Abstract

Mutations and non-paternity are two of the most influential factors that may bias conclusions in Y-chromosome (Y-c) population genetic and lineage studies. Previous studies have determined locus-specific mutation rates for 12 Y-ch loci. Herein, we characterize mutation rates of 46 STR loci based on 500 individuals separated by between 1-13 meioses, including eight previously reported loci for comparison to published rates. Non-paternity rates have been shown to vary among certain demographies, age groups, and time periods. To account for all of these factors, the present study reports a non-paternity rate for samples collected from over 100 different families that vary in age group, culture, country of origin and the number of meioses separating them.

Non-Paternity Rates

Accurate non-paternity rates are critical for population and family based studies and for the accurate reconstruction of recent genealogies. Systematic studies of non-paternity rates are limited (M.A. Jobling 2004). Early reports by Machinthe and Soman (1991) report rates between 1 to 20%. Most recently a study in Switzerland found a rate of 1% (Sane 1994). A study from Mexico established a rate of 12% (Cenda-Flores 1999). Several factors that increase the difficulty in determining an accurate non-paternity rate include variations in demographics, culture, and age groups (Cenda-Flores 1999, Jobling 2004, Machinthe and Soman 1991). It is also suggested that rates will vary over different time periods, with the rate decreasing back in time. Recently AABH accredited laboratories reported 28.79% of their cases as non-paternity (AAAB Annual Report Survey for Testing in 2002). However, it is expected that this rate is considerably higher than the non-paternity rate in the general population due to sampling bias.

In the present study, the derived results reflect from samples collected from families that vary in age group, culture, country of origin (see Figure 1) and the number of meioses separating them (see Table 1). It is presumed that the resulting rates represent an average for all of these factors. In addition, no selection for suspected non-paternity occurred. To preclude the necessity of obtaining a sample from both parents, Y-chromosome (Y-c) testing was applied in determining non-paternity rates in purported paternal lines.

Pedigree information was collected from each participant. The paternal line was extended using the Ancestral File database (see Family Search website). Only individuals without known non-paternal event were included in the analysis. Individuals shared a common ancestor along their strict paternal line were grouped into Common Lines (CL) and the number of meiotic events were counted. Equation 19b (Walsh 2001) was used to calculate the Most Recent Common Ancestor (MRCA) for each CL using the infinite alleles model and individual mutation rates (not based on pedigrees) from Huttinson (in prep). The cumulative likelihood (cam) was calculated for the expected CA that was indicated by the genealogy (see Table 1). This was converted to an Odds Ratio: OR=(1-cam)/(cam). The odds ratio is the likelihood of a more distant vs. more recent CA than indicated by the genealogy. For example CL 1.77 has two mutations in two meiotic events. The cumulative likelihood is 0.007 which converts to odds of 146:1 that there is a more distant common ancestor than 2 meioses separating these individuals. The cut off point of 35:1 was determined by a threshold between 1 and 2 mutations in 1 to 2 meioses. Observing two mutations in two meioses is unlikely, and the CL was determined to have a non-paternity. The non-paternity rate was calculated by dividing the number of pairs with a non-paternity by the average number of meiotic events between all pairs of individuals. The rate was determined to be 1.27%. For our calculations, we assumed only one non-paternity event per non-paternity pair. It is also assumed that incorrect genealogical data, either from Ancestral File or mistakes submitted from the participant are incorporated in our non-paternity rate. It is difficult to determine the exact generation in which a non-paternity occurs, therefore this frequency reflects an average over all generations separating the samples.

Conclusions

An accurate non-paternity rate is important for population study, recombination estimates, and estimating transmission probabilities of genetic disorders (Cenda-Flores 1999). Here we report a non-paternity rate that is lower than the previous estimate from Cenda-Flores et al. However, some recent studies have also detected lower rates, such as the Saus et al (1991) and Sykes 2000 (1.7%-generation) studies. Our rate was calculated based on samples originating in 18 world populations, predominantly from European countries and the continental United States. Individuals that had known non-paternity, or adopted in their pedigrees were excluded from this study. Thus, the 1.27% proposed rate describes a non-paternity rate for individuals that have no knowledge of a non-paternity event in their ancestry and it may be lower than that of the general population. These rates were calculated based on multigenerational pedigrees and 73 for paternal lines.

We report specific locus mutation rates for 36 Y-ch STR loci, the largest number of loci to date. We compared our mutation rates to those previously published by Heyer et al and Kayser et al. Our rate estimates are comparable to the 9 loci previously published. We estimate rates for additional 21 loci that are not published.

References