Low nonpaternity rate in an old Afrikaner family

Jaco M. Greeff\textsuperscript{a,}\textsuperscript{1}, Francois A. Greeff\textsuperscript{b}, Andre S. Greeff\textsuperscript{c}, Lucas Rinken\textsuperscript{b}, Dawid J. Welgemoed\textsuperscript{d}, Yolanda Harris\textsuperscript{d}

\textsuperscript{a}Department of Genetics, University of Pretoria, Pretoria, South Africa
\textsuperscript{b}Department of Medical Microbiology, MEDUNSA, University of Limpopo (retired), South Africa
\textsuperscript{c}Department of Haematological Pathology, MEDUNSA, University of Limpopo, South Africa

*Corresponding author. Department of Genetics, University of Pretoria, Pretoria 0002, South Africa. E-mail address: jaco.greeff@up.ac.za (J.M. Greeff).

Abstract

Extrapair paternity is a crucial parameter for evolutionary explanations of reproductive behavior. Early studies and human testis size suggest that human males secure/suffer frequent extrapair paternity. If these high rates are indeed true, it brings into question studies that use genealogies to infer human life history and the history of diseases since the recorded genealogies do not reflect paths of genetic inheritance. We measure the rate of nonpaternity in an old Afrikaner family in South Africa by comparing Y-chromosome short tandem repeats to the genealogy of males. In this population, the nonpaternity rate was 0.73\%. This low rate is observed in other studies that matched genealogies to genetic markers and more recent studies that also find estimates below 1\%. It may be that imposed religious morals have led to reduced extrapair activities in some historic populations. We also found that the mutation rate is high for this family, but is unrelated to age at conception.

Keywords

Extrapair paternity; Extrapair copulation; Religion; Pedigree; Genealogy; Mutation rate

1. Introduction

Genealogies drawn up from church records, birth certificates, and death notices are useful to identify founders of diseases (Austerlitz and Heyer, 1998, Hayden et al., 1980, Heyer et al., 1997b, Heyl, 1970 and Torrington and Viljoen, 1991); to identify disease-causing genes (Karayiorgou et al., 2004 and Vezina et al., 2005); to quantify the parameters that shape human life histories (Lummaa, 2007 and Voland, 2007); to determine the genetic heritage of individuals (Greeff, 2007); to estimate mutation rates of forensically important DNA markers (Heyer et al., 1997a, King and Jobling, 2009a and Pollin et al., 2008); and, as noted by King & Jobling (2009b), for fun. A weakness in genealogical hypotheses is that they can be incorrect due to unrecorded adoptions or extrapair paternities. A number of initial studies argued that extrapair paternity rates may be as high as 5\% or even 10\% (Baker and Bellis, 1995 and Macintyre and Sooman, 1991). This means that after 12 generations, only 54\% (\(=0.95^{12}\)) or even as few as 28\% (\(=0.9^{12}\)) of males will be genetically related to the original
male founder. Looking back in time, it can be shown that, with an extrapair paternity rate of $p$, a fraction

$$C = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} (1-p)^k.$$

of ancestors $n$ generations ago will be correctly identified. This would mean that as many as 40% of ancestors 10 generations ago, identified using genealogies, will be incorrect (Fig. 1A).

![Graphs showing the influence of extrapair paternity on accuracy of genealogies.](image)

**Fig. 1.** Influence of extrapair paternity on accuracy of genealogies. Distribution of the expected fraction of ancestors in the 10th generation, identified by genealogy, which is actually incorrect if extrapair paternity is 10% (A), 1.7% (B), and 0.75% (C). These distributions were obtained from 10,000 simulations in R. The filled circles indicate the means. (D) The expected fraction of ancestors that are correctly identified given a rate of extrapair paternity. From top to bottom for 5, 10, and 12 generations.

However, Anderson (2006) has shown that while extrapair paternity rates are indeed high in cases where paternity is contested (29.8%), it is substantially less in general, in the order of 1.7%. Similarly, Voracek, Haubner, and Fisher (2008) showed that more recent studies found
significantly lower estimates and an average of 3.1% extrapair paternity. These estimates would reduce the number of potentially misidentified ancestors substantially to 8% in 10 generations (Fig. 1B). It is also clear that the rates vary from population to population, with culture and specifically social norms regarding extramarital sex playing an important role (Anderson, 2006 and Bellis et al., 2005).

The Afrikaner population of South Africa descends mostly from Dutch, German, and French immigrants. Afrikaners have a well-recorded ancestry (De Villiers & Pama, 1966) and a vibrant genealogical community (Genealogical Society of South Africa: http://www.eggsa.org/Familia-index/general.htm; Genealogical Institute of South Africa: http://www.gisa.org.za/site/node/3), so it is potentially a rich field for doing family reconstitution from old records (Voland, 2007). The crucial question is how faithful the Afrikaner couples were to each other? During the 1980s, in cases where paternity was contested, 22% of South African Caucasians (i.e., including non-Afrikaans-speaking people) were excluded as fathers (Du Toit, May, Halliday, Schlaphoff, & Taljaard, 1989). This is the sixth lowest value of 31 studies worldwide (Anderson, 2006). Data recorded during 2007 for 72 Afrikaans-speaking Caucasians where paternity was contested also gave 22.2% nonpaternity events (A.S. Greeff & Y. Harris, unpublished data).

Despite the fact that, during the last two and a half centuries, Christianity, with its emphasis on monogamy, was seen as one of the cornerstones of the Afrikaner family and population (Giliomee, 2003), this has not always been the case. For instance, despite the fact that the population was strongly male biased, the reformed church of Stellenbosch district initially had three times more confirmed women than men (Giliomee, 2010). In the Stellenbosch of 1726, it was unusual for both members of a couple to be confirmed members of the reformed church (only 20%; Giliomee, 2010). In fact, a traveler referred to the Cape population as “an assembly of blind heathen” (Giliomee, 2010). During the first 150 years of the European settlement, there were very few European women, and illegitimate sex, especially with slaves, was far more frequent than later on (Giliomee, 2010). Female slaves were often part of the household, acting as child-minder, wet nurse, and sometimes even the mistress of the head of the house (Giliomee & Mbenga, 2007). Furthermore, the slave lodge acted as a brothel (Giliomee & Mbenga, 2007), suggesting frequent illegitimate sex at the Cape of Good Hope.

While these sexual activities of men could not have gone unnoticed, the children thus produced normally defaulted to the slave and Coloured communities (De Wit et al., 2010, Giliomee, 2010 and Quintana-Murci et al., 2010). Furthermore, the Roman Dutch law offered females divorce (although it was very infrequent) and partible inheritance as a formidable weapon to “control” their spouses (Giliomee, 2010). Under partible inheritance, the estate is divided between the remaining family members upon the death of the patriarch or divided between spouses after divorce. So, while men may have fathered illegitimate children with slaves in their youth, the recorded genealogical history of Afrikaners may in fact reflect genetic ancestry correctly.

YSTRs are short tandem repeats on the Y-chromosome that are inherited like surnames in Western societies—from father to son. Different alleles at a locus have different numbers of these tandemly repeated nucleotides. YSTRs normally do not code for genes and have a high
mutation rate that causes these loci to have highly variable numbers of repeats. Since the largest part of the Y-chromosome does not undergo recombination, the number of repeats at several loci can be counted and combined to form a composite genotype referred to as a haplotype. These haplotypes are numerous and thus unique enough that YSTRs are frequently used for forensics. Several studies have matched YSTR haplotypes to genealogies to record mutation rates (Heyer et al., 1997a and Pollin et al., 2008) and validate ancestries (Foster et al., 1998 and Kayser et al., 2007). YSTR haplotypes are also compared to surnames to understand the frequency of certain surnames (Immel et al., 2006, King and Jobling, 2009a and McEvoy and Bradley, 2006). In order to estimate the rate of nonpaternity events in the Afrikaner population, we matched YSTR haplotypes of one surname to the genealogies of the sampled individuals.

2. Methods

A recent study that disproved the global monophyly of the surname Greeff (Greeff, Greeff, Harris, Rinken, & Welgemoed, 2010) genotyped Greeff males with the AmpFEBSTR Yfiler kit that amplifies 17 variable YSTR loci (Applied Biosystems, Inc.). In short, men already on the South African Greeff website (http://www.greeff.info/tng01/), Greeffs on Facebook, and relatives were approached to contribute samples. In this fashion, 39 South African subjects signed an informed consent form and donated a mouth swab as well as relevant genealogical information. In addition to genotypes, the genealogies of many of the South African Greeffs were established from genealogical (De Villiers and Pama, 1966, Heese and Lombard, 1989 and Rinken, 2010) and church records. Most of the Greeff males stem from Matthias Greeff who arrived in South Africa from Magdeburg, Germany, before 1680. A few males were linked to another Greeff founder, Friedrich Greeff, who came from Brunswyk (Germany) 100 years later. Here we look at these data to estimate the rate of nonpaternity and the mutation rate. The binomial test in the statistics software R (R Development Core Team, 2010) was used to obtain 95% confidence intervals (CIs) for estimates.

If the Greeff haplotype is very common in Europe, several surnames may have carried the haplotype, and some extrapair paternity could potentially go undetected. Therefore, we estimated the frequency of the Greeff haplotype in the general European population by comparing our haplotype to those already recorded in the Y-chromosome haplotype reference database that contains tens of thousands of haplotypes (Willuweit & Roewer, 2007).

We also compared locus specific mutation rates to those observed in Goedbloed et al. (2009) using a Spearman rank correlation since the data were not normally distributed. To test for a relationship between age of a father and likelihood of a mutation, a generalized linear model was fitted to the data with binomial errors with number of mutations per conception as a response variable and father’s age at conception as the predictor variable. In cases where the exact position of the mutation was unknown, we took the average age as the predictor and the number of mutations out of the number of conceptions as the dependent variable. This will introduce some inaccuracy as the average age cannot take into account the variance in age on a particular stretch of lineage.
3. Results

3.1. Nonpaternity rate

The genealogies and genotypes are summarized in Fig. 2. Matthias Greeff’s tree is 12 generations deep, representing 103 recorded meioses and conceptions. Thirty sampled individuals could be linked to Matthias’s genealogy via records. Friedrich Greeff’s tree is six generations deep and reflects eight meioses and conceptions, linking two sampled individuals. One of Matthias’ presumed descendants, filled square in Fig. 2, can be excluded as such because of his anomalous haplotype. This gives a nonpaternity rate of 0.9% (one mismatch out of a total of 111 conceptions=103+8; 95% CI: 0.02%–4.92%). However, seven remaining South African Greeffs who are not currently linked to either of the Greeff trees are clearly descendants of Matthias (N=6) and Friedrich (N=1). These males should also be considered. We can do so by looking at the extremes (King & Jobling, 2009a), which are that each unconnected male meets up with the known tree in one generation, i.e., adding five conception events. At the other extreme, each unlinked sample meets up with the founder in 12 generations, adding another 5×12=60 conception events. We choose a middle road as follows, each of the 23 males at the bottom of the genealogy (the tips) contributed 103/23=4.48 conceptions to the genealogy. Therefore, we can expect another five unlinked

![Fig. 2](A) The ancestry of 30 genotyped descendents of Matthias Greeff. Smaller blocks are inferred haplotypes, whereas larger blocks are the genotyped haplotypes. Six males could not be linked to the genealogy but are clearly descendents of Matthias (inset). For cases where it is not clear where a mutation occurred along a lineage, haplotypes are indicated by a dot in a square. (B) The ancestry of two descendents of Friedrich Greeff. Further details similar to (A), except that there was only one unlinked individual. (C) Number of repeats at 17 microsatellites for each of the haplotypes, with the symbols in the left corresponding to those in (A) and (B). The filled block is the only person that requires a nonpaternity event to explain the genotype.
tipped males to add another 22.4 (=4.48×5) conceptions to Matthias' tree. Similarly, the one unlinked descendant of Friedrich represents another four conceptions. The corrected estimate of nonpaternity is thus 0.73% (=1/137; 95% CI: 0.02%–4.00%).

We did not find any matches out of 6656 European haplotypes on the database, giving an estimated frequency in Europe of 0 (95% CI: 0–5.541×10⁻⁴). There were no haplotypes on the database that could be linked to Matthias' haplotype by one mutational step either. Searching with only nine loci for which more data are available (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385a&b) gave 12 matches out of a sample of 41,559 Europeans in the database, giving a frequency of 2.887×10⁻⁴ (95% CI: 1.395×10⁻⁴–4.478×10⁻⁴). The Matthias haplotype is thus not frequent enough to be sampled from unrelated individuals.

3.2. Mutation rate

Considering the 97 meiotic events in Fig. 2A at 17 loci, nine mutations were recorded (97 because we need to exclude the six uncertain ancestors of the mismatched individual, filled square). However, 3891 and 389II are a composite YSTR, and the repeat addition in both can best be explained by a single mutation (Zhivotovsky et al., 2004). This overall rate of 4.85×10⁻³ (95% CI: 0.0021–0.0095; Fig. 3) is higher than the mean estimate 0.0028 measured by Goedbloed et al. (2009) based on more than 1700 father–son comparisons at the same 17 loci, but not significantly so (binomial test: p=.152). In line with findings that DYS458 has the highest mutation rate (Goedbloed et al., 2009), a third of the recorded mutations were at locus DYS458, giving a very high rate of 3.1×10⁻² (95% CI: 0.0064–0.0877). While the rank order of the mutation rates recorded here is correlated to that recorded by Goedbloed et al. (2009; Spearman: ρ=0.62; S=308.4477, p=.0077), the rates recorded here are higher (Fig. 3).

![Fig. 3. Comparison of the mutation rates observed in the Goedbloed et al. (2009) study to that in the current study at 17 YSTR loci (open circles). The heavy line shows the linear regression; and the thin line, the diagonal.](image-url)
Goedbloed et al. (2009) found that the men whose sperm had mutations were older on average (34.4 years) than men whose sperm did not have mutations (30.3 years). The males in our sample were 34.4±8.3 years (mean±S.D.) in age at childbirth, which is similar to the Goedbloed et al. (2009) mutation group. However, the generalized linear model found no increased likelihood of mutation with age ($p=.85$). In fact, the average age of males with mutations was lower than those without (34.8 viz. 33.1).

4. Discussion

4.1. Nonpaternity rate

We recorded a nonpaternity rate of 0.73%, which is the second lowest rate when compared to 22 studies on modern populations (Anderson, 2006). Only the Sephardic Kohanim Jewish priests have a lower rate (Anderson, 2006). The standard interpretation for a mismatch is that one of the males along the lineage has been cuckolded or an unrecorded adoption occurred. In the South African context, a secretive type of adoption also needs to be considered, namely, if a daughter got pregnant out of wedlock, the shame would be hidden by pretending that the child was that of the mother’s. Thus, it is essentially similar to parents adopting their grandchild. From a genealogical point of view, this will mean that one more generation needs to be squeezed in with the recorded parents set equal to maternal grandparents and the paternal grandparents are unknown. In fact, this grandparental adoption is the exact scenario of the mismatched individual's grandfather's sister (personal communication, G. Grundlingh). From the perspective of extrapair paternity rates, grandparental adoption will mean that the already low estimate of 0.73% may be even lower.

<table>
<thead>
<tr>
<th>Number of conceptions</th>
<th>% Nonpaternity</th>
<th>Region/Population</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>257</td>
<td>0.39</td>
<td>Sagueneay Region, Northeast Quebec</td>
<td>(Heyer et al., 1997a and Jobling et al., 1999)</td>
</tr>
<tr>
<td>75</td>
<td>0</td>
<td>Central Germany</td>
<td>Kayser et al. (2007)</td>
</tr>
<tr>
<td>±1200</td>
<td>0.08</td>
<td>Old Order Amish Pennsylvania</td>
<td>Pollin et al. (2008)</td>
</tr>
<tr>
<td>Minimum and maximum</td>
<td>1.28–3.26</td>
<td>British</td>
<td>King and Jobling (2009a)</td>
</tr>
<tr>
<td>66</td>
<td>1.5</td>
<td>American</td>
<td>Foster et al. (1998)</td>
</tr>
<tr>
<td>134</td>
<td>0.73</td>
<td>Afrikaner</td>
<td>Current study</td>
</tr>
<tr>
<td>Average</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparing our study to five other studies on historical populations suggests that a low nonpaternity rate of 0.73% is typical (Table 1). These low nonpaternity rates seem to be at odds with current sexual behavior (Simmons, Firman, Rhodes, & Peters, 2004), our very ancient behavior as reflected in our relative testis size (Harcourt, Purvis, & Liles, 1995), and the rate of molecular evolution of sperm and seminal fluid proteins (Wyckoff, Wang, & Wu, 2000). One explanation could be that religious norms and the dominance of the church in historic times reduced the extent of extrapair sexual behavior. Giliomee (2010) suggested that married women may have played an important role in enforcing these morals. These
specific norms must have been lacking throughout most of our evolutionary past that shaped our morphology. Recent low estimates (Voracek et al., 2008) even in relatively secular populations (Wolf, M., Musch, J., Enzmann, J. & Fischer J., unpublished) may be explained by the use of contraceptives (Voracek et al., 2008) rather than abstinence.

Since the rate of extrapair paternity is an important driver of sexual roles (Kokko & Jennions, 2008), it suggests that, at least in these historic populations, fathers could safely invest in their children.

Our study cannot pick up when wives cheated their husbands with the husband's brother or other patrilineal relatives of her husband. In isolated farm communities where families often clustered together, the only temptation may have come from relatives by marriage. Even so, we have presented strong evidence that, at least in this Afrikaner family, nonpaternity events are infrequent enough to result in less that 4% of incorrect ancestors 10 generations ago (Fig. 1C & D).

4.2. Mutation rate

The mutation rate in the descendents of Matthias was somewhat high. Although the average male was indeed of a similar age to males that are prone to mutations (Goedbloed et al., 2009), there was no evidence that older men had more mutations. A larger sample will be required to determine whether this increased rate is just a chance event.

Acknowledgments

We again thank the people who participated in the Greeff Y-DNA project. We thank Carina Schlebusch and Vinet Coetzee for insightful comments. We also thank three referees for their useful suggestions. The reanalysis of these data to estimate the rate of nonpaternity was approved by the University of Pretoria Ethics Committee (EC 110117-002).

References


