

Genedata Profiler™ – a collaborative, regulatory-compliant, integrated platform for omics-based patient and compound profiling



Mark Collins¹; Axel Schumacher²; Marc Flesch²; Harri Lempiainen³; Tamas Rujan³

¹Genedata Inc, Lexington, USA | ²Genedata GmbH, Munich, Germany | ³Genedata AG, Basel, Switzerland | Email: profiler@genedata.com



Abstract

Recent advances in next-generation sequencing (NGS) and other omics technologies for patient and compound profiling have created a variety of challenges for researchers, ranging from integration of large volumes of disparate genomic and phenotypic data, through generating meaningful conclusions from such data, to sharing data and results with collaborators across global locations. Furthermore, as organizations increasingly leverage confidential patient data (medical records, genetic information, etc.) for precision medicine initiatives, important regulations governing data security and privacy must be respected. Recent high profile decisions such as EU Safe Harbor changes create huge issues for use of patient data in non-clinical research environments. Addressing such issues must be part of any translational research informatics approach.

Genedata Profiler™ is a new translational research software platform developed in collaboration with leading pharmaceutical companies to effectively process, manage, and analyze omics and phenotypic data to the highest data quality and regulatory compliance standards. Strict user role management, audit trails, access authorization, data federation and a comprehensive reporting infrastructure are central components of the software to ensure conformity with the tightening legal framework governing patient data privacy and security.

In this poster we will discuss the key features of Genedata Profiler and illustrate the power of the platform using a case study that also demonstrates how Genedata integrates with tranSMART.

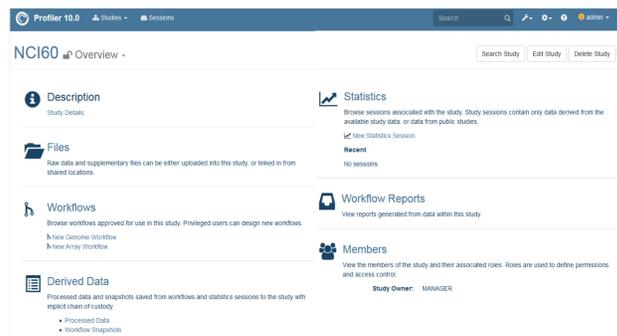
The challenges of translational research

Translational researchers face major challenges in drawing meaningful conclusions from the very large volumes of data they generate in order to deliver maximum value to the stakeholders in their organizations:

- **Infrastructure Challenge** – Facilitating collaboration within interdisciplinary teams
- **Regulatory Challenge** – Working with data from human samples in research environments
- **Data Challenge** – Efficiently managing, processing, and analyzing omics and other data

Genedata Profiler – a platform for translational research

Genedata Profiler provides a comprehensive and scalable translational research eco-system that addresses the **infrastructure**, **regulatory** and **data** challenges associated with patient profiling processes to accelerate translational research activities.



Study-centric, role based web UI facilitates collaboration, data, method and results sharing. The underlying enterprise architecture provides integration/federation with HPC file systems, internal and external databases and public data sources

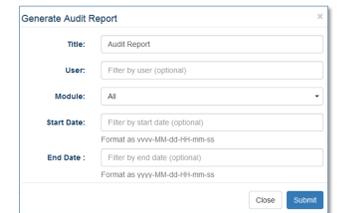
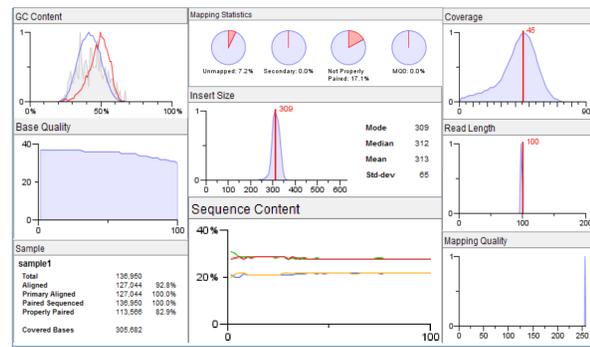
Key functionalities include:

- Secure, collaborative and regulatory-compliant data management infrastructure
- Standardized NGS, microarray, qPCR, and MS raw data processing pipelines
- Comprehensive, advanced statistical analysis and reporting
- Integration with internal, external and public genotype-phenotype data

Regulatory

Genedata Profiler facilitates compliance with the increasingly complex global regulatory landscape that surrounds use of patient data and specimens in both clinical trials and research studies. Specifically Genedata Profiler offers:

- **Role-based access control** to data to help comply with patient privacy regulations such as HIPAA and EU Safe Harbor
- **Full chain-of-custody** and controlled handover of data, results and reports to internal and external stakeholders
- **Audit Trails** to comply with electronic-records regulations (e.g. 21CFR.Part.11)
- **Method management system** that includes authorization and execution of SOPs, QC workflows, reporting, and approval processes



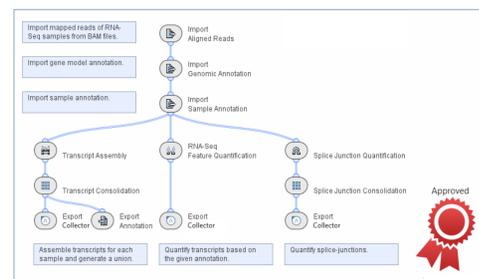
Audit logging is one of many optional features of Genedata Profiler which make the platform compliance-ready.

Configurable and dynamic quality metric plots illustrate information about the NGS reads that have been processed, such as sample & mapping statistics, coverage, mapping quality, read length, insert size, base quality, sequence- or GC-content.

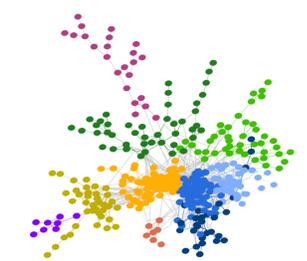
Data

The Genedata Profiler platform combines a powerful visual, workflow-based processing engine together with advanced interactive data analysis tools and intuitive visualizations. This allows the use of the latest NGS processing algorithms, for example, while meeting QC standards for processed data and ensuring pipelines can be used by non-experts.

- Share best practice NGS and other workflows throughout an organization
- Utilize existing in house and public scripts via the plug-in framework
- Unparalleled data visualization (Genome Browser)
- Comprehensive analytics from a rich and growing statistical toolbox, including external algorithms (e.g. R) as plug-ins



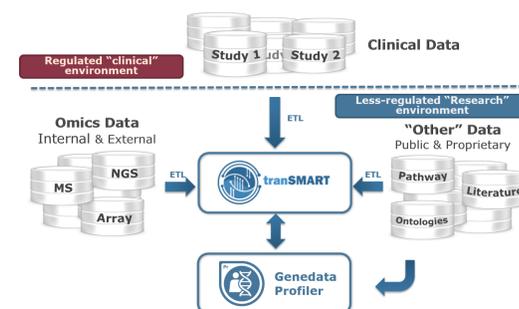
The processing framework offers a fully automatable graphical workflow system, a comprehensive reporting infrastructure and support for all major schedulers of computing clusters



Powerful analytics and visualizations coupled to easy integration of disease and genomic information from different studies allows innovative holistic data analyses, significantly reducing time and cost of NGS data analysis.

Case Study: Using Genedata Profiler to make tranSMART smarter

Genedata Profiler works together with several other software solutions such as tranSMART, an open-source data warehouse, to manage multi-omics, clinical and phenotype data sets in translational research. Integrating Genedata Profiler's infrastructure, regulatory and data processing platform with tranSMART addresses many of the commonly reported tranSMART limitations and significantly enhances an organization's capabilities for translational research.



- Enhanced raw data processing, management, and visualization across the workflow
- Advanced exploratory analytics and reporting
- Simplified curation and publishing of public and proprietary data

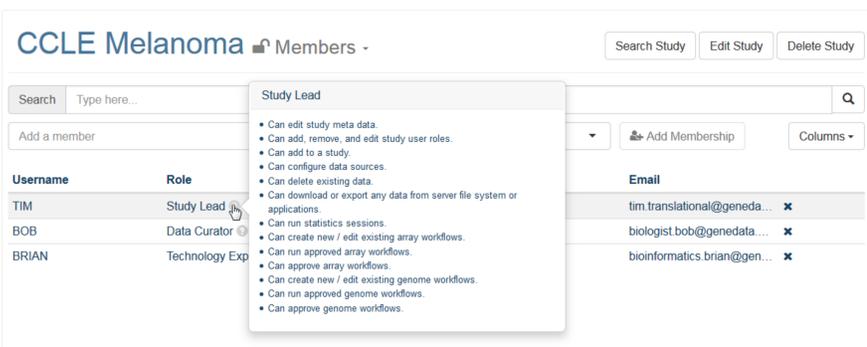
Summary

Genedata Profiler, an end-to-end solution for translational research, combines a high performance NGS raw data processing pipeline, sophisticated data analyses, and unparalleled data visualizations with an advanced distributed data management environment. It helps organizations to comply with regulatory data retention and privacy policies and provides integrated metadata management throughout the entire data lifecycle.

Infrastructure

Genedata Profiler comes with a rich set of enterprise-wide data and user management functionalities to help researchers overcome the infrastructure challenges associated with translational research and enables:

- Collaboration between different user roles throughout a global organization using comprehensive role-based access controls
- Integration, federation and curation of the wide variety of omic, phenotype and patient data from internal (e.g. tranSMART), external and public data sources
- Sharing of data, methods and results



Comprehensive, flexible role-based access controls support collaboration while maintaining regulatory compliance