

Genetic kinship in ancient human populations: estimating the reliability and efficiency of analysis methods

The study of genetic kinship allows anthropology to identify the place of an individual within which they evolve: a biological family, a social group, a population. The application of classical probabilistic methods (that were established to solve cases in legal medicine, such as Likelihood Ratios, or LR) to STR data from archaeological material has permitted the discovery of numerous parental links which together constitute genealogies both simple and complex. Our continued practice of these methods has however led us to identify limits to the interpretation of STR data, especially in cases of complex, distant or inbred kinship. The first part of the present work is constituted by the estimation of the reliability and the efficacy of the LR method in four situations: a large modern population with significant allelic diversity, a large modern population with poor allelic diversity, a large ancient population and a small ancient population.

Recent publications use the more numerous markers analysed using Next generation Sequencing (NGS) to implement new strategies in the detection of kinship, especially based on the analysis of chromosome segments shared due to common ancestry (IBD "Identity-by-Descent" segments). These methods have permitted the more reliable estimation of kinship probabilities in ancient material. They are nevertheless ill-suited to certain typical situations that are characteristic of ancient DNA studies: they were not conceived to function using single pairs of isolated individuals and they depend, like classical methods, on the estimation of allelic diversity in the population. We therefore propose the quantification of the reliability and efficiency of the IBD segment method using NGS data, focusing on the estimation of the quality of results in different situations with populations of different sizes and different sets of more or less heterogeneous samples.

This thesis work also proposes a method that would be adapted to the specific characteristics of ancient DNA (few individuals, ill-defined populations, technical errors...). We therefore present a method dedicated to the determination of kinship in isolated pairs of individuals when allelic variability is unknown and test that method to quantify its reliability and efficiency, as we have done with LR and IBD segments.

Reliability and efficiency were estimated using 727 modern individuals for whom more than 1000 familial links were declared and different genetic markers were studied (STR and SNP). These methods were then applied to 186 ancient individuals in order to establish their genealogies and propose solutions to the study of complex or distant kinship in an archaeogenetic context.

Popularised summary

The study of genetic kinship has always been of great interest to legal and forensic medicine, that sees an opportunity to confirm paternity, identify the victims of disasters or crimes and even, in particular situations, to participate in the identification of the authors of crimes. Genetic kinship determination methods were therefore conceived to solve legal cases before they were applied to the study of ancient DNA from the individuals recovered from archaeological sites that are hundreds or thousands of years old. With the apparition of new methods, and since the application of classical methods did allow the construction of ancient genealogies, it has become necessary to estimate the quality of determinations through empirical studies. This thesis work uses more than 1000 biological links identified in modern populations and implements three kinship determination methods, in order to estimate their reliability and efficiency and propose specific solutions for the study of ancient DNA.