In Finland, there are about 40 hereditary diseases that are characteristic of the Finnish population. The illnesses included in the disease heritage are caused by certain mutations that are more common in Finland than elsewhere in the world due to the so-called bottleneck effect. In the last 10,000 years, a relatively small number of settlers have moved to the Finnish territory. The individuals of this new population represented small and narrow genetic material, resulting in the regional enrichment of certain disease gene variants.

The bottleneck effect is beneficial today for determining the genetic heritage of diseases.

“Finland is the largest bottleneck population on the European scale. Few people have moved here over millennia. The genetic variants that have come to Finland with these immigrants can be a hundred times more common in Finland than elsewhere. This is not often the case in more admixed populations. In Finland, permanent settlements were first established in closer to the coastal areas and only much later in the east and in the north”, says Samuli Ripatti.

According to Ripatti, the internal migration that took place in the 16th century was a second bottleneck thanks to which large differences are discernible in the Finnish population between the east and the west. Although population isolates

Help from the Finnish genome for the prevention of cardiovascular diseases

Professor Samuli Ripatti’s group at the Institute for Molecular Medicine Finland and the Faculty of Medicine of the University of Helsinki studies the underlying mechanisms of cardiovascular diseases through genetic variation. The genetic heritage of the Finnish population provides a good opportunity for this.

“Defining the genetic background of the Finnish disease heritage has also inspired investigations of other diseases.”
also exist elsewhere, defining the genetic background of the Finnish disease heritage has also inspired investigations of other diseases.

“We are now able to understand the onset principles of many genetically inherited diseases, allowing the research data to be applied elsewhere. Although the diseases are different, they have genetic mechanisms that operate in the same way. Understanding this dynamic is a big thing and may provide new opportunities for developing new treatments.”

For example, when studying Parkinson’s disease, it is possible to determine whether the disease is more common in some parts of Finland. If this is the case, research into these parts of the country may produce new genetic information.

**Genetic variation of populations**

Ripatti is interested in population genetics and genetic variation in Finns. He has been a Professor of Biometry at the Faculty of Medicine of the University of Helsinki since 2013. Biometry is a field of statistics that focuses on the analysis of biological data. Ripatti’s research group combines statistical methods with sequence-level measurements of the human genome.

“Sequencing provides information on genetic variation, which may be rare in a population. Based on variation, it is possible to see the prevalence of certain genetic changes that modify disease risk in some areas of Finland or demonstrate predisposition for a certain illness.”

This allows the screening of valuable information about the health effects from the population’s genetic data. People at a high risk of getting sick can be found while also looking for ways to prevent diseases.

“We look at this from the point of view of diseases which touch many individuals. We study illnesses that are common in Finland, such as cardiovascular diseases and diabetes.”

Even though these diseases are affected by, for example, diet and other lifestyle choices, hereditary factors are also significant. That is why these conditions are referred to as common complex diseases. Thanks to the Finnish bottleneck effect, Ripatti’s group has identified genetic variants, genes that predispose you to cardiovascular diseases, in particular, and predictive and marker-controlling genes measured in blood. Ripatti uses high cholesterol levels as an example.

“Those with high cholesterol levels can be examined and their genomes sequenced in an efficient and easy manner.”

Cardiovascular diseases are the cause of one in three deaths in the world. The most affected areas are Central Asia and Eastern Europe. In the 1960s, Finland was the world leader in coronary artery disease mortality of middle-aged men. Upon entering the 2000s, men’s mortality had fallen to about one fifth of the highest level.

However, regional disparities in cardiovascular disease morbidity and mortality are high in Finland. The incidence of the diseases is much lower in Western and Southern Finland than elsewhere in the country. This large regional difference is of interest to researchers. So-called silenced genes that protect against diabetes and cardiovascular diseases have been found in the population of Western Finland.

**Silenced genes**

One interesting area of study is a gene being turned or becoming inoperative.
Such genes are referred to as silenced genes. A study led by Professor Aarno Palotie analysed over 80 mutations silencing an entire gene that are rare but more common in Finland than elsewhere in the world. The material was obtained from the genomes of more than 30,000 Finns.

In fact, Finns have more genes that silence the function of an individual gene than other populations.

“Gene variants that disrupt protein production are generally quite rare in human populations. However, the gene variants that disrupt protein production brought to Finland with settlers are more common here than in the rest of Europe and that is why studying the health effects of those variants that arrived here at that time is much easier in Finland than elsewhere.”

Gene knockouts whose inoperability does not cause health problems have been found in the population of Western Finland. On the contrary, they protect their carrier against diabetes or cardiovascular diseases.

“A gene variant that protects against diabetes has been found in Finland. Carriers of the variant have less incidence of diabetes compared to others. There are more carriers of the gene variant in Ostrobothnia than elsewhere in the world. This may benefit the pharmaceutical industry if it is possible to mimic the function of such a gene through molecular preparations.”

Another example is a gene that prevents the function of lipoprotein(a). Heart disease risk can be assessed by measuring lipoprotein(a) in the blood. Lipoprotein(a) or LPA is a member of the lipoprotein family that carries LDL cholesterol. The data contained, for example, genetic variants whose carriers lacked lipoprotein(a) produced by the LPA gene almost completely. People lacking lipoprotein(a) develop cardiovascular diseases less frequently than others.

“There are a few variants of the LPA gene that shut off its function. In such cases, there is less lipoprotein(a) in the blood. This results in less vascular disease. Lowering the protein level by pharmacological means would be possible and this could be beneficial in preventing coronary artery disease.”

The function of the USF-1 gene has also been studied in Finland. In humans, the gene affects blood fat levels and cholesterol. When the function of the gene was knocked out in mice, the level of the good HDL cholesterol in the blood increased.

“The sequenced sample material has been collected from Finnish patients and volunteers. The material has been used to calculate statistics on how prevalent each genetic variant is in Finland. Once a large enough database has been compiled, it is possible to determine what Finnish genomic variation is like in general.”

The SISu database currently has the protein-encoding variants of the genomes of some 10,000 Finns, and even the full genome of several thousand Finns has already been sequenced.

“The sequence data of the SISu database gives us the opportunity to supplement our material measured with more favourable genome microarrays than others with statistical imputation algorithms. We can now tell quite accurately what kinds of gene variants Finns have. For example, if one in a thousand carries a specific gene variant in Finland, it means that at least 20 people on average should have the variant in the current database.”

The database is already helping patient diagnostics.

“The variation data is in the database and the data is utilised all the time, especial-
ly in clinical genetics. The starting point is that the database provides further clarification for the treatment of a hospital patient. Therefore, if it is suspected that a variant in a gene may be the cause of a disease, the doctor will check the database to see how often this variant occurs in Finns. If it is common, it is unlikely that it would be the cause of a rare disease. If it is rare and its effect on gene function is significant, then the likelihood of the variant’s significance in the onset of a disease is also increased. This is a very concrete clinical application of the database.

FINNGEN records the genomes of half a million Finns

The SISu project data has been collected from research projects and patients. However, the project has focused solely on genomic data, meaning that the potential for utilising the data in health research is limited.

“A biobank sample should be collected from all of us”, says Ripatti. “Those with a predisposition for illness should be screened more closely.”

In Samuli Ripatti’s view, it is a great shortcoming that genomic data is not yet available in connection with medical examinations. It should be part of everyone’s routine check-up to allow concrete decisions for treatment.

“Finland would have excellent setting for this. We have well-functioning occupational and basic health care as well as good expertise in genetic research.”

As a continuation of SISu, the FinnGen project that will record the genomes of half a million Finns launched in August 2017. The project utilises samples collected by all Finnish biobanks. The data from genomes is combined with the information in national health care registers.

“Tools for risk assessments exist and statistical models have been developed for several diseases. Interpreting the data obtained from the genome as part of routine health care is the goal of the next few years. FinnGen contributes to this.”

Ripatti and his team participate in the development, implementation and testing of statistical algorithms.

“We develop prediction algorithms that evaluate, for example, the risk of cardiovascular diseases in a patient. We combine genomic data and lifestyle factors, based on which a prediction is made. Therefore, we are looking for ways to motivate the patients to change their lifestyles.”

Ripatti’s group also supplements genomic data with statistical algorithms. Due to Finland’s population history, it is possible to predict the missing genotypes in microarray data better and more precisely than almost anywhere else in the world. The algorithm works well in Finnish data because Finnish genomes are on average more alike than elsewhere. Sequence data is used to computationally supplement data collected using gene microarrays.

“If a gene microarray scanning the essential variation points has 500,000 genetic markers and we know 30 million genome variants of gene sequences in addition to these measurements, we can supplement the measurement done using the gene microarray into a full genome sequence with good statistical quality indicators. This allows the creation of more sufficiently reliable full genomes at a lower cost. For the time being, sequencing the entire genome is considerably more expensive.”

Data analysis environment needs improvement

Preserving genomic data in the future and designing its analysis environment are big issues in which the ELIXIR infrastructure plays a key role.

“We have an existing pool of data that can be used by authorised researchers. With large databases, it is necessary to provide researchers with a secure and efficient data analysis environment where the data can be analysed.”

Due to data protection, the best solution would be, for example, a remote desktop. At present, there are hundreds of copies of the population data of different countries and other research groups around the world. It is an enormous amount of data.

“On the other hand, in the future, we must have solutions suitable for analysing genomic data that enable the efficient and decentralised storage and analysis of enormous data sets. This challenge will not be resolved by the closed remote desktop solutions developed previously for much more modest data volumes. Instead, what is required are open computing environments that efficiently utilise cloud services and international cooperation. Considering this is quite essential.”

Ari Turunen

MORE INFORMATION:

Institute for Molecular Medicine Finland (FIMM)
The mission of the Institute is to advance new fundamental understanding of the molecular, cellular and etiological basis of human diseases. This understanding will lead to improved means of diagnostics and the treatment and prevention of common health problems. Finnish clinical and epidemiological study materials will be used in the research.

CSC – IT Center for Science
is a non-profit, state-owned company administrated by the Ministry of Education and Culture. CSC maintains and develops the state-owned, centralised IT infrastructure.

ELIXIR builds infrastructure in support of the biological sector. It brings together the leading organisations of 21 European countries and the EMBL European Molecular Biology Laboratory to form a common infrastructure for biological information. CSC – IT Center for Science is the Finnish centre within this infrastructure.

ELIXIR HEAD OFFICE
EMBL European Bioinformatics Institute

www.elixir-finland.org

ELIXIR FINLAND
Tel. +358 9 457 28215 ● e-mail: servicedesk@csc.fi

www.elixir-europe.org/about-us/who-we-are/nodes/finland

www.elixir-europe.org

www.elixir-finland.org