GA4GH 6th Plenary Meeting Report

INTRODUCTION

The GA4GH 6th Plenary Meeting in Basel, Switzerland brought together more than 430 individuals representing more than 200 organizations across 25 countries to learn about new deliverables being developed under GA4GH Connect — a multi-year strategic phase focused on connecting GA4GH development work to the immediate data sharing needs of the genomics and health community. An additional 200+ individuals joined the meeting by live stream on October 4 and 5.

GA4GH Work Streams met at various times throughout the meeting. Minutes from those discussions can be found here.

The GA4GH 6th Plenary Meeting was covered by *Bio-IT World* and by *GenomeWeb*. One of the newly announced deliverables, the Workflow Execution Service, was also covered by *Bio-IT World*.

OCTOBER 3

The first day of the GA4GH 6th Plenary Meeting focused primarily on topic-oriented workshops and ancillary meetings, with an introduction to the organization for those new to the community.

Welcome & Opening Remarks

Ewan Birney (EMBL-EBI)

Millions of genomes will soon be available from healthcare and large scale genomics initiatives around the globe. To make best use of this data for human health and medicine, activities must be interoperable and thus built on common standards.

Download Slides

Intro to GA4GH Connect

Peter Goodhand (OICR, GA4GH), Adrian Thorogood (McGill University, GA4GH), Melissa Konopko (Wellcome Sanger Institute, GA4GH), Rishi Nag (EMBL-EBI, GA4GH)

The Global Alliance for Genomics and Health (GA4GH) is the international standards organization for genomic data. It aims to accelerate progress in genomic science and human health by developing standards and framing policy for responsible genomic and health-related data sharing. Eight open, topic-oriented Work Streams are developing the standards and tools needed to achieve this mission, with direct input from GA4GH Driver Projects — real world genomic data initiatives that represent the broad genomics community and help develop and pilot GA4GH deliverables.

Download Slides

Workshops and Side Meetings

1. WORKSHOP: GENOMIC DATA COMPRESSION

Thomas Keane (EMBL-EBI), Marco Mattavelli (École Polytechnique Fédérale de Lausanne)

To discuss current approaches to genomic data compression and identify opportunities for collaboration.

Agenda & Slides

2. GENE WORKSHOP: TRANSCRIPT CHOICE

Andy Yates (EMBL-EBI), Robert Freimuth (Mayo Clinic), Fiona Cunningham (EMBL-EBI)

A forum for discussion of the current progress for the work that the EMBL-EBI and NCBI have done jointly to harmonise a subset of transcripts.

More Details

3. GENE WORKSHOP: GENE-DISEASE VALIDITY

Heidi Rehm (Broad Institute), Ellen McDonagh (Genomics England)

To align all groups defining the validity of gene-based relationship, which informs their use of clinical testing, including the formation of gene panels.

More Details

4. ELIXIR Meeting

Serena Scollen (ELIXIR), Gary Saunders (ELIXIR), Michael Baudis (University of Zurich)

Presentation and discussion of ELIXIR and GA4GH collaborations, including an in depth ELIXIR Beacon Driver Project session.

Agenda

5. WORKSHOP: DEMONSTRATING THE IMPACT OF GENOMIC DATA SHARING

Kathryn North (Murdoch Childrens Research Institute), Anna Middleton (Wellcome Sanger Institute)

To promote the positive impacts of genomic data sharing for the benefit of human health and medicine.

Agenda

OCTOBER 4

The second day of the GA4GH 6th Plenary Meeting was focused on enabling responsible genomic data sharing, with presentations from 13 active GA4GH contributors on work to date, a keynote talk from Andrew Morris (HDRUK), and panel discussions on the human right to benefit from data sharing and privacy and security.

National Initiatives Workshop

Kathryn North (Murdoch Childrens Research Institute), Mark Caulfield (Genomics England)

Participants discussed key challenges facing large scale genomics initiatives around the globe, including harmonizing minimal clinical datasets, automated consent, and ongoing communication between National Initiatives.

Agenda

GA4GH Connect: The Story So Far

Ewan Birney (EMBL-EBI)

The GA4GH Connect Strategic Roadmap lays out more than two dozen standards and deliverables to be added to the GA4GH suite to enable responsible genomic data sharing. These tools are designed to work together as an end-to-end solution for data sharing, but can also be used alone to solve a single challenge in an existing workflow. Each deliverable must pass through a product approval process to ensure it is fit for purpose and passes the ethics and security guidelines of the GA4GH foundational Work Streams.

Watch Video | Download Slides

Setting the Standard: The Path for International Data Sharing

Watch Video

1. htsget

Mike Lin (DNAnexus)

htsget provides a secure, consistent protocol for researchers to access data stored in different repositories — whether based in big public clouds or in more traditional infrastructure. It can be operated efficiently for very large datasets, and, because it uses the existing standards for transmitting data, it can be readily integrated into current pipelines and analytical methods. Users can employ htsget to download only the subsection of a genome sequence in which they are interested rather than the whole file, or they can download the entire genome as a series of “data slices” distributed across multiple disparate machines.

Watch Video | Download Slides | Documentation

2. ENA/EGA/EVA

Thomas Keane (EMBL-EBI)

Housed at EMBL-EBI, the European Nucleotide Archive (ENA), European Genome-phenome Archive, and European Variant Archive provide the basic infrastructure to enable global public sharing of genetic data. Collectively, they form a single GA4GH Driver Project and each has implemented a subset of the existing GA4GH standards. In particular, EGA and RD-Connect (a global platform for rare-disease research, of which EGA is a partner) have implemented htsget, ENA has implemented refget, EGA has implemented the GA4GH Data Use Ontology and the ELIXIR Access and Authentication Infrastructure.

Watch Video | Download Slides

3. refget

Andy Yates (EMBL-EBI)

The GA4GH refget API enables access to reference genomic sequences, without ambiguity from different sequence providers, using a checksum identifier based on the sequence content itself. This enables reliable, reproducible genomic analysis with clear provenance back to reference sequence.

Watch Video | Download Slides | Documentation

4. Workflow Execution Service (WES)

David Glazer (Verily)

The GA4GH WES API lets users run a single workflow (defined using CWL or WDL) on multiple different platforms, clouds, and environments, and be confident that it will work the same way. The API provides methods to request that a workflow be run, pass parameters to that workflow, get information about running workflows, and cancel a running workflow.

Watch Video | Download Slides | Documentation

5. Variant Interpretation for Cancer Consortium (VICC)

Alex Wagner (Washington University)

VICC is developing methods and tools for interoperability between knowledgebases containing clinical interpretations of cancer variants. Its ultimate clinical focus is for

consensus expert curated interpretations to be integrated into clinical reports to improve consistency and outcomes in precision medicine.

Watch Video | Download Slides

6. Data Breach Reporting

Kate Birch (Australian Genomics)

The GA4GH Data Breach Response plan aims to ensure that GA4GH can respond in a timely and constructive way to breaches associated with GA4GH standards. Australian Genomics ran a data breach drill to test the efficacy of and improve on the GA4GH Data Breach Response protocol.

Watch Video | Download Slides

7. Regulatory and Ethics Work Stream

Madeleine Murtagh (Newcastle University)

The GA4GH REWS toolkit currently includes a consent policy template/clauses, consent codes, access and discovery matrix, registered access, privacy and security guidance, genomic data lexicon, accountability policy, and ethics review equivalency policy. In addition to developing a regulatory and ethics review process of all GA4GH technical deliverables, the group is working on a survey (Your DNA, Your Say) to probe public views of genomic data sharing internationally and hosting a forum to provide expert guidance on GDPR for health data sharing.

Watch Video | Download Slides

8. Phenopackets

Jules Jacobsen (Queen Mary University of London)

Phenopackets is an open standard for sharing disease and phenotype information by linking detailed phenotype descriptions with disease, patient, and genetic information. It enables clinicians, biologists, and disease and drug researchers to build more complete models of disease.

Watch Video | Download Slides

9. Scaling Genomic Data Knowledge

Heidi Rehm (Broad Institute of MIT and Harvard)

Standards and consensus on knowledge representation are needed to support the interpretation of variants and determine which genes are valid disease genes, ready for clinical testing. The GA4GH Genomic Knowledge Standards Work Stream (GKS) is developing a framework of standards-based components that lower barriers for the exchange of genomic information and its translation into clinical practice. GKS is focused on genetic variant representation and annotation as well as gene annotation.

Watch Video | Download Slides

10. Search API

Harindra Arachchi (Broad Institute of MIT and Harvard)

The GA4GH Search API provides a flexible, extendable interface to query a federated network of patient genomic databases.

Watch Video | Download Slides

11. Data Use Ontology (DUO)

Melanie Courtot (EMBL-EBI), Moran Cabili (Broad Institute of MIT and Harvard)

DUO allows data holders to semantically tag datasets with restrictions about their usage, making them automatically discoverable based on the intended usage. It enables machine readable descriptions of data access requests and data use restrictions to be matched, and would ultimately alleviate the need for manual review when datasets are requested by researchers. DUO is a product of the Data Use and Researcher Identity work stream and is available at https://github.com/EBISPOT/DUO.

Watch Video | Download Slides

12. Researcher Identities

Stephanie Dyke (McGill University), Ilia Tulchinksy (Google)

The GA4GH Researcher Identity and Bona Fide Researcher Claims Standard will allow researchers and other users to establish identity and credentials claims with regards to their professional identity to acquire access across datasets.

Watch Video | Download Slides

13. Beacon API

Marc Fiume (DNAstack)

The Beacon API is a web-accessible service that users may query for information about a specific allele. V1.0,0 extends Beacon's functionality through support for additional types of genomic variants and improved metadata support.

Watch Video | Download Slides | Documentation

Panel Discussion

Watch Video

Keynote Address: Responsible Data Sharing

Andrew Morris (HDRUK)

Health Data Research UK (HDRUK) seeks to create a thriving, high-energy UK-wide network of inter-disciplinary research expertise that will disrupt traditional science by enabling new scientific discovery from large multi-dimensional datasets, apply cutting-edge technologies to enhance science, innovation and decision making, and improve healthcare for a population of 65 million people.

Watch Video | Download Slides

Responsible Genomic Data Sharing: Right to Benefit

1. Your DNA, Your Say

Anna Middleton (Wellcome Genome Campus)

Created by the GA4GH Participant Values Task Team on behalf of the Regulatory and Ethics Work Stream, Your DNA, Your Say is a global online survey gathering public attitudes towards genomic data sharing. The survey contains nine innovative films that explain what genomic data sharing is through the eyes of a little boy called Charlie and have been translated into ten languages with more forthcoming.

Watch Video | Download Slides

2. A View from Molecular Oncology: ICGC-ARGO

Peter Lichter (German Cancer Research Center)

The Pan-cancer Analysis of Whole Genomes Project (PCAWG) was an international collaboration to identify common patterns of mutation in more than 2,800 cancer whole

genomes from the International Cancer Genome Consortium (ICGC). Now ICGC-ARGO (Accelerating Research in Genomic Oncology) aims to expand that work by sequencing samples from at least 100,000 cancer patients in the context of clinical studies over the next decade.

3. Implications for Responsible Genomic Data Sharing

Rumiana Yotova (University of Cambridge)

The right to benefit from science was first set out in the 1948 Universal Declaration of Human Rights and later affirmed in the International Covenant on Economic, Social and Cultural Rights. This right has important implications for responsible genomic data sharing, namely, it imposes an obligation on States to enable access to scientific knowledge and data freely and without discrimination, to protect the freedom of scientific research, as well as to ensure that the right to benefit from science is exercised in a manner respectful of human rights, including the right to privacy and to prior informed consent.

Watch Video | Download Slides

Panel Discussion

Chair: Madeleine Murtagh (Newcastle University)

Speakers in the Right to Benefit session discussed issues of research across national borders and linkages between science and the commercial sector.

Watch Video

Responsible Genomic Data Sharing: Privacy and Security

1. Data Protection for Personalized Health (DPPH)

Jean Pierre Hubaux (École Polytechnique Fédérale de Lausanne)

DPPH is a Swiss response to growing concerns about health data privacy. it is a project whose goals are highly aligned with those of GA4GH, dimensioned for the Swiss scale, but with an international vision. With participation from five research groups across the ETH Domain and the Swiss Data Sciences Center, the project will (i) address the main privacy, security, scalability, and ethical challenges of data sharing for enabling effective P4

medicine; (ii) define an optimal balance between usability, scalability and data protection, and (iii) deploy an appropriate set of computing tools.

Watch Video | Download Slides

2. Making New Treatments Possible: The GDPR Code of Conduct for Health Research Initiative

Michaela Mayrhofer (BBMRI-ERIC)

The GDPR Code of Conduct for Health Research Initiative aims to (i) contribute to the proper application of the GDPR in the area of health; (ii) clarify and specify certain rules of the GDPR for scientific research in the area of health; (iii) demonstrate GDPR compliance and (iii) foster transparency and trust in the use of personal data in the area of health research.

Watch Video | Download Slides

3. Systemic Oversight of Big Data Health Research

Alessandro Blasimme (ETH-Zurich)

The health data ecosystem is evolving in the era of genomics and digital health, resulting in emerging challenges in areas such as consent, IRB/REC, and data protection. Systemic oversight is needed to foster an ethical common ground for a variety of governance actors and activities and to ensure distributed accountability throughout the data ecosystem.

Watch Video | Download Slides

Panel Discussion

Chair: Paul Flicek (EMBL-EBI)

Speakers in the Privacy and Security session discussed data access conditions, the responsibility to share data for the benefit of the community, balancing data encryption and protection with utility, and the ethics of sharing un-consented data.

Watch Video

OCTOBER 5

The final day of the Ga4GH 6th Plenary Meeting focused on “going global,” with the announcement of an open call for new GA4GH Driver Project applications to increase the organization’s international representation. Keynotes from Torsten Schwede (SIB-Swiss Institute for Bioinformatics) and Nicola Mulder (University of Cape Town) introduced the community to activities ongoing in Switzerland and on the African continent. Panel discussions focused on the patient perspective, alignment across standards organizations, and national initiatives.

Opening Remarks

Ewan Birney (EMBL-EBI), Peter Goodhand (OICR)

GA4GH will accept applications for new Driver Projects until November 1, 2018 with a specific focus on global representation, scientific merit, and capacity to contribute. Up to five new projects will be announced in January 2019.

Watch Video | Download Slides

Keynote Addresses: Going Global

1. Swiss Personalized Health Network

Torsten Schwede (SIB-Swiss Institute for Bioinformatics)

SPHN aims to bring Switzerland to the forefront of personalized health research by establishing nationwide interoperability of biomedical information. It establishes a dynamic network of interoperable and sharable resources (data, platforms, workflows, competences) distributed across the cantons and supported by national coordination.

Watch Video | Download Slides

2. H3Africa

Nicola Mulder (University of Cape Town)

A representative dataset of African genomes is critical both to address diseases that disproportionately affect the African continent as well as to improve research and

healthcare globally. Human Heredity & Health in Africa (H3Africa) aims to facilitate an African-based research approach to the study of genomic and environmental determinants of common diseases with the goal of improving the health of African populations.

Watch Video | Download Slides

Panel: Patient Perspectives

Opening Remarks

Heidi Rehm (Broad Institute of MIT and Harvard)

Rehm highlighted stories of patients catalyzing research on their diseases, as well as a series of platforms to assist in this pursuit.

Watch Video | Download Slides

1. Inspire2Live

Peter Kapitein (Inspire2Live)

The international Inspire2Live Patient Advocates organization inspires people (patients, clinicians and researchers), without limitations, to work together to get cancer under control by 2021 and to lead happy and healthy lives in harmony with cancer. Our Patient Advocates are leading in this process.

Watch Video

2. BLACKSWAN Foundation

Olivier Menzel (BLACKSWAN Foundation)

The BLACKSWAN Foundation was established in Switzerland in 2010 to contribute to the development of research on rare and orphan diseases worldwide. Besides its funding research mission, the Foundation organizes the bi-annual RE(ACT) Congress, an international reference point for rare disease experts as well as an international advocacy and awareness campaign called #RAREvolution to ensure rare diseases are recognized as an international public health and research priority.

Watch Video | Download Slides

3. Count Me In: Partnering with Patients to Accelerate Cancer Research

Corrie Painter (Broad Institute of MIT and Harvard)

Count Me In aims to generate a massive publicly available database of clinical, genomic, molecular, and patient reported data in cancer to accelerate discoveries and the development of new treatment strategies. Engaging with patients to directly partner with them on cancer research, CMI has launched open clinical-genomic databases to improve research on three specific diseases: Metastatic Breast Cancer, Angiosarcoma, Metastatic Prostate Cancer, and Gastroesophageal Cancer.

Watch Video | Download Slides

Panel Discussion

Chair: Heidi Rehm (Broad Institute of MIT and Harvard)

Speakers in the Patient Perspectives session discussed expanding patient engagement to more countries and across more disease areas and using these efforts to recruit for clinical trials, returning results to patients, integrating environmental data into patient surveys, and the right to try medications still in trial phase.

Watch Video

BRCA Exchange: Data Sharing for the Benefit of Science and Medicine

The BRCA Exchange aims to advance our understanding of the genetic basis of breast cancer, ovarian cancer and other diseases by pooling data on BRCA1/2 genetic variants and corresponding clinical

Watch Video

Panel: Advancing Standards in Genomic and Health Data

1. HL7

Robert Freimuth (Mayo Clinic)

Health Level 7 (HL7) International develops standards for transfer of clinical and administrative data between software applications used by healthcare providers and clinical systems. The HL7 Clinical Genomics Work Group (CG WG) facilitates the development of standards for the exchange of clinical and translational information related to an individual's genomic data and family health history, and (its linkage to) relevant clinical information. The HL7 CG and the GA4GH Genomic Knowledge Standards Work Stream are coordinating on the development of compatible standards that will enable the translation of genomic data and knowledge from the research to the clinical domain.

Watch Video | Download Slides

2. SNOMED CT

Ian Green (SNOMED CT)

Owned and governed by 36 Member countries, SNOMED International (formerly IHTSDO) is a not-for-profit organization charged with setting global standards for health terminology, a critical element of safe and effective healthcare. SNOMED CT is an ontological resource with comprehensive, scientifically validated clinical content that enables consistent representation of clinical content for processing in electronic health records. It maps to other international standards and is used in more than fifty countries.

Watch Video | Download Slides

3. Clinical Data Interchange Standards Consortium

Peter Van Reusel (CDISC)

CDISC is Global Standards Development Organization (SDO) founded in 1997 by volunteers and incorporated as a non-profit organization in 2000. It aims to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.

Watch Video | Download Slides

4. ISO/TC 215 Genomics & NCI Terminology Standards for Health Informatics

Bron Kisler (NIH)

ISO Technical Committee 215 (TC215): Health Informatics was founded in 1998 to achieve standardization in the field of health informatics to facilitate the coherent and consistent interchange and use of health-related data, information, and knowledge to support and enable all aspects of the health system.

Watch Video | Download Slides

5. Industry Perspective: Roche Sequencing Solutions

Chris Bakan (Roche)

6. UK's 100,000 Genomes Project: Clinical and Research Uses

Augusto Rendon (Genomics England)

There is a gap between the number of sequences collected by Genomics England to date and the number of sequences that have been analysed. This gap is due in part to "fragile interoperability," wherein a specification is either insufficient or not enforced, making it difficult to share data across platforms and make it available for analysis.

Watch Video | Download Slides

Panel: National Initiatives

Opening Remarks

Kathryn North (Murdoch Childrens Research Institute)

With GA4GH Connect, the Global Alliance for Genomics and Health is bringing together leading genomic data initiatives as Driver Projects to help develop and pilot standards most needed by the community. Input from national genomic data initiatives around the globe are critical to this endeavour. Under the banner of GA4GH, Australian Genomics and Genomics England have already begun to share knowledge and resources and seek to include more initiatives in this endeavour to exchange learnings for the benefit of all.

Watch Video | Download Slides

1. Genomics England

Mark Caulfield (Genomics England)

The 100,000 Genomes Project is a UK Government project to sequence whole genomes through the National Health Service to enable scientific discovery and medical insights, with an initial focus on rare diseases and some common types of cancer. On October 2, 2018, the UK Government announced a plan to expand this project to collect 5 million genomes.

Watch Video | Download Slides

2. Australian Genomics

Kathryn North (Murdoch Childrens Research Institute)

Australian Genomics is a national research collaboration of clinicians, researchers and diagnostic geneticists working together to provide evidence for the equitable, effective and sustainable delivery of genomic medicine in healthcare. Australian Genomics unites 80 partner institutions across the country.

Watch Video | Download Slides

3. *All of Us*

Chris Lunt (NIH)

The *All of Us* Research Program is a historic, longitudinal effort to gather data from one million or more people living in the US to accelerate research and improve health. By taking into account individual differences in lifestyle, socioeconomics, environment, and biology, researchers will uncover paths toward delivering precision medicine.

Watch Video | Download Slides

4. Japan Agency for Medical Research and Development (AMED)

Hidewaki Nakagawa (RIKEN)

The Japan Agency for Medical Research and Development (AMED) promotes integrated research and development in the field of medicine, from basic research to clinical trials. AMED is currently exploring the possibility of a Japanese genome umbrella project that would bring together several large research projects from around the country, including the Tohoku Medical Megabank Project and the BioBank Japan Project.

Watch Video | Download Slides

5. Plan France Médecine Génomique 2025

Anne Cambon-Thomsen (CNRS, Inserm and University Toulouse III Paul Sabatier)

Under the leadership of the Alliance of French Institutions for Life Sciences and Health (AVIESAN), Plan France Médecine Génomique 2025 aims to integrate genomic sequencing into generic healthcare and develop a national genomic medicine sector by 2025. It has three major objectives: (i) set up the tools for a genomic healthcare pathway, (ii) ensure these developments take place in a safe technical and ethical framework, and (iii) implement monitoring and management tools.

Watch Video | Download Slides

6. The Finnish National Initiative

Ilka Lappalainen (ELIXIR-Finland)

With a population of 5.4 million, Finland provides national healthcare to each citizen. Over the past 15 years, healthcare funding has increased from 7% to 9.5% of national gross income. A number of national registries, biobanks and digital health information are already available and could be connected using a national security identifier. An effective national genome strategy would have a significant impact on healthcare, promote innovation, and make the data collected through healthcare data available for research in a secure manner.

Watch Video | Download Slides

Closing

Ewan Birney (EMBL-EBI)

Birney closed the 6th Plenary Meeting of the Global Alliance for Genomics and Health with an announcement of the next meeting, which will take place in the US from October 21 – 23, 2019.

"This is a community that I think can change the world and that I'm very proud to be a part of."

Watch Video

6th PLENARY PROGRAMME COMMITTEE

Thank you to the Programme Committee who made this meeting possible:

- Kathryn North (Murdoch Childrens Research Institute; Committee Chair)
- Dixie Baker (Martin, Blanck & Associates)
- Michael Baudis (University of Zurich)
- Robert Freimuth (Mayo Clinic)
- David Glazer (Verily)
- Peter Goodhand (Ontario Institute for Cancer Research)
- Madeleine Murtagh (Newcastle University)
- Gunnar Ratsch (ETH Zurich)
- Heidi Rehm (Broad Institute of MIT and Harvard)
- Torsten Schwede (SIB-Swiss Institute for Bioinformatics)
- Serena Scollen (ELIXIR)

FUNDING PARTNER RECOGNITION

Thank you to the GA4GH 2018 Funding Partners whose support made GA4GH 6th Plenary possible.

Host Institutions

Core Funders

Funding for this meeting was made possible in part by a supplement to U41HG006834 from the National Institutes for Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention by trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

CanSHARE

wellcome

6th Plenary Local Co-Hosts

elixir

20 YEARS SIB Swiss Institute of Bioinformatics

Gold-level Funding Partners



Silver-level Funding Partners





Bronze-level Funding Partners