Beyond One Million Genomes

The Genome of Europe (GoE)

Realising a population genomic reference cohort of at least 500,000 citizens across Europe by 2022

Beyond One Million Genomes
B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

Table of Contents

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

Summary

- As part of the European 1+ Million Genomes (1+MG) initiative, targeting the creation of an overall cohort of at least 1 million citizens and patients of which the genome is determined by 2022, we here propose to establish The Genome of Europe (GoE) Project that will contribute at least 500.000 genomes to the 1+MG ambitions.
- While 1+MG targets both clinical data derived from patients and data collected from the general citizen population (not specifically targeting diseased individuals), the GoE project will specifically target the latter and help create an important subset of the collection aimed for by 1+MG.
- The Genome of Europe multi-country project brings together European countries to build a high-quality European network of national genomic reference cohorts of at least 500.000 citizens, selected to be representative for the European population by 2022.
- All countries involved generate a national genomic reference dataset via Whole Genome Sequencing and establish a (Genome of [Country] collection) based on their own national population genomic reference cohort, all according to jointly established '1+MG-proof'[1] guidelines.
- Each national dataset will in its own right form a unique national reference collection that will benefit national personalised healthcare and prevention strategies. Collectively, crosslinked via the 1+MG initiative, the national collections will establish a world-class European reference data resource (The Genome of Europe) for research and innovation of healthcare.
- To drive this crucial European digital and personalised health project, we propose countries to make their national (Genome of [country]) project part of the prioritisation process for the **European Recovery and Resilience Facility** (RRF[2]).
- Next to a description of the Genome of Europe overall project and its relation to the 1+MG initiative, we offer a country-level description of the project (Annex 1). Countries may choose to use this in their national preparation/fund-raising process.

[1] '1+MG-proof' here means: Compliant to 1+MG ELSI framework (incl. consent), standards & minimal clinical phenotypes, data quality and interoperability, connected national infrastructure.

[2] See document attached distributed through DG-CNECT as part of a Jan 2021 workshop on RRF Multi-Country Projects

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

Introduction

Building on the European 1+ Million Genomes (1+MG) initiative, a multinational driving project is being prepared to develop a reference genome cohort of at least 500.000 citizens representative for the European population. **The Genome of Europe (GoE) multilateral action plan** is set up in a coordinated effort to 1) establish comparable reference genome collections built on national population cohorts of citizens in participating countries, and 2) to connect these data into a European cohort, thus establishing a collective European reference dataset. The project will boost 1+MG-proof genomics data collection across European countries. Therefore, it will contribute to realising the ambitions set out in the 1+ Million Genomes declaration, signed by 23 European countries so far and firmly supported by the European Commission.

The 1+MG initiative is preparing for a collective 'trust framework' to enable effective and secure access to genomic data repositories. However, the recent 1+MG survey shows that it is difficult to make datasets or sequences that have already been collected in (institutional) databases accessible, according to this 1+MG trust framework, e.g. due to lacking proper citizen consent, or not meeting data quality standards. This frustrates effective clinical care and hampers the implementation of a genomic medicine and health approach in national health care and public health systems.

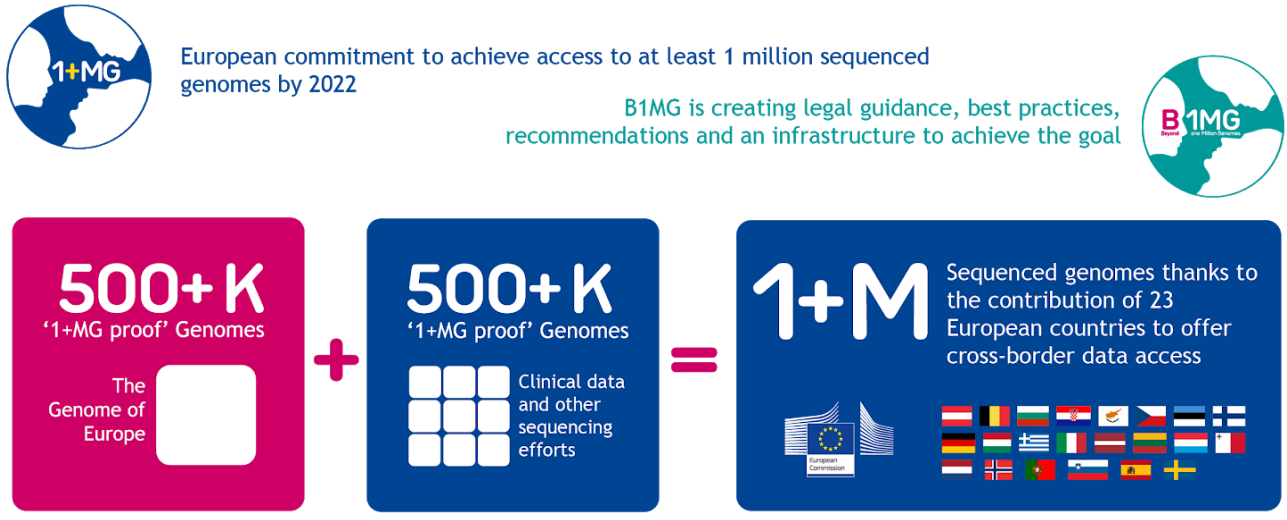

Figure 1. GoE reference collection in the context of 1+MG

With two years to go to reach the 1+MG goal of a collective and accessible reference cohort of at least one million European citizens that have their genome determined, prospective, new data generation must be planned and completed according to the 1+MG trust framework as soon as

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

possible. The GoE is set up to drive such prospective genome data generation across the 1+MG coalition. It is not expected to realise the full million genomes aspired by 1+MG, but will make a substantial contribution by adding to the already existing data a dedicated cohort of at least 500.000 citizens, selected to represent the general European population as a reference cohort (see Figure 1). As such, the GoE project perfectly fits the roadmap adopted by the 1+MG signatory countries, and it will closely liaise with the B1MG Project funded by the European Commission mid-2020 to set the guidelines and recommendations for '1+MG proof' data collection. It will incentivise the development and implementation of national genomic medicine & health strategies. Moreover, the project will help build a strong expertise base and public trust base in participating countries, and it will connect digital health infrastructures across Europe.

Part of the European 1+ Million Genomes (1+MG) initiative, supported by the B1MG project

The European 1+ Million Genomes declaration has been signed by 23 countries so far and the 1+MG initiative is firmly supported by the European Commission. Building the 1+MG cohort of at least one million European citizens that have their genome determined, via Whole Genome Sequencing (WGS), will be a solid basis for the development of personalised medicine and health approaches. Researchers and clinicians will be able to analyse and compare people's genetic and clinical information. This will help them detect illnesses earlier, predict the development of the disease, and decide on the best ways to improve health. Improvements in personalised medicine and health will make national health systems more efficient and cost-effective. New insights into how genes affect susceptibility to disease or a person's response to a drug will stimulate innovation and new products across the healthcare industry. 1+MG will thus benefit citizens in all countries through the development of more targeted personalised health-support, in terms of personal prevention, diagnosis, medication and/or healthcare treatment. It will also realise an independent position for Europe compared to other similar efforts elsewhere in the world.

The goal of the 1+MG initiative is to make the personal genomic datasets accessible in a secure manner for collective diagnostic purposes and prevention, and for research and innovation. To this end, a 1+MG trust framework will enable the effective and secure cross-border access to repositories of personal genomic datasets among participating countries. This trust framework covers collective agreements on ELSI aspects such as GDPR-compliance and personal consenting by participating citizens, what standards and minimal clinical phenotype to include, and quality criteria to perform genome sequencing. Besides, a federated data infrastructure is being designed to safeguard that national genomic data collections will be stored safely in the respective countries, and will be made accessible for collective analysis across borders rather than being collected in a central European repository.

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

The 1+MG initiative will give a boost to digital innovation of healthcare and help prepare countries to create a European Health Data Space. The European Commission states in its Communication COM(2018)233 on the digital transformation of health and care: *"We need to better coordinate these existing initiatives to reach the necessary critical mass at EU level and match similar initiatives in other world regions. Significant breakthroughs can be achieved by linking Europe's fragmented resources through secure cross-border digital infrastructures while ensuring full compliance with data protection legislation and ethical principles. Ensuring interoperable standards for genomic and other data is also critical for an effective sharing of datasets."*

Starting from a voluntary coordination mechanism, the representatives of the signatory countries now constitute a European Commission Special Group (1+MG Group) to oversee the implementation of a collective Roadmap for 1+MG. This Roadmap has been developed and adopted by the 1+MG signatory countries to realise the ambitions set out in the 1+MG Declaration. Under the governance of the 1+MG Group, and actively supported by the European Commission (DG-CNECT, DG-SANTE, DG-RTD), 11 international expert Working Groups are active in defining the requirements and above-mentioned trust framework for '1+MG proof' collection and access to genomic datasets. Four disease areas have been selected as exemplary 1+MG use cases: Rare Diseases, Cancer, Common & Complex Diseases (incl. prevention and pharmacogenomics) and Infectious Diseases (starting with Covid-19).

Since June 2020, the European Commission has provided funding to establish the Beyond 1 Million Genomes (B1MG) project as a Coordination and Support Action to strengthen the 1+MG initiative. Coordinated by the European life science data infrastructure ELIXIR, the B1MG project assembles a consortium of expert organisations across Europe to support the 1+MG Working Groups and help reach the 1+MG targets. Along with the 1+MG ambitions, the B1MG project is helping to create a network of genetic and clinical data collections across Europe. It has now taken the initiative to help set up the GoE project as part of the overall 1+MG initiative. While 1+MG targets both clinical data derived from patients as well as data collected from the general citizen population, the GoE project will specifically target the latter and help create an important subset of the collection aimed for by 1+MG.

A co-funding strategy is needed to realise GoE

The Genome of Europe as a multi-country project can of course only be realised via a multilateral approach and co-funding strategy. Within the 1+MG framework, a feasible funding strategy will need to be worked out. A range of major stakeholder organisations at European and national level will be involved to provide the necessary funding, expertise and capacity. Typical stakeholders/beneficiaries to be involved at national level will be public health authorities, biobanks, research institutes and agencies, hospitals, medical laboratories and biotech companies (including SMEs).

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

While the European Commission could help to plan and coordinate the data collection up to the standards defined in the 1+MG policy initiative, the signatory countries will need to find a mechanism to realise their own national (Genome of [Country]) cohort and data generation projects. NB: If prioritised on a national level, EU funds that will be made available through the Recovery and Resilience Facility (RRF) might complement available national funds. Upon request, the 1+MG Coordination Group can provide a template that countries may choose to use and adapt as part of their national prioritisation process. The GoE project has already been presented by the European Commission (CNECT) to the national RRF coordinators as a strong example of a Multi-Country Project.

Liaising with 1+MG national representatives, the B1MG project team will explore further possibilities to establish a co-funding strategy for GoE, also looking at other European funding instruments such as Horizon Europe/Digital Europe/EU4Health.

The Genome of Europe: general approach and planning

The goal of GoE is to build a collective reference genome cohort (GoE Cohort) of European citizens, selected (on a national basis) to mirror the genetic composition of the European population. Based on their voluntary participation, selected citizens are characterised by determining their genome (via whole-genome sequencing (WGS)). In addition, a specified set of associated health-related information will be collected, together hereafter termed 'personal genomic dataset'.

All countries participating in GoE build their own national reference collection of personal genomic datasets (Go[Country]). Each country establishes a population cohort that reflects the genetic composition of their national population per country, including both healthy and diseased individuals. In many countries, regional or national population cohorts and biobanks will offer an excellent basis for their national reference genome cohort because of the immediate availability of DNA samples and (longitudinal) health-related information built up over the years through extensive questionnaires and/or further deep health profiling data.

GoE targets a European reference collection of at least 500.000 European citizens. Population genomic analysis will help determine the recommended size and composition of the national cohort per country. Provided active personal consent of participating citizens is given, the generated personal genomic datasets will be made accessible across borders in a secure and GDPR-compliant manner for collaborative innovation of healthcare across participating countries in Europe. The 1+MG initiative offers the guidelines and the trust framework to make this effective. To make national cohorts comparable, collection of personal genomic datasets will be done in a '1+MG-proof' manner, i.e. compliant to the 1+MG trust framework: incl. ELSI aspects (explicit personal consent), data standards & minimal clinical phenotypes, data quality and

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

interoperability, and the requirement for a federation of securely connected national data infrastructures.

Together, this establishes a unique, specific health data resource for the population of European citizens. The GoE cohort will be of high value to all participating countries as a European reference resource for collective research, diagnosis and innovation of healthcare and prevention. GoE will incentivise the development and implementation of national genomic medicine & health strategies. Moreover, it will help build a strong expertise base and broad public trust base for genome-based health approaches in participating countries, and connect digital health infrastructures across Europe.

GoE outline and planning

The GoE focuses on prospective (new) generation of personal genomics datasets of a highly comparable nature to provide a high-quality European network of national population genomic reference collections. All countries involved thus build their own national collection according to jointly established guidelines. Each national genomic dataset will in its own right form a unique national reference dataset that will immediately benefit national personalised healthcare and prevention strategies. Collectively, crosslinked via the 1+MG initiative, the national collections will establish a world-class European reference data resource.

To build the most effective national (and European) genome data collection that will both serve current clinical care as well as boost research into prevention strategies in the shortest possible term, several aspects need to be considered. These include usefulness as a reference dataset for all/most clinical disciplines and prevention strategies, representation of minority populations with a migration background, rich and standardised phenotype information; immediate availability of (material for) DNA (extraction) and phenotype information, preferably in bulk for parallel processing. At the same time, we want to promote the implementation of whole-genome sequencing (WGS) in clinical genetic and cancer diagnostic practice where this is of added value above current diagnostics. Such clinical implementation trajectories will result in steadily growing datasets of individual patients with WGS data over time, which can be added to the national reference dataset, but as a special category because of their biased (i.e. specific disease-related) background.

The GoE initiative, therefore, targets genomic data collection from regional or national population cohorts and biobanks that have been established in most countries. These cohorts and biobanks are often built up in a research setting and/or included in national public health strategies. The availability of such collections is crucial for an effective and rapid realisation of GoE sequencing. Crucially, these cohorts and biobanks represent the broader healthy population of a country, manifold with a broad consent in place to share data for research and innovation of healthcare. In addition, a deep health and lifestyle data profile is often available for the citizens involved. Building a GoE-national reference genome dataset offers the basis to build

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

country-specific health risk profiles for national populations using Polygenic Risk Score (PRS-)based analyses. The national collections thus constitute an immediate national asset to build national prevention programmes for common diseases such as diabetes, Alzheimer, osteoarthritis and cancer.

In several countries, Whole Genome Sequencing (WGS) or Whole Exome Sequencing (WES) is gradually being adopted as part of standard clinical care in certain disciplines, e.g. for patients in oncology, and rare diseases. In these cases, WGS or WES data generation is already being funded as part of the regular healthcare reimbursement scheme. As described above, countries may choose to add personal genomics datasets generated in such clinical/healthcare settings as a special category to their national reference dataset. Crucial to the inclusion in the 1+MG initiative will be that also these datasets are gathered according to the 1+MG trust framework.

GoE planning: 2021 - 2022

The general scheme of GoE is as follows:

- 2021: Preparation phase at European and country-level, including securing of budgets. Countries prepare national data generation projects, setting up national partnerships of public and private organisations to be involved, preparing the selection of citizens to be included, including GDPR-proof consent for using data across borders, and setting up the logistics for sampling and data management. In parallel, 1+MG Working Groups define requirements and guidelines at ELSI, standards, quality and technical/infrastructural level as the 1+MG trust framework. The required multilateral (co-)funding strategy is worked out and budgets for 1+MG proof data generation are secured, possibly involving European RRF-based funding to match national funding. At country-level, required investments are defined (capital investment required for buying technology, hiring service providers or paying for staff costs). Collective negotiations with sequencing service providers may lead to gains in cost efficiency.

- 2022: Data collection phase. Execution of [Go-country]-level data generation projects, performing Whole Genome Sequencing of selected national population cohorts. Data generation and data stewardship are performed according to 1+MG standards. National infrastructure nodes are being established and prepared for secure cross-linking to the other national GoE repositories set up in the 1+MG framework. By the end of 2022, the GoE cohort of personal genomic datasets of at least 500.000 citizens is realised. In parallel, research and healthcare innovation consortia are being prepared/established to help analyse the GoE Cohort data, and design subsequent health care and research programmes which make use of such reference genomic datasets.

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724

GoE coordination: 1+MG Coordination Group, supported by the B1MG project team

The 1+MG country representatives have mandated the B1MG project team to look at the long term sustainability of the 1+MG effort. Within that remit, the B1MG team coordinated by ELIXIR will work in close collaboration with the 1+MG Coordination Group and the 1+MG National Mirror Groups to maintain an updated oversight of the GoE development across countries. B1MG will promote sharing the lessons learnt and best practices that could accelerate the project's outcomes. This activity will contribute to the B1MG long term sustainability deliverable.

Recovery & Resilience Facility (RRF)

This proposal clearly fits within the granting framework of the European Recovery & Resilience Facility (RRF):

- Building a national strategy to be developed around genomic medicine & health, driven by building a Go-[country] dataset and in close coordination with the other 1+MG signatory states, meets the country-specific RRF recommendations in the fields of digital and health policies by boosting and linking genomic data resources, developing person-centred care, improvement of public health and strengthening national health care systems. The initiative provides the basis for future accessibility of genomic medicine & health for every citizen.

- Economy of scale: development of the national data infrastructure for this data, in connection with the European 1+ Million Genomes initiative, is a direct investment in the digital transformation of health: this initiative is a crucial use case for digital innovation around privacy-sensitive health data, contributing to the European Health Data Space.

- Economic resilience and preparedness for future pandemics: active participation in the 1+MG initiative prepares our national healthcare system for future healthcare crises by building on the European digital health infrastructure and establishing the knowledgebase for national health-susceptibilities in relation to future pandemics.

- Investments in the life science sector place our country in a strong international position and will attract innovative companies and talent, as already demonstrated in the United Kingdom and Finland, among others. The investments are of a non-structural nature and can be implemented during the term of the RRF (2021-2024), in particular in 2021-2022.

B1MG has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 951724