



HAL
open science

PrgmNr 2515 -Sequencing the general population at a fine-geographic scale: the French POPGEN project View session detail

Anthony Francis Herzig, Gaëlle Le Folgoc, Marcel Goldberg, Marie Zins,
Emmanuelle Génin

► To cite this version:

Anthony Francis Herzig, Gaëlle Le Folgoc, Marcel Goldberg, Marie Zins, Emmanuelle Génin. PrgmNr 2515 -Sequencing the general population at a fine-geographic scale: the French POPGEN project View session detail. American Society of Human Genetics (ASHG), Oct 2021, Virtual congress, France. <hal-05316495>

HAL Id: hal-05316495

<https://hal.science/hal-05316495v1>

Submitted on 15 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

PrgmNr 2515 - Sequencing the general population at a fine-geographic scale: the French POPGEN project

[View session detail](#)

Author Block: A. Herzig¹, G. Le Folgoc¹, M. Goldberg², M. Zins², E. GÃ©nin³; ¹Inserm, Univ Brest, EFS, UMR 1078, GGB, Brest, France, ²Inserm-Paris Saclay Univ. UMS 011, UniversitÃ© de Paris, Villejuif, France, ³Inserm UMR1078, Brest, France

Disclosure Block: A. Herzig: None.

Population-level genome-sequencing projects have been undertaken by many nations. The availability of these, often immense, datasets has accelerated methodological development and provided new avenues for genetic epidemiological investigation. France will soon join the growing list of nations where such datasets have been compiled through the POPGEN project; a pilot study of the National Genomic plan led by Inserm. Whilst the envisaged dataset may not rank highest in terms of size, it promises to break new ground with its meticulous selection criteria.

Using questionnaires completed by volunteers from the Constances cohort, we selected participants based on the birth-places of their four grand-parents. As participants are adults over 35 years of age (mean of 59, range of 36-80), the average birth year of their grandparents was 1906. Our key assumption was that an individual with all four grand-parents born within a small locality should be a reliable representative of that region. By using such a selection criterion, we will obtain insights on the genetic diversity between French regions at the start of the 20th century when population movement remained limited. We selected 15,000 individuals respecting a geographic distribution based on historical population proportions in France from the 1901 census. Hence, POPGEN will capture a complete genetic cross-section of the French population. The connection between each participant and a geographical origin will permit a high resolution for analyses of genetic fine-structure.

The selected individuals were sent saliva collection kits through the post, a design aimed to best avoid recruitment bias and maximize participation. A pilot study successfully showed the appropriateness of this DNA collection method for whole-genome sequencing (WGS) as well as the possibility for parallel sequencing of salivary microbiomes. Using a smaller but similarly selected group of 856 French individuals with WGS data, we established the potential for using POPGEN as a population specific reference panel for imputation. Finally, we present analyses of demographic patterns observable through the multi-generational birth-place data and how we predict these could manifest in the genetic dataset that will materialize in 2022.