

Herzig A, Le Folgoc G, Goldberg M, Zins M, Génin E

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

American Society of Human Genetics, October 18-22, 2021, Virtual congress

ABSTRACT

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.

KEYWORDS: -

FOR OTHER INFORMATION, CLICK [HERE](#)