

THE de.NBI SERVICE CENTER HD-HuB

Heidelberg Center for Human Bioinformatics

The HD-HuB Service Center integrates the expertise of four major institutions: The German Cancer Research Center (DKFZ), European Molecular Biology Laboratory (EMBL), Heidelberg University, and Charité/Berlin Institute of Health (BIH).

In addition, the de.NBI-epi partner project consisting of partners from DKFZ, Saarland University, and MPI for Informatics has been associated with HD-HuB since November 2016. The goal of the HD-HuB Service Center is to provide access to bioinformatics services, compute resources and training for the management, analysis and interpretation of sequencing and image-based data to facilitate data-driven discoveries in key fields of human biology ranging from basic to clinical research. HD-HuB is a node of the de.NBI cloud within the de.NBI network.

TOPICS OF THE SERVICE CENTER HD-HuB:

- Human genomics and epigenomics
- Human microbiomics
- Systematic phenotyping of human cells

SELECTION OF TRAINING COURSES CARRIED OUT BY HD-HuB:

- Microscopy Image Analysis
- Data Interpretation of RNASeq and Bisulfite Sequencing Data in Cancer Research
- Data Interpretation of Whole-Genome and Exome Sequencing Data in Cancer Research
- Software Carpentry

CURRENT PROJECTS:

- **HD-HuB 1 at DKFZ Heidelberg and EMBL Heidelberg:** Human genetics and genomics
- **HD-HuB 2 at EMBL Heidelberg:** Human microbiomics
- **HD-HuB 3 at Heidelberg University and DKFZ Heidelberg:** Systematic phenotyping of human cells
- **de.NBI-epi 1 at Heidelberg University and de.NBI-epi 2 at Saarland University:** Analysis tools for epigenomic data



Peer Bork

*Coordinator
EMBL Heidelberg*
bork@embl.de



Roland Eils

*Deputy coordinator
Charité/BIH*
roland.eils@charite.de

↓ **Members of the Service Center HD-HuB.** From left to right: Y. Assenov, N. Habermann, S. Yakhnin, U. Trojahn, G. Zeller, M. Smith, R. Russell, P. Bork, C. Lawerenz, I. Buchhalter, B. Rauscher, K. Rohr, P. Kensche, J. Puyol, M. Sharan, N. Ishaque, J. Korbel, T. Wollmann, M. Lang, P. Lutsik, K. Nordstroem, M. Gunkel (October 2018)



THE de.NBI PROJECT HD-HuB 1

Human Genetics and Genomics

The HD-HuB 1 project at Heidelberg University, the German Cancer Research Center (DKFZ), European Molecular Biology Laboratory (EMBL) and Charité / Berlin Institute of Health (BIH) offers tools, services and training courses in the area of human genetics and genomics.

↓ The HD-HuB 1 team at Heidelberg University, DKFZ, EMBL, and BIH.



Rob Russell

robert.russell@bioquant.uni-heidelberg.de

Project leader



Ivo Buchhalter

i.buchhalter@dkfz.de



Wolfgang Huber

whuber@embl.de



Jan Korbel

korbel@embl.de



Roland Eils

roland.eils@charite.de



Naveed Ishaque

naveed.ishaque@charite.de

COLLECTION OF SELECTED TOOLS AND SERVICES OFFERED BY HD-HuB 1:

- **Butler** - A computational framework for orchestrating genomics analysis workflows on the cloud.
- **rhdf5/Rhdf5lib** - The package provides an interface between HDF5 and R. In particular HDF5 is being used to address the rapid growth in the size of single-cell datasets.
- **biomaRt** - Provides an R interface to BioMart services.
- **DELLY** - Workflow for the discovery of germline and somatic structural variants
- **DESeq2 & DEXSeq** - Suite of R/Bioconductor packages for differential expression analysis of high-throughput sequencing assays
- **PanCancer alignment workflow** - Based on the Pan-Cancer project for the alignment of NGS data to reference genomes
- **ACSEq** - Detection of somatic copy number variations from matching tumor/control whole genome sequencing data pairs
- **SNV calling pipeline** - Construction of high confidence somatic SNVs based on set from tumor and control samples
- **Indel calling pipeline** - Detection of high confidence indels (1-20 bp) from tumor and control samples
- **Roddy** - Framework for large scale NGS processing
- **OTP** - Automation platform for management and analysis of NGS data
- **Mechismo** - Rapid predictions of how genetic variants impact on biomolecular function employing pre-computed collections of protein-protein, protein-DNA/RNA, protein-chemical interactions.

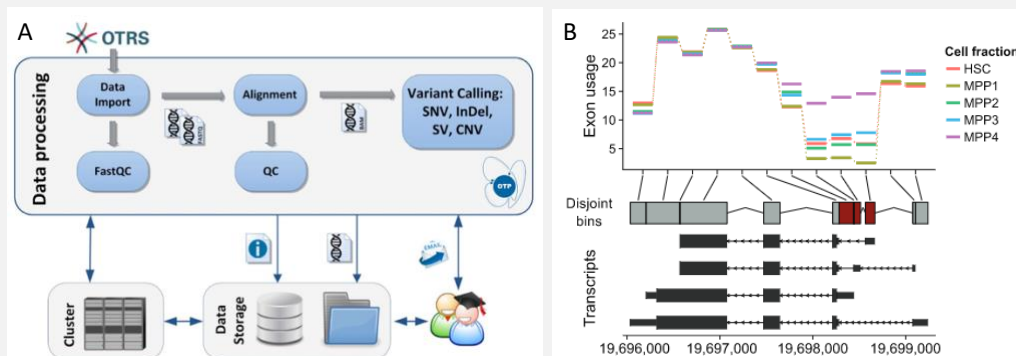


Figure 1. A - Overview of data processing in OTP. For details please refer to J Biotechnol. 2017 Nov 10; 261:53-62. doi: 10.1016/j.jbiotec.2017.08.006. **B** - Visualisation of differential exon usage in the Apoe gene as quantified by the DEXSeq method (y-axis of the top panel). For details see Nat Methods. 2015 Feb;12(2):115-21. doi: 10.1038/nmeth.3252.

THE de.NBI PROJECT HD-HuB 2

Human Microbiomics

The **HD-HuB 2** project at the European Molecular Biology Laboratory (EMBL) covers the whole range of analysis tasks typically applied in the field of microbiomics.

▼ The HD-HuB 2 team at EMBL.

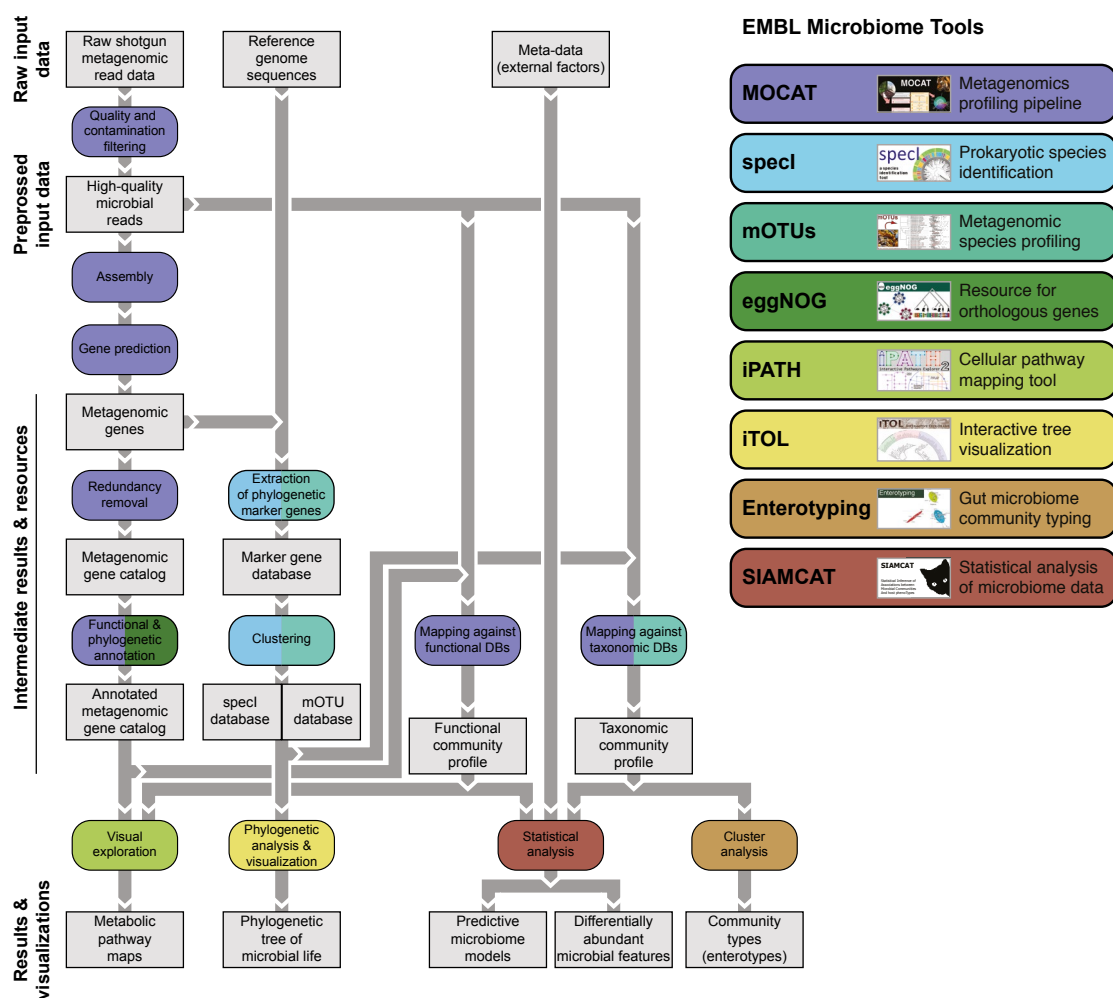


Peer Bork
Project leader
bork@embl.de



Georg Zeller
zeller@embl.de

COLLECTION OF TOOLS AND SERVICES OFFERED BY HD-HuB 2:



THE de.NBI PROJECT HD-HuB 3

Systematic Phenotyping of Human Cells

The HD-HuB 3 project at DKFZ and Heidelberg University offers tools, services and training for the analysis and interpretation of microscopy image data for automated phenotyping of human cells.

↓ The HD-HuB 3 team at DKFZ and Heidelberg University.



Michael Boutros
m.boutros@dkfz.de
Project leader



Oliver Stegle
o.stegle@dkfz.de



Holger Erfle
holger.erfle@bioquant.uni-heidelberg.de

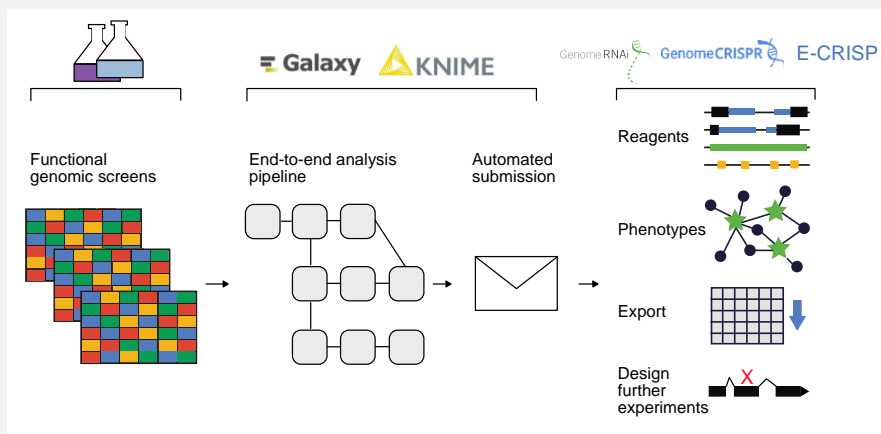


Karl Rohr
karl.rohr@bioquant.uni-heidelberg.de

COLLECTION OF TOOLS AND SERVICES OFFERED BY HD-HuB 3:

- **KNIME-based cellular phenotyping of microscope image data** – a data mining platform including manifold libraries for image processing (based on Fiji/ImageJ) and data exploration. The platform provides access to workflows and pipelines for (large-scale) automated phenotype analysis of microscope image data.
- **Galaxy Image Analysis** – Workflows and tools for the analysis of microscopy image data and cellular phenotyping using the Galaxy platform.
- **GenomeRNAi** – Provides RNA interference (RNAi) phenotype data extracted by manual curation from the literature, or submitted by data producers directly. In addition to the RNAi screening data of human and Drosophila cells, it contains RNAi reagent information, along with an assessment as to their efficiency and specificity.
- **GenomeCRISPR** - Database for high-throughput screening experiments performed by using the CRISPR/Cas9 system. A dynamic web interface guides users through the process of finding information about published CRISPR screens. The database contains detailed data about observed hits and phenotypes. Moreover, it provides knowledge about performance of individual single guide RNAs (sgRNAs) used under various experimental conditions.
- **E-CRISP** - E-CRISP is a computational tool to design and evaluate guide RNAs for use with CRISPR/Cas9. The design process incorporates different parameters of how CRISPR constructs can be used in experimental applications, such as knock-out and tagging experiments.

↓ Processing workflows of human imaging data.



THE de.NBI PROJECT de.NBI-epi 1

Analysis of DNA Methylation

The de.NBI-epi 1 project at the German Cancer Research Center (DKFZ) focuses on methods to analyze data from whole-genome bisulfite sequencing (WGBS) and array-based hybridization to Illumina Infinium methylation arrays.

↓ *The de.NBI-epi 1 team at DKFZ.*



Benedikt Brors

b.brors@dkfz.de

Project leader

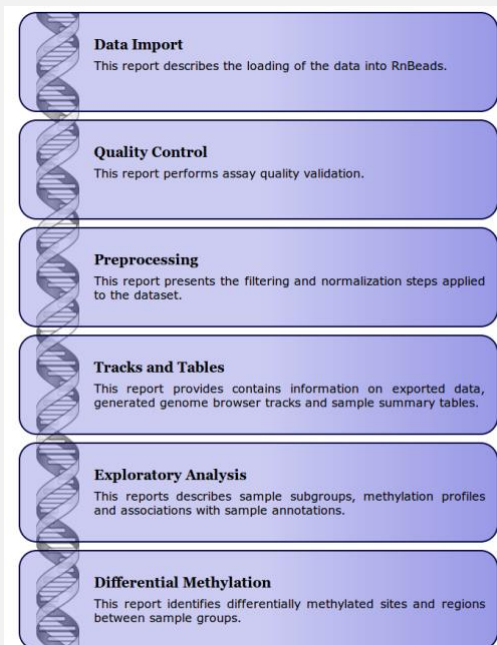


Yassen Assenov

y.assenov@dkfz.de

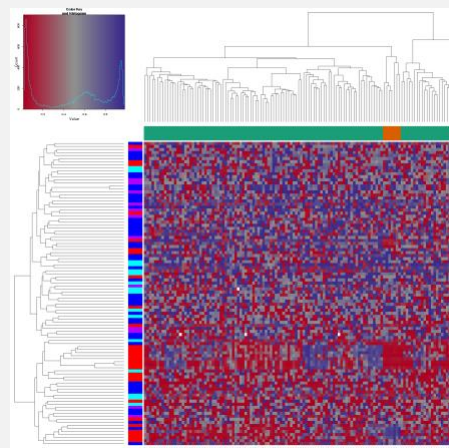
COLLECTION OF TOOLS AND SERVICES OFFERED BY de.NBI-epi 1:

- **RnBeads** – A R/Bioconductor library for analysing data from Illumina Infinium methylation arrays (Infinium 450k and EPIC arrays). It includes methods for data import, different normalization methods, quality assessment, exploratory analysis and calling of differentially methylated nucleotides.



← **RnBeads analysis workflow** (Assenov *et al.*, 2014).

↓ **RnBeads DNA methylation heatmap.**



THE de.NBI PROJECT de.NBI-epi 2

Analysis of Epigenomic Data

The de.NBI-epi 2 project at Max Planck Institute for Informatics and Saarland University provides a series of high-end tools for analysis of epigenomic data with focus on DNA methylation data.

↓ The de.NBI-epi 2 team at Max Planck Institute for Informatics and Saarland University:



Jörn Walter

j.walter@mx.uni-saarland.de

Project leader



Thomas Lengauer

lengauer@mpi-inf.mpg.de



Michael Scherer

mscherer@mpi-inf.mpg.de

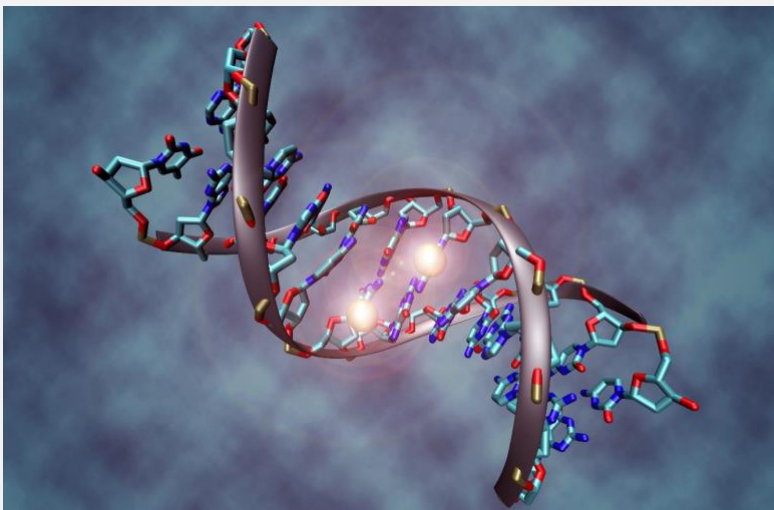


Karl Nordström

karl.nordstroem@uni-saarland.de

COLLECTION OF TOOLS AND SERVICES OFFERED BY de.NBI-epi 2

- **BioQ Analyzer HiMod** – Processing and primary analysis of data obtained in standard targeted bisulfite sequencing experiments, including Tet-assisted (TAB-Seq) and chemical modification-based bisulfite sequencing (Ox-BS) approaches.
- **MeDeCom** – A reference free method for decomposition of multi-cellular, heterogeneous DNA methylomes into quantified latent components approximating distinct cell population.
- **EpiExplorer** – Integration of multiple epigenetic and genetic annotations and making them explorable via an interactive interface.
- **RnBeads** – R package for comprehensive analysis of DNA methylation data obtained by Infinium microarrays and bisulfite sequencing protocols. RnBeads implements a series of QC steps and comprehensively outputs a number of analyses levels as annotated and readable hypertext report. Rnbeads is the standard preparation tool for DNA-methylation data obtained by the DEEP (German Epigenome Programme).
- The **DeepBlue Epigenomic Data Server** provides a data access hub for large collections of epigenomic data. It organizes the data using controlled vocabularies and ontologies. The data is stored in our server, where the users can access the data programmatically or by our web interface.



← DNA segment with highlighted methylation sites