

Applied High-throughput Analysis (C003695)

Course size (nominal values; actual values may depend on programme)

Credits 6.0 Study time 180 h Contact hrs 60.0 h

Course offerings and teaching methods in academic year 2018-2019

A (semester 1)	English	seminar: practical PC room classes	20.0 h
		lecture	35.0 h

Lecturers in academic year 2018-2019

De Meyer, Tim	LA26	lecturer-in-charge
Vandesompele, Jo	GE31	co-lecturer

Offered in the following programmes in 2018-2019

	crdts	offering
Brugprogramma Master of Science in Bioinformatics (main subject Engineering)	6	A
Master of Science in Bioinformatics (main subject Bioscience Engineering)	6	A
Master of Science in Bioinformatics (main subject Engineering)	6	A
Master of Science in Bioinformatics (main subject Systems Biology)	6	A
Master of Science in Biochemistry and Biotechnology	6	A
Exchange programme in Biochemistry and Biotechnology (master's level)	6	A
Exchange Programme in Bioinformatics (master's level)	6	A

Teaching languages

English

Keywords

omics data analysis, massive parallel sequencing, microarrays, kwantitatieve en digitale PCR, RNA-seq

Position of the course

This is an advanced course master of bioinformatics that focuses on the full data-analytical pipelines to (pre)process high throughput omics data. This will include the experimental protocols, as potential sources of variation/bias, image (microarray/sequencing) and mass spectrometry data analyses leading to raw sequence/intensity data, quality control of the raw data, discussion and application of normalization and statistical analysis tools, quality control of data normalization/analysis and gene set/pathway analysis. The different steps will be described and illustrated by means of real examples in the areas of genomics, transcriptomics, and epigenomics. Importantly, the full data analytical pipeline is studied, focusing on conceptual insight rather than the theoretical details of the studied algorithms.

Contents

Each theoretical part will be followed by a practical session in which a data set will be analyzed making use of the state-of-the art tools.

Part 1: Introduction

- History and evolution of high throughput omics
- Ethics of omics research, scientific integrity
- The advent of novel omics technologies
- Omics data repositories en databronnen (TCGA, GEO, ArrayExpress, SRA, ...)

Part 2: Microarrays

- Examples, quality control, applied data normalization and statistical analysis of:

- expression microarrays (single and two-channel)
 - comparative genomic hybridization arrays (CGH)
 - SNP arrays / DNA methylation arrays
 - Tiling arrays
 - Pathway / gene set analysis tools
- Part 3: Quantitative sequencing
- Subtypes and experimental procedures: expression (RNA-seq), enrichment (ChIP-seq, MBD-seq/MethylCap-seq), bisulfite sequencing for DNA-methylation, metagenomics (species diversity)
 - Raw data formats (e.g. FASTQ) and quality control
 - Sequence alignment: standard mapping (e.g. BOWTIE), splice site aware mapping and isoform calling for RNA-seq (e.g. STAR) and associated aligned data formats (e.g. SAM/BAM)
 - Manipulation of large (sequencing) data sets (e.g. SAMTOOLS)
 - Data summary using existing annotation (e.g. gene level, species level) or *de novo* (peak calling, particularly for ChIP-seq, e.g. MACS)
 - Application of data normalization strategies (library size, trimmed mean of m-values method, etc.)
 - Statistical analysis: identification of differential expression/enrichment (e.g. EdgeR, DESeq)
 - SNP and mutation calling, GWAS
 - Pathway / gene set analysis tools
- Part 4: High-throughput PCR based methods
- qPCR: quantification models, error propagation, normalization, experimental design, assay design
 - dPCR: Poisson correction, advanced copy number analysis, power analysis, assay design

Initial competences

Basics molecular biology, statistical data analysis and scientific programming.
Conceptual insight in general(ized) linear models is a plus.

Final competences

- 1 Understanding the application domains of high-throughput omics research, the generation of these data and the implications regarding bias and variance.
- 2 Knowledge of the different steps of the full omics analytical pipeline and how they are linked in a multidisciplinary fashion.
- 3 Be familiar with relevant omics databases, and its impact on scientific integrity.
- 4 Have insight into the fast evolution of omics technologies and associated required bioinformatics solutions, and how this will/may have major implications for society for several important application domains such as medicine, food production and ecology.
- 5 Critical selection, evaluation and application of methods for high-throughput data (pre) processing.
- 6 Divide the omics analytical pipeline into separate steps and critically select the most appropriate methodology to solve each of these steps, taking into account both the practical applicability and the limitations/consequences of each methodology.
- 7 Report the outcome and particularities of a high-throughput omics analytical project to a non-specialist.
- 8 Apply fundamental knowledge of different subdomains (statistics, (bio)informatics) to plan and perform a complex data-analytical project at a professional level, taking into account financial, computational and other limitations.
- 9 Infer knowledge from complex high-throughput data using quantitative methods.
- 10 To critically select and evaluate methods for high-throughput data-analysis.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Group work, lecture, seminar: practical PC room classes

Learning materials and price

- Lesslides beschikbaar op Minerva
- Broncode voorbeelden worden via Minerva ter beschikking gesteld van de student.

References

Course content-related study coaching

- Practica worden begeleid door assistenten.
- Bijkomende informatie via Minerva.

Evaluation methods

end-of-term evaluation and continuous assessment

Examination methods in case of periodic evaluation during the first examination period

Open book examination, oral examination

Examination methods in case of periodic evaluation during the second examination period

Open book examination, oral examination

Examination methods in case of permanent evaluation

Assignment

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible

Extra information on the examination methods

- assignment: completion of a high-throughput omics data analytical project starting from several sources of raw data and communication of the results in a report for the non-specialist (e.g. clinician)
- open book examination: well defined problems associated with individual steps in the omics data analytical pipelines will be presented,
- oral examination: feedback regarding the assignment will be provided, and methodological aspects will be discussed

Calculation of the examination mark

- Assignment report + discussion during oral examination: 50% of total mark
- Open book examination: 50% of total mark