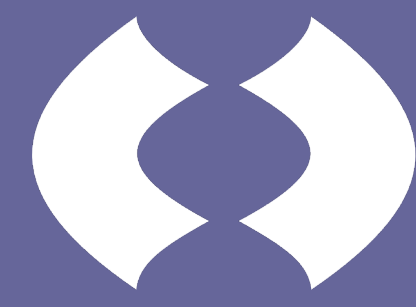


BIOMOLECULAR DATA SCIENCE IN PNEUMOLOGY

Lennart Bartels & Inken Wohlers
Data Science, Research Center Borstel



Forschungszentrum Borstel
Leibniz Lungenzentrum

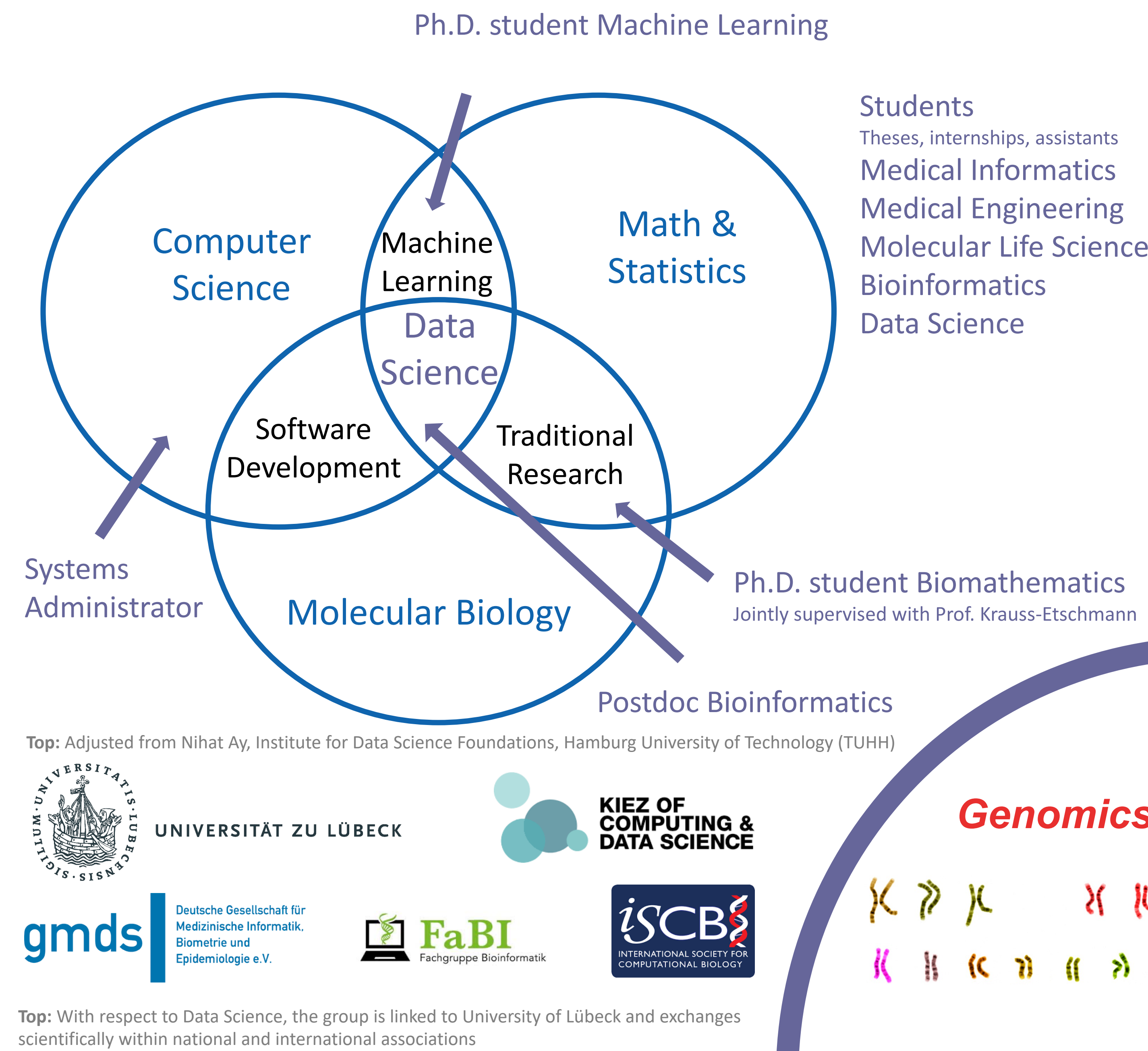
OBJECTIVES

- Bioinformatics consulting
- Computing & OMICS data storage infrastructure
- Standardized OMICS data analyses
- Computational and collaborative research projects

EXPERTISE

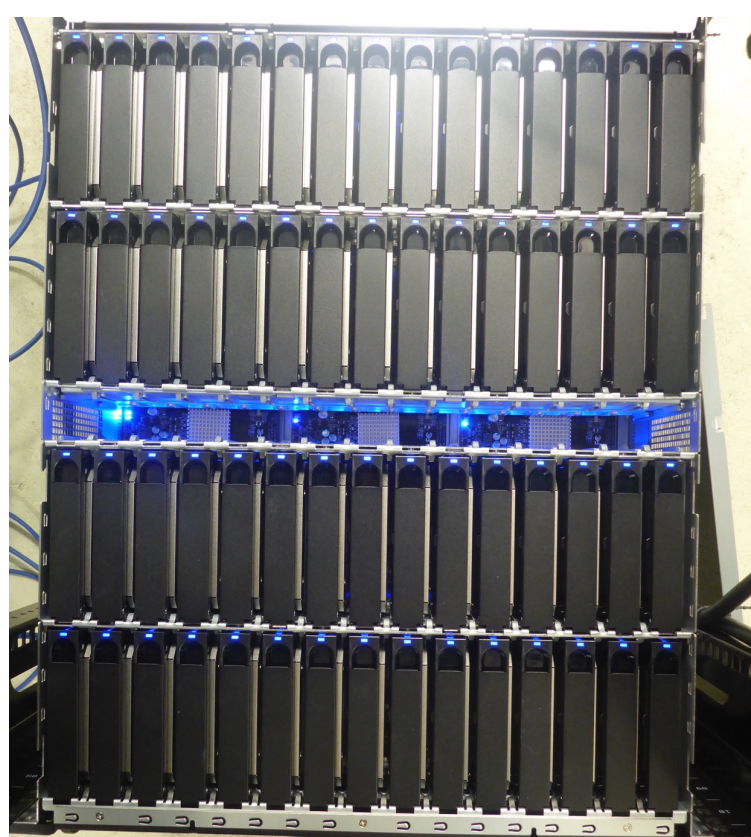
- Sequencing data analyses
- Genomics & Transcriptomics
- Systems Genetics
- Long read sequencing-based Epigenomics & Epitranscriptomics

DATA SCIENCE GROUP SETUP

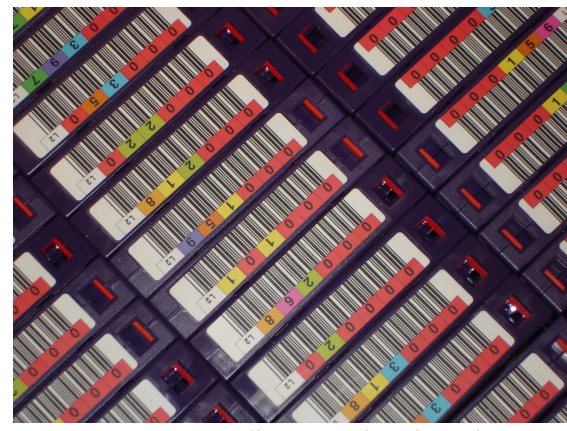


COMPUTING & STORAGE

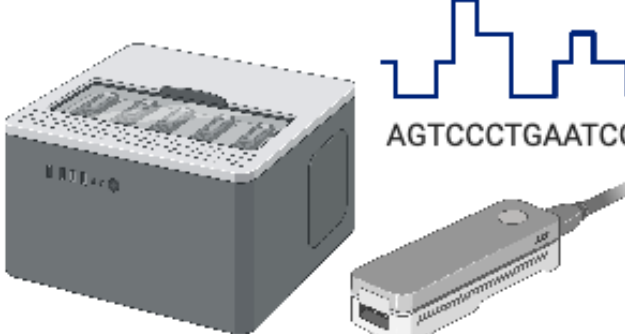
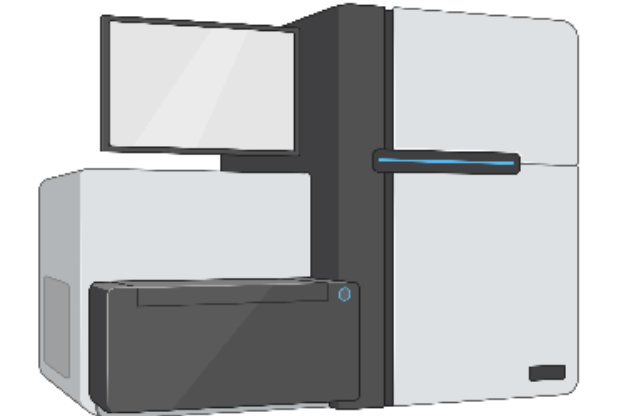
- Harmonizing and extending high performance computing infrastructure at Research Center Borstel
- Setup of OMICS data storage



Top: Research center Borstel has a server room solely dedicated to OMICS data storage and analyses, with a new, extendable 1 PB Synology NAS (HD6500) for data storage



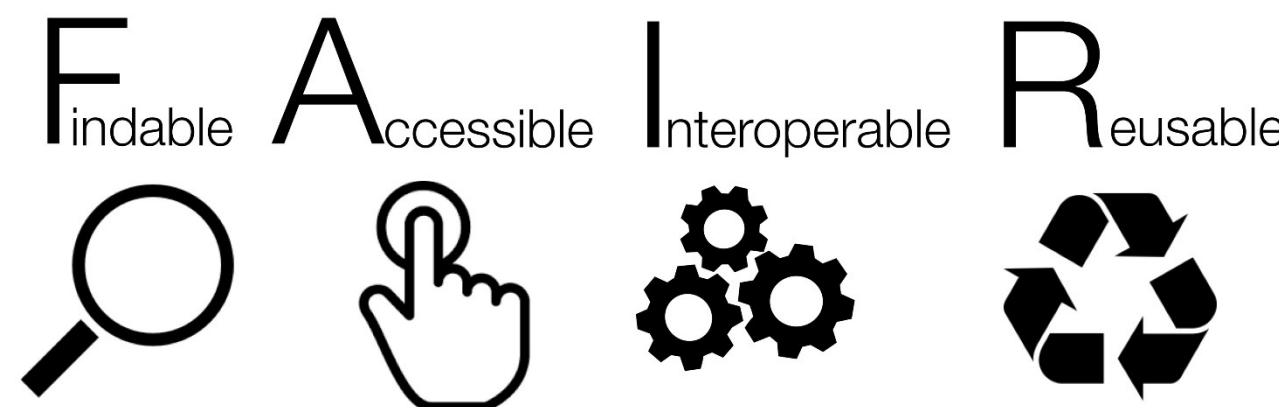
Left: Raw OMICS data backup on tape for long-term data retention



Left: The compute cluster uses slurm for job scheduling



Left/bottom: The group is in the process of setting up a FAIR, center-wide data management for OMICS data.



Top: Illumina, ONT and PacBio sequencers on campus generate big data

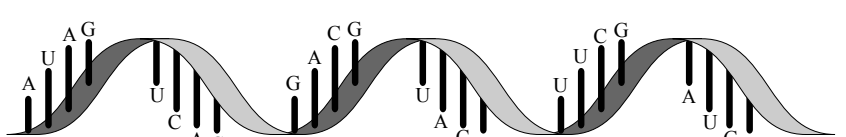
OMICS

The collective characterization and quantification of many biological entities as well as their function, structure and dynamics through high-throughput technologies

Genomics



Transcriptomics

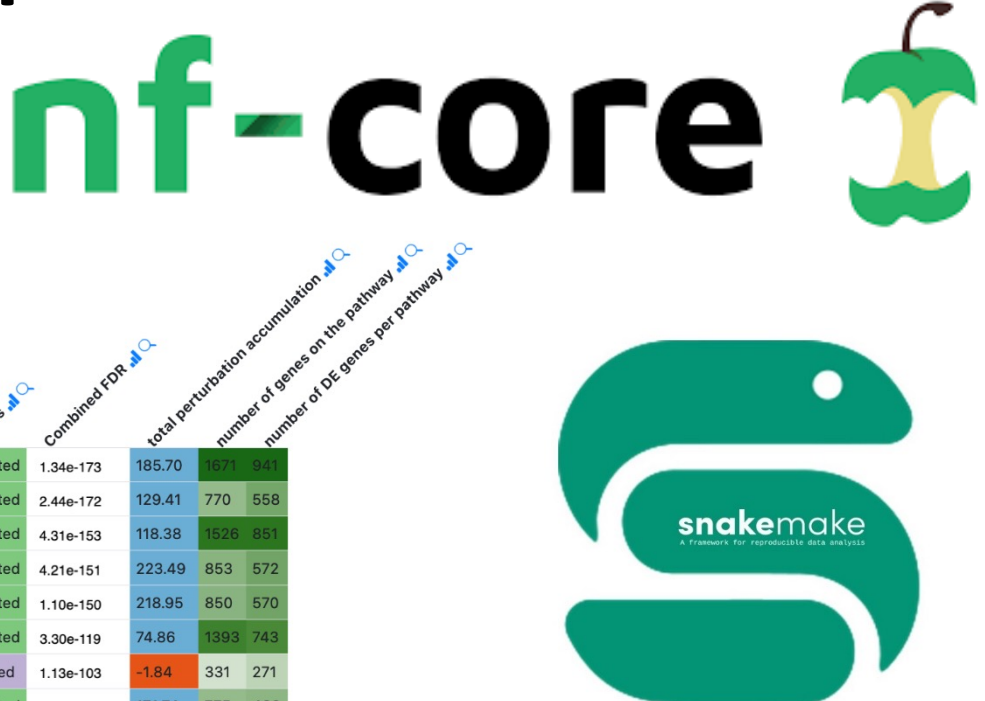


Proteomics

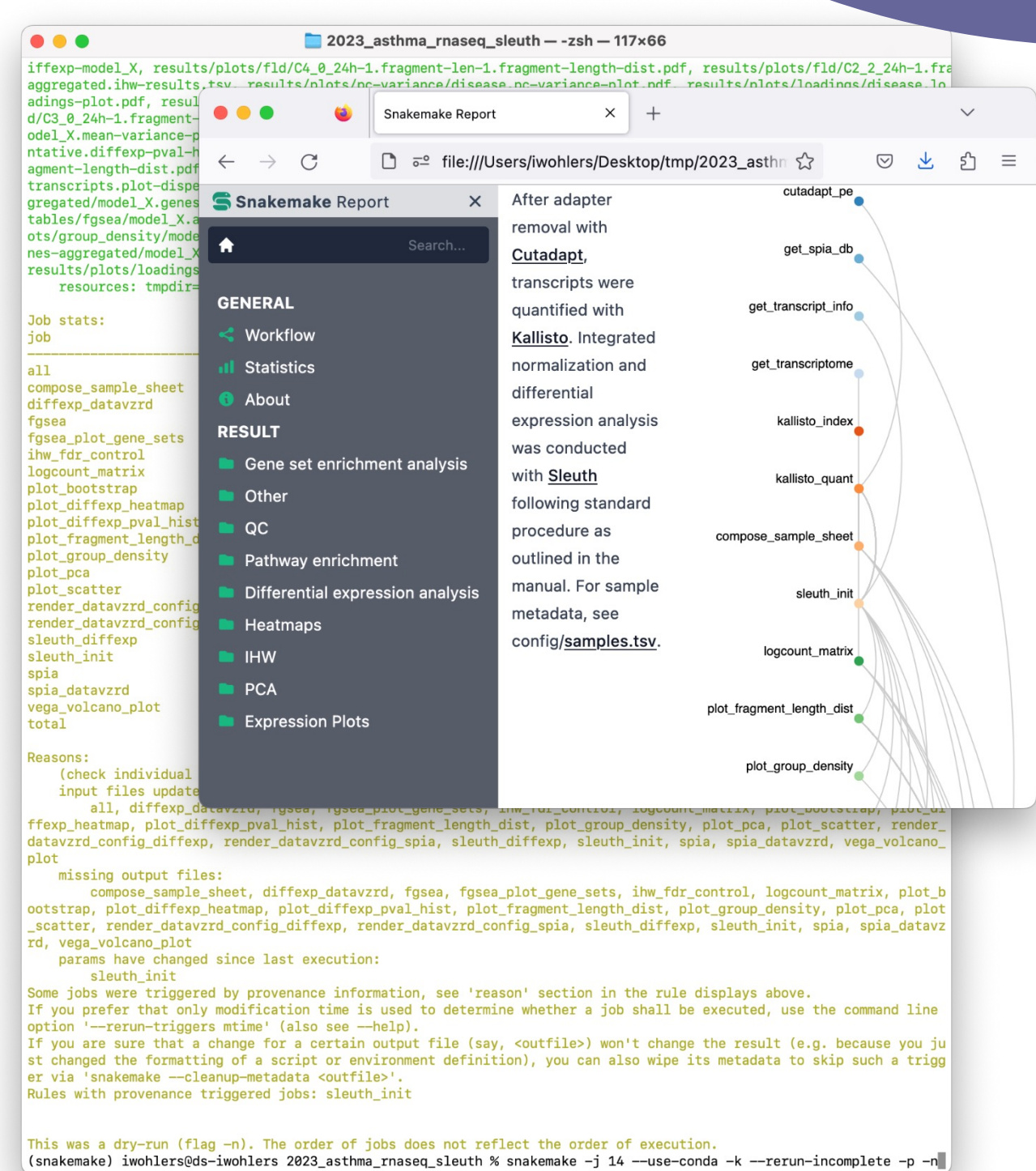


STANDARDIZED ANALYSES

- Reproducible, community supported workflows e.g. from nf-core or from the snakemake workflow catalog
- Available: bulk RNA-Seq data analyses using workflows rna-seq-kallisto-sleuth and rna-seq-star-deseq2

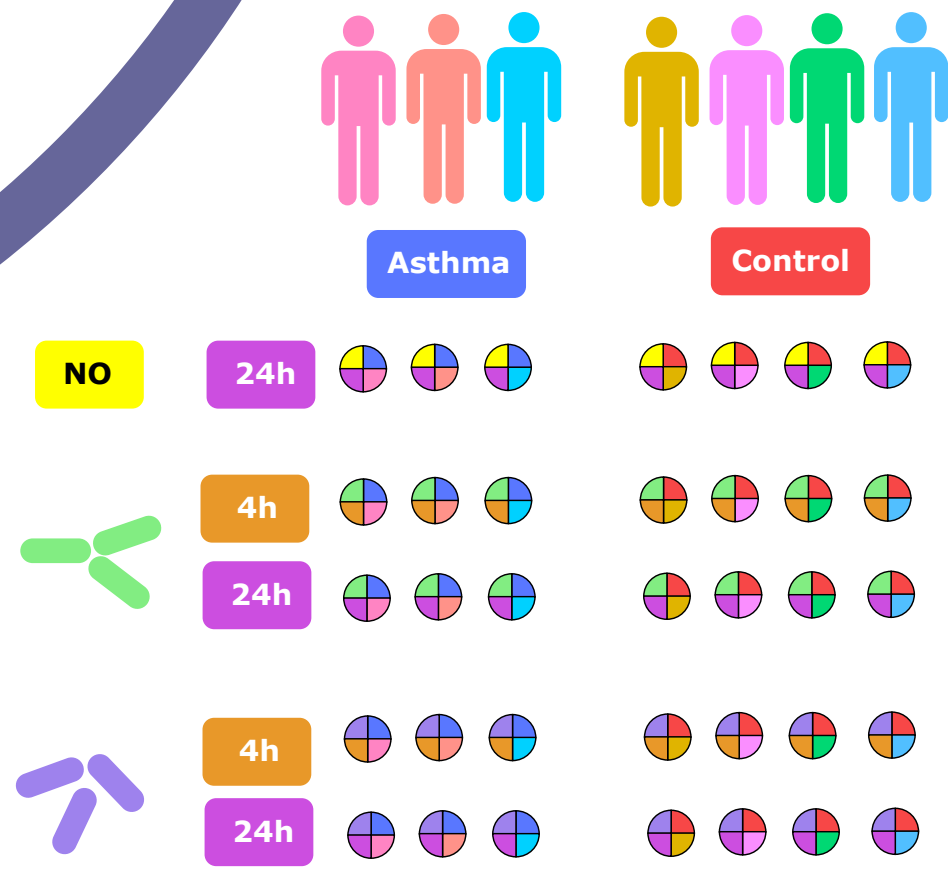


Left / Bottom: HTML result tables generated by the RNA-Seq workflow

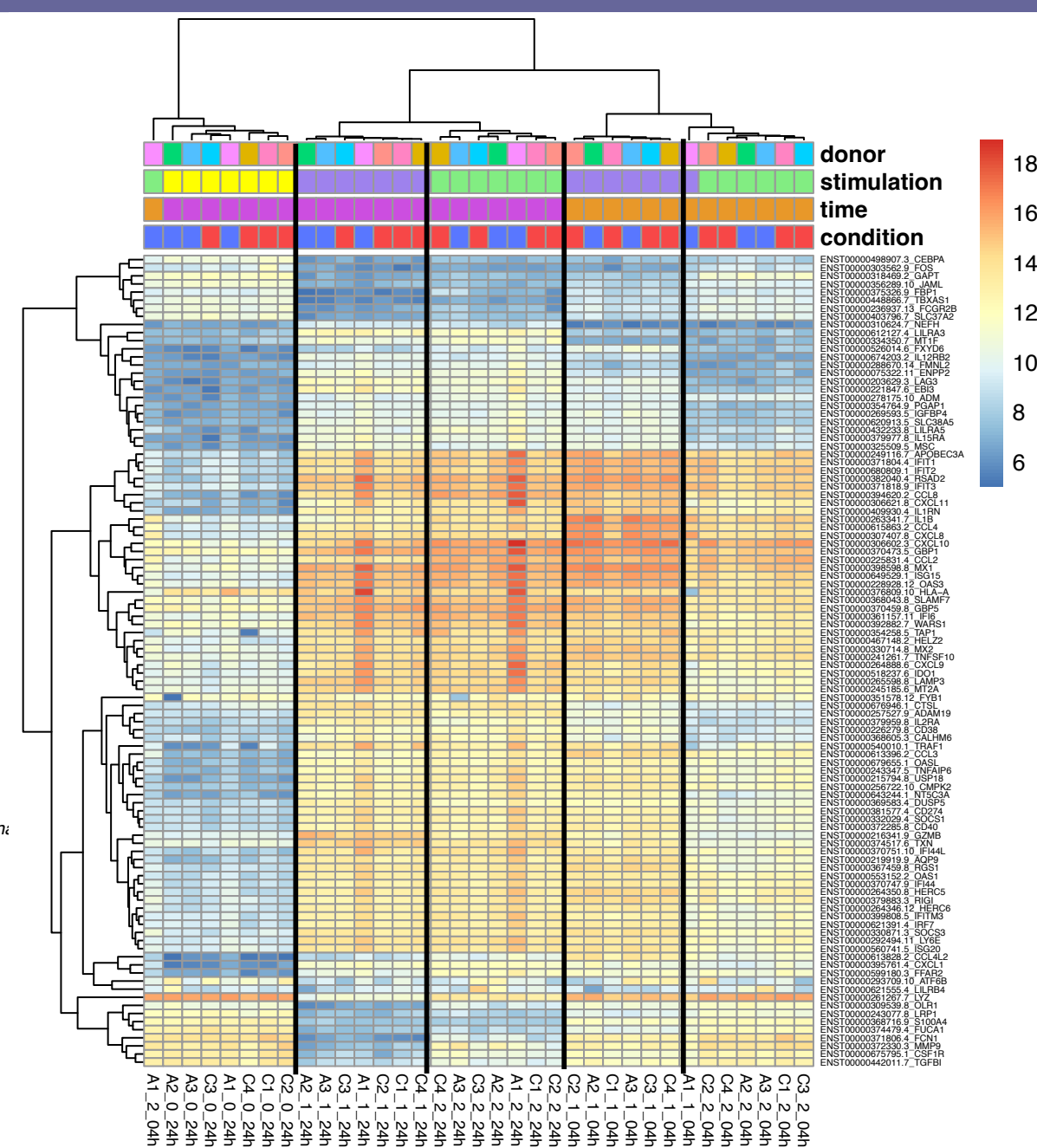


Top: Execution of snakemake workflow and a snakemake results report, which can be uploaded to the Zenodo research data repository for obtaining a citable persistent digital object identifier (DOI)

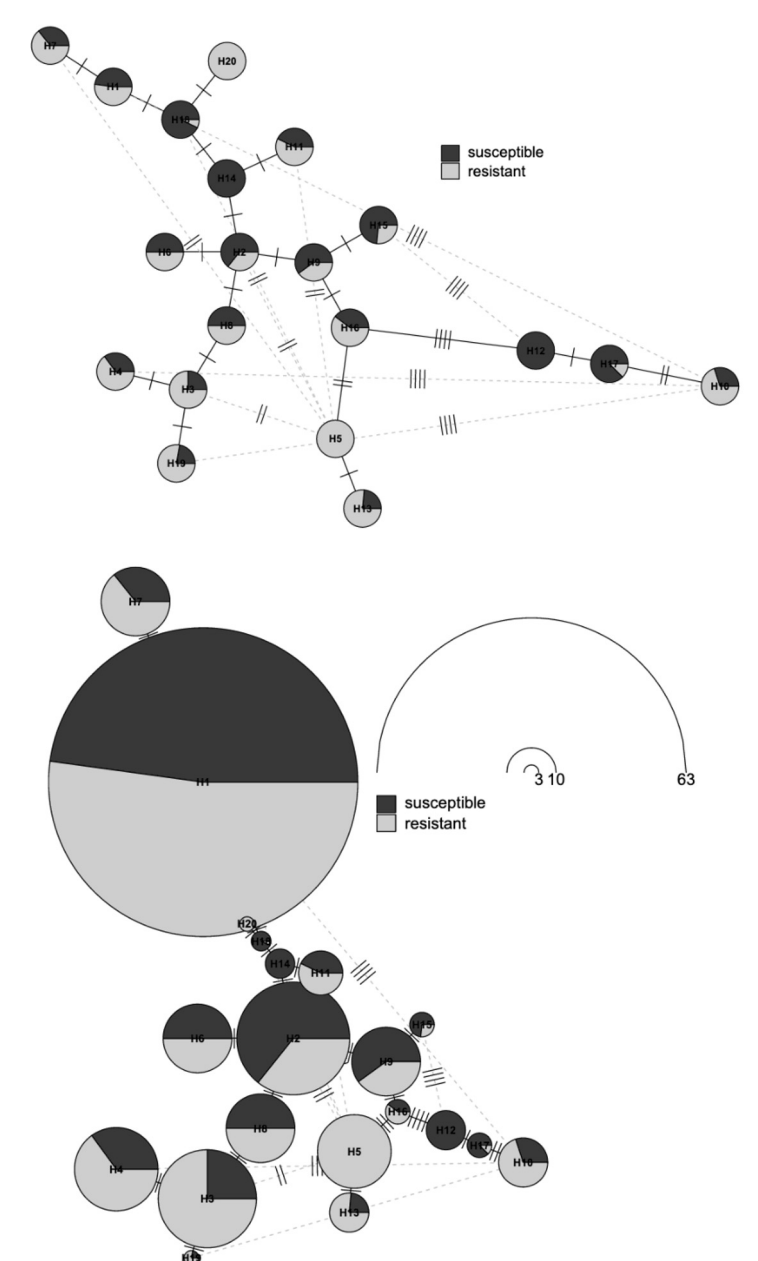
RESEARCH PROJECTS



Top: Collaboration with Research Center Borstel priority area *Chronic Lung Diseases* – Archaea as possible immune system trainers in the context of asthma, investigated via transcriptome analysis of *methanospira* stimulated PBMCs



Right: Collaboration with Research Center Borstel priority area *Infections* – Network of *Haemophilus influenzae ftsI* gene-related haplotypes together with antibiotic resistance and susceptibility of corresponding isolates, investigated within an interactive, web-based application developed in the Data Science group



- RNA-Seq data analysis
- Nanopore-based methylation analyses
- Nanopore-based Epitranscriptomics
- Association and polygenic scores
- Deep learning-based phenotype prediction
- *De novo* assembly
- Genomic haplotype-based characterization

Collaborations: Research Center Borstel
Leibniz Lung Center



UCCSH

Universitäres Cancer Center
Schleswig-Holstein



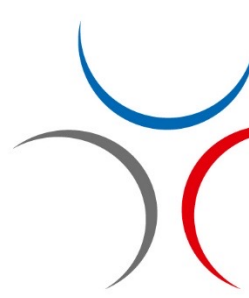
UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR
NEURO- UND BIOINFORMATIK



Leibniz-Institut
für Präventionsforschung
und Epidemiologie – BIPS



جامعة خليفة
Khalifa University



PMI Precision Medicine
in Chronic Inflammation