The German Human Genome-Phenome Archive

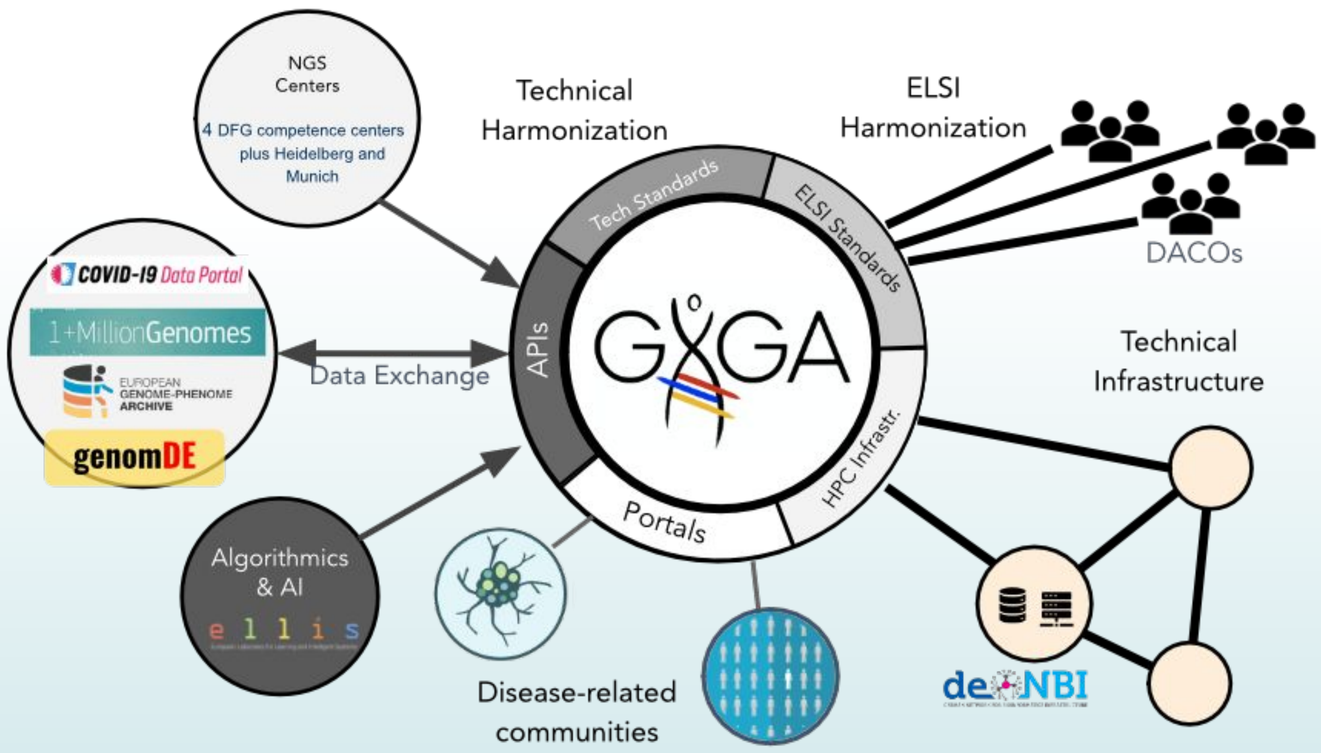
A national infrastructure for secure archival and community-driven analysis of omics data

Kübra Narci [1,2],on behalf of the GHGA consortium

1 Computational Oncology Group, Molecular Precision Oncology Program, NCT Heidelberg and DKFZ, Heidelberg, Germany.;2 German Human Genome-Phenome Archive (GHGA, W620), Deutsches Krebsforschungszentrum, Heidelberg, Germany

What is GHGA?

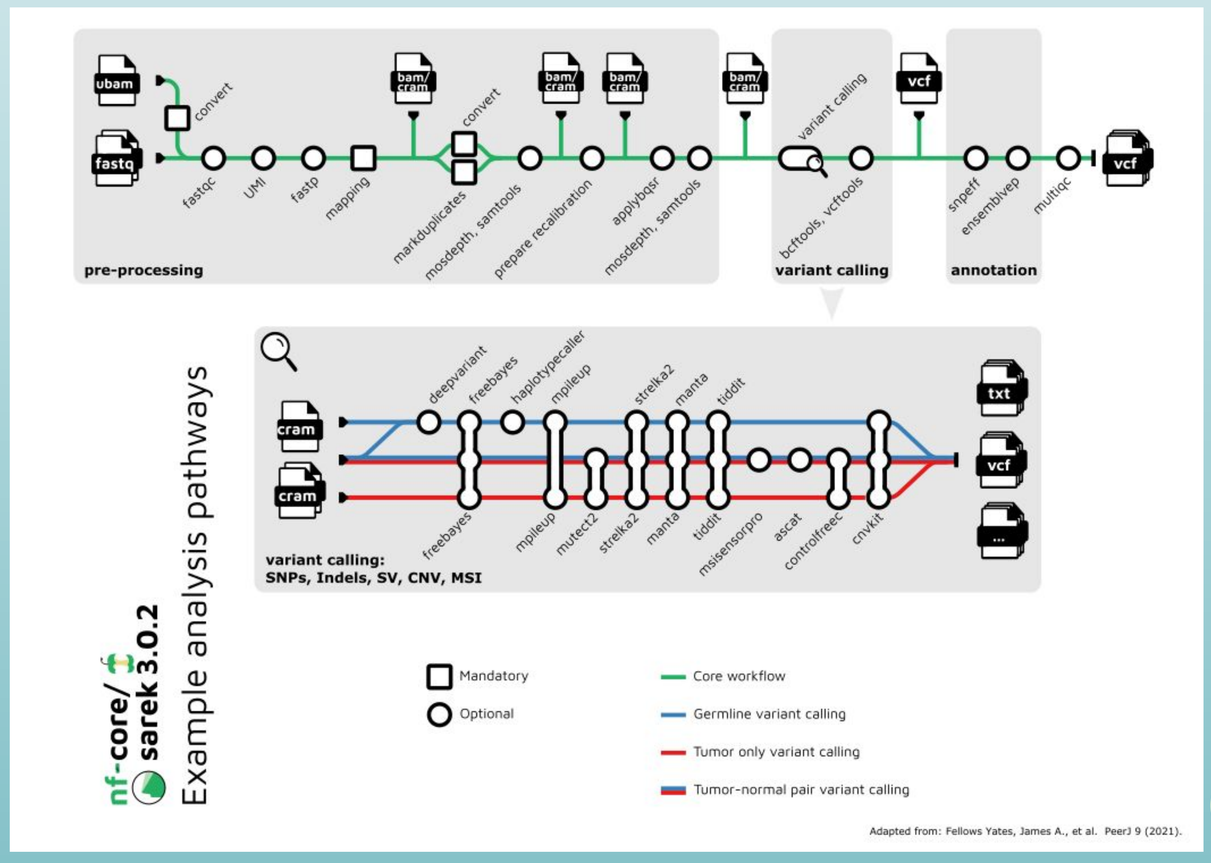
GHGA is a newly established consortium with the aim to build a national federated infrastructure to store and share human omics data.
It is funded as part of the National Research Data Infrastructure (NFDI e.V.), via the DFG and It will act as the national node for the federated European Genome-Phenome Archieve (fEGA) .
With 46 participants from 21 institutions, GHGA is organised in 6 data hubs across Germany combining leading institutions in genomic medicine, major omics data producers and HPC centers committed to provide scalable infrastructure.

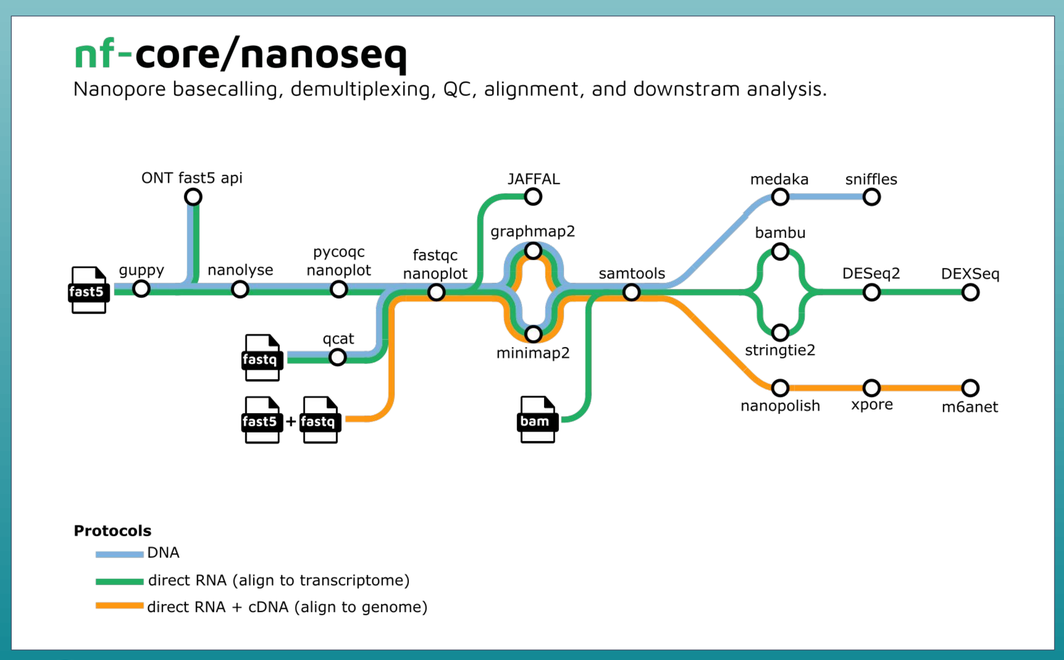


Standardized and harmonized Next Generation Sequencing (NGS) analysis workflows

GHGA is going to be more than an archive. With additional functionalities, we aim to bring the algorithms to the data.

- Unified processing and standardized workflows using platforms like nf-core for large-scale omic data to democratize data processing and metadata.
- Processing existing and relevant datasets relating to cancer and rare disease in the same way to establish a cohort catalog with comparable and ready-to-use data for the research community.
- Automated and continuous benchmarking of workflows to guarantee not only the technical but also scientific high standards using GiaB or CHM data.





Examples of the GHGA workflows co-produced with the (nf-core) community:

- nf-core/SAREK: Short-read DNA variant calling
- nf-core/nanoseq:Nanopore long-read DNA-seq analysis
- nf-core/scrnaseq: 10x single cell RNA-seq analysis
- gagneurlab/DROP: Rare disease RNA-seq analysis

What we are currently working on:

- Research diagnostic analysis
- Somatic variant calling
- RNA-seq quantification
- Benchmark analysis

Want to know more? Go to: www.ghga.de or get in touch under ghga-office@dkfz-heidelberg.de