

DNA DOE PROJECT – OVERCOMING CHALLENGING DNA SAMPLES

Kevin Lord

DNA Doe Project

Since our first identification of Marcia King (aka Buckskin Girl), announced in early April 2018, we have now identified approximately 50 Jane and John Does utilizing genetic genealogy. In addition to the vast genealogy expertise of our volunteers, we have also gained a wealth of experience in working with challenging samples to obtain the data necessary for genealogical analysis.

Many of the cases brought to us consist of remains that are decades old or older. In the case of remains found in Hudson OH in 2019, we were able to successfully identify the John Doe as someone who had passed away pre-Civil War and had been buried for nearly 170 years before being unearthed during the development of a new subdivision.

In this talk I will detail some of the technologies and methods that allowed us to obtain usable data for upload to genealogical databases in this extremely challenging case. This began with reviewing particular QC metrics to determine whether the sample was likely to be successful. We then employed whole genome enrichment to overcome bacterial contamination and utilized whole genome sequencing to obtain as much relevant data as possible. Finally, cutting-edge bioinformatics techniques were utilized to obtain the data required for upload to the genealogical databases. An example of this is imputation, a process developed originally for genetic association studies, which allowed us to recover genotypes from low-coverage sequencing data.

The use of genetic genealogy in human identification is rapidly becoming more prevalent. Bioinformatics has become increasingly critical as we tackle more and more compromised DNA samples from unidentified remains. Our hope is that by providing these details, those in the field will know what is now possible and be armed with the information required to make more informed decisions when deciding how to proceed with their cases.