de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

eliXir GERMANY

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

BiGi
Ressource Center for microbial genome, biotechnology and medical research at Justus Liebig University Giessen

Oliver Schwengers, Burkhard Linke, Max Pfister, Jochen Blom, Alexander Goesmann
Bioinformatics and Systems Biology, Justus Liebig University Giessen, Giessen 35392, Germany   Contract Nr. W-de.NBI-010

Short description of the project

The Bielefeld-Giessen Resource Center for Microbial Bioinformatics (BiGi) provides bioinformatics software tools, pipelines and resources for microbial research. Focused on analyzing microbial genomes, metagenomes, and related datasets, BiGi at JLU Giessen supports scientists with software, workflows, and computational cloud infrastructure tailored to microbial bioinformatics needs. In addition, BiGi offers training programs, workshops and user support to help the scientific community to use the various bioinformatics solutions effectively in their research. Thus, de.NBI users as well as the larger international microbial researcher community strongly benefit from a bioinformatics triplet of a broad set of well-maintained and constantly improved software solutions, an accessible powerful and growing cloud-computing infrastructure, and a regular offer of accompanying workshops and training courses.

de.NBI services

ASA³P – WGS data to reports
ASA³P is a high-throughput processing and characterization pipeline for large cohorts of bacterial WGS isolates data. Using this pipeline, more than **12,000** bacterial genomes have been analyzed in the context of the German Center for Infection Research (DZIF), alone.
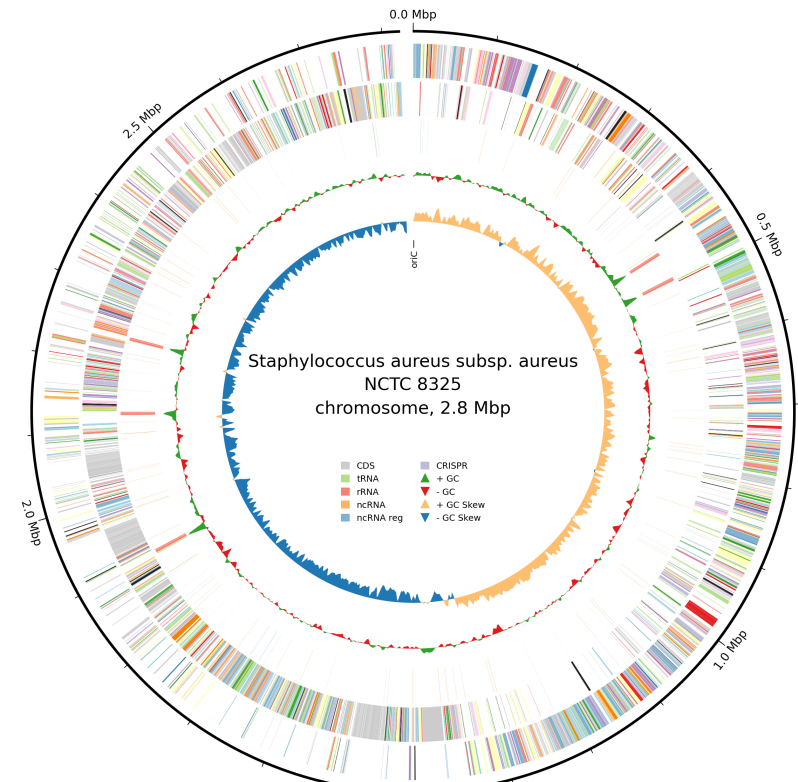
Bakta – comprehensive genome annotations
Bakta is our new de.NBI flagship tool for the fully-automated, comprehensive but rapid annotation of bacterial genomes, MAGs and plasmids. Bakta provides FAIR annotations via a novel alignment-free sequence identification approach. It is actively used (**>77,000** Conda downloads) and cited (**>470**) wordwide by a broad and growing user community.

BakRep – large-scale genome repository
BakRep is a new large-scale bacterial genome repository. Currently, for more than **661,000** genomes robust taxonomic classifications, detailed MLST subtypings, and comprehensive genome annotations have been conducted in a standardized and reproducible manner. Results integrating taxonomic information, annotation features and original submission metadata are publicly available and easily accessible via a flexible search engine. The number of available genomes is synchronized with the international AllTheBacteria project and thus, will soon be expanded to more than **2,400,000** genomes.
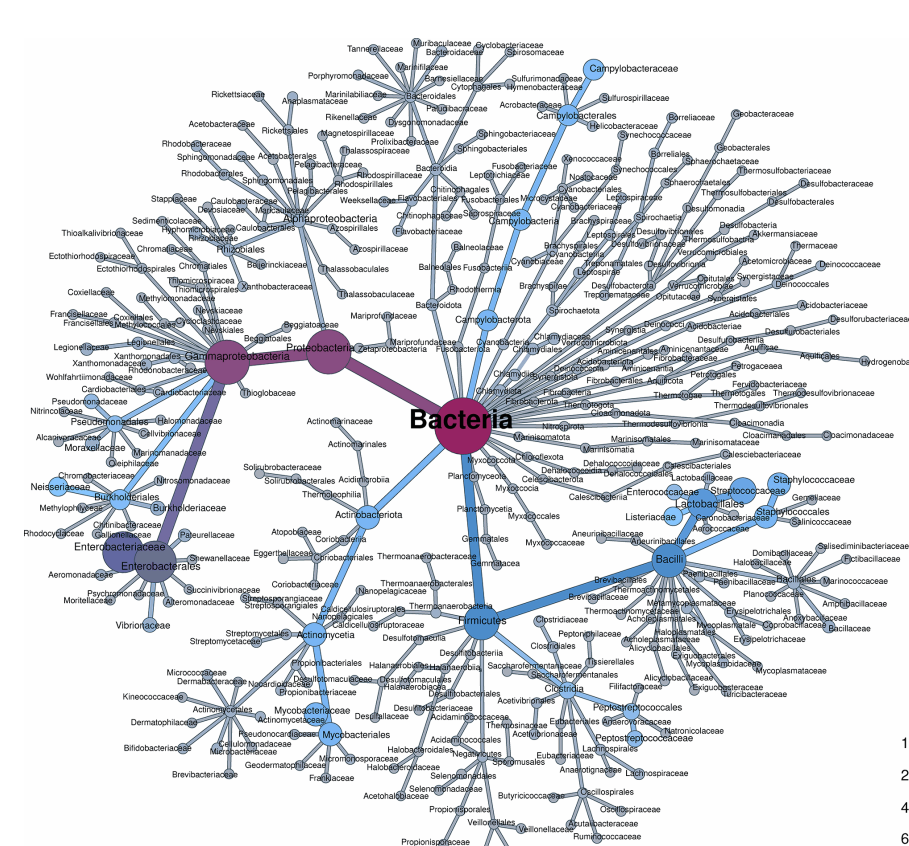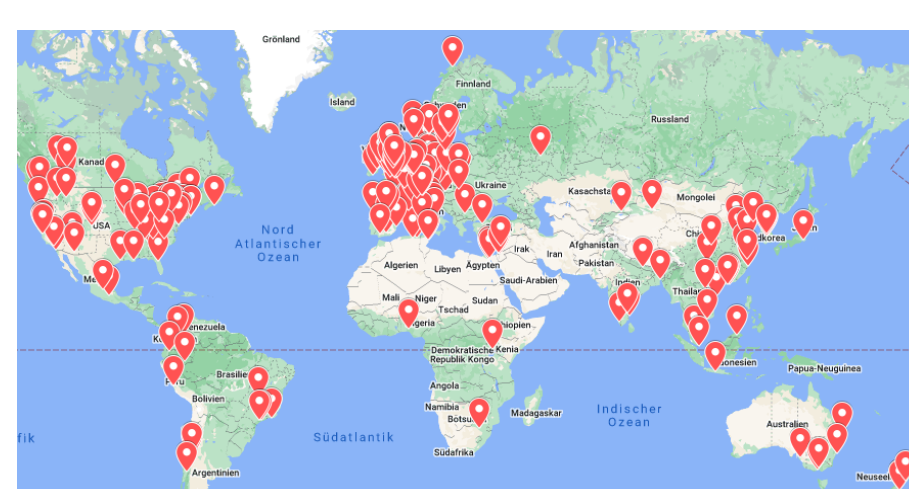
EDGAR – comparative genomics
For the comparative analysis of bacterial genomes, the de.NBI flagship tool EDGAR (**>1,200** private projects) was constantly improved and enriched with new features, *e.g.* faster orthologs detection, a scalable cloud backend, and gene annotations with functional categories. More than **25,000** genomes are analyzed using EDGAR, anually.
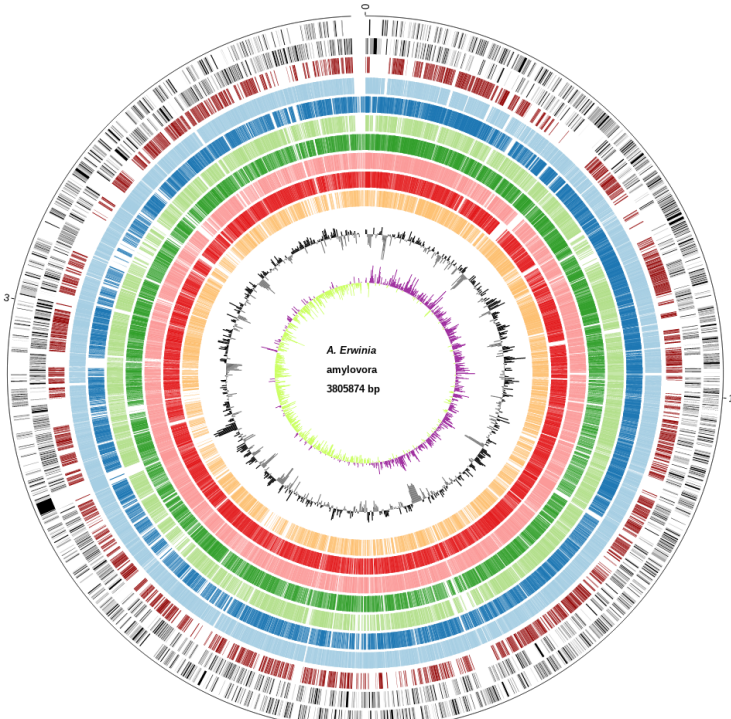
WASP – single cell RNA-Seq processing
WASP is a versatile, web-accessible single cell RNA-Seq processing platform designed for the management, analysis and interpretation of scRNA-seq high-throughput data. WASP addresses all aspects from quality control, demultiplexing and reference alignment to statistical evaluation supporting various sequencing platforms and protocols.
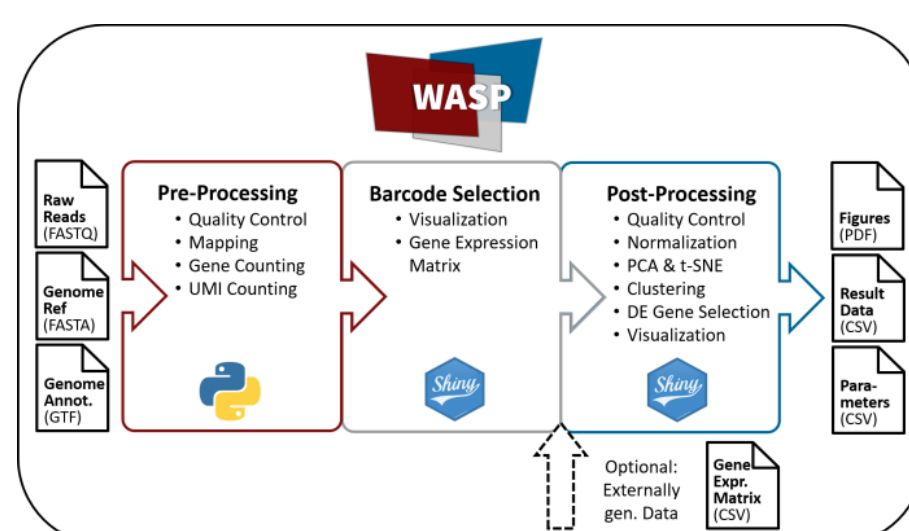
Further services
In addition, the service center BiGi at JLU Giessen provides further specialized bioinformatics services for the analysis of microbial sequence data:

Curare, **GenExVis**, **Platon**, **ReferenceSeeker**, **MGX**, ...

More at https://computational.bio/software

Progress report

Services & Tools
The various existing services have constantly been expanded and maintained following the feedback of our vast user community. In addition, with **Bakta** a novel flagship service has been established and **BakRep** will soon be added to our service repository.

Cloud
A new datacenter container became operational in 2023/12 to expand de.NBI cloud infrastructure integrating new hardware and improving data resiliency. Thus, cloud resources were significantly expanded, and software upgrades, including **Kubermatic**, enhanced management capabilities.

de.NBI & ELIXIR
The integration of de.NBI into the ELIXIR Germany node poses a major milestone for the de.NBI network. BiGi at University Giessen is engaged with one ELIXIR officer and - since January 2024 - with the chair of the ELIXIR board.

General information on the project

Currently, **6** scientists in Giessen are directly funded by de.NBI. These employees provide project work to the extent of **5** full time equivalents (2 tenured). They are involved in service, training, cloud operations as well as ELIXIR coordination.

de.NBI Training and education

Genomics Training Course
The topic of this annual workshop is sequence analysis with a focus on short-read data, genome annotation and comparative genomics. Participants get hands-on experience in the usage of BiGi tools such as ASA³P, Bakta and EDGAR.

Introduction to Deep Learning
This novel trainings course is planned for Dec 2024 and offers a comprehensive introduction to essential concepts and practical applications of modern Deep Learning approaches addressing experienced bioinformatics researchers.

Galaxy Training Course - RNAseq Analysis
This workshop held in cooperation with the RNA Bioinformatics Center (RBC) in March 2024 introduced users to the Galaxy platform using the European Galaxy Server – the biggest instance in Europe with more than 85,000 users and free access to a huge computational cloud infrastructure, databases and 3,200 bioinformatics tools. The trainings course offered demonstrations and hands-on sessions with step-by-step instructions of the Galaxy Training Material.

Publications

- **EDGAR**: Dieckmann et al., 2021: Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microbial Genomics, DOI:10.1099/mgen.0.000685

- **Bakta**: Schwengers et al., 2021: Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microbial Genomics, DOI:10.1099/mgen.0.000685

- **WASP**: Hoek et al., 2021: WASP: a versatile, web-accessible single cell RNA-Seq processing platform. BMC Genomics, DOI: 10.1186/s12864-021-07469-6

- **Curare**: Blumenkamp et al., 2024: Curare and GenExVis: a versatile toolkit for analyzing and visualizing RNA-Seq data. BMC bioinformatics, DOI:10.1186/s12859-024-05761-2

- **BakRep**: Fenske et al., 2024: BakRep-A searchable large-scale web repository for bacterial genomes, characterizations and metadata. Microbial Genomics, DOI:10.1099/mgen.0.001305

www.denbi.de

JÜLICH Forschungszentrum

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung