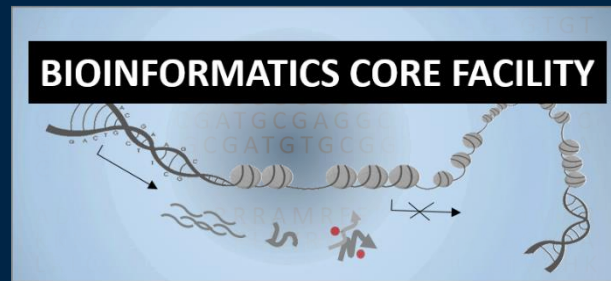


INTRODUCTION TO THE BIOINFORMATICS CORE FACILITY



Paludan group meeting
February 2023

ABOUT ME

Academic appointment

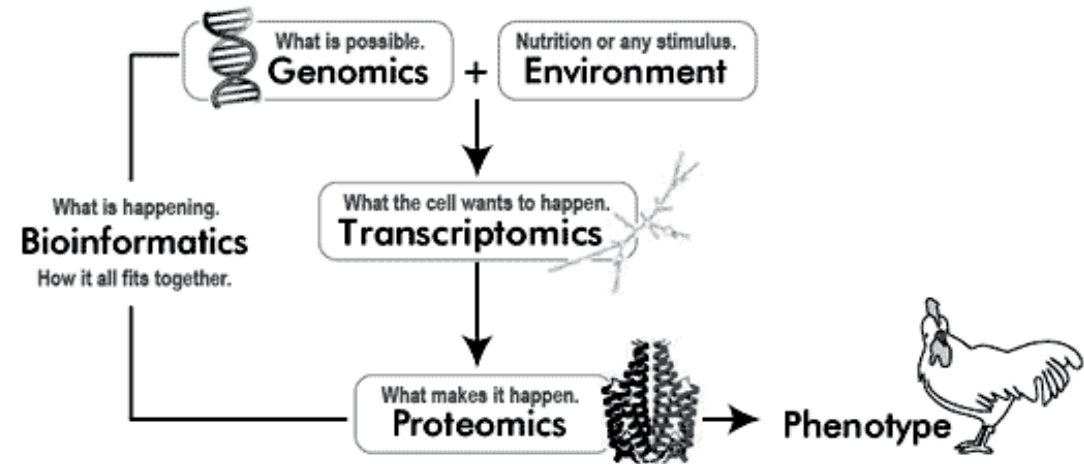
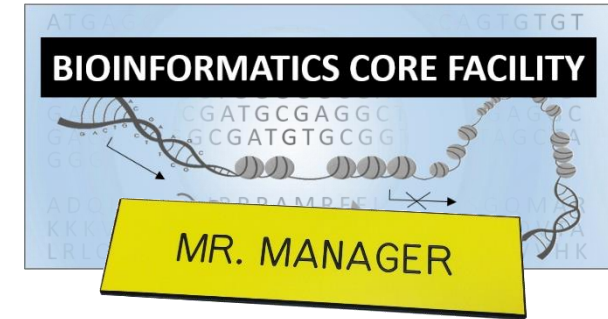
- Associate Professor in Bioinformatics

Educational Background

- MSc. Molecular Biology (2010)
- PhD. Neuroscience (2014)

Research Focus

- Functional Genomics
- Psychiatric Genetics
- Precision Psychiatry
- Systems Biology



ABOUT ME

Academic appointment

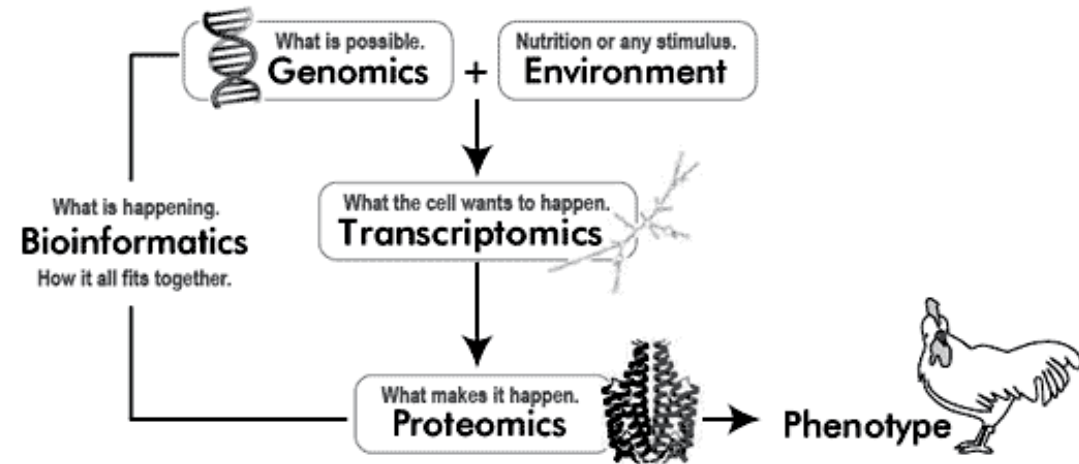
- Associate Professor in Bioinformatics

Educational Background

- MSc. Molecular Biology (2010)
- PhD. Neuroscience (2014)

Research Focus

- Functional Genomics
- Psychiatric Genetics
- Precision Psychiatry
- Systems Biology



ABOUT ME

Academic appointment

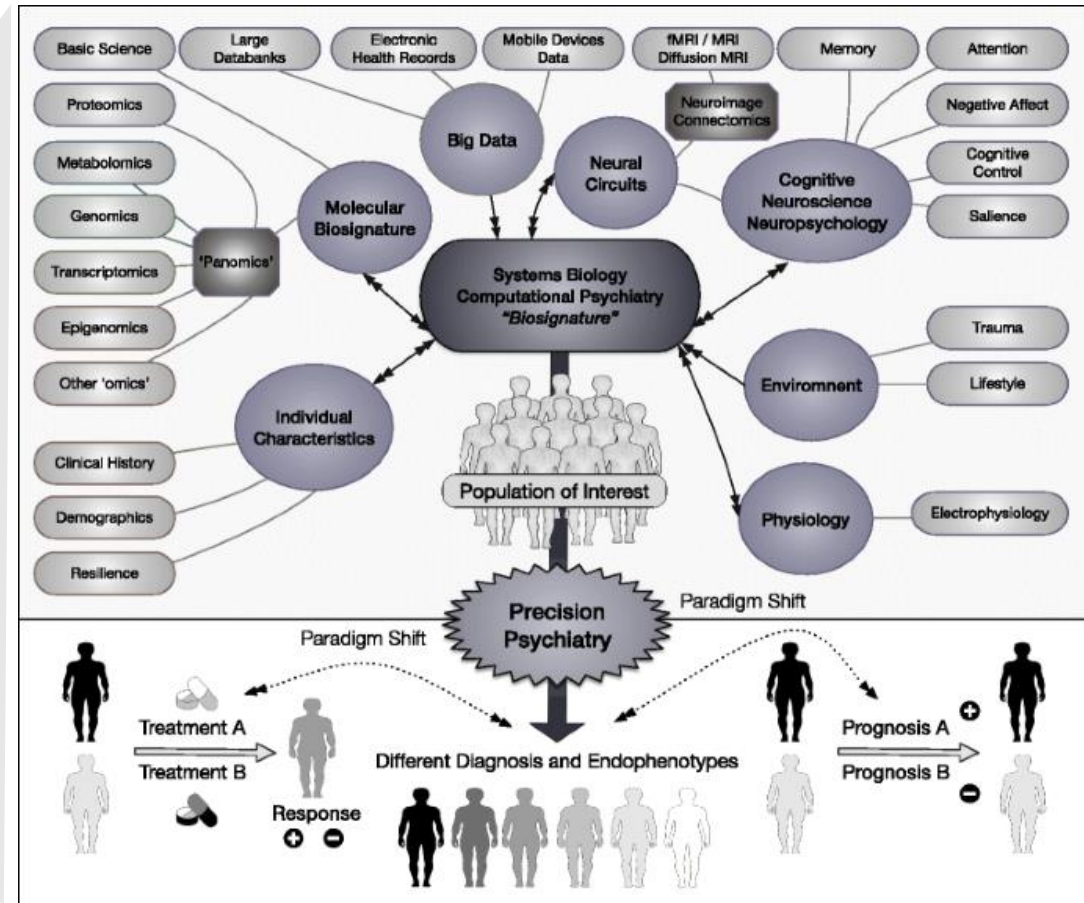
- Associate Professor in Bioinformatics

Educational Background

- MSc. Molecular Biology (2010)
- PhD. Neuroscience (2014)

Research Focus

- Functional Genomics
- Psychiatric Genetics
- **Precision Psychiatry**
- Systems Biology



ABOUT ME

Academic appointment

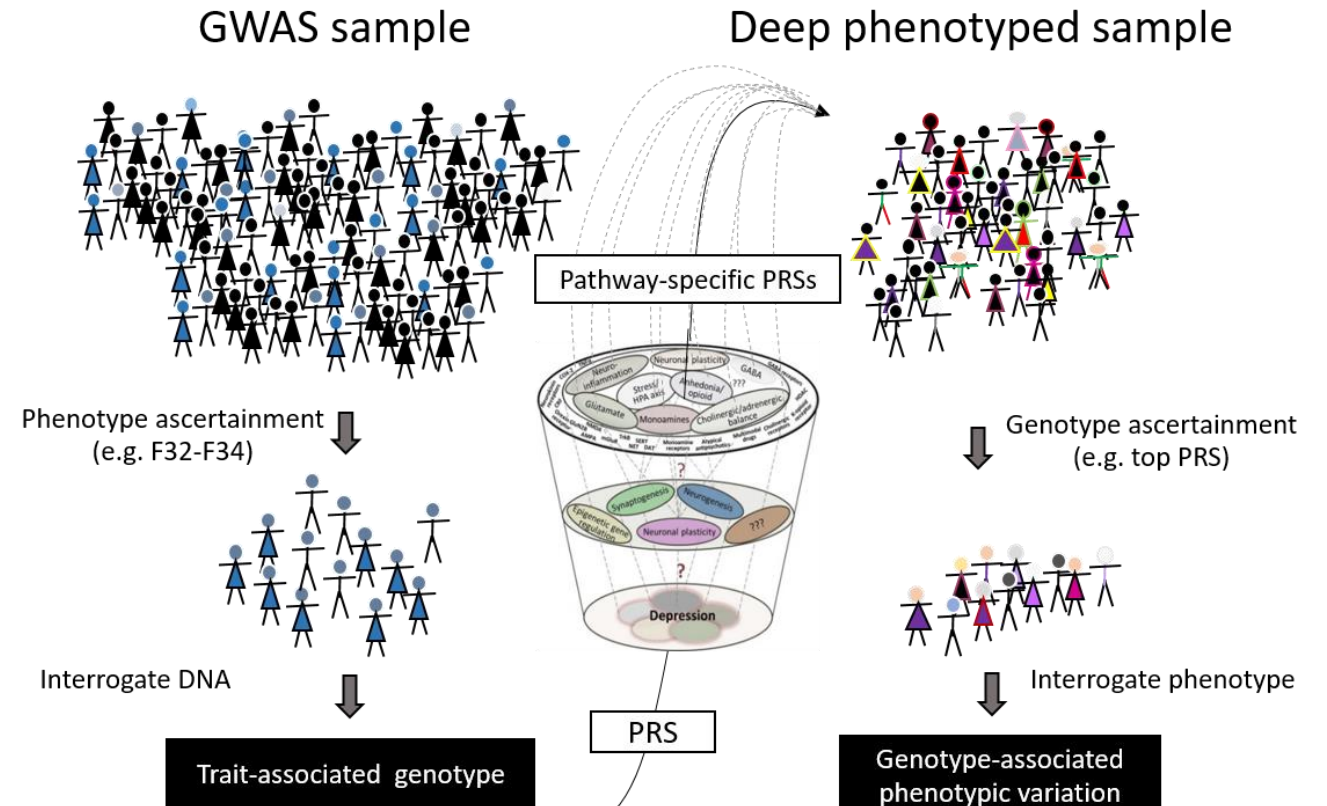
- Associate Professor in Bioinformatics

Educational Background

- MSc. Molecular Biology (2010)
- PhD. Neuroscience (2014)

Research Focus

- Functional Genomics
- Psychiatric Genetics
- **Precision Psychiatry**
- Systems Biology



ABOUT ME

Academic appointment

- Associate Professor in Bioinformatics

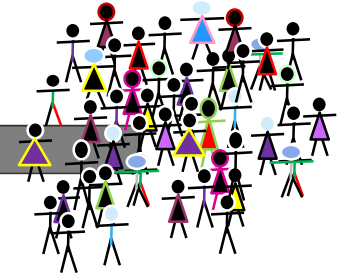
Educational Background

- MSc. Molecular Biology (2010)
- PhD. Neuroscience (2014)

Research Focus

- Functional Genomics
- Psychiatric Genetics
- **Precision Psychiatry**
- Systems Biology

DEEP PHENOTYPED COHORTS



PHENOTYPING INDEX

BEHAVIOR

- General**
- Demographics, Health (incl. smoking)
 - Personal and family history of mental health problems
- Metacognition**
- MAIA-2 (Multidimensional Assessment of Interoceptive Awareness)
 - Superiority illusion
- Personality/Cognition**
- NEO-PI 240 items (NEO Personality Inventory)
 - WAIS-4 (Wechsler Adult Intelligence Scale)
 - CSI (Cognitive Style Indicator)
 - REI (Rational Experimental Inventory)
 - BPS (Boredom Proneness Scale)
 - BIS (Barratt Impulsiveness Scale)
 - CFQ (Cognitive Failures Questionnaire)
 - ESS (Epsworth Sleepiness Scale)
- Emotion (regulation)**
- CTQ (Childhood Trauma Questionnaire)
 - DERS (Difficulties in Emotion Regulation Scale)
 - MDI (Major Depression Inventory)
 - PSS (Perceived Stress Scale)
 - SEC (College Student's Stressful Event)
 - APS (Arousal Predisposition Scale)
 - The Snaith-Hamilton Pleasure Scale
 - CES-D (Center for Epidemiological Studies Depression Scale)
 - AQ (Autism Questionnaire)
 - GAD-7 (General Anxiety Disorder)
- Perception/distraction**
- VFQ-25 (Visual Functioning Questionnaire)
 - OSS (Olfaction Sensitivity Scale)
 - CSS (Chemical Sensitivity Scale)
 - NSS (Noise Sensitivity scale)
 - GOLD MSI (Musical Sophistication)
- Social**
- SNS (Social Network Score)
 - TAS-20 (Toronto Alexithymia Scale)
- Dreaming**
- Dream Recall Questionnaire
- Mindfulness**
- Freiburg Mindfulness Inventory
- Memory**
- MMQ (Multifactorial Memory Questionnaire)

GENOMICS

- Infinium Global Screening SNP array**
- Polygenic risk scores (PRS) and pathway-specific PRS (pPRS)

EPIGENOMICS

- Infinium Methylation EPIC BeadChip**
- genome-wide methylation status → Methylation risk scores (MRS/pMRS)

PROTEOMICS

- 2D gel protein gel coupled with AI**
- proteome-wide profiling → quantitative and qualitative changes

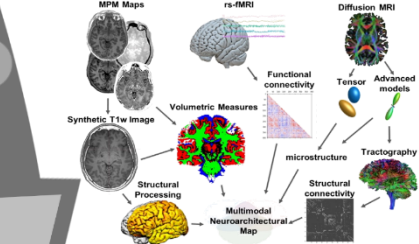
METABOLOMICS / LIPIDOMICS

- LC-MS/MS**
- Quantification of >10,000 proteins, lipids and small molecules

IMMUNOMICS

- Multiplex immunoassays**
- Immune signatures
 - Other targeted molecular biomarkers (e.g. hormones)

IMAGING



INTERMEDIATE PHENOTYPES

DIAGNOSIS

National Danish Registers

ABOUT ME

Bioinformatic support for AU-BIOMED

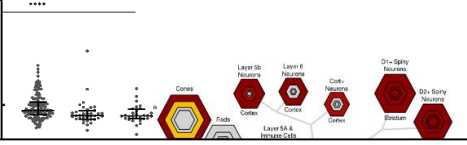
- Served as a bioinformatics go-to person for years
- Broad bioinformatics expertise vs field-specific expert
- Widely involved in research at the department across research fields

BIOINFORMATICS CORE FACILITY

MR. MANAGER

Enhanced Production of Mesencephalic Dopaminergic Neurons from Lineage-Restricted Human Undifferentiated Stem Cells

Muyesier Maimaitili, Muwan Chen, Fabia Febraro, Noémie Mermet-Joret, Johanne Lauritsen, Ekin Ucuncu, Ida Hyllen Klæstrup, Per Qvist, Sadegh Nabavi, Marina Romero-Ramos, Mark Denham
doi: <https://doi.org/10.1101/2021.09.28.462222>

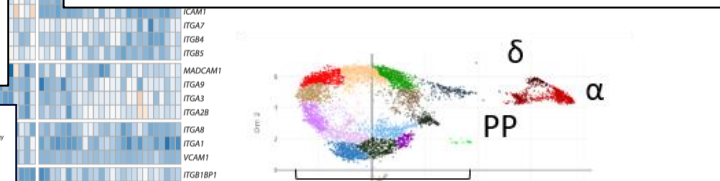


Plasma ACE2 predicts outcome of COVID-19 in hospitalized patients

Tue W. Kragstrup, Helene Søgaard Singh, Ida Grundberg, Ane Langkilde-Lauesen Nielsen, Felice Rivellese, Arnav Mehta, Marcia B. Goldberg, Michael R. Filbin, Per Qvist, Bo Martin Bibby
Published: June 4, 2021 • <https://doi.org/10.1371/journal.pone.0252799>

Single-cell transcriptomics reveals altered pancreatic cell composition and insulin deficits in *SorCS2*^{-/-} mice

Oleksandra Kalnytska¹, Vanessa Schmidt¹, Per Qvist² and Thomas Wilnow^{1,2}
¹Max Delbrück Center for Molecular Medicine, Berlin, Germany
²Department of Biomedicine, Aarhus University, Aarhus, Denmark



SorCS2 binds progranulin and regulates motor axon outgrowth

Thomasen PB^{1,2}, Salašová A^{1,2}, Login H^{1,2}, Beel S³, Tranberg-Jensen J^{1,2}, Qvist P², Ovesen PL^{1,2}, Nolte S^{1,2}, Nejsum LN⁴, Chao MV⁵, Dasen J⁵, Van Damme P³, Kjaer-Soerensen K⁶, Oxvig C⁶, Nykjaer A^{1,2,7}.

Sortilin and SorCS2 are involved in neurotrophin signalling in Schwann cells but are not required for motor nerve regeneration.

Sort1 and *Sorcs2* in peripheral nerve regeneration.
Maj Ulrichsen^{1,7}, Nádia P. Gonçalves^{1,7}, Simin Mohseni², Ebbe Toftgaard Poulsen³, Per Qvist^{4,5}, Simone Hjaerensen⁸, Thomas Lisle^{1,7}, Simon Molgaard⁷, Jan J. Enghild³, Anders Nykjaer^{1,7}, Simon Glerup⁷, Christian B. Vægter^{1,7*}



Size-Selective Phagocytic Clearance of Fibrillar α -Synuclein through Conformational Activation of Complement Receptor 4

Kristian Juul-Madsen, Per Qvist, Kirstine L. Bendtsen, Annette E. Langkilde, Bente Vestergaard, Kenneth A. Howard, Martxel Dehesa-Etxebest, Søren R. Paludan, Gregers Rom Andersen, Poul Henning Jensen, Daniel E. Otzen, Marina Romero-Ramos and Thomas Vorup-Jensen

ARTICLE

SORLA is required for insulin-induced expansion of the adipocyte precursor pool in visceral fat

Vanessa Schmidt¹, Carla Horváth², Hua Dong², Matthias Blüher³, Per Qvist^{4,5}, Christian Wolfrum², and Thomas E. Willnow^{1,4}

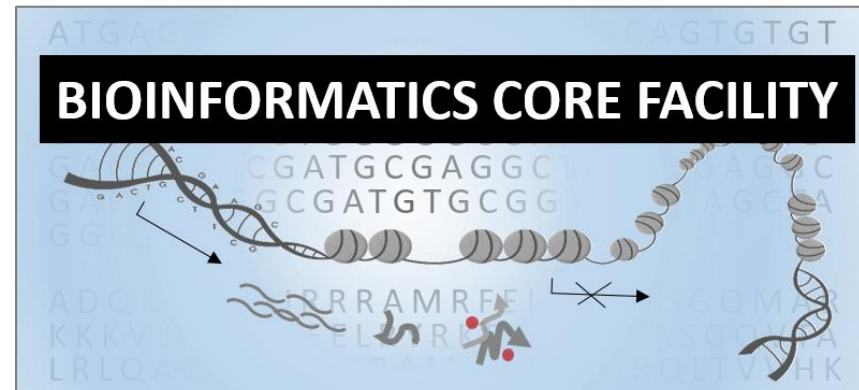
Genomics of JAK-STAT signaling in venous thromboembolism

Stine R Haysen¹, Per Qvist^{1,3,4}, Ane Langkilde-Lauesen Nielsen¹, Tue W Kragstrup^{1,2}
¹Department of Biomedicine, Aarhus University, Denmark.
²Department of Rheumatology, Aarhus University Hospital, Denmark.
³Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus 8000, Denmark.
⁴Centre for Genomics and Personalized Medicine, CGPM, Aarhus University, Aarhus 8000, Denmark.

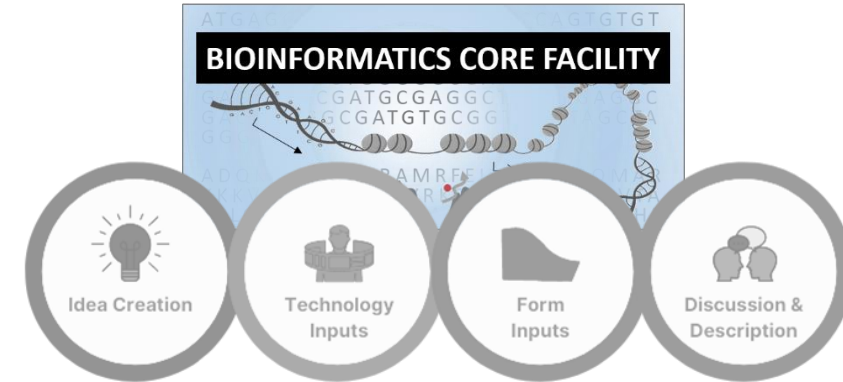


THE BIOINFORMATICS CORE FACILITY

- Established **autumn 2022**
- **Aim:** to provide basic bioinformatics training, services as well as full-scale scientific collaborations for all scientists at Aarhus University as well as external academic and industrial partners.
 - One priority is to provide biologist-friendly solutions in order to implement the use of OMICs and other large-scale data at the Department of Biomedicine .
- **Core activities:**
 - Data analysis
 - Consultation
 - Training
 - Research collaboration
 - Knowledge dissemination
 - Fundraising

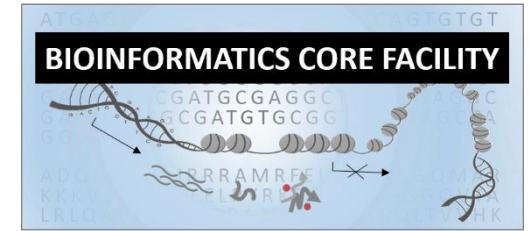


SERVICES



- **Consultation**
 - Experimental design, bioinformatics methods and computational tools
- **Data analysis**
 - NGS analyses (e.g. RNAseq (bulk, spatial and single cell), ChIPseq, DNaseq (WGS and WES))
 - DNA analyses (e.g. Functional annotations and geneset enrichment/association analyses)
 - RNA analyses (e.g. Expression profiling, target prediction, splicing)
 - Sequence analysis (e.g. variant detection, SNP annotation, Transcription factor binding sites)
- **Data formatting/structuring**
 - Preparation of data for public repositories (Open Science)
- **Data mining**
 - OMICs data
- **Knowledge dissemination**
 - Assistance in preparation of manuscripts, presentations and grant proposals
- **Training**
 - Workshops on Bioinformatics topics and data access/analysis
 - Open Office
 - Teaching
 - Supervision

ACTIVITIES SINCE LAUNCH



USERS

- 26 – roughly 2 visits/enquiries per week since launch

Data analysis

- Bulk RNAseq (Human/virus hybrid genome)
- CRISPR off targets indels in WES data
- Gene set association / enrichment

Consulting

- Bulk RNAseq (difficult tissues – weird species)
- Single and spatial transcriptomics
- Proteomics
- Databases and tools
- Biologist-friendly data analysis pipeline/platforms
- High Performance Computing

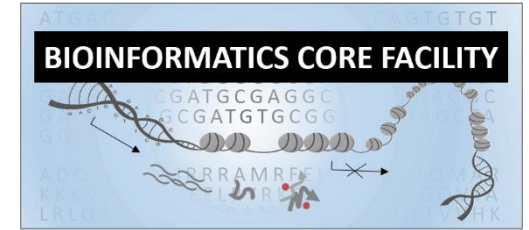
Training

- Workshop
 - Bulk RNAseq analysis using Galaxy
- Supervision/mentoring
 - 2 Research Year students
 - 2 PhD students

Data mining

- Extraction and re-analysis of deposited data
 - Olink COVID19 data
 - GWAS
 - miRNA targets
 - sc data

PRICING MODEL



- Come by for a chat
...that is for **free**
- Think Bioinformatics fee into grant proposals
- We are not competing with other AU/non-AU service providers - we complement

TRAINING

- Seminars
- Workshops
- Q&A
- Open office
- Supervision
- etc

Free

PRODUCTS

- Data formatting
- Data processing
- Mapping
- Annotation
- DE analyses
- Visualization
- etc

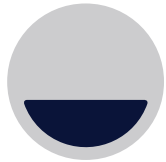
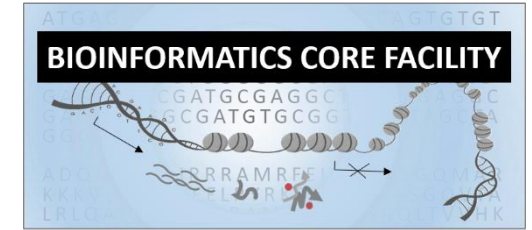
DKK 450/h

PROJECTS

- Customized requests
- Biological interpretation
- Field-specific platforms
- Downstream targeted approaches
- etc

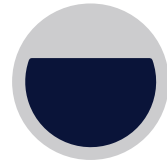
- Collaboration
- Estimated cost
- Outsource

RESOURCE ALLOCATION



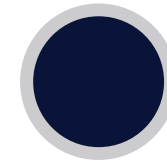
Hours - Days

- GWAS lookup
- Data access
- Data formatting
- Simple visualizations



Days - Weeks

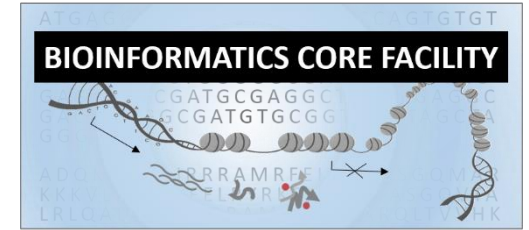
- Mapping
- Annotation
- PCA
- Clustering
- Gene set association
- Gene set enrichment
- Variant detection



Weeks - Months

- Customized tasks
- Biological interpretation
- Field-specific platforms
- Project-specific targeted approaches
- Data integration

PRICING EXAMPLES



#1

Bulk RNAseq dataset analysis

PRODUCT

- Project information meeting
- Transfer of raw data
- Quality assessment/trimming etc
- Sequence alignment
- Feature quantification
- DEG analyses
- Basic geneset enrichment analysis
- Basic visualization
- Project report / report meeting

~15 hours for a project with <20 libraries

~DKK 6,750

#2

Data mining collaboration

PROJECT

- Project information meeting
- Contribution to grant proposal
- Literature and data review
- Extraction of OMICS data from multiple sources
- Re-analysis of OMICS data
- Collective reporting of OMICS data incl. customized visualizations
- Contribution to interpretation/manuscript

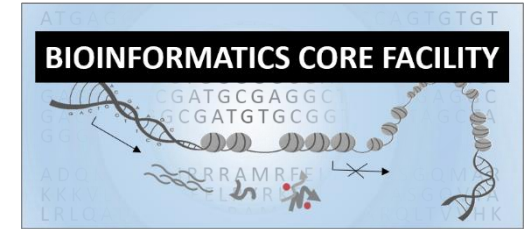
~60-100 hours distributed over multiple weeks

~DKK 15,000

ACTIVITIES SINCE LAUNCH

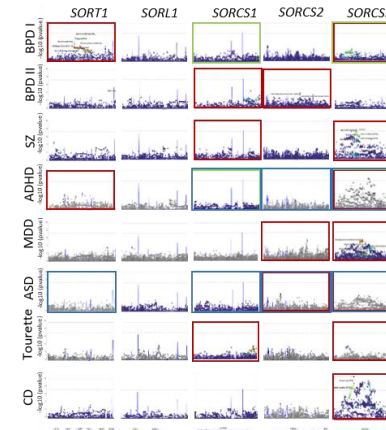
USERS

- 26 – roughly 2 visits/enquiries per week since launch



Grant proposals

- Co-writer on 4 grant proposals
 - 2 grants (~DKK 60K)
 - incl generation of supporting data



BIOINFORMATICS TRAINING

-
- HPC Type 1 (UCloud)
 - HPC Type 2 (Computerome 2)
 - HPC Type 2 (GenomeDK)
 - HPC Type 2 (Sophia)
 - HPC Type 3 (Hippo)
 - LUMI

UNIX / Bash environment

R / Python programming

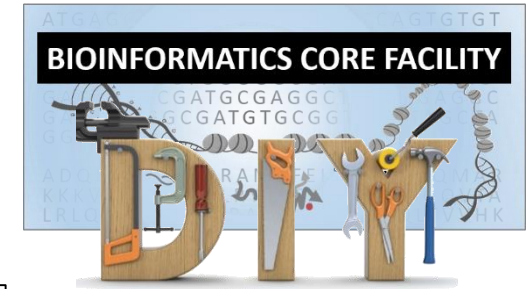
Graphical user interface



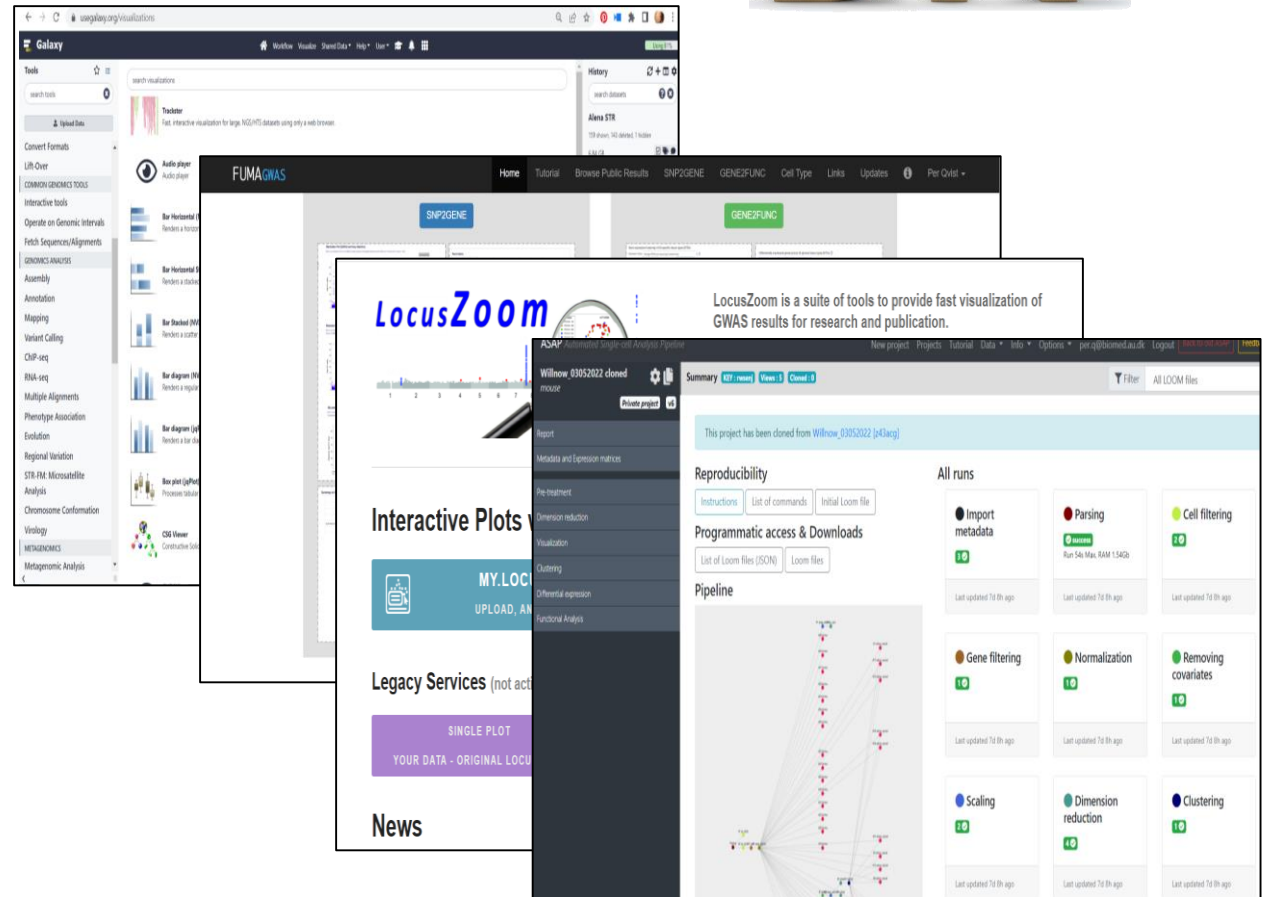
Drag and drop

Graphical user interface

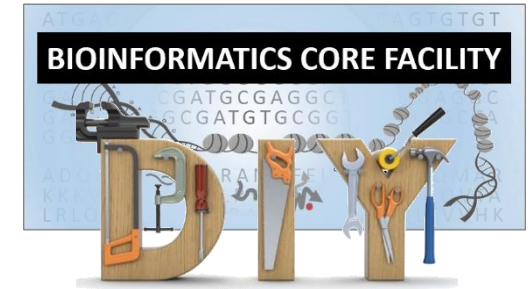
BIOINFORMATICS TRAINING



- We provide training in the use of pipelines and tools....not coding
- Many biologist-friendly tools available
 - From GWAS to single cell OMICs
 - Most are for free...
 - Require very limited introduction
 - No/only minor programming skills needed
- HPCs
 - Ucloud
 - GenomeDK



BIOINFORMATICS TRAINING



- Many biologist-friendly tools available
 - From GWAS to single cell OMICs
 - Most are for free...
 - Require very limited introduction
 - No/only minor programming skills needed

Any lab can perform basic bioinformatics analyses using these tools
– the Bioinformatics Core Facility can get you started

AARHUS UNIVERSITY
Department of Biomedicine

Bioinformatics Core Facility

Services Pricing Activities Publications and acknowledgements Contact About us

Department of Biomedicine > Bioinformatics Core Facility > Activities

Bioinformatics Core Facility

- > Services
- > Pricing
- > **Activities**
- > Publications and acknowledgements
- > Contact
- > About us

Activities

Upcoming and past events

- > **Workshop - Analysis of bulk RNA-sequencing data using Galaxy and UCloud**
A physical workshop with the subject of analyzing bulk RNA-sequencing data. The participants will be presented to the computational workflow of analyzing bulk RNA-sequencing data including file formats, file information, data processing, result generation, and interpretation. The workshop will include a live analysis and description for participants to do their work themselves - without the need of coding experience. Find slides from the workshop [here](#).
19th January 2023
- > **PhD Course - Single cell and spatial OMICs (Basic Course)**
Core staff will contribute the upcoming course in single cell and spatial OMICs offered by AU HEALTH PhD school. Read more about the course [here](#).
May 2023
- > **Seminar - Introduction to the Bioinformatics Core Facility and its use in single cell analyses**
OMICs focus group meeting 3. Find slides from the seminar [here](#).
October 2022

Revised 24.01.2023 - Per Qvist

DEPARTMENT OF BIOMEDICINE

Aarhus University
The Skou Building
Høegh-Guldbergs Gade 10
DK 8000 Aarhus C
Denmark

ABOUT US

Profile
Staff
Contact
Vacant positions
Join the department

AU DEGREE PROGRAMMES

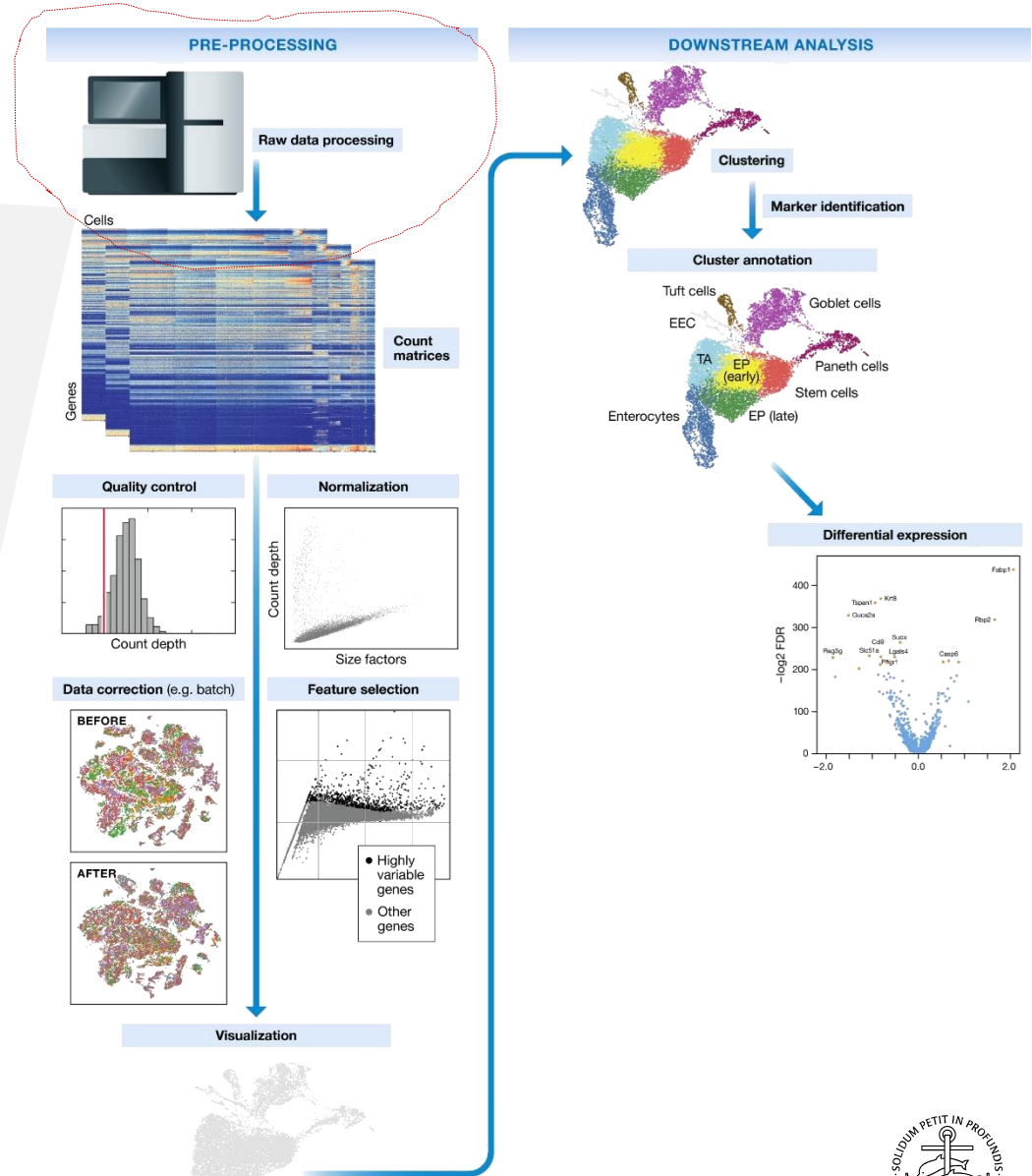
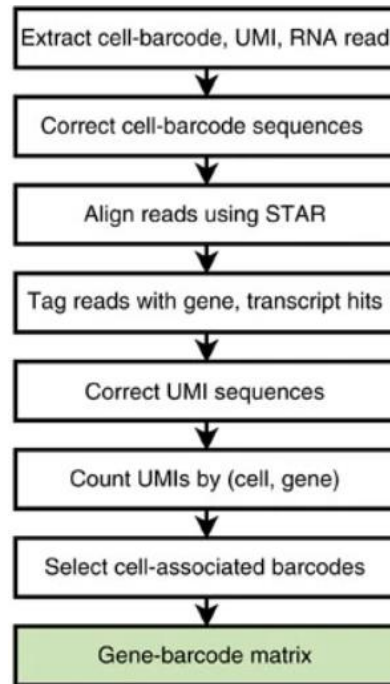
Bachelor
Master
Engineering
PhD

SHORTCUTS

Core-facilities
Nobel Prize winner from
Biomedicine

EXAMPLE: CODING-FREE SC PIPELINE

- **Chromium Single Cell Software Suite**
 - Cell Ranger
- **Galaxy**
 - An open, web-based platform for accessible, reproducible, and transparent computational biological research
 - RNA STAR
 - Hisat
 - Limma



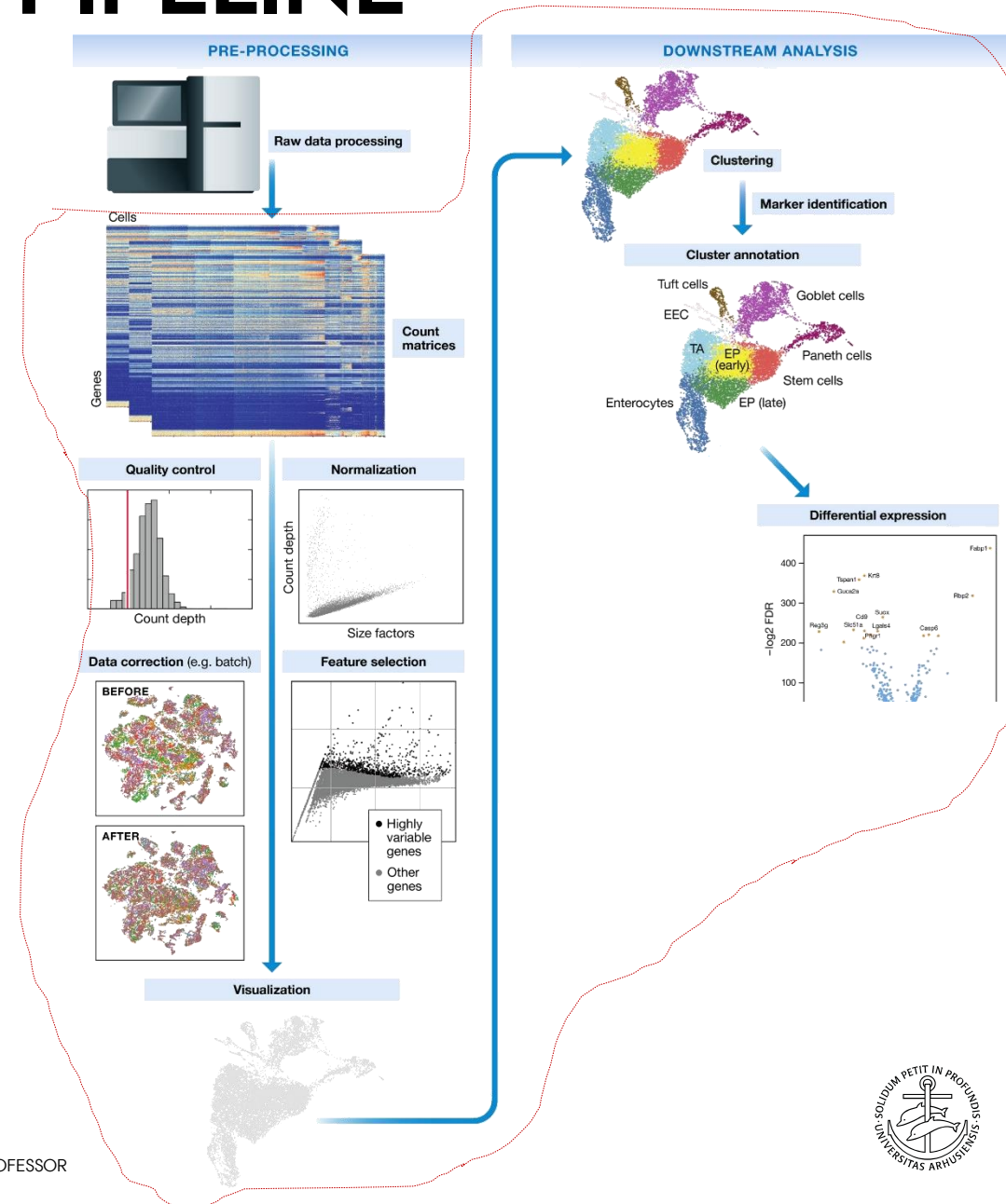
Example: Coding-free sc pipeline

The screenshot displays the Galaxy Single Cell Omics interface. On the left, a sidebar lists various tools under categories like Annotation, Mapping, RNA-Seq, and Single-cell. The main content area features a 'Welcome to the world of Single Cell Omics' message and a 'Content' section with links to 'Get started', 'Training and Workshops', and 'Material'. A 'Get started' section provides instructions for new users. The 'Material' section lists lessons such as 'Introduction to Transcriptomics', 'Plates, Batches, and Barcodes', and 'Understanding Barcodes'. On the right, a 'History' panel shows an empty history with a message: 'This history is empty. You can load your own data or get data from an external source.' The central tool configuration window for 'RNA STARSolo mapping, demultiplexing and gene quantification for single cell RNA-seq (Galaxy Version 2.7.8a)' is open. It includes sections for 'Custom or built-in reference genome', 'Reference genome with or without an annotation', 'Select reference genome' (set to 'A. mellifera genome (apiMel3, Baylor HGSC Amel_3.0)'), 'Gene model (gff3,gtf) file for splice junctions' (with a message 'No gff3 or gtf dataset available'), 'Length of the genomic sequence around annotated junctions' (set to 100), 'Type of single-cell RNA-seq' (set to 'Drop-seq or 10X Chromium'), and 'Input Type' (set to 'Separate barcode and cDNA reads').

• <https://singlecell.usegalaxy.eu/>

EXAMPLE: CODING-FREE SC PIPELINE

- Chromium Single Cell Software Suite
 - Loupe
- ASAP
 - A collaborative portal to analyze single-cell transcriptomics data
 - [Seurat](#)
 - [Scanpy](#)
 - Other toolkits
 - From Matrix (H5/loom) to functional analysis of sc data



EXAMPLE: CODING-FREE SC PIPELINE

- Report
- Metadata and Expression matrices
- Pre-treatment
 - > Parsing
 - > Cell filtering
 - > Gene filtering
 - > Normalization
 - > Removing covariates
 - > Scaling
- Dimension reduction
- Visualization
- Clustering
- Differential expression
- Functional Analysis

New Clustering

Method: **Hierarchical Clustering** (selected from dropdown: K-Means, Hierarchical Clustering, K-Means, SC3, Scanpy [beta], Seurat)

Input matrix: **Select** (Please select one or several)

Parameters: Number of clusters: 2:10, Algorithm: Hartigan-Wong

Cluster Cancel

Summary KEY: x43acj Views: 4 Cloned: 1

Reproducibility: Instructions, List of commands, Initial Loom file

Programmatic access & Downloads: List of Loom files (JSON), Loom files

Pipeline:

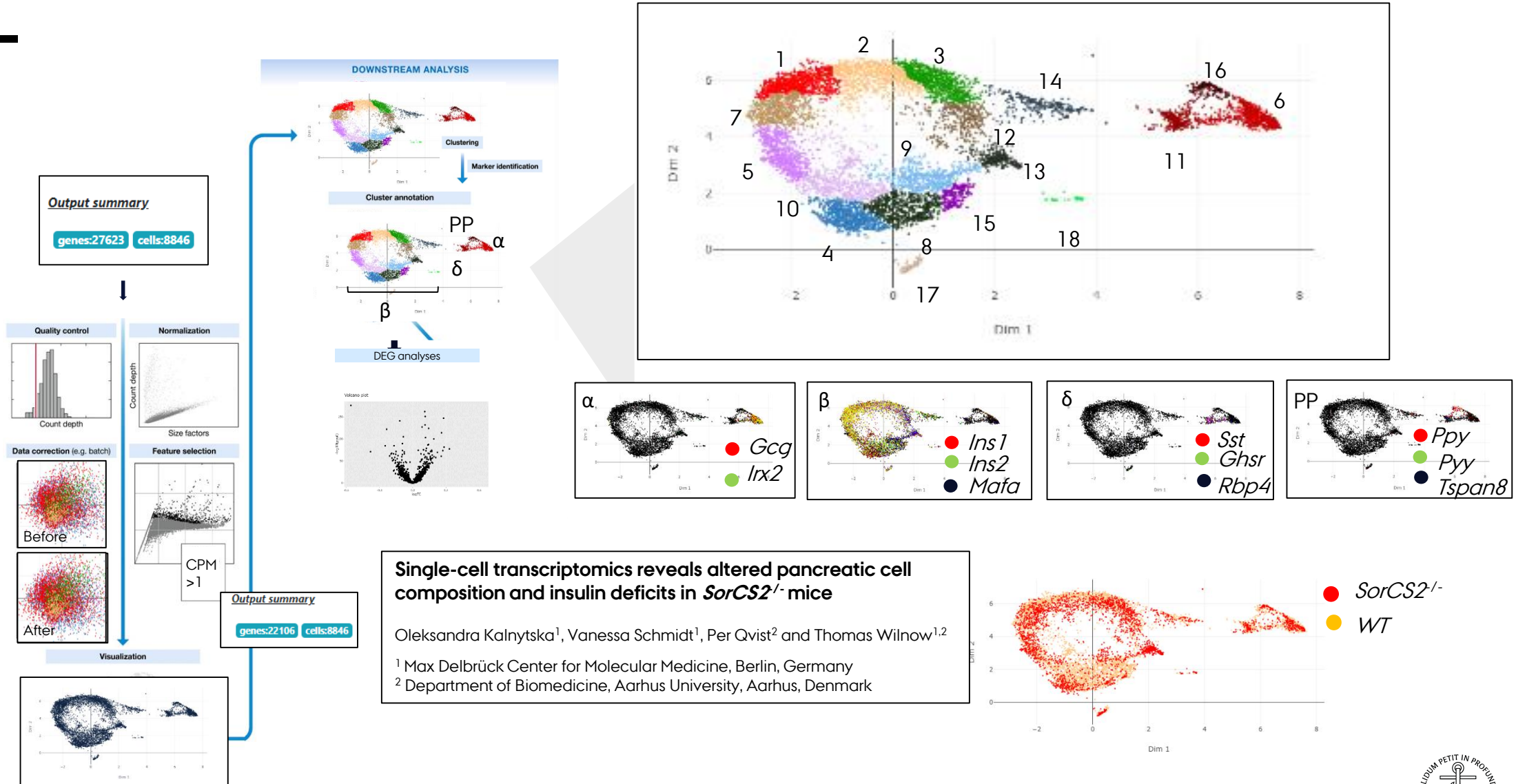
All runs:

- Import metadata (Last updated 173d 13h ago)
- Normalization (Last updated 174d 9h ago)
- Clustering (Last updated 173d 7h ago)

Dataset:

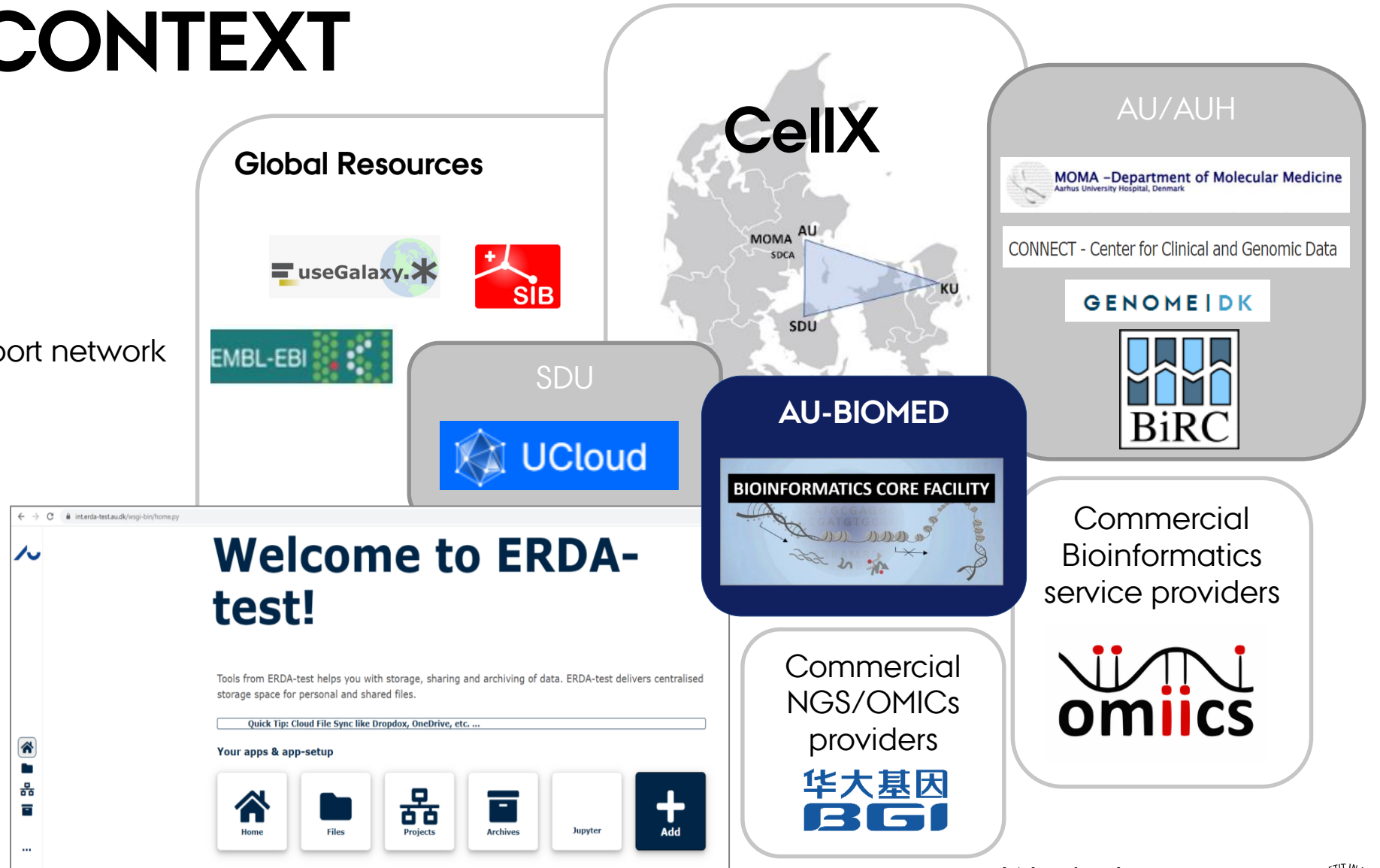
Easy to share and deposit project as H5/loom file
- all scripts saved - highly reproducible
= great for interactions with the core facility

EXAMPLE: CODING-FREE SC PIPELINE



THE CORE CONTEXT

- We are part of a larger support network



We don't compete
– we complement!

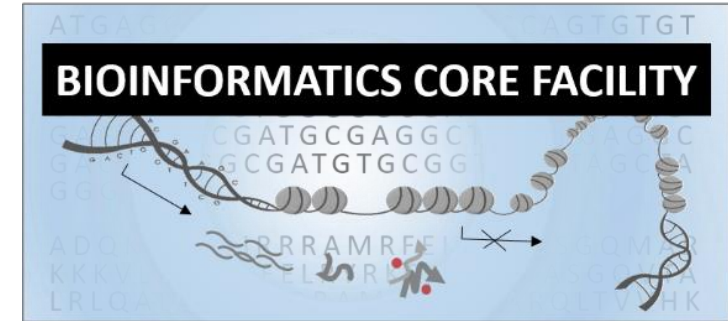
CONTACT

- **Address**

- The Skou Building 1116, room 456
Høegh-Guldbergsgade 10
8000 Aarhus C
Denmark



<https://biomed.au.dk/bioinformatics-core-facility>



- **Open Office**

- You are free to come by for an initial discussion of your projects always and we attempt to keep the core open every Monday 10.30-12.00.

However, we encourage you to call/write first and make an appointment.

- **Core Facility Manager** (and currently only staff :/)

- Per.q@biomed.au.dk
- 50192006





AARHUS
UNIVERSITY