Krzysztof A. Cyran
Silesian University of Technology, Institute of Informatics

Bayesian inference applied in branching process-based estimate of the admixture of Neandertal mtDNA in a gene pool of anatomically modern humans 30,000 years ago

Novelty of the model proposed, as compared to the model presented at the EMBO|EMBL Symposium [3] and in [4], is an integration of the results from the Neanderthal genome project (NGP) in modeling the purifying effect of the genetic drift on Neanderthal mtDNA. For estimating parameters of the new model the intensive stochastic simulations have been performed in 7FP ArSInformatiCa project. Such forward-time simulations of supercritical branching processes constitute serious algorithmic challenges for the efficient memory management, because the nonextinct processes, which model the evolution of a population of anatomically modern humans since the time of the root of anatomically modern human mtDNA polymorphism (about 150 000 years ago), tend to grow to very large sizes. However, the advantage of such modeling, as compared to backwardtime methods, is its greater flexibility, including capacity for stochastic and time-inhomogeneous demographic effects ([5]). Therefore, the results obtained using this method explain more accurately the evolutionary pattern of human genetic variation at the cost of a larger computational effort.

In the talk, an original modeling of the long-term genetic drift by slightly supercritical BP ([3]) will be combined with Bayesian inference model using prior distributions based on data from NGP, which yielded among others: draft sequence of Neanderthal nuclear DNA genome ([8]), complete Neanderthal mtDNA genome ([7]), and analysis of targeted retrieval of five Neanderthal specimens ([1]). The talk will also refer to works of Wall et al. [11] indicating ancient admixtures in multiple human populations.

The results of NGP indicate a slight admixture of the Neanderthals in the genome of H. sapiens, supported also by the fact that the speech-related gene FOXP2 with the same mutations as in modern humans was discovered in ancient DNA in the El Sidron 1253 and 1351c specimens ([10]). It suggests that Neanderthals might have shared some basic language capabilities with H. sapiens what could support genetic exchange to the extent estimated preliminarily ([3]) based only on mtDNA record. In May 2010, the Planck Institute's team, led by Svante Pääbo, has sequenced a draft genome sequence of the Neanderthals, although with a relatively small coverage ([8]). Pääbo [10] and his team showed that using this data we can now for the first time determine which substitutions occurred recently by positive selection in the period shortly after the divergence of modern humans from the Neanderthal lineage. They demonstrated that it is now possible to ask detailed questions about the relationship of the Neanderthals to present-day humans in different parts of the world.

In this context, the Bayesian inference model to be presented in the talk, by

integrating the effect of genetic drift on mtDNA lineage and results of NGP regarding nuclear DNA, yields more accurate estimates of the Neanderthal contribution to H. sapiens gene pool in Europe some 30,000 years ago, than a solely mtDNA-based model ([3], [5], [9]). The discussion to be presented will also extend the analysis given in [6] by inclusion of the recent works on genetic drift in populations governed by a Galton-–Watson branching processes applied by Burden and Simon [2] to analyze root of mtDNA polymorphism of modern humans.

Acknowledgments. The author would like to thank Prof. Marek Kimmel from Department of Statistics at Rice University in Houston, TX, USA for long discussions and advise concerning the branching processes and their applications in biology and genetics.

References

[1] A. W. Briggs et al., *Targeted retrieval and analysis of five Neandertal mtDNA genomes*, Science 325 (2009), 318–321.

[2] C. J. Burden, H. Simon, *Genetic drift in populations governed by a Galton–Watson branching process*, Theor. Pop. Biol. 109 (2016), 63–74.

[3] K. A. Cyran, *Branching process model for estimating the hypothetical Neandertal contribution to the Upper Paleolithic H. sapiens gene pool*, EMBO(1)|EMBL(2) Symposium on Human Variation: Cause and Consequence, EMBL Advanced Training Centre, Heidelberg, Germany, 2010.

[4] K. A. Cyran, M. Kimmel, *Alternatives to the Wright–Fisher model: The robustness of mitochondrial Eve dating*, Theor. Pop. Biol. 78:3 (2010), 165–172.

[5] K. A. Cyran, *Artificial Intelligence, Branching Processes and Coalescent Methods in Evolution of Humans and Early Life*, Publishers of the Silesian University of Technology, Gliwice, Poland, 2011.

[6] K. A. Cyran, M. Kimmel, *Estimating Neanderthal contribution to the Upper Paleolithic H. sapiens gene pool using genetic drift effect in branching process model*, Recent Advances in Information Science: Proc. 3rd European Conference on Computer Science, Paris, 2012, 57–62.

[7] R. E. Green et al., *A complete Neandertal mitochondrial genome sequence determined by highthroughput sequencing*, Cell 134 (2008), 416–426.

[8] R. E. Green et al., *A draft sequence of the Neandertal genome*, Science 328 (2010), 710–722.

[9] M. Kimmel, D. Axelrod, *Branching Processes in Biology*, Springer-Verlag, New-York, 2015.

[10] S. Pääbo, *The Neandertal Genome*, EMBO|EMBL Symposium on Human Variation: Cause and Consequence, EMBL Advanced Training Centre, Heidelberg, Germany, 2010.

[11] J. D. Wall, K. E. Lohmueller, V. Plagnol, *Detecting ancient admixture and estimating demographic parameters in multiple human populations*, Mol. Biol. Evol. 26 (2009), 1823.