the aforementioned papers or information you can retrieve from:

<http://www.montefiore.ulg.ac.be/~chaichoompu/CK/?Courses>