Genedata Profiler™: A Collaborative, Scalable, Regulatory-compliant, Computational Infrastructure for Efficient Multi-omic Profiling of Patients in Global Clinical Trials

Tim Zeppenfeld[1], Michael Remmert[1], Johannes Wollnik[1], Tamas Rujan[2], Marc Flesch[1]
[1]Genedata GmbH, Munich, Germany | [2]Genedata AG, Basel, Switzerland | Email: profiler@genedata.com

Abstract

Modern clinical trials are increasingly biomarker-based and make use of multi-omic profiling of patients for recruitment and monitoring during the trial. Frequently, less than 20% of patients respond to new immunotherapies, therefore biomarkers predicting patient response can significantly increase the chance of approval of new immunotherapies. This increasing use of multi-omic data in clinical trials poses significant challenges to pharma companies, which must generate timely results from huge volumes of data while ensuring full regulatory (GCP) compliance. Critical challenges range from moving proven algorithms and analysis pipelines from research into clinically actionable results, through lack of standardization and appropriate GCP controls, to lack of scalability and suitable analytics. We report here how Genedata Profiler™, an enterprise software platform for multi-omic patient profiling, solves the challenges of processing, analyzing, and managing large volumes of multi-omics data from clinical trials, at scale, and in a regulated (GCP) environment. We also demonstrate that by establishing this infrastructure, organizations can bridge the clinical-research divide to provide harmonization of data, processes, and procedures for translational researchers.

The Challenges of Multi-omics in Clinical Trials

Many leading biopharma companies are developing novel targeted protein therapies, in particular utlilizing the power of the patient´s own immune system to fight a broad range of cancers. These therapies are highly selective and show impressive results in target "enriched" patient populations. As a result, clinical trials for such immunoregulatory therapies involve in-depth multi-omic profiling of patients entering the trial, which poses considerable challenges to organizations that wish to conduct these biomarker-based trials as efficiently as possible. These challenges fall into three main categories:

Infrastructure Challenge
Federating globally distributed NGS and clinical data

- Scale to tens of thousands of genomic profiles
- Integrate with existing and new data sources (e.g. other omics data) at scale
- Ensure compatibility with clinical data reporting standards, e.g. CDISC

Regulatory Challenge
Complying with patient data privacy rules & GxP regulations

- Ensure GCP compliance with a qualified infrastructure
- Support comprehensive data provenance, chain of custody, data security and privacy
- Meet requirements for validation of data, processing and analysis

Data Challenge
Efficiently managing, processing, and analyzing omic & other data

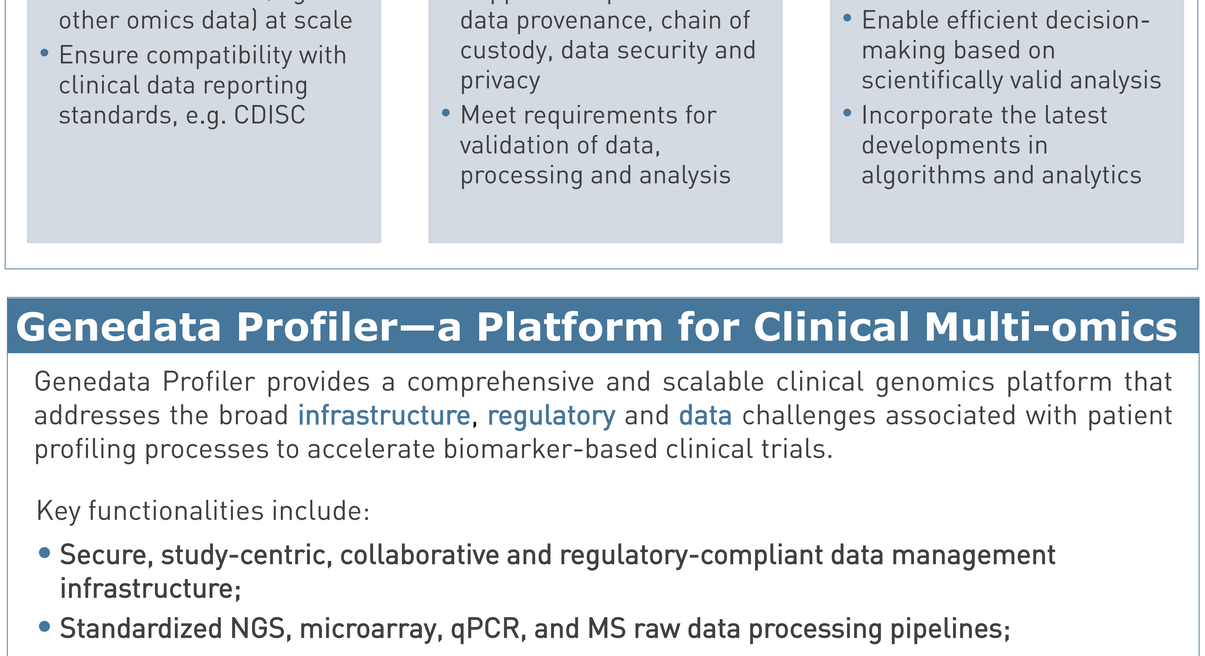
- Scale NGS pipelines from research use to deliver clinically actionable results
- Enable efficient decision-making based on scientifically valid analysis
- Incorporate the latest developments in algorithms and analytics

Genedata Profiler—a Platform for Clinical Multi-omics

Genedata Profiler provides a comprehensive and scalable clinical genomics platform that addresses the broad infrastructure, regulatory and data challenges associated with patient profiling processes to accelerate biomarker-based clinical trials.

Key functionalities include:

- Secure, study-centric, collaborative and regulatory-compliant data management infrastructure;
- Standardized NGS, microarray, qPCR, and MS raw data processing pipelines;
- Regulatory submission-ready data reporting to CDISC SDTM-PGx format;
- Comprehensive method lifecycle management;
- Integration with major HPC environments;
- Wide range of data analyses via a rich statistical, visualization and reporting module;
- Out-of-the-box integration with public data repositories;
- Rich, open APIs for integration with internal and external data and applications;
- Cloud-ready.

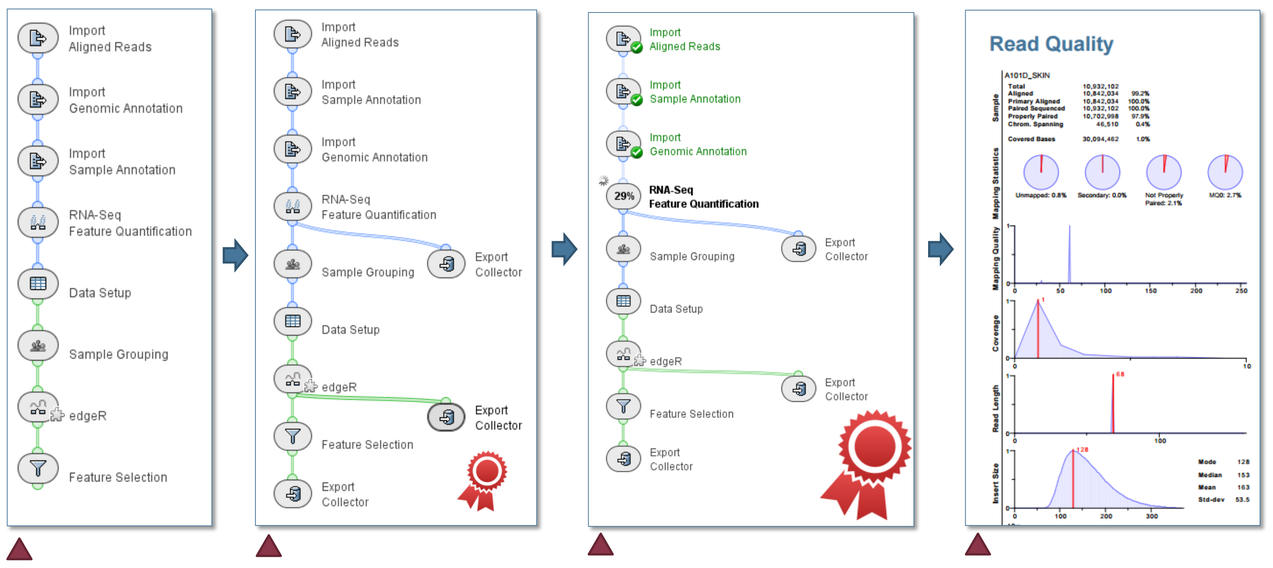


Study-centric, role-based data management system facilitates collaboration, data, method, and results sharing. A rich web UI offers a one-stop-shop for all study-related raw data, methods, processed data, results and reports.

Method Lifecycle Management of NGS Processing Pipelines

Genedata Profiler significantly reduces the time and risk of scaling NGS processing pipelines from biomarker R&D for use for patient stratification in trials with a comprehensive method management system that allows users to:

- Build processing and analysis workflows using over 120 pre-built activities;
- Automate and deploy best-practice workflows throughout an organization with full method lifecycle management, including versioning, approval and access control;
- Easily integrate state-of-the art algorithms;
- Maintain quality and auditability with detailed quality reports.



Rapidly build NGS and other omics processing pipelines with a full-featured visual editor

Test and lock down parameters and approve for use by others, e.g. in clinical genomics

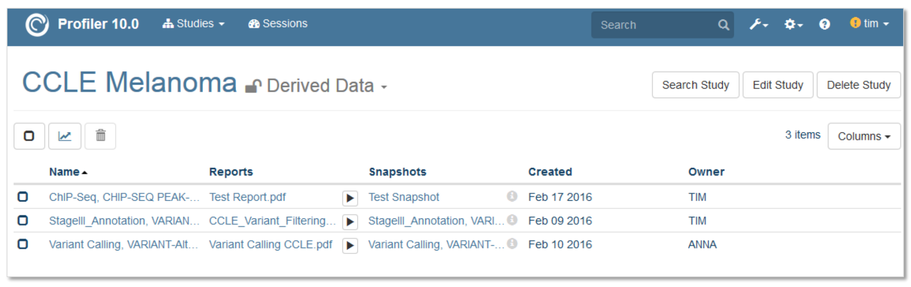
Non-expert users may run the standardized, approved workflow or use in fully automated mode

The quality report includes metrics, all parameters, version number, and details of who ran the workflow and when, and is linked to the data for full GCP compliance

Ensure Compliance with Data Lifecycle Management & Access Control

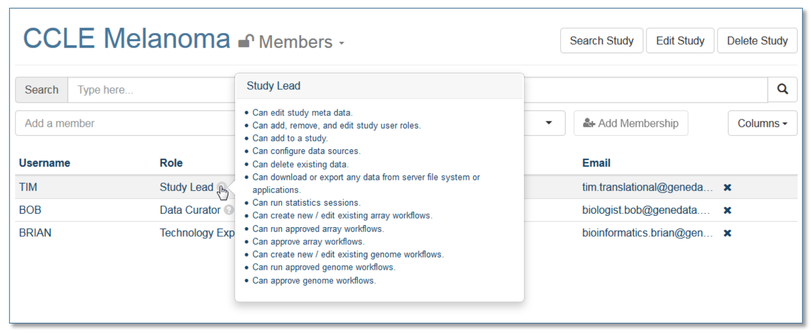
Information systems that process, manage and analyze data from clinical trials are subject to regulatory and data privacy rules. Genedata Profiler fulfills all the requirements for use in GCP-controlled environments, through implementation of the following core functions:

- Data provenance to ensure knowledge of which data is being used and from where;
- Chain of custody of data offers full linkage of data, from raw data to results;
- Comprehensive role-based access controls provide fine-grained permissions for access to data and the operations that can be performed on that data;
- Audit trails document changes to data, by whom, when and reasons for change.



Genedata Profiler ensures GxP compliance by maintaining a link between raw data, analyses, derived/processed data, and results on a per study basis.

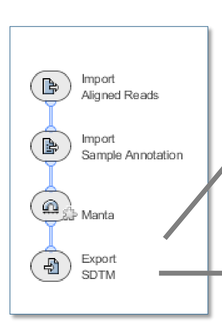
Comprehensive, flexible role-based access controls support collaboration while maintaining regulatory compliance. Data can be shared with translational researchers, ensuring harmonization of clinical and translational research programs



Submission-ready Data—CDISC SDTM-PGx

Genedata Profiler uniquely provides a submission-ready system of record reporting for genomics data to the CDISC SDTM-PGx standard.

Export SDTM activity can be added to any processing pipeline to automatically output data in SDTM-PGx format for inclusion in a regulatory submission



gene_fusion.desc

# CDISC Term	Type	Description
DOMAIN	String	# Required; CDISC domain
PFSEQ	String	# Required; sequence number;
STUDYID	String	# Required; identifier
VISITNUM	String	# Required; patient
PFREFID	String	# Required; identifier for

gene_fusion.tsv.gz

#DOMAIN	PFTEST	PFGENRI	PFGENLOC	PFGENSR	PFSTRESC	PFSTRESU
PF	Illumina kit 4711	ENSG00000153207.10	hg19:chr1	247094431	AHCTF1--NAAA;-;198;38;1	GeneFusionName;Strand;Score;SpanCount;
PF	Illumina kit 4711	ENSG00000124209.3	hg19:chr20	56886178	RAB22A--MYO9B;+;1193;11;6	GeneFusionName;Strand;Score;SpanCount;

Summary

Genedata Profiler allows the genomic profiling of patients to be "operationalized", allowing biopharma companies to conduct biomarker-based trials efficiently, e.g. for immunotherapeutic drugs. Genedata Profiler is regarded by clients as the only software platform capable of performing this process at scale in a GCP environment, with the ability to harmonize with translational research as required.