GBIO0002 – Genetics and Bioinformatics

Montefiore Institute - Systems and Modeling

GIGA-R Medical Genomics

bio3.giga.ulg.ac.be

kristel.vansteen@uliege.be

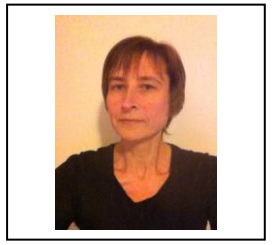
fdequiedt@uliege.be

Administration

- Course instructors

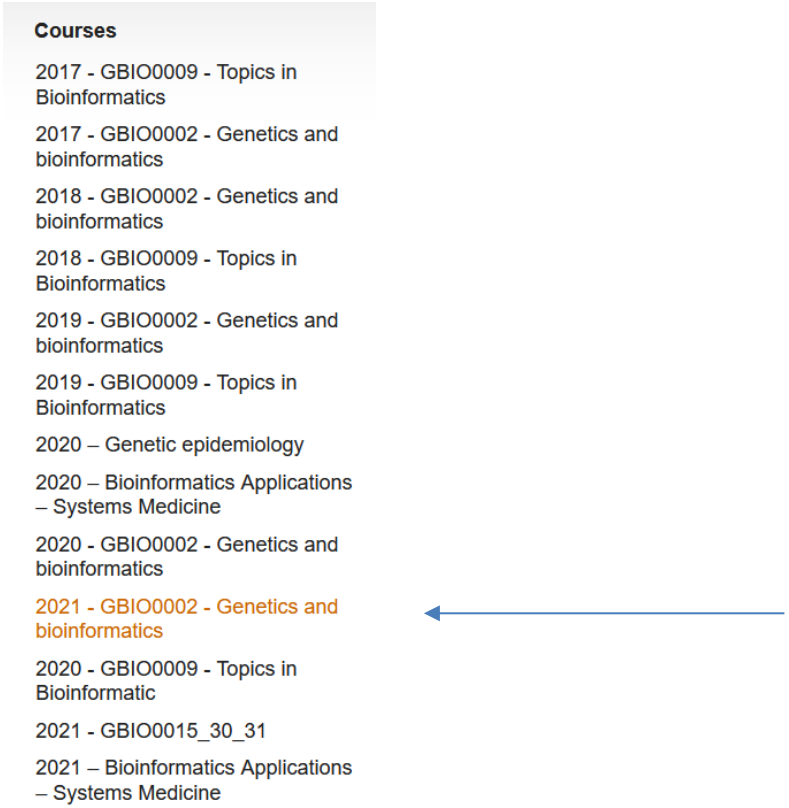
Prof. Kristel Van Steen
- Office: level +5, B34 (GIGA tower)
- E-Mail: kristel.VanSteen@uliege.be
- bio3.giga.ulg.ac.be/

Prof. Franck DEQUIEDT
- Office: level +5, B34 (GIGA tower)
- E-Mail: fdequiedt@uliege.be



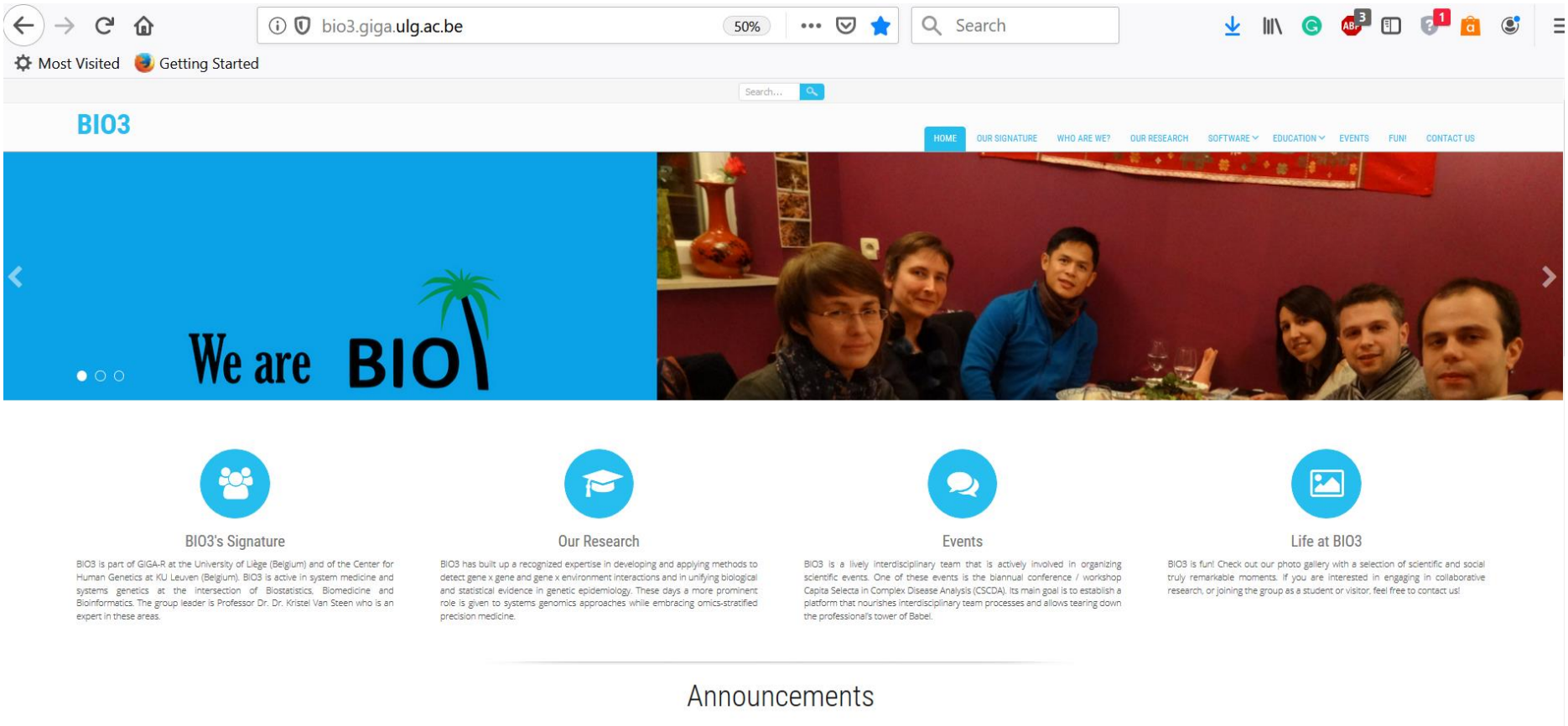
Administration

- Course website 2012-2021:



http://bio3.giga.ulg.ac.be/archana_bhardwaj/?Courses

Administration



http://bio3.giga.ulg.ac.be/ [research BIO3]

Events

Upcoming

- TranSYS Summer School 2: Mark your calendar: 22-26 November 2021 – Paris, France (https://h2020transys.eu/); Click here for meeting links
- MLFPM Summer School (Machine Learning Frontiers in Personalized Medicine) 3: Click here for meeting links.

EUREGIONAL MICROBIOME CENTER

http://bio3.giga.ulg.ac.be/ [events with BIO3 involvement]

Integrative Omics

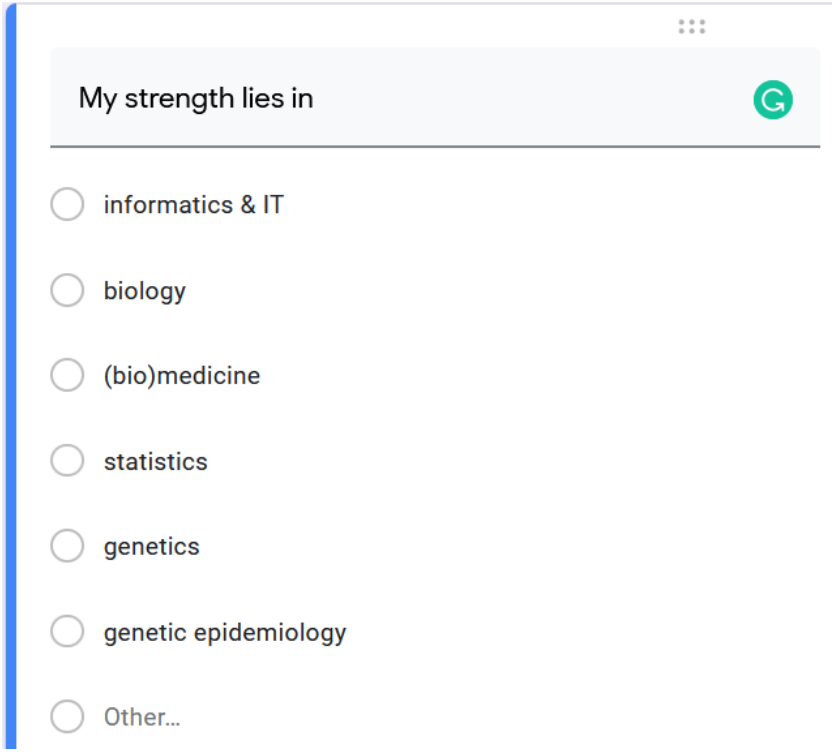
> “. In 2010 think-tank activities started around a 2nd research line in parallel to the previous one: The research line on integrative analyses involves the development and implementation of integrative omics approaches for complex disease traits dissection and preventive or diagnostic medicine. BIO3’s activities have resulted in a book chapter on “Perspectives on Data Integration in Human Complex Disease”, highly appreciated conference programs for Capita Selecta in Complex Disease Analysis (CSCDA), and machine learning based (integrative) network construction methods [6-8] . Instrumental was (and still is) the COST Action BM1204 on “An integrated European platform for pancreas cancer research: from basic to clinical and public health interventions for a rare disease” (2012-2016), which is further accelerating the conception of several manuscripts related to data integration and pancreas cancer. For this Action, Kristel Van Steen has been working group leader on “Integration of omics data”.
>
> 6. Bessonov, K. & Van Steen, K. Practical aspects of gene regulatory inference via conditional inference forests from expression data. Genet Epidemiol 40, 767-778 (2016).
>
> 7. Gadaleta, F., Bessonov, K. & Van Steen, K. Integration of gene expression and methylation to unravel biological networks in glioblastoma patients. Genet Epidemiol 41, 136-144 (2017).
>
> 8. Gadaleta, F. & Van Steen, K. Discovering main genetic interactions with LABNet LAsso-based network inference. PLoS One 9, e110451 (2014). ”

- Genomic Medicine – Bridging research and the clinic (Slovenia, 3-7 May 2016): VAN STEEN – BIG DATA INTEGRATION CHALLENGES

http://bio3.giga.ulg.ac.be/ [educational talks]

Administration

- Complete online form: https://docs.google.com/forms/d/e/1FAIpQLScwmFy_aAkFidJR_qKQio2_W5CklfBK7mUQ2D2F386s729ywQ/viewform



Uniqueness of "this" course

The course "Genetics and Bioinformatics" is unique in the following aims, goals and properties:

- Two standalone disciplines, namely "Genetics" and "Bioinformatics" are integrated into one course. Special care is given to the "integration" aspects which underpin this course. We do so in multiple ways:

  - There are bridging classes between the two disciplines (e.g., TA, invited speakers)
  - During the bridging classes software applications are shown, either on analytics seen in class, and/or by building upon these. You are invited to follow on your own computer.
  - Bioinformatics parts build on Genetics concepts that have been introduced in a foregoing class (there is a logical flow)

- The learning outcomes are well explained and relate to the following:

  - understanding key concepts, terminology and their context;
  - interpreting findings / analysis results (not yet carrying out multiple types of analyses that require a multi-faceted programming skills set).

  The learning outcomes form the basis for the exam.

  For instance for the last part, screenshots of analyses runs are shown (as seen in class), which need to be interpreted.

- Keeping these learning outcomes in mind, interim evaluations are also done in a unique way, for this course:

  - You are given the choice to select a classical style of homework assignment (i.e. questions & answers) or a literature style homework assignment (i.e. students can suggest papers or select one of the papers proposed to capture the essence, to link back to the course notes materials, to be inspired to look at the broader context and read (or look up) additional materials).

- This system was inspired by implemented systems in the United States, and gives you the freedom to select the assignment that best matches your background.
  - Indeed, this course has a history of having a heterogeneous student population and bachelor students may have more difficulties in finding their way in scientific literature in English than master students can.
  - Furthermore, the fact that students suggest their own papers adds an extra layer of accommodation towards the student.
  - Furthermore, guidelines are provided in class about “how to critically evaluate a paper”, as we experienced that often students lack knowledge about the basic underlying principles to critically evaluate a paper.

- As a consequence, **reading papers and processing these** in view of the course (including forming evidence-based opinions), is crucial to this course.
  - It brings you into the real-life experience of a bioinformatician in that there is an abundance of information out there and that an abundance of materials can be found even when modifying just a few conditions underlying a bioinformatics analysis.
  - Anyone working in a bioinformatics environment will testify that a big part of the time is spent on browsing the literature for pros and cons of adopted methods or on identifying new routes of analyses that can increase optimality.
  - In this course, you are exposed to such a situation, in a mini-version. Furthermore, you will need to work in groups on assignments, also mimicking real-life situations for most bioinformatics professionals, adding on an extra component of “project management & communication” to the course.

- Clear selection of subtopics: avoiding not seeing the trees for the forest
  - Bioinformatics and Genetics are broad topics
  - An overview of multiple Bioinformatics topics, possibly unlinked to each other, is given in the course GBIO0009.
  - For GBIO0002, we have made the decision to consider a red thread via "genome-wide association studies" (GWAS).
    - Analyses and post-analyses of GWAS link to multiple sub-disciplines using different data sources, including gene expression and protein interactions.
    - These links are covered in the course, from a "genetics perspective" (f.i. including the coverage of underlying data generating technologies) and from a "bioinformatics" perspective (which is taken to be an "data analytic" perspective).

Course schedule

Information about the organization of feedback sessions to homework evaluations will follow

COURSE SCHEDULE-IN-A-GLANCE & COURSE NOTES (check for updates):

Dates	Topics	Materials	Class links & supplementary info
Sep-14	Meet & Greet	Lecture 1_ Course Administration and Expectations Lecture 1_ HW Organization Lecture 1_Critical Evaluation of a Paper Supporting Book References Important Links Scientific Communication Self-study_ Software Introduction (author: Archana Bhardwaj) DATA_FILES Specific_R_CODE_REPORT	First class (Prof Van Steen) via https://global.gotomeeting.com/join/916883981 Video : Introduction by Prof Van Steen Class 1_Admin (ac year 2020-2021) Video : Software (environments) R_PLINK
PART I : Genomic Association Studies			
Sep-21	Genetics and Genetic Markers: from "DNA blueprint" to variation	Lecture 2 - principles of genetics	Prof Dequiedt
Sep-28	Genetic mapping using GWAS as high-level analysis: Why, What, How?	Lecture 3 - GWAS **Supporting docs** (exam material): Bush et al. (2012) - Computational biology introduction to GWAS Balding et al. (2016) - Statistical introduction to GWAS	Prof Van Steen

Course schedule

- Practical highlights

  - "Theory" classes: Bioinformatics (KVS); Genetics & Genomics (FD)
    - Concepts and contexts in GWAs and post-GWAs
    - Interpreting analysis findings: discussing different viewpoints
    - Slides as supporting framework ("syllabus")
    - Part of "syllabus": provided "supporting docs"; in class (on computer) reading of exerts of discussion papers
    - "Background reading" is not part of the "syllabus" but should help you to better understand the in-class course notes

- "Practical classes": bridging classes as explained before
  - Software installation instructions: prior to the class

    Depending on the student population, in some years, we have had the request from some students to be able to do "real" practical analyses (data → problem → install and apply/write software code → results → interpretation).

  - This aspect is a crucial component of the "topics in bioinformatics" (GBIO0009) class, to which GBIO0002 is a predecessor
    - GWAS → GWAIS → Microbiome interaction networks
    - Via group work, apply existing analysis pipelines on real-life data; combine all subtopic activities into a big analysis workflow
    - Theory in function of the expected outcomes
    - Not organized for small numbers of students

- Homework assignments:

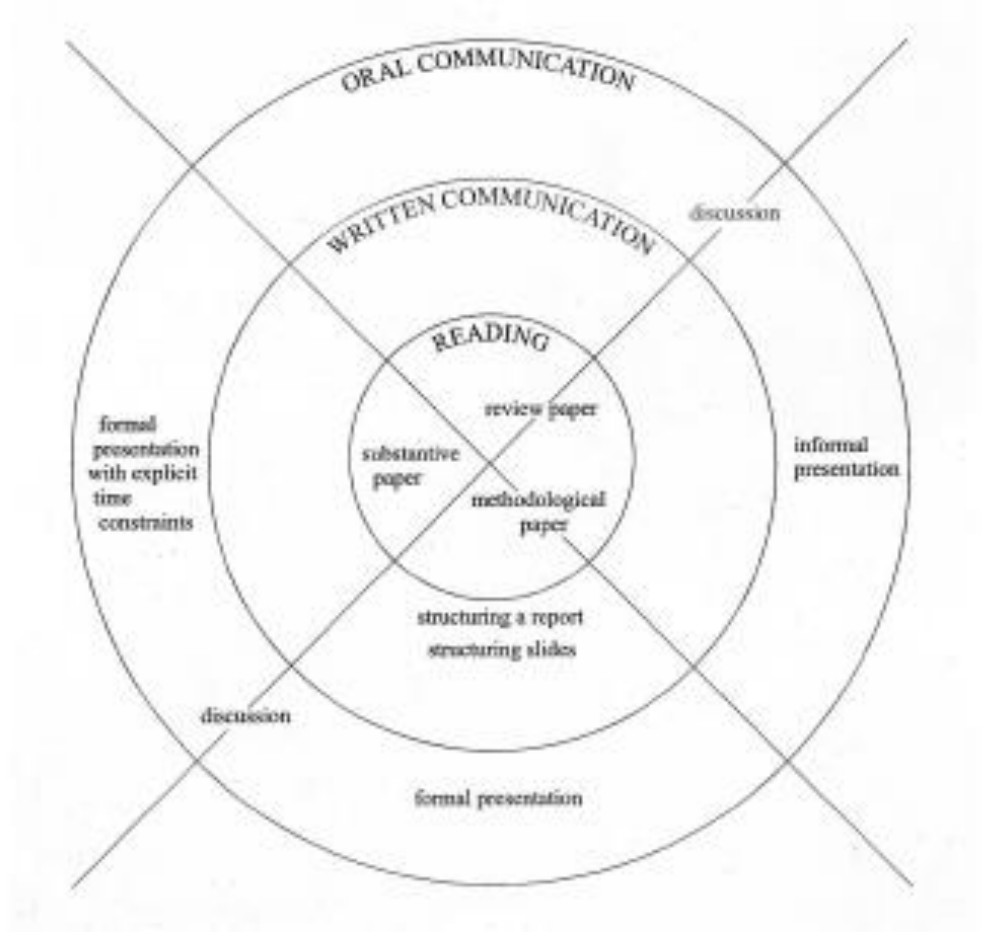
  - Most time-consuming part of this course
  - Link to the theory AND practical classes
  - Organized around “group work”

Effective Reading

"Sure the consultants are paid twenty times more than you. They look more impressive when presenting your ideas."

Effective Presenting

Guidelines on increasing your communication skills (supporting doc)



Components of Scientific Communication

- Homework assignments:
  - Group formation:
    - questionnaire will be used to control “group formation”
    - max 8 groups
    - the aim is to form interdisciplinary groups:
      - mix of bachelor / master levels
      - mix of informatics / biology / biomedical … strengths

    Motivation: enormous **opportunities in heterogeneity**

  - coverage of papers from different angles
  - acquire knowledge outside initial comfort zone
  - promote listing to each other (reformulate questions such that the other understands)
  - increase exploitation potential of complementary skills

- Homework assignments:
  - Selection of two homework styles
  - At the end of the year, each group should have selected minimally 1 Genetics Literature Style homework and minimally 1 Bioinformatics Literature Style homework. Depending on HW1 choices, you may have limited choice for HW2 (X)

Group	Assignment	Genetics		Bioinformatics	
		Q&A	Literature	Q&A	Literature
1	1	X		X	
	2		X		X
2	1	X			X
	2		X	X	
3	1		X		X
	2	X		X	
4	1		X	X	
	2	X			X

- Homework assignments:

  - Q&A:
    - Classical questions and answers
    - Group report

  - Literature Style
    - Read and process a given paper or suggested paper
    - Follow “critical evaluation of a paper” while reading the paper
    - Follow additional guiding questions to process the paper in the context of this course
    - Group slides (serving as a report) & presentation

      ✓All students are present
      ✓Tutors are present

Critical evaluation of a paper (as supporting doc)

Introduction

1. Did the author(s) indicate why the study was undertaken?
2. Was the background information provided adequate to understand the aims of the study?

Methods

1. Has the source of the data been clearly given?
2. Were the methods described in sufficient detail for others to repeat or extend the study?
3. If standard methods were used, were adequate references given?
4. Have the author(s) indicated the reasons why particular procedures were used?
5. Have the author(s) indicated clearly the potential problems with the methods used?
6. Have the author(s) indicated the limitations of the methods used?
7. (Have the sources of drugs been given?)
8. Have the author(s) specified the statistical procedures used?
9. Are the statistical methods appropriate?

Results

1. Were the experiments/calculations done appropriate with respect to objectives of the study?
2. Do the results obtained make sense?
3. Do the legends to the figures describe clearly the data obtained?
4. Are the data presented in tabular form clear?
5. Has the appropriate statistical analysis been performed on these data?

Discussion

1. Were the objectives of the study met?
2. Do the author(s) discuss their results in relation to available information?
3. Do the author(s) indulge in needless specualtion?
4. If the objectives were not met, do the author(s) have any explanation?

References

1. Do the author(s) cite appropriate papers for comments made?
2. (Do the author(s) cite their own publications needelessly?)

Abstract

1. Is the abstract intelligible?
2. Does the abstract accurately describe the objectives and results obtained?
3. Does the abstract include data not presented in the paper?
4. Does the abstract include material that cannot be substantiated?

- Course language:

  - Teaching staff are mixed francophone and non-francophone
  - Slides may be in English or French although the main course language is English

  - Bridging class tutors (including TA's) are selected upon their combined "genetics & bioinformatics" skills, and their fluency in English. Just like you may speak with an accent, English speaking tutors may also have an accent that requires some adjustments.

Guidelines on homework organization (as supporting doc)

GBIO0002: Genetics and Bioinformatics

Style 1: Literature project

This involves processing and analyzing a paper from the literature that extends or provides additional background on the course materials. In particular, you are asked to summarize the objectives, methods and results of the paper, while further browsing the internet for additional information or supporting materials.

Do not copy the paper, but show you have understood the main ideas of the paper and "discuss" the paper. Such a discussion could include thoughts on what was the key idea, strengths or weaknesses of the methods/experiments, comments on the writing, ways to extend the work, flaws in the argument/data/experiments, etc. Anything is fine, as long as it demonstrates some real thought.

For review papers you can focus on one or a few subtopics to be worked out in greater depth, by following up on referenced work and/or searching for additional explanations on the internet.

A selection of papers will be provided, but if you have another interesting paper to discuss, please send your suggestion the PI to which the paper relates (Genetics: Prof Dequiedt; Bioinformatics: Prof Van Steen). The course instructors will then decide whether the paper is eligible or not.

All literature projects will be presented and discussed in class. No report is needed: instead send your well-organized slides to kristel.vansteen@uliege.be with the subject title "GBIO0002 slides".

Course evaluation

- What will be evaluated?
  - At the end of the course, you have acquired knowledge about **genetics** (in particular genomics, transcriptomics, technology-related aspects), about GWAS and related **bioinformatics concepts** and applications, and about a selection of state-of-the-art, yet basic, **analytic tools**.
  - You will be evaluated about key concepts related to **genetics & bioinformatics** and the **analytic approaches** presented during the course (incl. pros and cons, general contexts).
  - You will be presented with a **few analysis results to interpret**.

Course evaluation

- How will be evaluated?

HW1		**HW2**		**Exam (written)**
Genetics	**Bioinformatics**	**Genetics**	**Bioinformatics**	
15	15	15	15	40

- You cannot pass without homeworks
- No final grade without exam
- Homeworks not handed in in time (without eligible notice – e.g. medical condition) == ZERO

- Second term exam:
  - Less then 30/60 for homeworks: exam is organized around retaking worst homeworks on Genetics & Bioinformatics and writing a report (see next). Written exam can be recycled.
  - Less than 20/40 for exam: exam is organized around retaking the written exam of the first term. Homework scores can be recycled.

- Evaluation criteria – presentation (individual adjustments to group scores)

Criterium	Key words
Clarity	Concepts, slides content, slides composition, “new” terms are clearly explained during the presentation
Illustrations on slide	Not too much – not too little; not only copy and paste from course but incorporate novel relevant illustrations; supportive illustrations
Presentation Skills	Eager beaver (a person who is very enthusiastic about doing something), attract the attention of the audience
Understanding	Presentation content as presented is understood: adequate reply to questions and comments (incl. those from fellow students)
Group dynamics	Balanced partitioning of tasks (pre-, during, post-presentation)

- Evaluation criteria – “reports”

Applies to “slides reports”, and potentially “Q&A reports” and 2nd term exams.

- Ability to formulate the research problem and to sketch the context (introductions, data description, tool description, etc)
- Presentation summary of the workflow (e.g., methods, technological -analysis - experimental section)
- Discussion (e.g., analysis tools, quality of the analysis, validity of results – when put in a broader context, …)
- Creative input (e.g., stuffing, conclusion section)
- General structure of the report (sectioning)

Effective Reading

Why?

Your teachers give you a pile of papers / book chapter to read.

Ouch…

Efficient reading skills will be helpful in multiple ways: knowledge gain, insight in writing styles, structuring thoughts, distinguishing main and secondary issues, …

What are different types of scientific literature?

- Primary (authors carried out the work)
  - Examples: monographs, theses or dissertations, conference papers and reports
  - Peer-reviewed journal
  - Particular format

- Secondary (work of others; target: others in the field)
  - Examples: review journals, monographic books and textbooks, handbooks and manuals
  - More flexible style: still scientific and fully referenced

Human Genetics (2019) 138:293–305
https://doi.org/10.1007/s00439-019-01987-w

REVIEW

How to increase our belief in discovered statistical interactions via large-scale association studies?

K. Van Steen[1,2] · J. H. Moore[3]

Received: 26 July 2018 / Accepted: 20 February 2019 / Published online: 6 March 2019

Abstract

The understanding that differences in biological epistasis may impact disease risk, diagnosis, or disease management stands in wide contrast to the unavailability of widely accepted large-scale epistasis analysis protocols. Several choices in the analysis workflow will impact false-positive and false-negative rates. One of these choices relates to the exploitation of particular modelling or testing strategies. The strengths and limitations of these need to be well understood, as well as the contexts in which these hold. This will contribute to determining the potentially complementary value of epistasis detection workflows and is expected to increase replication success with biological relevance. In this contribution, we take a recently introduced regression-based epistasis detection tool as a leading example to review the key elements that need to be considered to fully appreciate the value of analytical epistasis detection performance assessments. We point out unresolved hurdles and give our perspectives towards overcoming these.

What are different types of scientific literature?

- Tertiary (work of others; target: interdisciplinary audience, public)
  - Examples: science magazines, newsletters, science articles in newspapers, introductory textbooks and encyclopedias
  - Popular rather than a scientific style; reduced/short bibliography
- Grey (limited distribution, difficult accessing)
  - Examples: technical reports, journals published by special interest groups, abstracts of conference papers and conference proceedings that are only made available to conference participants, working papers, some online documents

An efficient algorithm to perform multiple testing in epistasis ...
https://www.ncbi.nlm.nih.gov › pmc › articles › PMC3648350
by F Van Lishout - 2013 - Cited by 23 - Related articles
Apr 24, 2013 - Here, m and n refer to the number of SNP pairs and the number of top pairs to retain ..... In this paper we have presented the epistasis screening software MBMDR-3.0.3. .... [6]Department of Systems Biology, University of Vic, 08500 Vic, Spain ... Calle ML, Urrea V, Vellalta G, Malats N, Van Steen K. Improving ...

Model-Based Multifactor Dimensionality Reduction for ... - ORBi
https://orbi.uliege.be › handle
by ML Calle - 2008 - Cited by 43 - Related articles
Author, co-author : Calle, M L []. Urrea, V []. Vellalta, G []. Malats, N []. Van Steen, Kristel - mailto [Université de Liège - ULiège > Dép. d'électric., ... Publication date : 2008. Publisher : Department of Systems Biology, Universitat de Vic,. Report number : Technical Report No. 24. Permalink : http://hdl.handle.net/2268/144460 ...

[PDF] Travelling the world of gene-gene interactions ...
https://www.semanticscholar.org › paper › Travelling-the-world-of-gene-gen...
ML Calle, V Urrea, N Malats. Technical Report n. 24. Department of Systems Biology, Universitat de Vic,; 2008. VIEW 11 EXCERPTS. HIGHLY INFLUENTIAL ...

[PDF] VAN STEEN - Statistical Genetics Research Group
www.statgen.ulg.ac.be › VAN_STEEN_GxG_and_GxE_INTERACTIONS
WELBIO, GIGA-R Medical Genomics (BIO3), University of Liège, Belgium ..... Calle, M. L., Urrea, V., Vellalta, G., Malats, N. & Van Steen, K. (2008a) Model-Based. Multifactor ... 24, Department of Systems Biology, Universitat de Vic, http://www.recercat.net/handle /2072/5001 [technical report, first mentioning MB-MDR]. • Calle ...

[PDF] Improving strategies for detecting genetic patterns of disease ...
public-files.prbb.org › publicacions
by ML Calle - 2008 - Cited by 89 - Related articles
Oct 6, 2008 - M. L. Calle1,*,†, V. Urrea1, G. Vellalta2, N. Malats3 and K. V. Steen4. 1Department of Systems Biology, Universitat de Vic, Carrer de la ...

Mluz Calle | PhD Mathematics | University of Vic, Vic | UVIC ...
https://www.researchgate.net › ... › Department of Systems Biology
Mluz Calle of University of Vic, Vic (UVIC) | Read 64 publications | Contact Mluz Calle. ... Department of Systems Biology; Vic, Spain .... points (P = .015) and restricted a dominant T cell response to HIV Gag p24 (P = .038). ..... Calle ML, Urrea V, Malats N, Van Steen K. mbmdr: an R package for exploring ..... Technical Report.

"ML Calle, V Urrea, N Malats. Technical Report n. 24. …UVIC"

M.Luz Calle - Citas de Google Académico - Google Scholar
scholar.google.com › citations
ML Calle, V Urrea, G Vellalta, N Malats, KV Steen ... Statistical Papers 45 (2), 139-173, 2004 ... Department of Systems Biology, Universitat de Vic,, 2008.

Model-Based Multifactor Dimensionality Reduction for ...
https://onlinelibrary.wiley.com › doi › full
by T Cattaert - 2011 - Cited by 51 - Related articles
Sep 8, 2010 - (Calle et al., 2008a, 2008b) and are graphically displayed in Figure 1. ..... Table S2 reports the specific power to detect the functional pair(s), both with .... The work of M. L. Calle and V. Urrea has been supported by Grant MTM2008-06747-C02-02 from the Ministerio de Educación y .... Technical Report No.

(PDF) Participant's Case Studies (Day 2) | Kristel Van Steen ...
https://www.academia.edu › Participant_s_Case_Studies_Day_2_
17 CSCDA 2010 Leuven, 25-27 August 2010 [2] Mukherjee B, Chatterjee N (2008) .... Dr. Gut is author of over 100 research papers, inventor of 24 patents or patent ..... [2] Calle, M.L., Urrea, V., Vellalta, G., Malats, N. & Van Steen, K. (2007) ..... Luz Calle⋆ , N´ uria Malats† ⋆ Department of Systems Biology, Universitat de Vic ...

Comparison of genetic association strategies in the presence of rar…
https://cyberleninka.org › article › n
Similar topics of scientific paper in Biological sciences , author of scholarly ..... Technical Report 24 Department of Systems Biology, Universitat de Vic, Vic, Spain. 2. Calle ML, Urrea V, Vellalta G, Malats N, Steen KV: Improving strategies for ...

Some results may have been removed under data protection law in Europe. Learn more

Why is it useful to regularly read scientific documents?

- To gain knowledge (scientific knowledge, opinions, strategies)
- To stay on top of your field as well as linked fields (intro, discussion)
- To learn about journal styles / slang
- To become an expert in sifting through literature
- To learn about written communication

How to read a scientific article?

- Skim the article and identify its structure
- Distinguish the main points
- Generate the questions and be aware of your understanding
- Draw inferences
- Take notes as you read …

Skim the article and identify its structure

- Features of abstracts:
  - Purpose / rationale (why?)
  - Methodology (how?)
  - Results (what was found?)
  - Conclusion (what do the results mean?)

Skim the article and identify its structure

- Features of introductions:
  - Triggering interest
  - Providing enough information to understand the article
    - Broad: What is known?
    - Specific: What is not known?
    - Focus: What are the questions addressed?

Skim the article and identify its structure

- Features of methods:
  - Which experiments / tools were used to address the questions?
  - Most difficult to read especially when not well structured
  - Should provide the reader with information about the design of the experiment such that the validity of them can be evaluated
- Features of results and discussion:
  - Statements of what was found and reference to (visual) data [Figures, Tables] -- results
  - Comparisons to other results, interpretations, opinions -- discussion

Distinguish the main points

- Document level
  - Title, abstract, keywords
  - Visuals (captions)
  - Introduction
- Paragraph level
  - First few sentences in a paragraph
  - We hypothesize, we propose, we introduce, we develop, data suggests, in contrast to, surprising, …

Generate questions and be aware of understanding: active reading

- Before and during reading:
  - Who are these authors? What journal is this? Might I question the credibility of the work? Have I taken the time to understand all the terminology? Have I gone back to read an article or review that would help me understand this work better? Am I spending too much time reading the less important parts of this article? Is there someone I can talk to about confusing parts of this article?
- After reading:
  - What specific problem does this research address? Why is it important? Is the method used a good one/ the best? What are the specific findings? Am I able to summarize them in a few sentences? Are the findings supported by persuasive evidence? Is there an alternative interpretation not addressed? How are the findings unique/new/unusual or supportive of other work in the field? How do these results relate to my work? Applications? Interesting additional experiments to address the questions?

Draw inference: improve understanding and recall information

- Rely on your prior knowledge, world experience, materials provided in the paper, to draw inferences.
  - We learn about some things by experiencing them first-hand, but we gain other knowledge by inference — the process of inferring things based on what is already known.

Take notes as you read

- Details will slip away, eventually …
  - Stuff your (electronic) notebook, keep records of all of your scientific reading with summaries of their importance.
  - Time spent doing this will be regained when writing background, related work or literature review sections.

Be critical of published data/results!

- A lot of data is at your disposal but are they thrust-worthy?
  - Private data collections (curated according to standards?)
  - Public data collections (curated uniformly?)
  - Publications (source or summary data provided?)
  - Computerized databanks (block-chained or not?)

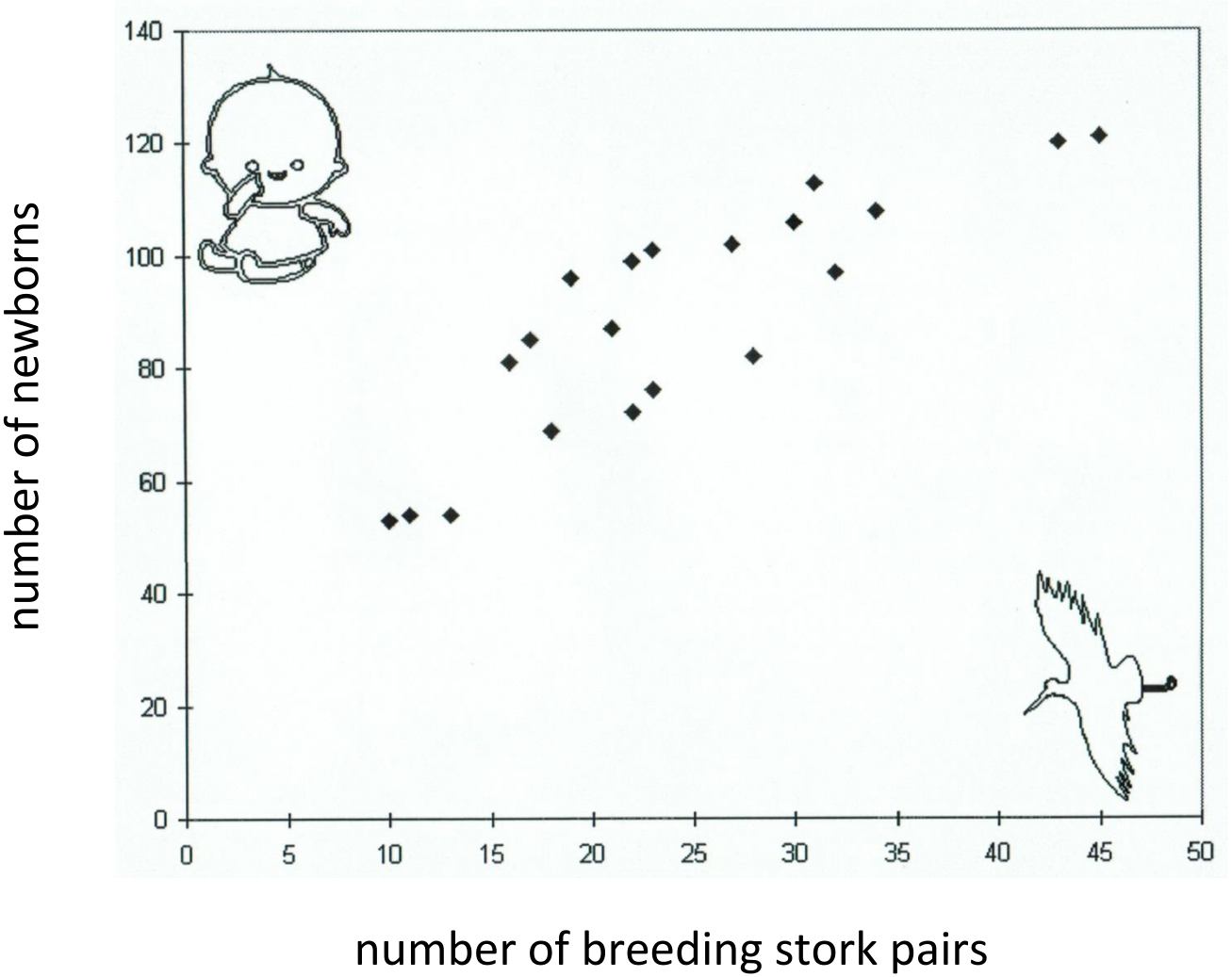
Errors will almost surely exist

- Apart from sampling errors, measurement error may arise:
  - mistakes in conceptualization
  - structural characteristics of the data collection process
- Relevant questions include:
  - How large are the errors?
  - What is the probability for a given error range?
  - Do errors cluster towards the end of a distribution?
  - In which direction does the error go?

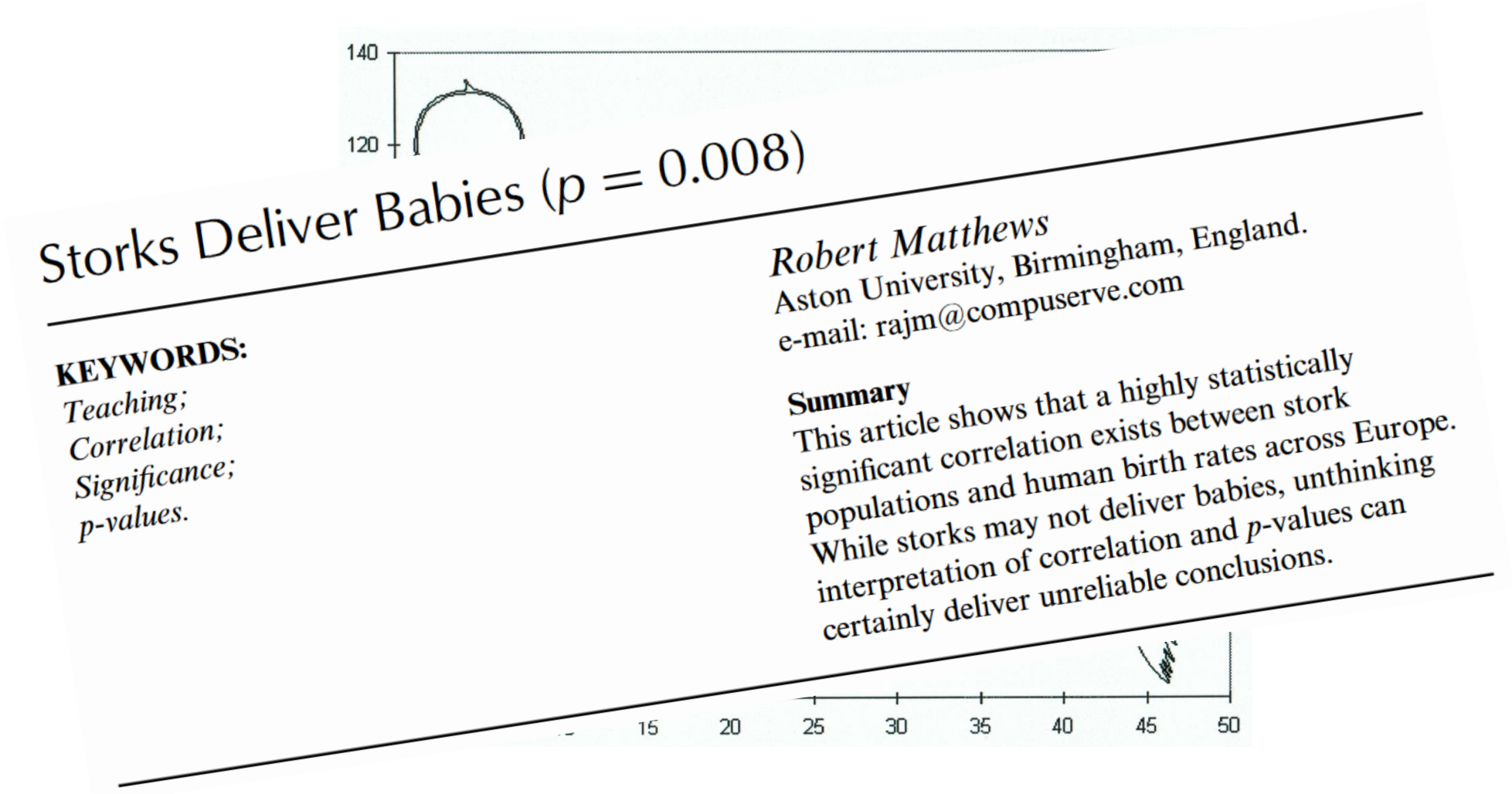
In general: “better” science through “better” data

(www.nature.com/openresearch/)

Beware of jumping to conclusions: causation versus association



Beware if jumping to conclusions: causation versus association



Questions?