de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

dkfz.
GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

HD-HuB

Heidelberg Center for Human Bioinformatics
Systematic Phenotyping of Human Cells

WP9 Systematic Phenotyping
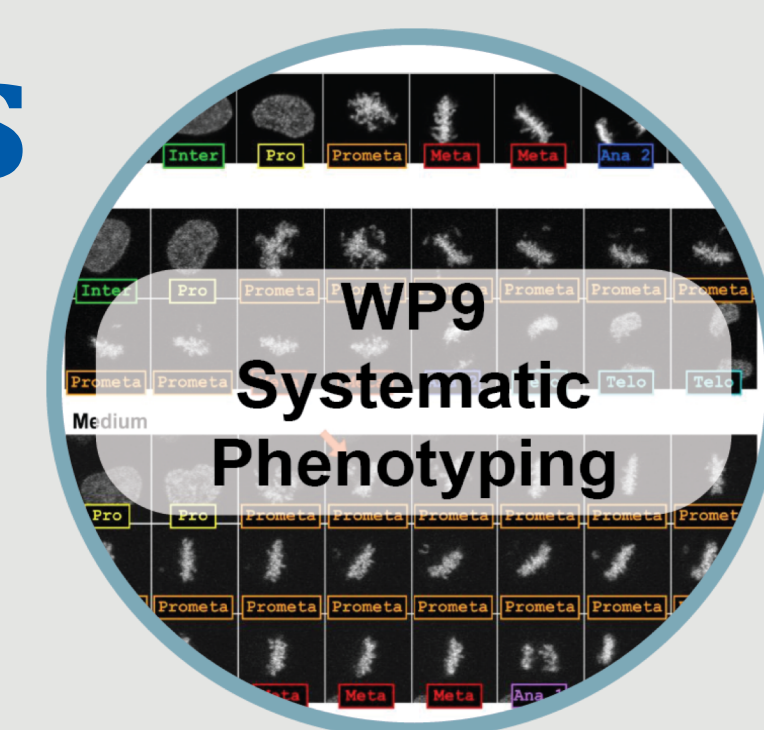
Manuel Gunkel[1], Thomas Wollmann[2], Erica Valentini[3], Benedikt Rauscher[3], Ulrike Hardeland[3], Oliver Stegle[4], Holger Erfle[1], Karl Rohr[2], Michael Boutros[3]

[1]Advanced Biological Screening Facility, BioQuant, Uni Heidelberg; [2]Biomedical Computer Vision Group, BioQuant, Uni Heidelberg; [3]Signaling and Functional Genomics, DKFZ, Heidelberg; [4]Computational Genomics and Systems Genetics, DKFZ, Heidelberg

Fkz 031A537A/C

Short description of the project

Integration and provision of pipelines for automated analysis and visualization of large-scale high-content phenotype screening data of human cells

Our aim is to set up, run, and maintain a web-based workflow unit for high-content screening and image processing in close collaboration with other partners in HD-HuB. Users will be supported by a web-based software repository providing workflows for high-content data analysis, or by customized solutions. We will further develop the GenomeRNAi database to integrate phenotypic data, and extend the analysis workflows to provide a comprehensive and user-friendly pipeline from data acquisition to visualization and data mining. Major objectives include (i) web-based solutions providing access to workflows and pipelines for large-scale phenotype analysis, (ii) development of workflows for high-content screening (HCS) (e.g., bioimage data analysis, visualization, and interfacing to GenomeRNAi), and (iii) development of novel tools to visualize multi-phenotype data in GenomeRNAi.
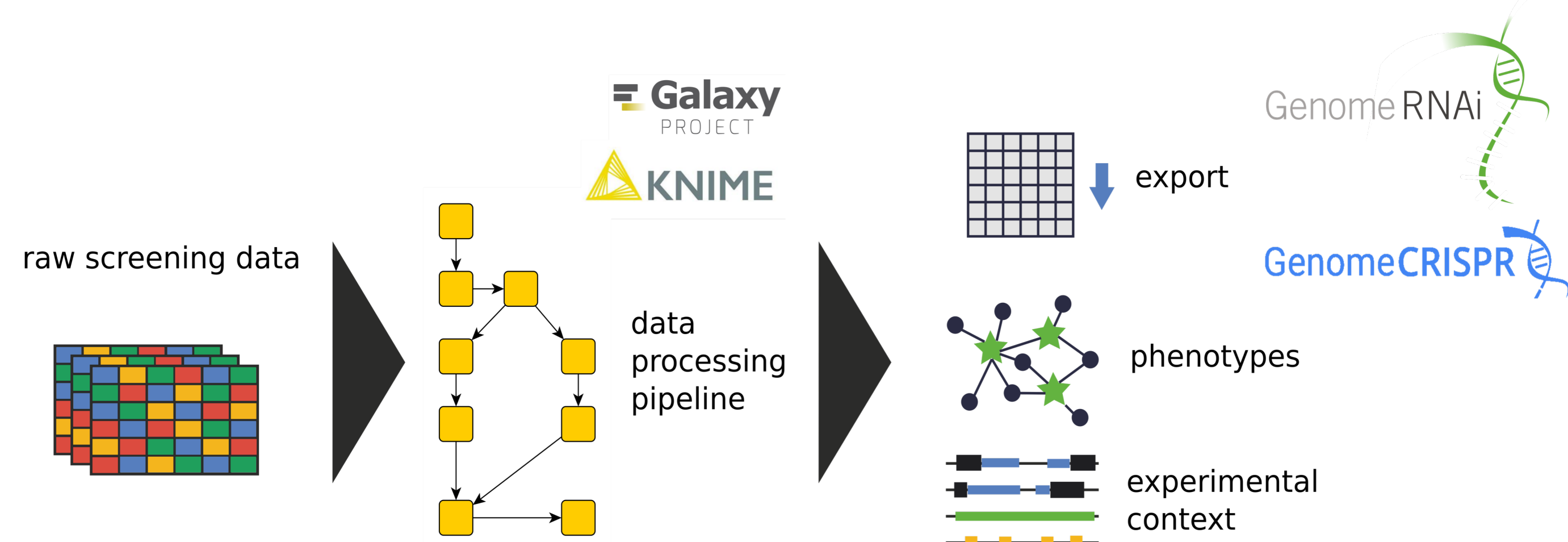
de.NBI services

The phenotype tools, workflows, and resources for all services are developed and provided by HD-HuB, and comprise the platforms KNIME cellular phenotyping of microscope imaging data, Galaxy Image Analysis for microscopy data and automatic cellular phenotyping, as well as the phenotype databases GenomeRNAi and GenomeCRISPR. Detailed information and contact data can be found on the HD-HuB homepage (*www.hd-hub.de/services*).

Service	Type	Citations	Users	Number of downloads	Number of samples / projects	Disk/File usage (TB)
KNIME Cellular phenotyping of microscope imaging data	Workflows	27	67	292	12 projects	19
Galaxy Image Analysis	Workflows	139	>25	172,809	1,986,839 images	16.7
GenomeCRISPR	Database	24	5,221	(*)	(*)	(*)
GenomeRNAi	Database	145	5,154#	(*)	(*)	(*)
E-CRISP	Webservice	299	36,994	(*)	(*)	(*)

Reporting Period: 1.1-3.12.2019; # 1.1.-31.10.2018   (*) KPI not applicable to this service

General information on the project

The team working on Systematic Phenotyping of Human Cells comprises experts in the development of workflows and pipelines for image data analysis using KNIME (Manuel Gunkel; Erfle group, Heidelberg University) and Galaxy (Thomas Wollmann; Rohr group, Heidelberg University) as well as in the development of databases and webservices (Benedikt Rauscher, Florian Heigwer and Erica Valentini; Boutros group, DKFZ). These specialists work in close collaboration on the development, integration, support and virtualization of the services and provide training and specific bioinformatics expertise to the community.

Progress report

KNIME Cellular phenotyping of microscope imaging data
Workflows for cellular phenotyping have been developed for the KNIME Analytics Platform, published via HD-HuB, GitHub and KNIME Hub (hub.knime.com) and have been successfully tested in the de.NBI Cloud. These workflows have found use in several projects and applications, also spawned by the training activities within de.NBI and HD-HuB. Additionally, plugins and nodes for the KNIME platform have been developed which are capable of controlling multiple microscope systems in order to provide feedback driven integration of image processing and image acquisition. These extensions facilitate microscopy based targeted and multiscale acquisition of phenotypic data.

Galaxy Image Analysis for microscopy data and automatic cellular phenotyping
The platform Galaxy Image Analysis was developed for web-based analysis of microscopy image data and systematic phenotyping of human cells using Galaxy. Workflows and tools were developed, integrated in Galaxy Image Analysis, and deployed by Bioconda. Selected workflows and tools were integrated within ELIXIR services like Galaxy Europe. Galaxy Image Analysis was used for several cooperation projects, for example, in an interdisciplinary cooperation on joint exploitation of microscopy and mass spectrometry image data. Results were presented at several conferences, and training activities were performed.
The tools from the service *Galaxy Image Analysis* are provided through the de.NBI cloud (*denbi-galaxy.bioquant.uni-heidelberg.de*). The tools are also available at the de.NBI und ELIXIR platform *usegalaxy.eu.*

Phenotype databases GenomeCRISPR and GenomeRNAi; E-CRISP webservice
To address the increased popularity of the CRISPR/Cas9 method as a method for functional perturbation screens, the database GenomeCRISPR was developed to provide phenotype data from high-throughput CRISPR/Cas9 screens. New screens are continuously added to both databases. Phenotype databases are used by both experimental biologists and bioinformaticians worldwide. They are publicly available through the websites *genomernai.org* and *genomecrispr.org*. GenomeCRISPR data have, e.g., been used to build integrated maps of cancer interactions (published in February 2018). The webservice E-CRISP (*e-crisp.org*) uses fast algorithms to design and evaluate guide RNAs for use with CRISPR/Cas9. E-CRISP analyzes target specificity of the putative designs and assesses their genomic context for genomes of more than 30 organisms. New functionalities and requests of the community are steadily integrated.

de.NBI Training and education

The two day training course „Microscopy Image Analysis Course 2019" with 20 participants was conducted by Karl Rohr and Thomas Wollmann, with contribution by Manuel Gunkel. Online training material for *Galaxy Image Analysis* was provided via the Galaxy Training Network and for KNIME via the KNIME Hub.

Publications

Quantification of telomere features in tumor tissue sections by an automated 3D imaging-based workflow. (2017) Gunkel M, Rohr K, Erfle H, Rippe K et al., *Methods*
GenomeCRISPR - a database for high-throughput CRISPR/Cas9 screens. (2017) Rauscher B, Heigwer F, Breinig M, Winter J, Boutros M. *Nucleic Acids Res*
Workflows for Microscopy Image Analysis and Cellular Phenotyping. (2017) Wollmann T, Erfle H, Eils R, Rohr K, Gunkel M. *J Biotech*, de.NBI Special Issue
Phenotype Databases for Genetic Screens in Human Cells. (2017) Rauscher B., Valentini, E., Hardeland, U., Boutros, M. *J Biotech*, de.NBI Special Issue
Toward an integrated map of genetic interactions in cancer cells. (2018) Rauscher B, Heigwer F, Henkel L, Hielscher T, Voloshanenko O, Boutros M. *Mol Syst Biol.*
Bioconda: sustainable and comprehensive software distribution for the life sciences. (2018) Grüning B, Wollmann T, Köster J, et al., *Nature Methods*
Accessible and reproducible mass spectrometry imaging data analysis in Galaxy. (2019) Föll MC, Wollmann T, Rohr K, Grüning BA, Schilling O et al., *GigaScience*

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

PTJ
Projektträger Jülich
Forschungszentrum Jülich

www.denbi.de