

# Institute of Biomedical Informatics



WISSEN TECHNIK LEIDENSCHAFT





Leila Taher

Fernando Fernandez Mendoza



Gerhard Thallinger

Bakk::Mas Days 2024

https://www.bioinfo.tugraz.at/



Bioinfo

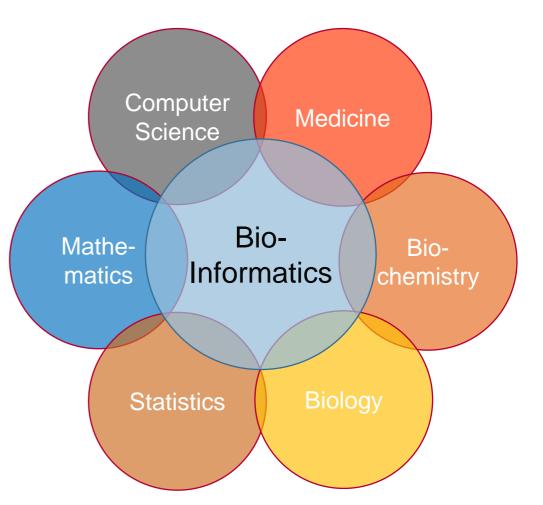
## What is "Bioinformatics"

"Ask five bioinformaticians for a definition and you'll get six different ones"

### **Biology and Informatics = Bioinformatics**

- Theoretical biology
- Computational biology
- Systems biology
- (Theoretical) ecology
- Biomathematics
- Biostatistics
- Computational omics
- Computational evolutionary biology
- Applied bioinformatics

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger

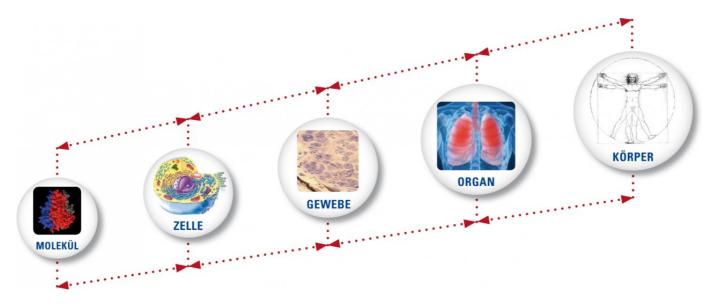




Bioinfo

### What is "Bioinformatics"?

Elucidation of biomolecular relations in organisms at the cell-, tissueand organ level ...



... by developing and applying methods from computer science, mathematics and statistics on "wet-lab" generated data.

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger



Algorithms and methods

Tools and

databases

Bioinfo

### What is "Bioinformatics"?

Data analysis and interpretation

"Bioinformatics is the field of science in which biology, computer science, and information technology merge into a single discipline."\*

\* NCBI, https://www.ncbi.nlm.nih.gov/Class/MLACourse/ Modules/MolBioReview/bioinformatics.html

> 12.03.2024 Institute of Biomedical Informatics

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger



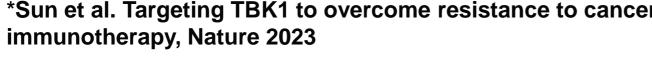
Bioinfo

### **Bioinformatics** applications

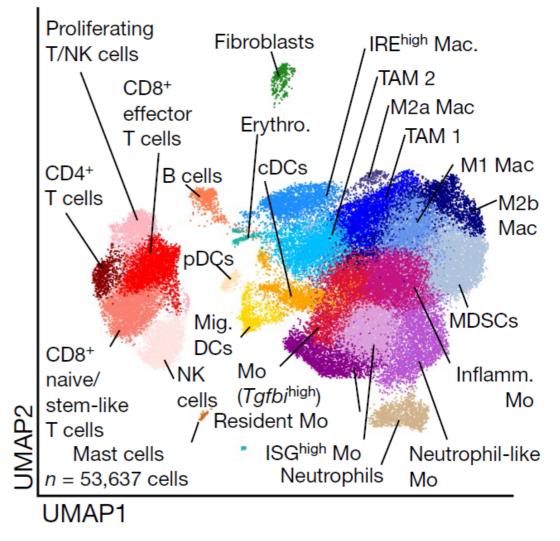
### Fight resistance to cancer immunotherapy\*

- Single-cell sequencing data analysis
- Flow cytometry data analysis

## \*Sun et al. Targeting TBK1 to overcome resistance to cancer









Bioinfo

# **Bioinformatics applications**

### **Completion of the human genome\***

- Long-read sequencing
- Genome assembly
- Genome annotation and visualization

\*Nurk et al. The complete sequence of a human genome, Science 2022

9 40 50 60 70 80 90 100110120130140150160170180190202 chr19 chr20 EUR SAS EAS AMR Contromoric satellite Segmental duplication CHM13 exclusive gene density GRCh38 gene density Fixed GRCh38 gaps and issue

internet and the second s

12.03.2024 Institute of Biomedical Informatics

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger

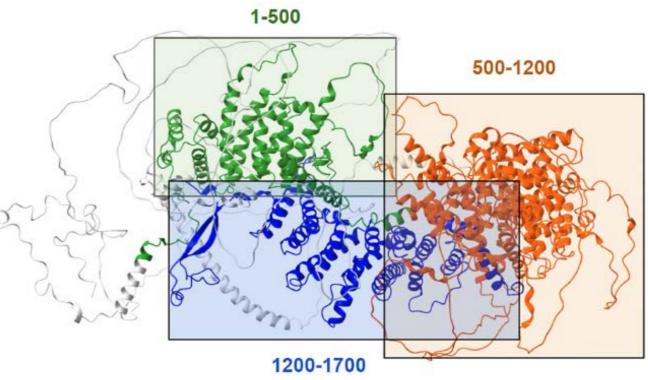


Bioinfo

### **Bioinformatics applications**

# 3D Structure database of all known proteins\*

- Deep-learning for structural prediction
- Database application
- Web-frontend



\*Varadi et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with highaccuracy models, Nucleic Acids Res 2022

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger

Curriculum

**Bachelor**:



Bioinfo

### Bioinformatics in the BME curriculum I

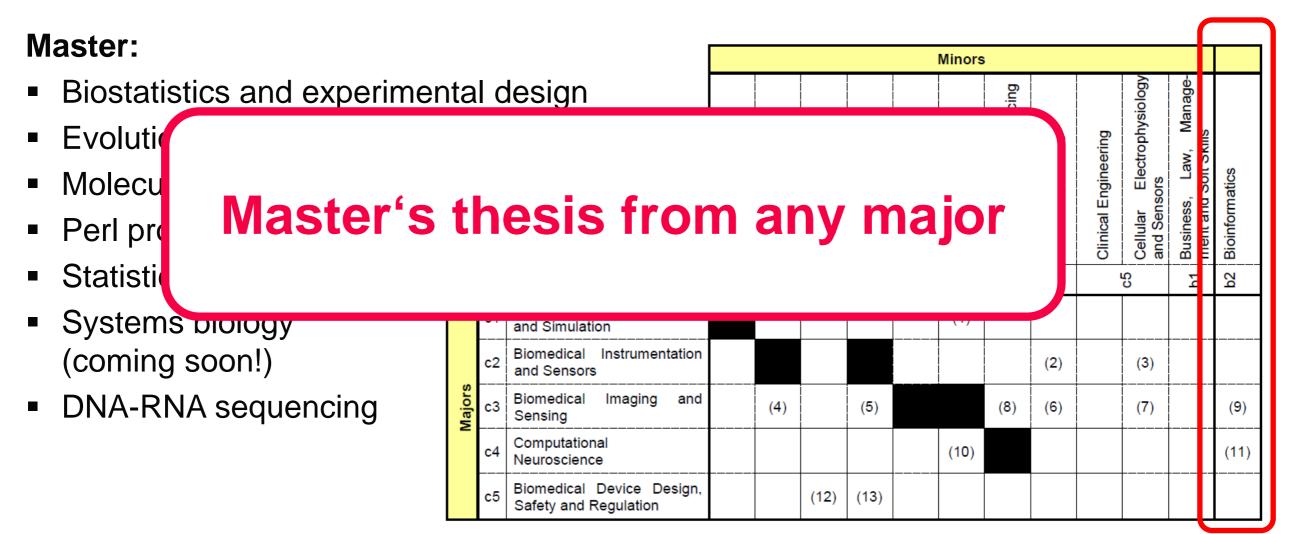
- Grundlagen der Informatik
- Grundlagen der Molekular- und Zellbiologie
- Algorithmen in der Bioinformatik
- Verfassen wissenschaftlicher Arbeiten

Curriculum



Bioinfo

## Bioinformatics in the BME curriculum II



Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger





### Thesis topics I

- Characterization of transposons as drivers of genome evolution
- Generation and/or analysis of Oxford Nanopore sequencing data
- Identification of structural variants associated with different cancer types
- Comparative transcriptomic analysis of cancer tumours and derived cell lines
- Implementation of models for simulating different population genetics scenarios
- Investigation of the relationship between enhancer RNA (eRNA) secondary structure and DNA enhancer function
- Application of deep learning in the analysis of omics data
- Long-read transcriptomics (ISO-seq) of fungi

Theses

Bioinfo



- Genome assembly and annotation
- Summarization of functional enrichment analyses
- Visualization of genome annotation
- Investigation of the lipidome using mass spectrometry
- Integrative analysis of data from different omics technologies with applications in biotechnology
- Characterization of the microbiome
- Comparative genomics of microorganisms
- Expansion of the regulatory network of a yeast
- Analysis of the alterations of the genome and epigenome of mammalian cells during fermentation

Theses

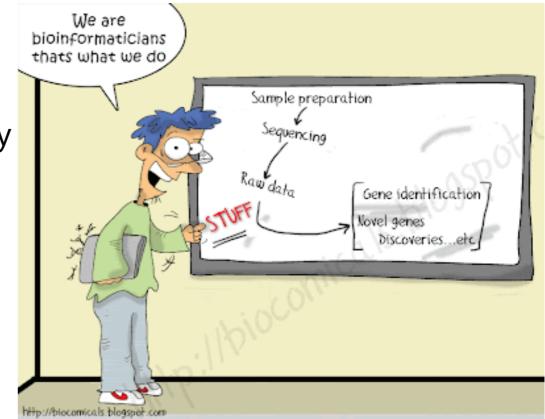


Bioinfo

### Bioinformatics core skills

- Proficiency in at least one programming language.
  R, Python and/or Perl knowledge is a plus
- For specific projects you may need to be familiar with other programming languages
- (Good) understanding of molecular cell biology
- Linux (command line)
- Algorithms
- Databases
- Statistics
- Time management
- Readiness for self-study
- Patience

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger







Bioinfo

### What you should expect

- No pure literature work (there is always a practical part)!
- Frequent (weekly or biweekly) meetings between you and your supervisor
- Thesis supervision agreement
- Research proposal and short presentation at the beginning of your project
- Final presentation shortly before submitting your thesis
- Regular participation in our weekly lab meetings

Conclusions



Bioinfo

### Collaborations

### **Collaboration with:**

- Harvard Medical School (Boston)
- Sloan Kettering Cancer Center (New York)
- New York University, Department of Medicine (New York)
- => Stay abroad possible

### Thesis combined with a Fem In Tech scholarship:

- Female students only
- 2-6 month scholarship
- Attractive remuneration

Conclusions

Bioinfo

### Contact

Talk to us. If you are interested in a particular bioinformatics topic that is not part of the list, we will do our best to design a project around it!

mailto:Bakk-Mas@genome.tugraz.at





Leila Taher



Fernando Fernandez Mendoza



Gerhard Thallinger

12.03.2024 Institute of Biomedical Informatics

Bakk::Mas Days 2024 Leila Taher, Fernando Fernandez Mendoza, Gerhard Thallinger

