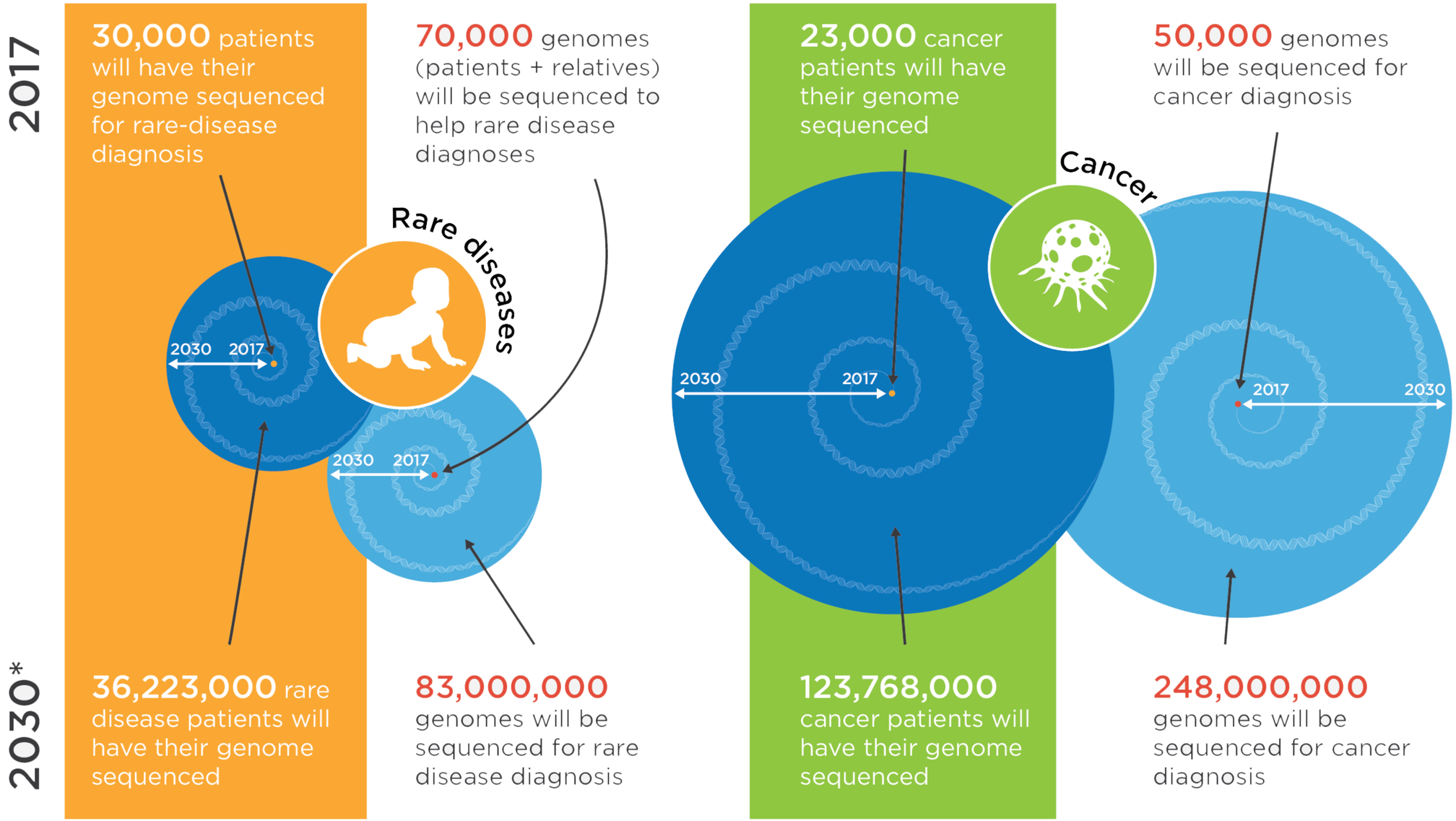
Beacon

Ethical & Legal Aspects of a Genomic Data Discovery Protocol

Michael Baudis | DSI Ethics | 2022-08-25

Global Alliance
for Genomics & Health



* Projected figures, based on current data and known status of genomics initiatives worldwide.

Limited Population Diversity in Cancer Studies

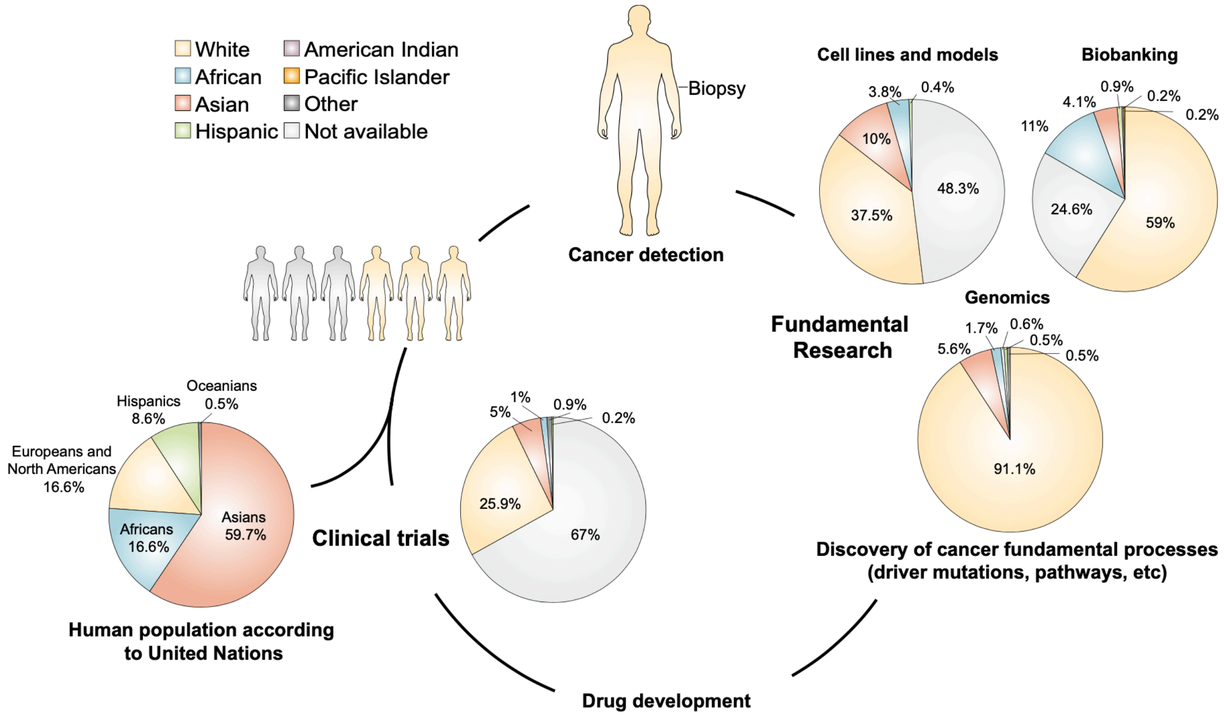


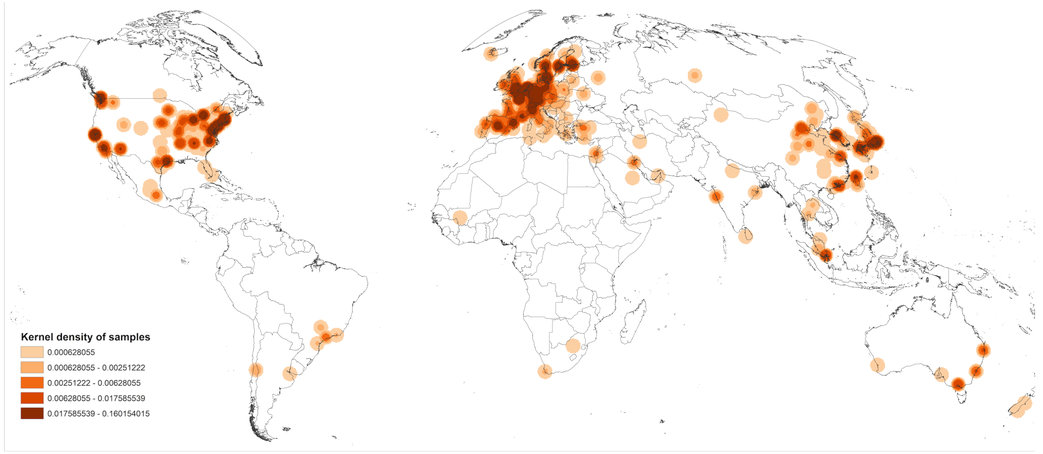
Figure 1. Racial/Ethnic disparities in cancer research. Racial/ethnic inclusion was studied in several aspects of oncological research, from cell lines and patient-derived xenografts to biobanking, genomics and clinical trials.

Guerrero S, López-Cortés A, Indacochea A, et al. Analysis of Racial/Ethnic Representation in Select Basic and Applied Cancer Research Studies. *Sci Rep*. 2018;8(1):13978.

Publication Landscape of Cancer CNV Profiling

Publication statistics for cancer genome screening studies. The graphic shows our as- sessment of publications reporting whole-genome screening of cancer samples, using molecular detection methods (chromosomal CGH, genomic array technologies, whole exome and genome sequencing).

For the years 1993-2018, we found 3'229 publications reporting 174'530 individual samples in single series from 1 to more than 1000 samples. Y-axis and size of the dots correspond to the sample number; the color codes indicate the technology used.



progenetix

The vision: Federation of data

GENOMICS

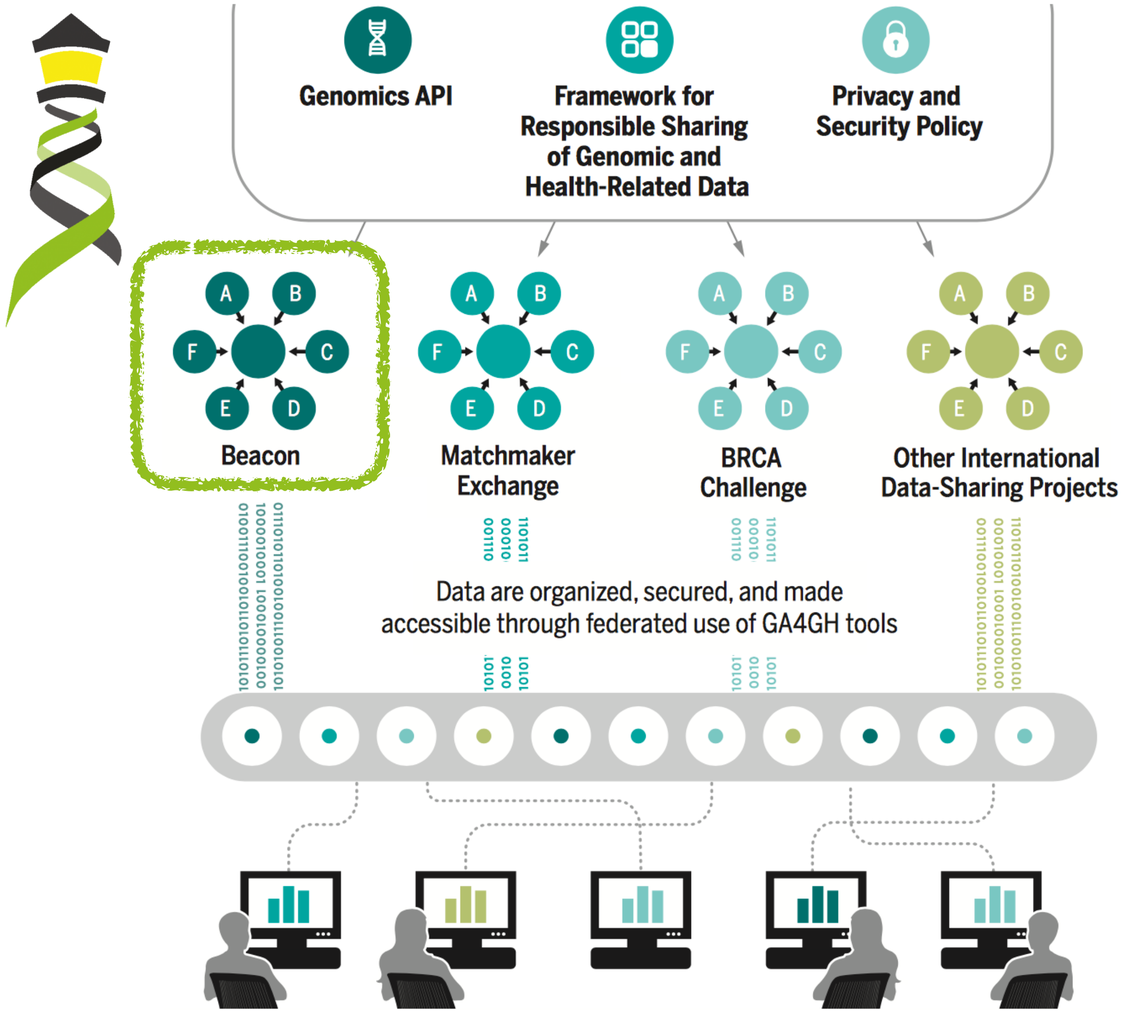
A federated ecosystem for sharing genomic, clinical data

Silos of genome data collection are being transformed into seamlessly connected, independent systems

The Global Alliance for Genomics and Health*

SCIENCE 10 JUNE 2016 • VOL 352 ISSUE 6291

A federated data ecosystem. To share genomic data globally, this approach furthers medical research without requiring compatible data sets or compromising patient identity.



DNASTACK

Global Alliance for Genomics & Health

The Global Alliance for Genomics and Health

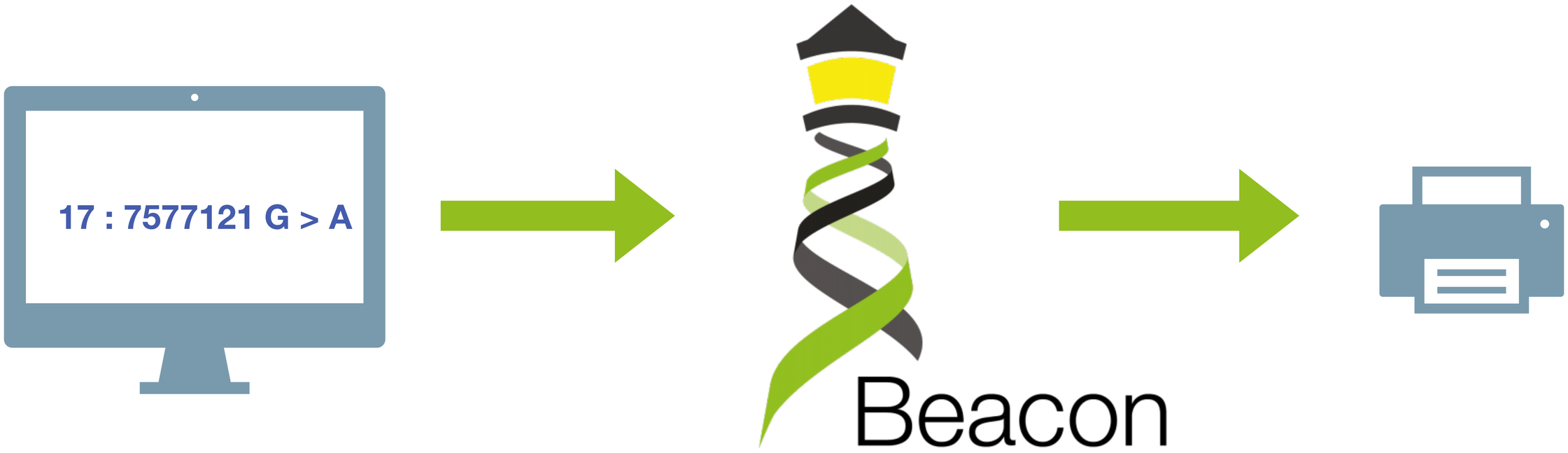
Making genomic data accessible for research and health

- January 2013 - 50 participants from eight countries
- June 2013 - White Paper, over next year signed by 70 “founding” member institutions (e.g. SIB, UZH)
- March 2014 - Working group meeting in Hinxton & 1st plenary in London
- October 2014 - Plenary meeting, San Diego; interaction with ASHG meeting
- June 2015 - 3rd Plenary meeting, Leiden
- September 2015 - GA4GH at ASHG, Baltimore
- October 2015 - DWG / New York Genome Centre
- April 2016 - Global Workshop @ ICHG 2016, Kyoto
- October 2016 - 4th Plenary Meeting, Vancouver
- May 2017 - Strategy retreat, Hinxton
- October 2017 - 5th plenary, Orlando
- May 2018 - Vancouver
- October 2018 - 6th plenary, Basel
- May 2019 - GA4GH Connect, Hinxton
- October 2019 - 7th Plenary, Boston
- October 2020 - Virtual Plenary, June 2021 - Virtual Connect ...
- October 2021 - Virtual Plenary ...
- September 2022 - 10th Plenary, Barcelona

22 SEPTEMBER 2022 | BARCELONA, SPAIN

GA4GH 10th Plenary

Global Alliance for Genomics & Health



A ***Beacon*** answers a query for a specific genome variant against individual or aggregate genome collections

YES | **NO** | **\0**

Have you seen this variant? It came up in my patient and we don't know if this is a common SNP or worth following up.

A Beacon network federates *genome variant queries* across databases that support the ***Beacon API***

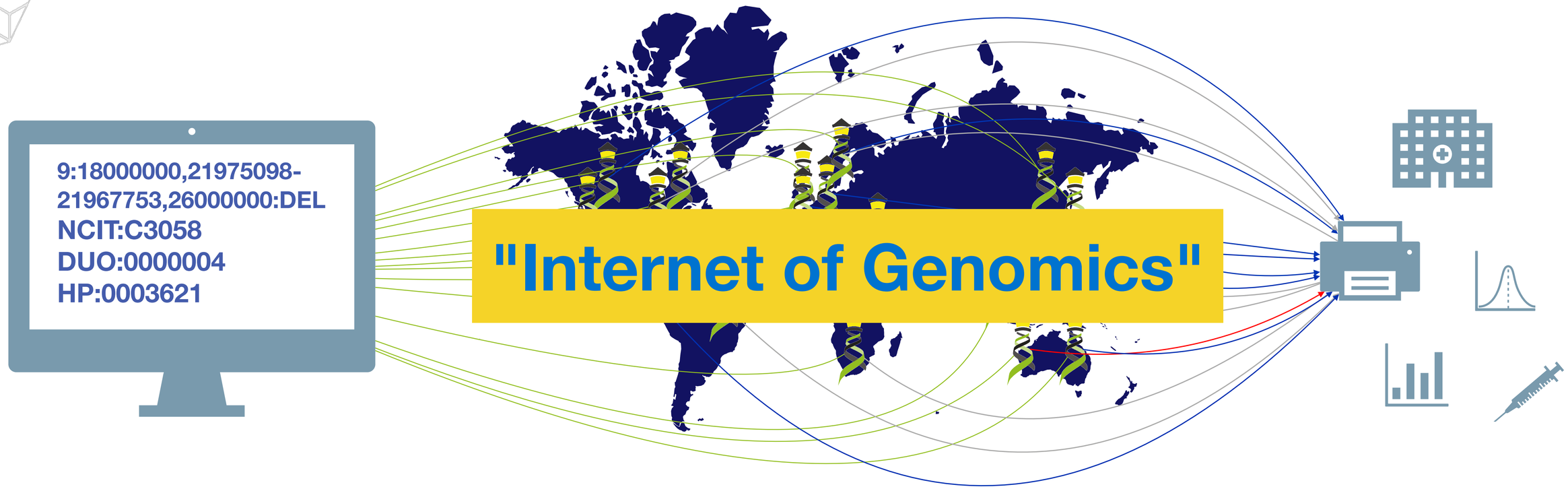
Here: The variant has been found in **few** resources, and those are from **disease** specific **collections**.

ELIXIR ALL HANDS 2022

Have you seen deletions in this region on chromosome 9 in Glioblastomas from a juvenile patient, in a dataset with unrestricted access?

Beacon *v2* API

The Beacon API v2 proposal opens the way for the design of a simple but powerful "**genomics API**".



Have you seen deletions in this region on chromosome 9 in Glioblastomas from a juvenile patient, in a dataset with unrestricted access?

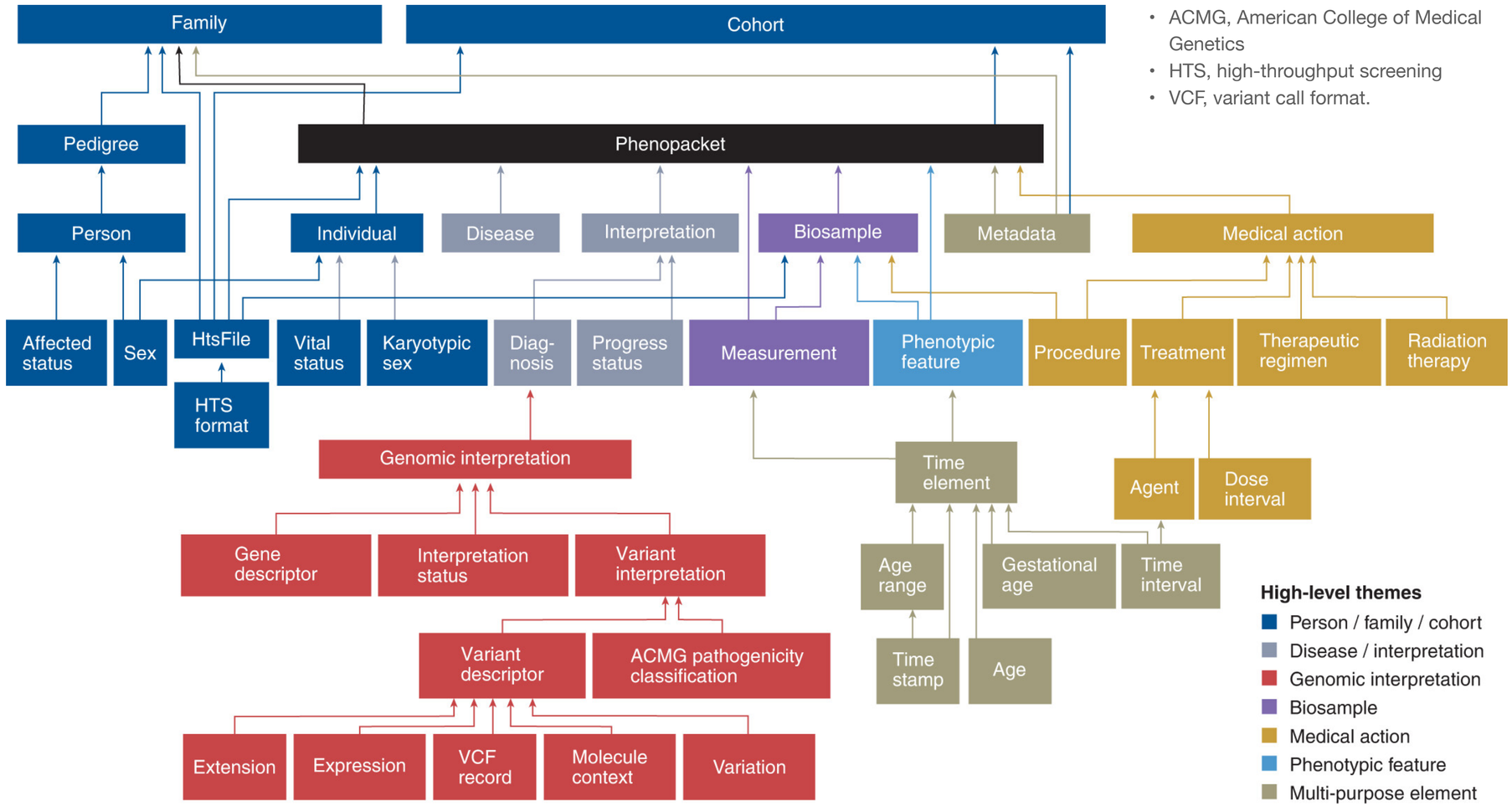
Beacon *v2* API

The Beacon API v2 proposal opens the way for the design of a simple but powerful "**genomics API**".

The GA4GH Phenopackets v2 Standard
A Computable Representation of Clinical Data

The GA4GH Phenopacket schema consists of several optional elements, each containing information about a certain topic, such as phenotype, variant or pedigree. An element can contain other elements, which allows a hierarchical representation of data.

For instance, Phenopacket contains elements of type *Individual*, *PhenotypicFeature*, *Biosample* and so on. Individual elements can therefore be regarded as **building blocks** of larger structures.



Jacobsen JOB, Baudis M, Baynam GS, Beckmann JS, Beltran S, Buske OJ, Callahan TJ, *et al.* 2022.
The GA4GH Phenopacket Schema Defines a Computable Representation of Clinical Data.
Nature Biotechnology 40 (6): 817–20.

Onboarding

Demonstrating Compliance

- Progenetix Beacon+ has served as implementation driver since 2016
- Beacon v2 as service with protocol-driven registries for federation
- GA4GH approved Beacon v2 in April 2022

Global Alliance for Genomics & Health

Genomic Data & Privacy - Key Areas

- **Re-identification**
  - identification of an individual based on sets of genomic variants they (or close relatives) carry - so one needs some genome data first
  - information to be gained is circumstantial (e.g. their genome is in a particular disease related dataset)
  - currently only risk with some practical use (e.g. **long-range familial attacks**)
- **Genotype-to-Phenotype (G2P) attacks**
  - determination of some disease risk or phenotypic features from a genome itself
  - needs access to genome data which is illegal in many jurisdictions (but technically more & more feasible)
  - real-world use cases are limited but abuse through wrong perception of utility
- **Genomic Determinism**
  - assignment of individual abilities and personal development trajectories from genomic profiling
  - topic of (some good, most bad) SciFi
  - but: **Wehret den Anfängen**!

Technical, legal and ethics aspects of genomic data sharing

DSI Proposal Summary

- The area of "personalized" or "precision" health relies on the use of molecular characteristics - e.g. inherited genomic variations or mutations in a tumor genome - to tailor individual recommendations in such diverse areas as nutrition, pharmaceutic selection and dosing or use of preventive screening, up to the tailoring of antineoplastic therapies based on the molecular profiling of a patients cancer cells. While it has been widely recognized that many of these applications can provide personal health benefits and have a positive impact on society at large, large questions remain regarding the secondary use of genomic data, the potential dangers arising from collection, storage and sharing of genomic information as well as the associated legal regulations.

- The need to have access to an ever increasing set of genomic data for biomedical research and treatment decisions is being addressed on an international level through such efforts as the Global Alliance for Genomics and Health **GA4GH** - of which **UZH is a founding member** - or European B1MG initiative. The Swiss Personalized Health Network **SPHN** provides the technical and and logistic backbone for the exchange of genomic and clinically derived information for research applications and clinical re-use, while not directly generating data itself.

- Issues related to the generation, handling and application of personal genomic information span a large set of academic specialities, from **medicine**, **life science** research, **computational science** and **data security** to **ethics**, **law**, **psychology** as well as **communication** and **media sciences**. The topic of "**Technical, legal and ethics aspects of genomic data sharing**" is a natural fit for the Digital Society Initiative, offering participation and networking opportunities for a wide range of potential stakeholders from different faculties to develop projects in an area with **scientific relevance** and large **public interest**.

Technical, legal and ethics aspects of genomic data sharing

DSI Proposal Summary

- The area of "personalized" or "precision" health relies on the use of molecular characteristics - e.g. inherited genomic variations or mutations in a tumor genome - to tailor individual recommendations in such diverse areas as nutrition, pharmaceutic selection and dosing or use of preventive screening, up to the tailoring of antineoplastic therapies based on the molecular profiling of a patients cancer cells. While it has been widely recognized that many of these applications can provide personal health benefits and have a positive impact on society at large, large questions remain regarding the secondary use of genomic data, the potential dangers arising from collection, storage and sharing of genomic information as well as the associated legal regulations.

- The need to have access to an ever increasing set of genomic data for biomedical research and treatment decisions is being addressed on an international level through such efforts as the Global Alliance for Genomics and Health **GA4GH** - of which **UZH is a founding member** - or European B1MG initiative. The Swiss Personalized Health Network **SPHN** provides the technical and organizational backbone for the exchange of genomic and clinically derived information for research applications and clinical re-use, while not directly generating data itself.

- Issues related to the generation, handling and application of personal genomic information span a large set of academic specialties, from **medicine**, **life science** research, **computational science** and **data security** to **ethics**, **law**, **psychology** as well as **communication** and **media sciences**. The topic of "**Technical, legal and ethics aspects of genomic data sharing**" is a natural fit for the Digital Society Initiative, offering participation and networking opportunities for a wide range of potential stakeholders from different faculties to develop projects in an area with **scientific relevance** and large **public interest**.

Genomic Data & Privacy - Key Areas

- **Re-identification**
  - identification of an individual based on sets of genomic variants they (or close relatives) carry - so one needs some genome data first
  - information to be gained is circumstantial (e.g. their genome is in a particular disease related dataset)
  - currently only risk with some practical use (e.g. **long-range familial attacks**)
- **Genotype-to-Phenotype (G2P) attacks**
  - determination of some disease risk or phenotypic features from a genome itself
  - needs access to genome data which is illegal in many jurisdictions (but technically more & more feasible)
  - real-world use cases are limited but abuse through wrong perception of utility
- **Genomic Determinism**
  - assignment of individual abilities and personal development trajectories from genomic profiling
  - topic of (some good, most bad) SciFi
  - but: **Wehret den Anfängen**!

Ethical & Legal Aspects of Genomics Data Sharing

... with special consideration of the Beacon protocol in the Swiss context

- Genomic / "-omics" data are integral part of biomedical research projects and clinical procedures (e.g personalized cancer therapies, rare diseases...)
- Traditional "data access through research agreement" approaches do not accommodate **federated discovery** & analysis approaches needed in many genomics approaches (population variability, rare variants ...)
- The GA4GH Beacon v2 protocol allows genomic variant & biomedical data discovery on varying granularity (from "Boolean" to "document)
- Beacon is being implemented and considered in many international projects and resources, including Swiss ones (e.g. SPHN & UZH projects)
- What specific regulations & considerations for the different types of data & levels of granularity in the **Swiss** context?

Ethical & Legal Aspects of Genomics Data Sharing

... with special consideration of the Beacon protocol in the Swiss context

- Genomic / "-omics" data are integral part of biomedical research projects and clinical procedures (e.g. personalized cancer therapies, rare diseases...)
- Traditional "data access through research agreement" approaches do not accommodate federated discovery & analysis approaches needed in many genomics approaches (population variability, rare variants ...)
- The GA4GH Beacon v2 protocol allows genomic variant & biomedical data discovery on varying granularity (from "Boolean" to "document")
- Beacon is being implemented and considered in many international projects and resources, including Swiss ones (e.g. SPHN & IBD projects)
- What specific regulations & considerations for the different types of data & levels of granularity in the Swiss context?

Getting the discourse going ...

... and feed back to stakeholders

John Yuyi, NYT 2018-02-09

Welcome to *openSNP*

openSNP **lets customers of direct-to-customer genetic tests publish their test results, find others with similar genetic**

For Genotyping Users

Upload Your Genotyping File

Upload your raw genotyping

Phenotypes are the

openSNP gets the latest

Home | Family tree | Discoveries | DNA | Research

MyHeritage DNA

Valentine's Day
DNA SALE

Only **59€** ~~89€~~ per kit
When ordering 2+ kits

Order now

Shipping not included
Ends February 14th

23andMe

Welcome to you

saliva collection kit

Find out what your DNA says about you and your family.

- See how your DNA breaks out across 31 populations worldwide
- Discover DNA relatives from around the

ancestry

SUBSCRIBE SIGN IN

THE AVERAGE BRITISH PERSON'S DNA IS ONLY 36% BRITISH

GROW YOUR TREE

Find your ancestors in

ancestryDNA

Discover