

A Brief Introduction

The German Human Genome-Phenome Archive (GHGA)

Flashlight Talk, CSAMA
2023

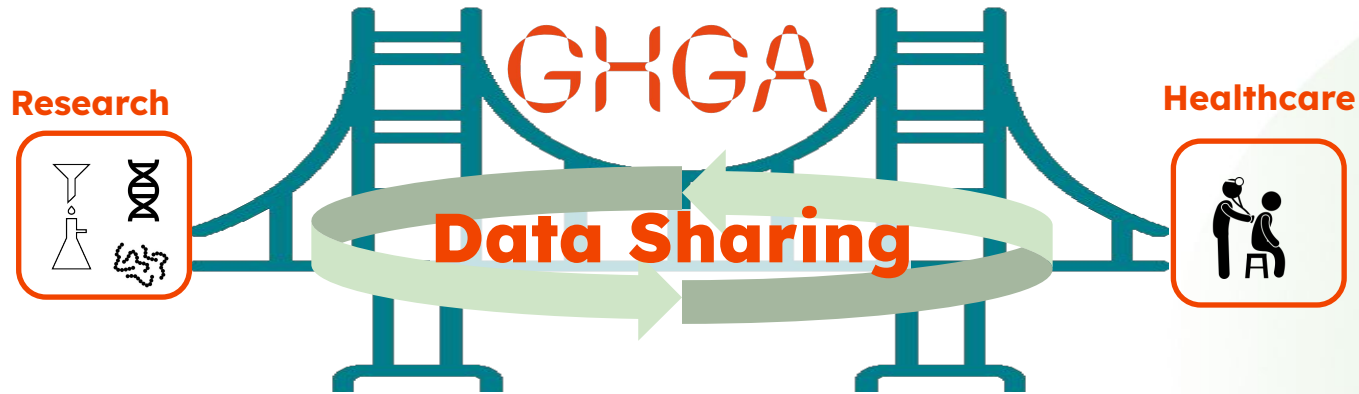
Julia Philipp
Training Coordinator GHGA

In cooperation with

Omics as Driver for Research & Healthcare

- Technological revolution made genome sequencing affordable
- Omics data will revolutionise healthcare - improving the prevention, diagnosis and treatment of certain diseases
- Genomics is increasingly part of the becoming standard care
 - More data is being produced, but locally stored
 - missing exchange between research and health care (data/knowledge)

The GHGA Vision: Enabling Genomic Medicine

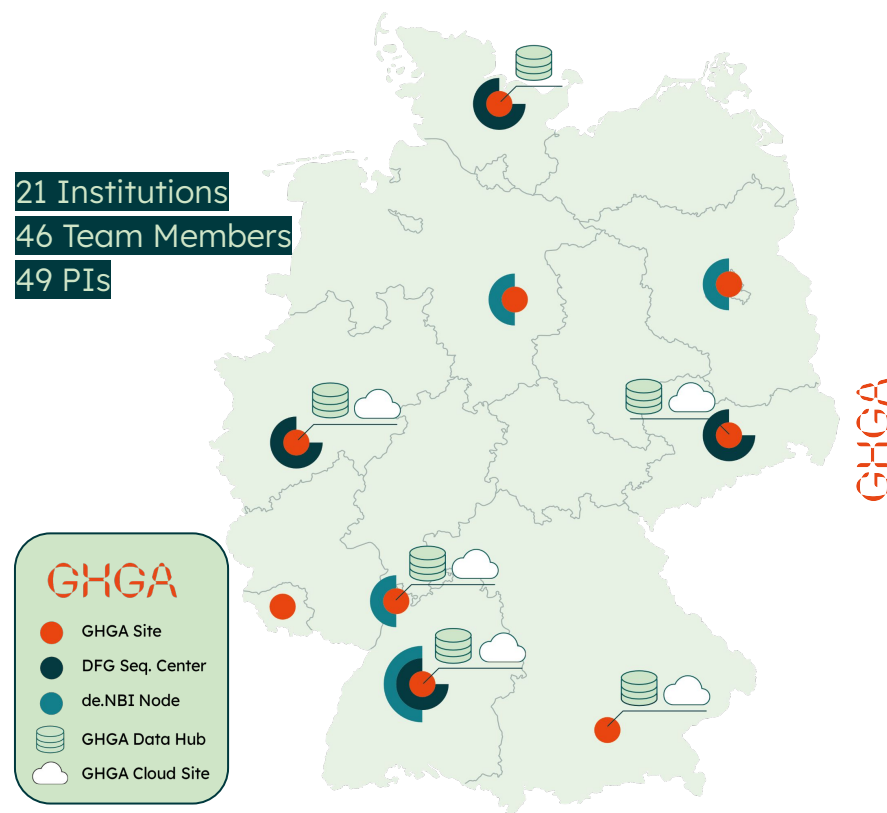


Main Aims:
Establishment of a nationally coordinated, interdisciplinary infrastructure integrating genome research and healthcare

– From Researchers for Research–

Who we are

- NFDI consortium
- Network of data hubs co-located with major academic sequencing centers
- Connected to national cloud infrastructure (de.NBI cloud) for large-scale analyses
- German national node within the federated European Genome-Phenome Archive (EGA)



Who we are

- NFDI consortium
- Network of data hubs co-located with major academic sequencing centers
- Connected to national cloud infrastructure (de.NBI cloud) for large-scale analyses
- German national node within the federated European Genome-Phenome Archive (EGA)



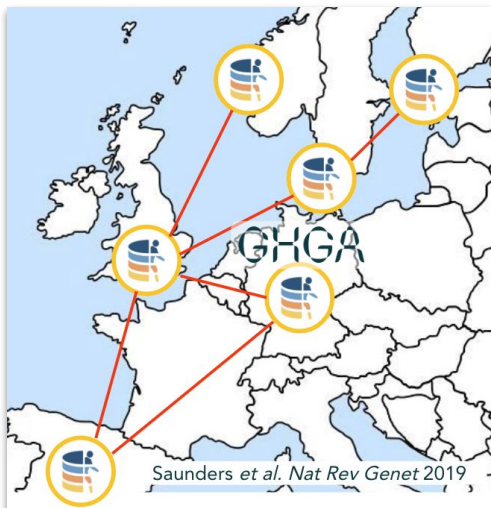
GHGA

GHGA as national node within a European data sharing infrastructure: federated EGA and GDI

GHGA as national node within a European data sharing infrastructure: federated EGA and GDI



Federated EGA Network



- European research community promotes EGA federation to enable international data sharing.
- Exchange of technologies, standards, methods and best-practice approaches
- **Linking national datasets via common metadata schemes**
- GHGA is integrated into the new **“GDI - The European Genomic Data Infrastructure”** project
⇒1+MG Initiative and EHDS

GHGA

Connection to other international key activities

Deutschland ist jetzt offizieller Partner des europäischen Großprojekts **“1+Million Genomes Initiative”**. Bundesforschungsministerin Anja Karliczek und Bundesgesundheitsminister Jens Spahn unterzeichneten heute in Berlin gemeinsam die Deklaration „Towards access of at least 1 million sequenced Genomes in the EU by 2022“. 16.01.2020



Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.

Projects that are planning to use GHGA

- In general: GHGA will be open for data submissions for all German research institutions
- Development with “Pioneer Projects” to develop infrastructure according to needs of the communities



Data Types in GHGA

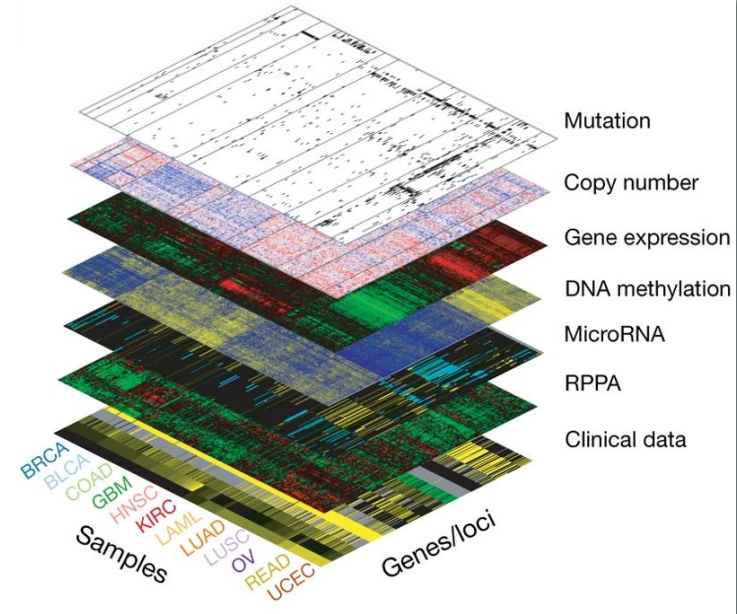
- HUMAN Genome, Exome, Epigenome, Transcriptome (including bulk and single cell data), ...
- **Connection to clinical metadata** (“Phenome”) is key!
- **Access to raw data** is often essential for research

Disease Communities:

- Initial Focus on **Rare Diseases** and **Cancer**

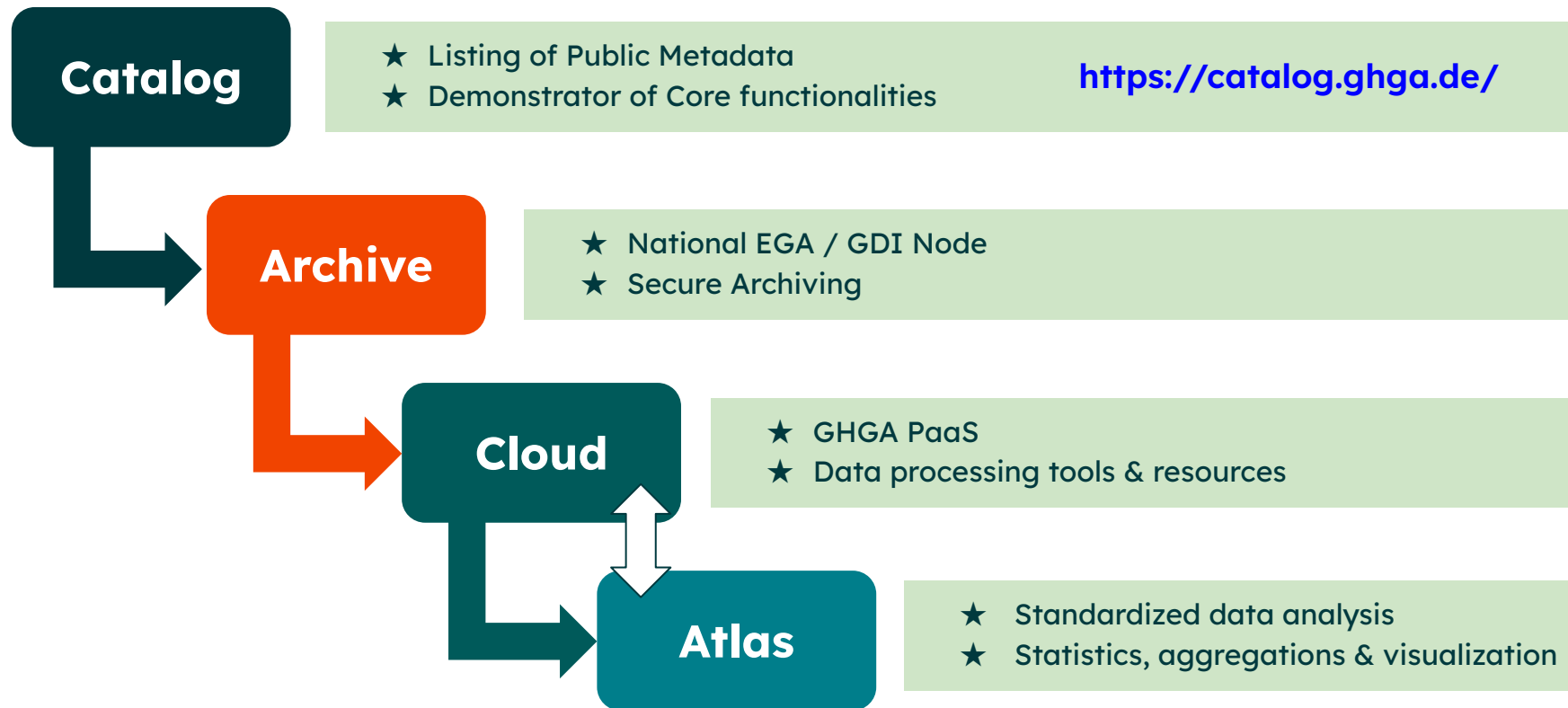
Regulated Data Access:

- Access only under “Controlled Access” with clear contractual regulations according to **GDPR and patient consent**



Source: ICGC/TCGA; The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 45, 1113–1120 (2013)

GHGA Phases and Features





Total Datasets: 63

Study ID☐ EGAS00001004813 **62**☐ EGAS00001006680 **1****Project**☐ MASTER_remastere
d **62**☐ MNP2 **1**

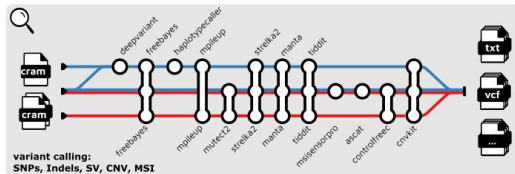
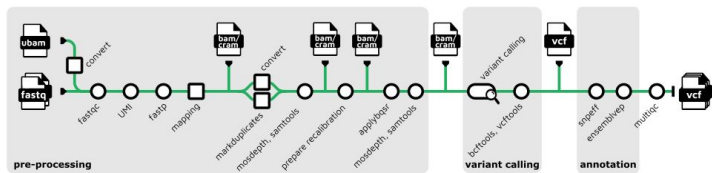
Dataset ID	Title	
EGAD00001008861	Dataset for head and neck cancer RNA	▼
EGAD00001008863	Dataset for hepatopancreaticobiliary malignancy RNA	▼
EGAD00001008859	Dataset for soft tissue tumor RNA	▼
EGAD00001008862	Dataset for neuroendocrine adrenal tumor RNA	▼
EGAD00001008893	Dataset for neuroendocrine adrenal tumor WHOLE GENOME	▼

Workflow Standardization

More info: ghga.de/resources/data-analysis

Short-read variant calling

- **Germline (Nahnsen)**
- **Somatic (Hübschmann)**



☐ Mandatory

☐ Optional

— Core workflow

■ Germline variant calling

— Tumor only variant calling

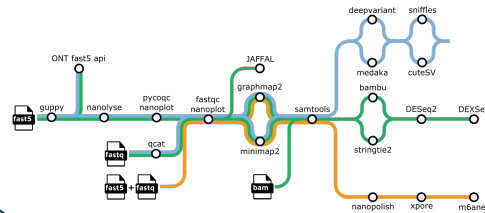
 Tumor-normal pair variant calling

Adapted from: Fellows Yates, James A., et al. *PeerJ* 9 (2021).

Long-read sequencing (Ossowski)

nf-core/nanoseq

Nanopore basecalling, demultiplexing, QC, alignment, and downstream analysis



RNA-seq for RD (Gagneur)

Aberrant Expression



Aberrant Splicing



Mono-allelic Expression



RNA-seq Variant Calling



Integration into communities

- Co-development with nf-core

nf-core 

- Training (Stanford, Broad)
- Benchmarking (NGS-CN)



Training Resources

Live & On Training
ghga.de/resources/training



[HOME](#) [ABOUT US](#) [IMPACT](#) [NEWS & EVENTS](#) [RESOURCES](#)

Training

We are passionate about supporting our users and communities in all topics relevant to sharing data in the field of biomedical research and health care, as well as related bioinformatic methods. Therefore we made it our mission to provide training and learning opportunities in these areas.



Sequencing Techniques and Bioinformatic Analyses

Material covering information from DNA and RNA sequencing methods to bioinformatic analyses and analysis workflows.

[Learn more](#)



Ethical and Legal Implications

Learn more about ELSI topics, e.g. consent in general, but also GHGA tools that can help you assess existing consent forms.

[Learn more](#)



FAIR and Metadata

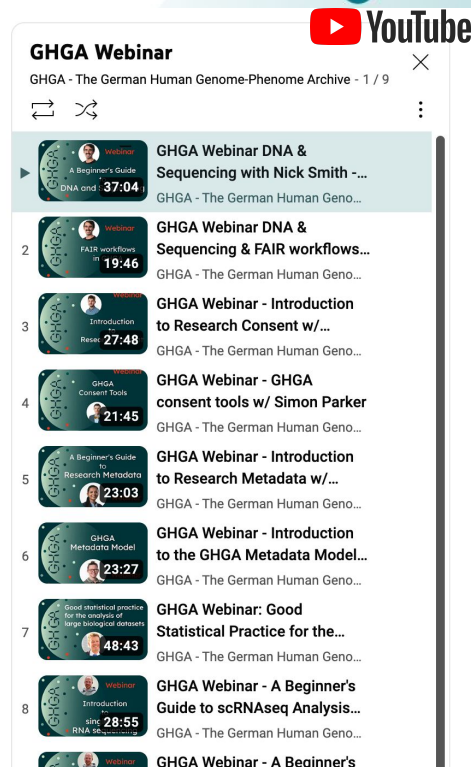
Interested in Research Data Management? Want to learn more about research metadata? Curious about what the FAIR principles are and how to apply them in the context of biomedical research?

[Learn more](#)



Sign up to our newsletter for news about future events like our monthly webinars!

Youtube
<https://t1p.de/rgsse>



Upcoming GHGA events



German Conference on Bioinformatics | GCB2023

12 - 14 September 2023, Deutsches Elektronen-Synchrotron DESY Hamburg

11 September 2023, Hamburg
**NGS Harmonization &
Standardization Workshop**

WS4) Standardizing and harmonizing NGS analysis workflows

Organizers: Dr. Florian Heyl (German Cancer Research Center, Workflow Coordinator for GHGA); Dr. Kübra Narci (German Cancer Research Center); Dr. Paul Menges (German Cancer Research Center); Dr. Christian Mertes (Technical University of Munich); Dr. Nicole Schatłowski (University of Tübingen); Dr. Julia Philipp (European Molecular Biology Laboratory, Training Coordinator for GHGA)

Description: With increasing numbers of human omics data, there is an urgent need for adequate resources for data sharing while also standardizing and harmonizing the processing of the data. Within the federated European Genome-Phenome Archive (EGA), the German Human Genome-Phenome Archive (GHGA) strives to provide (i) the necessary secure IT-

Upcoming GHGA events

ghga.de/events



Webinar

Introduction to
Benchmarking of NGS
Workflows

20.06.2023
16:00 (CEST)



Kübra Narci
DKFZ

GHGA

GHGA



Jobs

If you are interested in joining GHGA, please get in touch! We are constantly looking for talents in software development and cloud computing with an interest in setting up a state-of the art infrastructure for genome research. Please contact contact@ghga.de for further information.

GHGA

THE
GERMAN
HUMAN
GENOME -
PHENOME
ARCHIVE

