

## Exam GBIO0015+30+31 - September 2021

1. Take the two provided case-control datasets. Both provided sets are cleaned, and LD pruned. For convenience, consider the completed data (so no missing data on SNPs)
  - a. Info about the data is on the course website
  - b. To get access to the data and servers, contact [diane.duroux@uliege.be](mailto:diane.duroux@uliege.be)
  - c. References linked to your assignment are included in the TABLE of the original assignment.
2. If estimated computation time surpasses the time available to you, contact [diane.duroux@uliege.be](mailto:diane.duroux@uliege.be) prior to reducing the dataset.

### Robin LIBERT – encoding for GWAIS

3. Implement the RFselect way of recoding SNP pairs (we are only interested in the recoded SNPs).
4. Implement the EDGE encoding (this is a second way to recode SNPs).
5. Choose an AI/Machine Learning method that can take such encodings to detect epistasis in the two data sets.
6. Compare results and discuss.

### Nicolas WLASOW-WLASOWSKI - GWGGI:

2. Perform GWGGI GWAIS, using “default” options for parameters or as recommended in GWGGI
3. If you have chosen TAMW before, then also consider using LRMW (or vice versa)
4. Discuss differential findings.

### **For the evaluation, prepare:**

- A slides presentation of your work (progress)
- Add a few slides about practical issues you encountered (issues that are intrinsic to the analysis methods used)
- For the chosen epistasis detection method, fill in the attached Table (see example in column B)
- Annotate your code so someone else can understand how it works
- Examples of final presentations are on our server (see links on the course website)

Send everything to BOTH [diane.duroux@uliege.be](mailto:diane.duroux@uliege.be) and [kristel.vansteen@uliege.be](mailto:kristel.vansteen@uliege.be) by August 20<sup>th</sup>

### BEFORE the exam, make sure that you

- include your final presentation in your folder (on the cluster)
- include a README.txt file (see for example, the one in Christophe or Ba Thien's folders)

**The (online) exam will take place on September 1<sup>st</sup>** and will exist of a 30 min presentation of your work, followed by 15-max30 minutes of questions:

- Robin: 10-11 am; Nicolas: 11 am-12 pm
- LINK: <https://global.gotomeeting.com/join/534351157>