The FinnGen research project: Combining genomics and health record data to understand disease mechanisms

Aarno Palotie, MD, PhD, is the research director of the Human Genomics Program at the Institute for Molecular Medicine Finland (FIMM). He is a faculty member at the Center for Human Genome Research at Massachusetts General Hospital in Boston and an associate member of the Broad Institute of MIT and Harvard. He is also the chief scientific officer of the large FinnGen project that collects genome and health record data from 500,000 Finnish participants.

Aarno Palotie, MD, PhD

He has a long track record in human disease genetics, having held or currently holding professorships and group leader positions at the University of Helsinki, UCLA, Wellcome Sanger Institute, the Broad Institute of MIT and Harvard, and Massachusetts General Hospital. He has also been the director of the Finnish Genome Center and Laboratory of Molecular Genetics in the Helsinki University Hospital. He has served on numerous national and international boards, including the FIMM board.

Dr. Palotie has extensive experience in establishing, running, and overseeing infrastructures in both research and clinical settings. In addition to running clinical laboratories and the Finnish Genome Center, he served as the director of Medical Sequencing and was a member of the sequencing committee at the Sanger Institute. He also established and ran the tissue array unit at UCLA and has been a key player in planning the National Genome Strategy and national biobanking strategies in Finland. He has published over 450 original publications, reviews, and book chapters.

Thermo Fisher Scientific: What is your research goal?

Dr. Palotie: Our research goal is to improve our understanding of disease mechanisms, specifically in chronic diseases that are affecting a large part of the population, because that's the clue toward new therapies; we need to understand more about the disease. And how we do that is we use special populations like the Finnish population, which has a special population history, providing certain benefits. We use electronic health record systems, which are specific for the Nordic countries, that record them from birth to death. And we use large data sets, building large biobanks, combining health data and genome data to improve this knowledge of the disease background.

Thermo Fisher: Tell us a little bit about the FinnGen research project. What are the challenges? Accomplishments?

Palotie: The FinnGen research project is a project where we want to combine the electronic health record data of 500,000 individuals with their genome data. The background is related to the Finnish population history, which has evolved from a small founder population and is enriched for low-frequency and rare genetic variants, in such a way that we have the statistical power to analyze them. What's unique for the Nordic countries is that we

> "The Axiom array was chosen because of its flexibility, the reputation of its quality, and its use in other large-scale projects."

ThermoFisher SCIENTIFIC

applied
biosystems

have health record information from birth to death, and this is something that provides a special opportunity for a large biobank study like this.

> "What's unique for the Nordic countries is that we have health record information from birth to death, and this is something that provides a special opportunity for a large biobank study like this."

Thermo Fisher: What is the state of precision medicine, and what does it mean to you?

Palotie: I think that precision medicine as a term is a hype. It's used in very many contexts, sometimes appropriately and sometimes just to promote our own opportunities and maybe funding resources. For me, it means that we try to develop the means and tools to treat patients; to use the information in a more individualized manner.

Thermo Fisher: How does Thermo Fisher Scientific fit into your workflow?

Palotie: When we were thinking of the design on the FinnGen project, we were looking for an option where we would have an efficient, streamlined genotyping capability. Thermo Fisher was actually the only company that provided us a trustworthy, industry-level throughput or pipeline, so that we could just provide our samples to Thermo Fisher and then receive high-quality genotypes.

> "[Thermo Fisher] provided us a trustworthy, industry-level throughput or pipeline, so that we could just provide our samples … then receive high-quality genotypes."

Thermo Fisher: How did you come to decide on using microarrays, and more specifically, on the Applied Biosystems™ Axiom™ platform?

Palotie: For the FinnGen project we were thinking about how genotype data is produced. For array design, the genome-wide association study (GWAS) design was an obvious option. The Axiom array was chosen because of its flexibility, the reputation of its quality, and its use in other large-scale projects.

Thermo Fisher: Other than the microarray itself, what else are you looking for in a genotyping provider?

Palotie: When starting a large project like FinnGen and deciding on the genotype provider, we look for reliability, quality, timelines, commitment to R&D, flexibility, and a rigorous output regarding, as I said, timelines and quality.

Thermo Fisher: What is the current status and future direction of the FinnGen study?

Palotie: FinnGen is now in its eleventh month. Every six months, we have a data freeze. We are just approaching the second one, which will consist of 100,000 individuals. The next one, six months from now, will have 150,000, and so on.

> "When starting a large project like FinnGen, we look for reliability, quality, timelines, commitment to R&D, flexibility, and a rigorous output..."

We are currently collecting a large number of prospective samples for the FinnGen project. We've collected 300,000 new samples from hospitals, and we are proceeding very rapidly there. We have roughly 100,000 new samples collected within a year. And we have a special focus on hospitalized patients in order to have the maximum number of disease cases in our study. This is something that makes us slightly different from the UK Biobank, which is a working-age population cohort. In our case, we enriched for disease cases and the older generation so that we would have a long history of health record data available for the study.

Find out more at **thermofisher.com/scientistspotlight**

ThermoFisher SCIENTIFIC