Simulating genetic patterns in **European human evolution**

Dr Mathias Currat

Scientia

SIMULATING GENETIC PATTERNS IN EUROPEAN HUMAN EVOLUTION

Dr Mathias Currat is a population geneticist and computer scientist interested in human evolution. Here he describes a new project combining computer simulations and molecular genetic data to understand the role played by various demographic and migratory events that established the human European gene pool.

For starters, could you tell us a little more about your background and what attracted you to the field of human evolution?

Having studied biology at the University of Geneva, I started to investigate more specifically human evolution when I joined the department of anthropology for my Master degree. I was attracted by the multidisciplinary frame of mind of the research carried out in that field, as well as by the new computational methods developed by my host lab, which is at the cutting edge in this domain. I was fascinated by the crossover between various disciplines such as biology, computer science and history, which fulfilled my hunger for learning. I consider myself fortunate to learn new things on a daily basis through my research and teaching in subjects as diverse.

You hold degrees in both biology and computer science. Do you consider population genetics an 'ideal cross-over' between the two?

Nowadays, computer skills are a primary tool in most areas of biology and more generally of science. This is particularly true in the field of population genetics which is mainly based on theoretical models, mathematical and statistical developments which can be implemented in computer programs. Moreover, the accumulation of genetic and genomic data over the last decades made their analysis more complex and it is now almost impossible to avoid computational methods to store and analyse these large datasets. In that context, computer simulations have become an indispensable tool because they allow studying complex models. For instance, they allow me to incorporate in the models both the effects of population demography and the molecular evolution of DNA, which is essential when analysing genetic and genomic data.

SERIAL SPLATCHE is a new simulation modelling technique you developed to tackle your main research questions. In what way is it different from existing simulation techniques?

SERIAL SPLATCHE is derived from SPLATCHE, a program developed initially with my colleagues Prof. L. Excoffier (University of Bern) and Dr N. Ray (University of Geneva), which allows making spatially-explicit simulations. It simulates the movements of people and hence of their genes through time within a given geographic area. There are few programs of that kind available, and SPLATCHE is the most versatile of them, allowing the simulation of realistic scenarios, such as the effects of migration, demographic growth, competition for the resources, mixing between species, mutation, and many other factors. Moreover, SERIAL SPLATCHE is the only spatially explicit program capable of simulating ancient DNA data retrieved from fossil bones, all other approaches including only DNA from modern populations.

You attempt to combine ancient and modern DNA data. A problem with using ancient DNA is that the sample size is usually limited. How can your approach overcome this problem?

Because of the scarcity of ancient DNA, methods specifically designed to handle it, such as SERIAL SPLATCHE are needed to extract maximum information from it on the evolution of our species. Our approach tries to overcome the limited sample size of ancient DNA by combining into the analyses the large databases of modern DNA in a joint fashion. In that way, we hope to increase the power of our estimations by adding direct picture of the past diversity through ancient DNA to the detailed knowledge of genetic diversity of today's humans.

You mention that to understand how and when genes diffused in Europe can have important implications in biomedical studies. Can this help in the understanding and possibly the treatment of certain present-day genetic defects in humans?

Our simulation approach aims at better understanding their evolution and why their occurrence may be so different among populations. Genetic diseases such as cystic fibrosis or iron overload are more frequent in Western Europe than elsewhere. Their negative effects could have been overcome by demographic processes, thus explaining their higher frequency in some populations. Large demographic growths as the arrival of the first Homo sapiens 40,000 years ago and the Neolithic transition 10,000 years ago could possibly have spread some disease-associated variants in European populations. Spatially-explicit computer simulation has precisely described this purely demographic process affecting genetic data. Studying the behaviour of diversity during complex demographic events may serve to understand basic patterns for the spread of genetic diseases in certain populations.

COMBINING MODERN AND ANCIENT DNA ON A SPATIALLY EXPLICIT SCALE

The genetic diversity of human population is shaped by various processes, such as natural selection and demographic events. The research of Dr Mathias Currat of the University of Geneva is centred on the use of simulation models in reconstructing human evolution on the European continent and determining the importance of demographic events, thereby using a combination of ancient and modern DNA data in a spatially defined context.

A modelling approach in human evolution

The composition and genetic diversity of the contemporary human population is shaped by a myriad of processes. Natural selection is clearly one of the major structuring force. New mutations arise in time; unfavourable ones might become extinct whereas others have more chance to be passed on. Past demography and large population migrations also leave their imprint on the gene pool. By statistically analysing these data we can attempt to unravel the evolution of populations and determine the main processes that formed them.
Different approaches exist to tackle these questions. One common approach is comparing the genetic material present amongst existing populations. Populations are defined on the basis of geographic or cultural criteria. Another approach is the use of models. By modelling evolutionary scenarios, we can obtain a theoretically expected genetic diversity that can then be compared to real data. We can use this method to confront evolutionary scenarios with observed data, and try to explain which scenarios are more likely to have shaped present genetic diversity.

A challenging aspect of human population genetics is that there are so many factors acting simultaneously on the gene pool. Considering all these factors together leads to immensely complex models. Fortunately, sophisticated computer programs exist that allow for simulating quite complex models. Thereby we can generate an expected result of genetic diversity under various evolutionary scenarios, combining demographic and genetic parameters and accounting for archaeological and environmental information. The goal of these models is not to reproduce the actual past (which is impossible), but to understand the main processes that may have shaped the genetic diversity we observe today.

Incorporation of ancient DNA data

Thus far, mainly contemporary genetic material has been used to study human genetic diversity. Recent advances in DNA extraction, recuperation (sequencing) and analysis techniques have given a spectacular boost to the field, enabling the analysis of DNA from ancient fossils. This ancient DNA or aDNA in short, may be broadly defined as the retrieval of DNA sequences from museum specimens, archaeological finds or fossil remains. This has opened up whole new avenues of paleogenetic research.
By recovering genetic data from different periods in the past (heterochronous genetic data), we can follow changes in genetic diversity over time. aDNA samples have strongly enlarged our knowledge of our own evolution, and new insights include the identification of the Neanderthal portion of our genome and the discovery of a new relative of ours, the Denisovan.

Climatological conditions in Europe, especially the northern and central areas, have promoted the preservation of prehistoric fossils. Therefore, the majority of aDNA data so far has been recovered from this continent, thereby providing a great potential for reconstructing the processes that lead to its genetic diversity. 'Despite several decades of multidisciplinary research on the settlement history of Europe, there are still many open questions and hot debates about the genetic impact of the main demographic and migratory events which shaped the diversity of the populations living today on this continent', says Dr Currat.

For instance, it remains unknown what the genetic legacy is of the first anatomically modern humans (AMHs – Homo sapiens) that entered Europe around 45,000 ago and rapidly colonised the continent and what is of the subsequent settlement episodes. Another pressing question considers the extent of the interaction with the Neanderthal population that preceded AMH. Then there are debates regarding the roles played in shaping genetic diversity of population migrations during the Last Glacial Maximum (between 26,500 and 19,000 years ago) and the Neolithic transition (between 10,000 and 5,000 years ago).
The results of ancient mitochondrial DNA (mtDNA) are particularly challenging to interpret, as they suggest a general pattern of

genetic discontinuity between Neolithic and contemporary populations of the same area. This is in conflict with the view that our present diversity has been largely formed by demographic events during the Neolithic or even before, as analysing DNA from contemporary populations seems to indicate. These are the questions Dr Currat and his collaborators hope to answer in a project funded by the Swiss National Science Foundation.

SERIAL SPLATCHE, a powerful simulation tool

This project, entitled 'Reconstructing Europeans' genetic evolution through computer simulations and heterochronous molecular data' aims to synthesise and integrate data from various sources to gain a better understanding of European evolution. The first problem is the lack of tools capable of simultaneously considering aDNA analyses and demography and population migrations. As Dr Currat explains: 'Our work tries to fill this lacuna by developing a new, highly flexible, spatially-explicit approach which is specifically designed to study aDNA at the population level, including the incorporation of spatial elements (geographic locations, movement of populations)'.

This approach has taken shape in the form of SERIAL SPLATCHE, a newly developed simulation method developed by Dr Currat and collaborators. SERIAL SPLATCHE has several unique advantages over existing models. First of all, it is capable of accounting for the geographic characteristics of the study area and exactly positioning populations, samples and migrations. Secondly, it is at present the only spatially-explicit model that can integrate data from both modern and ancient DNA. Thirdly, SERIAL SPLATCHE enables the simulation of complex interactions between populations, such as competition for resources and the mixing of populations.

The project consists of three main research lines. The first line, carried out by PhD student Nuno Silva, focuses on the discrepancies that arose from the mtDNA results. The second line, for which PhD student Jérémy Rio is responsible, aims to optimise the amount of information that can be extracted from aDNA sequences. In the third line, Masters student Nicolas Broccard studies the extent to which local historic processes may have affected genetic diversity without having erased the signature of more ancient demographic events.

Prospects and other applications

The simulation approach is already rendering surprising results, which could not have been obtained in other ways. Simulation results indicate a much larger contribution of pre-Neolithic hunter-gatherers than assumed thus far and underline the potential impact of ancient demographic events on the present genetic diversity. 'We believe that our work holds the promise of solving current interrogations and long-standing debates on Europeans' genetic evolution', says Dr Currat.

Although the principal aim of this project is to improve our understanding of European human evolution and the processes that lead to today's genetic diversity, there are also applications in other fields. SERIAL SPLATCHE is a powerful and flexible simulation method, which can in principle be applied to any area around in world and at any scale. Dr Currat explains that it can also be applied to other species: 'as long as their mode of reproduction and dispersion is relatively similar to that of humans'.

The results of Dr Currat's studies also have important implications for biomedical studies. Certain genetic diseases or traits, such as iron overload, cystic fibrosis and lactase persistence occur more in Europe than elsewhere in the world. By understanding how and when genes dispersed in Europe we can obtain an understanding of the spreading and evolution of these diseases or traits, and why their occurrence is so varied amongst different populations.

Meet the researcher

Dr Mathias Currat
AGP Lab
Anthropology Unit
Department of Genetics & Evolution
University of Geneva

Dr Mathis Currat is Senior Lecturer of the Anthropology Unit, Department of Genetics & Evolution of the University of Geneva. He has studied biology and computer science. In 2004, he obtained his PhD in science from the University of Geneva. He also used to work as a freelance web developer and expert. His research specialises in the ancient history of human settlement and evolution in Europe.

CONTACT

E: mathias.currat@unige.ch
T: (+41) 22 379 69 79
W: http://www.mathiascurrat.com/
W: http://ua.unige.ch/en/personne/mathiascurrat/

KEY COLLABORATORS

Prof. Alicia Sanchez-Mazas, University of Geneva, Switzerland
Dr Estella Poloni, University of Geneva, Switzerland
Prof. Joachim Burger, University of Mainz, Germany
Prof. Daniel Wegmann, University of Fribourg, Switzerland
Dr Nicolas Ray, University of Geneva, Switzerland
Prof. Laurent Excoffier, University of Bern, Switzerland
Dr David Roessli, University of Geneva, Switzerland
Nuno Silva (PhD student)
Jérémy Rio (PhD student)
Claudio Quilodran (PhD student)
Nicolas Broccard (Masters student)

FUNDING

Swiss National Science Foundation
Swiss National Science Foundation (grant 31003A_156853)

REFERENCES

Currat M and Silva NM, Investigating European genetic history through computer simulations, Human Heredity, 2013, 76, 142–153.

Currat M and Excoffier L, Strong reproductive isolation between humans and Neanderthals inferred from observed patterns of introgression, Proceedings of the National Academy of Sciences USA, 2011, 15129–34.

Ray N, Currat M, Foll M and Excoffier L, SPLATCHE2: a spatially-explicit simulation framework for complex demography, genetic admixture and recombination. Bioinformatics, 2010, 26, 2993–2994.

Currat M, Ruedi M, Petit RJ and Excoffier L, The hidden side of invasions: massive introgression by local genes, Evolution, 2008, 62, 1908–1920.

Currat M and Excoffier L, The effect of the Neolithic expansion on European molecular diversity, Proceedings of the Royal Society London B., 2005, 272, 679–688.

UNIVERSITÉ DE GENÈVE