

# BEST PRACTICE COMPUTATIONAL APPROACH FOR KINSHIP ANALYSIS OF LOW COVERAGE/ANCIENT DNA SAMPLES

Emil Nyerki<sup>1</sup>

<sup>1</sup>Department of Pediatrics, University of Szeged Albert Szent-Györgyi Medical Center  
Faculty of Medicine, Szeged, Hungary

<sup>2</sup>Research Centre for Archaeogenetics, Institute of Hungarian Research, Budapest, Hungary

harmadikemil@gmail.com

## ABSTRACT

Current state of art kinship analysis is capable to infer relatedness up to the 5-6th degree from deeply sequenced DNA if the proper reference population is known. Low coverage, partially genotyped, degraded archaic (or forensic) DNA and often unavailable or unknown reference population poses additional challenges. Accordingly up to now kinship analysis from low coverage archaic sequences is mainly possible up to the second degree with large uncertainties. We performed extensive simulations to assess the main factors of bias. Accordingly, here we offer a best practice guideline to overcome the difficulties associated with aDNA kinship analysis. We validated our proposed methodology on experimental modern and archaic data with widely different genome coverages using samples with known family relations and known or unknown population structure as well. With a proposed practical workflow we provide the necessary additional tools to calculate the corrected kinship coefficient using the widely used genome data formats. Our methodology allows to reliably identify relatedness up to the 4-5th degree from variable/low coverage archaic (or badly degraded forensic) WGS genome data.

Emil Nyerki was supported by the ÚNKP-21-3-SZTE-67 New National Excellence Program of the Ministry for Innovation and Technology, from the source of the National Research, Development and Innovation Fund.

*Keywords: bioinformatics, archaeogenetics, simulations*