Computational Science: Bioinformatics

Research Group for Biomedical Informatics & Centre for Bioinformatics

16 August 2022

Torbjørn Rognes
Department of Informatics, UiO
torognes@ifi.uio.no

UiO : University of Oslo

Computational Science at IFI

Study programme (master):

- Computational Science (CS)

Study programme options at IFI:

- CS: Bioinformatics
- CS: Imaging and Biomedical Computing

ML Research groups:

- BMI – Biomedical informatics
  (Biomedisinsk informatikk)
- DSB – Digital Signal Processing and Image Analysis
  (Digital signalbehandling og bildeanalyse)

Centre:

- SBI – Centre for Bioinformatics

Scientific staff at BMI & SBI

Please contact any of the scientific staff if you are interested in a master thesis within bioinformatics.

Torbjørn Rognes
BMI leader & Professor

Ole Christian Lingjærde
Professor

Geir Kjetil Sandve
Professor

Rolf Skotheim
Professor (20%)

Knut Liestøl
Professor emeritus

Eivind Hovig
SBI leader & Professor (70%)

Trine B. Rounge
Assoc. prof SBI

Lex Nederbragt
Senior lecturer SBI (30%)

Jonas Paulsen
Assoc. prof SBI

Oleksandr Frei
Assoc. prof SBI

Jon K. Lærdahl
Assoc. prof SBI (20%)

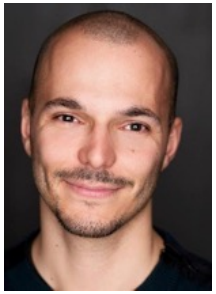
Simula: Xing Cai, Are Magnus Bruaset, Aslak Tveito

Postdocs, PhD students & Engineers


Chakravarthi Kanduri


Ivar Grytten

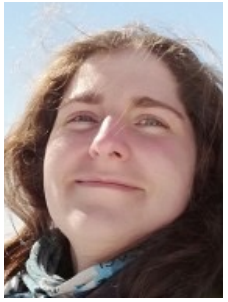

Knut D. Rand


Enrico Riccardi


Paula Istvan


Bayram Cevdet Akdeniz


Christian D. Bope


Jeanne Chèneby


Einar Birkeland

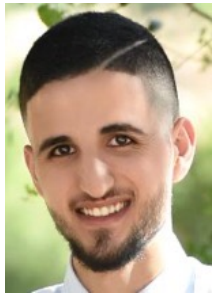

Jean-Marc Costanzi


Milena Pavlovic


Cecile Bucher-Johansen


Lonneke Scheffer


Sumana Kalyanasundaram


Ghadi S. Al Hajj


Jonas Meier Strømme


Ekaterina Avershina


Maja Jacobsen


Sveinung Gundersen


Morten Johansen

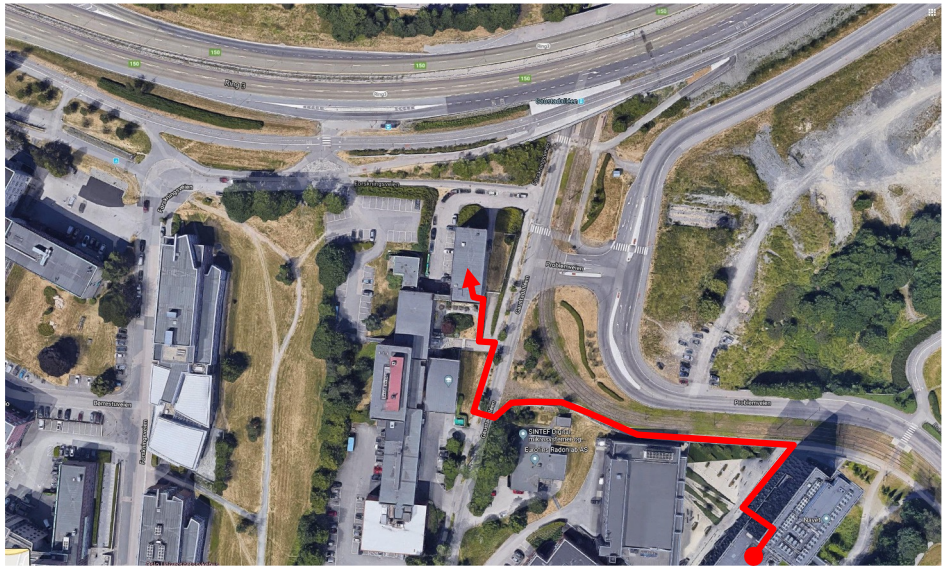

Nazeefa Fatima


Jushua Baskaran


Lev Mozhaev

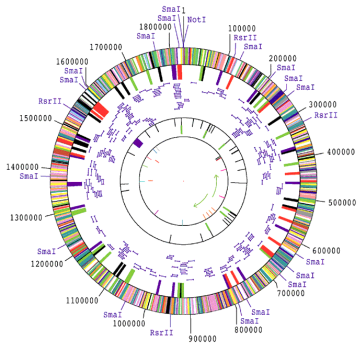
Where are we located?

Floors 1 and 2 in the former NEMKO building, Gaustadalléen 30, D



BMI/SBI

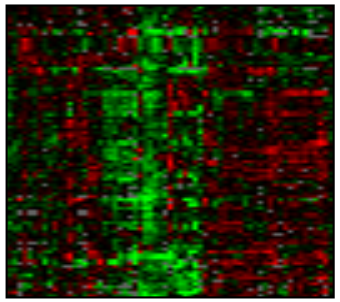
What is bioinformatics?



An interdisciplinary subject:
Computer science/maths/statistics + biology/medicine

Definition: Bioinformatics is the development and use of computational and mathematical methods to gather, process and interpret molecular biological data.

Aim of research: To increase our understanding of biological mechanisms while developing better theories and methods in computer science and statistics.

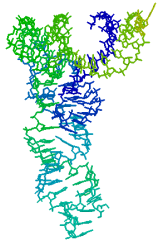
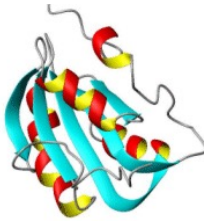
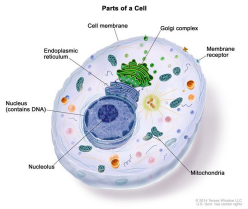


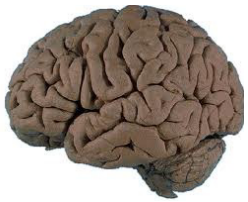
Methods: Machine learning, statistics, graphs, algorithms, visualisation, parallelisation, efficient data structures, ...

Data: Genomes, genes, DNA, RNA, proteins, immune receptors, gene expression levels, 3D molecular structures, cells, organisms, populations, gene variants and mutations, patient info, ...

```
P13051|UNG_HUMAN/1-313     144 K D V K V V I L G Q D P Y H G P N Q A H G L C F S V Q R P V P P P P S L E N I Y K E L S T D I E D F V H 195
P12887|UNG_YEAST/1-359     152 N K V K V V I I G Q D P Y H N F N Q A H G L A F S V K P P T P A P P S L K N I Y K E L K Q E Y P D F V E 203
P12295|UNG_ECOLI/1-229      54 G D V K V V I L G Q D P Y H G P G Q A H G L A F S V R P G I A I P P S L L N M Y K E L E N T I P G F T R 105
P16769|UL114_HCMVA/1-250    81 E Q V R V V I V G Q D P Y C D - G S A S G L A F G T L A G R P P P P S L N N V F R E L A R T V D G F Q R 131
```

Bioinformatics at many levels



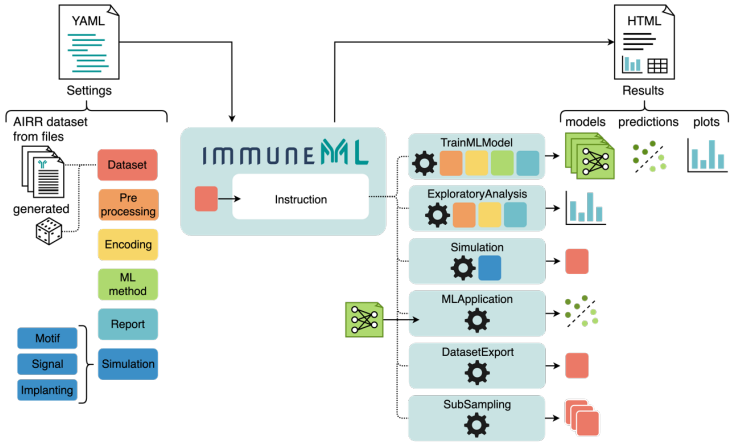

DNA	RNA	Protein	Cell	Organ	Individual	Population	Biosphere
Genomics	Transcript-omics	Proteomics	System biology	Neuro-informatics	Precision medicine	Population genetics	Metagenomics
Genome assembly	RNomics	Structural biology	Cell simulation	Organ modelling/ simulation	Variant detection	Epidemiology	Evolutionary biology
Genefinding	Microarrays	Drug design	Metabolism studies	Meta-genomics			Phylogenomics
Annotation	RNA-folding	MS analysis		Cancer genomics			
”Read mapping”	RNA-seq	Binding site analysis					
ChIP-Seq	Structure biology	Interaction networks					

Main research areas of BMI/SBI

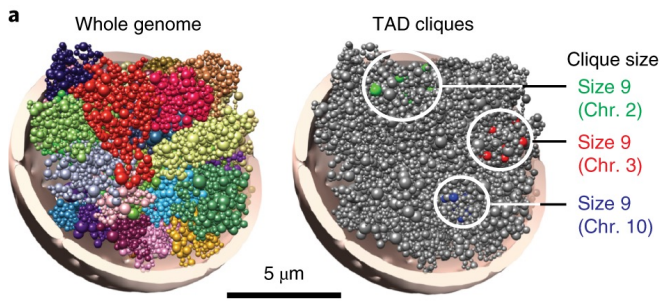
Advanced use and development of new methods and tools in computer science and statistics to analyse molecular data

Research projects

- ML for Immune receptor repertoire analysis + Genome statistics and visualisation (GK Sandve)
- Cancer genomics (OC Lingjærde)
- Cancer sequencing & Precision medicine (E Hovig)
- Microbiome bioinformatics and efficient sequence comparison (T Rognes)
- Microbiomes and cancer (T Rounge)
- Modelling of 3D organisation of DNA in the nucleus (J Paulsen)
- Genome heterogeneity in cancer (RI Skotheim)
- Genome assembly and representation (AJ Nederbragt)
- Genetics of Psychiatric Disorders (O Frei)
- Medical statistics in cancer biology (K Liestøl)



immuneML



DNA 3D organisation

Major collaborators and partners

- Oslo University Hospital (adjunct positions)
- OsloMet
- Akershus University Hospital (Ahus)
- Cancer Registry of Norway
- Big Insight (SFI)
- KG Jebsen Celiac Disease Research Centre
- Centre for Computational Inference in Evolutionary Life Science (CELS) ("Endringsmiljø" / Strategic Research Initiative)
- PharmaTox ("Endringsmiljø" / Strategic Research Initiative)

CS Programme structure

4. semester	Master's thesis	Master's thesis	Master's thesis
3. semester	Master's thesis	Master's thesis	Master's thesis
2. semester	IN4200 – High-Performance Computing and Numerical Projects / Mandatory course	Elective course	Elective course
1. semester	FYS-STK4155 – Applied Data Analysis and Machine Learning / Mandatory course + HSE courses	MAT4110 – Introduction to Numerical Analysis / Elective course	Elective course
	10 ECTS credits	10 ECTS credits	10 ECTS credits

Alt: 10 credits elective course

Alt: 10 credits master thesis

Mandatory, 2 out of these 3 courses:

- IN4200 – High-Performance Computing and Numerical Projects
- FYS-STK4155 – Applied Data Analysis and Machine Learning
- MAT4110 – Introduction to Numerical Analysis

Recommended elective courses

CS: Bioinformatics

- IN4030 – Introduction to bioinformatics (spring)

- IN4050 – Introduction to artificial intelligence and machine learning (spring)
  Be aware of 3 credits overlap between FYS-STK4155 and IN4050, usually solved by a small special curriculum.

- IN-BIOS5000 – Genome Sequencing Technologies, Assembly, Variant Calling and Statistical Genomics (autumn 2022)

- BIOS-IN5410 – Bioinformatics for Molecular Biology (autumn)

Other relevant courses

- IN4020 – Database systems (spring)
- IN4110 – Problem solving with high level languages (autumn)
- IN4120 – Search technology (autumn)
- IN5050 – Programming Heterogeneous Multi-Core Architectures (spring)
- IN5320 – Development in Platform Ecosystems (autumn)
- IN5400 – Machine Learning for Image Analysis (spring)
- IN5490 – Advanced Topics in Artificial Intelligence for Intelligent Systems (autumn)
- IN-STK5000 – Adaptive methods for data-based decision making (autumn)
- BIOS-IN5010 – Computational physiology (summer)
- STK4051 Computational Statistics (spring)
- STK-IN4300 – Statistical Learning Methods in Data Science (autumn)
- TEK5020 – Pattern recognition (autumn)
- MEDFL5155 – Introduction to statistics and bioinformatics for the analysis of large-scale biological data (autumn, 5 credits)

Please see the Centre for Bioinformatics' studies page:

https://www.mn.uio.no/sbi/english/studies/

Master thesis at BMI/SBI

- Look for available master projects on website:
  https://www.mn.uio.no/ifi/english/research/groups/bmi/master-projects/
- Presentation of new projects in September/October
- Contact potential supervisors
- Usually 60 credits (0+10+20+30 credits in each semester)
- Make contact early!

Deadlines:

- 1 December 2022: Study plan, select project and supervisor
- 1 June 2023: Essay
- 15 May 2024: Master thesis

Master theses submitted 2020-2022

- Tool to remove specific organisms from microbiome sequencing data (2022)
- A Web Application for MirMachine, a MicroRNA Annotation Tool (2022)
- An assessment of methods for determining and visualizing differential Hi-C contact maps (2022)
- Developing a Toolset for 3D Genome Visualisation and Analysis (2022)
- Comparing Artificial Neural Networks and Microbial Genome Representation Methods for Taxonomic Classification (2021)
- Developing a pipeline for deep analysis of cancer causing genes in the gut microbiome (2021)
- Interpretability of convolutional neural networks when applied to receptor sequence data (2020)
- Algorithms for exploring heterogeneity in prostate cancer patients from RNA-seq data (2020)

Buddy

- Behnoosh Ashrafi
- behnoosh.ashrafi@fys.uio.no
- Buddy for all CS-students at IFI (both CS:Bioinformatics & CS:Imaging and biomedical computing)
- She started as a student at CS:Bioinformatics 1 year ago.