Am. J. Hum. Genet. 66:1417–1419, 2000

Report

Surnames and the Y Chromosome

Bryan Sykes and Catherine Irven

University of Oxford, Institute of Molecular Medicine, Oxford

A randomly ascertained sample of males with the surname "Sykes" was typed with four Y-chromosome microsatellites. Almost half the sample shared the same Y-chromosome haplotype, which has not been observed in control samples either from the same geographic region or from the United Kingdom as a whole. This points to a single surname founder for extant Sykes males, even though written sources had predicted multiple origins. The distribution of other Sykes Y-chromosome haplotypes were not significantly different from those in controls and may be accounted for by the historical accumulation of nonpaternity during the past 700 years, in which case the average rate estimate is 1.3%/generation. If this pattern is reproduced with other surnames, it may have important forensic and genealogical applications.

Since the English Mediaeval period it has been common practice for children to take their father's surname. It is also a biological necessity that a son inherit the Y chromosome of his father, so it follows that males sharing the same surname might also share the same haplotype in the nonrecombining segment of the Y chromosome. However, this would only hold if (i) there had been no subsequent Y-chromosome divergence, (ii) there were a single common ancestor for each surname and there had been no subsequent surname adoption by unrelated males, and (iii) there had been no nonpaternity events involving unrelated Y chromosomes.

In order to assess the correspondence between surname and Y-chromosome haplotype, a sample of males with the surname "Sykes" was ascertained from published lists compiled from electoral rolls and other registers. From the geographic distribution of the 9,885 registered U.K. voters with that surname, it was clear that the highest residential concentrations of Sykes are in the counties of West Yorkshire, Lancashire, and Cheshire. This matches the earliest occurrences of the name, during the 13th and 14th centuries, in the villages of Flockton, Slaithwaite, and Saddleworth, close to Huddersfield, West Yorkshire (Redmonds 1992, pp. 53–54). A postal request for a cheek-cell sample was sent to 269 male Sykes chosen at random from the three counties. A total of 61 replies (response rate 22.7%) were received, and DNA from 48 (78.6%) of this latter group was successfully extracted and genotyped. Two control groups were established. The first was a sample of 139 native English males from all over the country ("English"), and the second, to control for abundant local haplotypes, was a group of 21 unrelated male neighbors recruited by the Sykes volunteers ("Neighbor"). Y chromosomes from all three samples were genotyped at four microsatellite loci in common use. (table 1).

There is a highly significant association between Y-chromosome haplotype distribution and the Sykes sample, an association due entirely to haplotype 15-23-11-14 ($P_i = 1.4 \times 10^{-10}$, where P_i is the probability that the Sykes and the English samples come from the same population divided by the probability that the Sykes and the English samples come from different populations). This haplotype is not encountered in either the Neighbor or English controls and suggests a single patrilineal ancestor for these males. This is surprising, given the general locality category of the name "Sykes," which means "spring," "stream," or "boundary ditch" and is thought to have had several origins in Yorkshire (Redmonds 1992, pp. 53–54).

Haplotype 15-23-11-14 is carried by 43.8% of the Sykes sample. Only two haplotypes, carried by four Sykes males, are one mutational step removed from 15-23-11-14. If a stepwise mutational model (Kimura and Ohta 1978) is assumed, this suggests that there has been

Received May 27, 1999; accepted for publication January 6, 2000; electronically published March 17, 2000.

Address for correspondence and reprints: Dr. Bryan Sykes, Institute of Molecular Medicine, Oxford OX3 9DS, United Kingdom. E-mail: bryan.sykes@imm.ox.ac.uk

0002-9297/2000/6604-0025$02.00

Table 1

Haplotype Occurrences in the Three Population Samples

Haplotype[a]	No. (%) in Sykes	Neighbor	English
13-22-10-12	2 (4.2)	0	1 (.7)
13-22-11-12	0	0	3 (2.2)
14-22-10-12	0	2 (9.5)	0
14-22-10-13	0	2 (9.5)	9 (6.5)
14-23-10-13	1 (2.1)	0	9 (6.5)
14-23-10-14	0	1 (4.8)	1 (.7)
14-23-11-13	1 (2.1)	2 (9.5)	26 (18.7)
14-23-11-14*	1 (2.1)	0	3 (2.2)
14-24-10-13	0	3 (14.2)	22 (15.8)
14-24-11-13	5 (10.4)	1 (4.8)	36 (25.9)
14-24-11-14	0	0	6 (4.3)
14-25-10-13	0	1 (4.8)	5 (3.6)
14-24-11-13	3 (6.3)	1 (4.8)	4 (2.9)
15-21-11-12	0	1 (4.8)	0
15-22-10-12	0	1 (4.8)	0
15-23-9-13	0	1 (4.8)	0
15-23-10-13	3 (6.3)	1 (4.8)	3 (2.2)
15-23-10-14*	3 (6.3)	1 (4.8)	4 (2.9)
15-23-11-14	21	0	0
15-24-10-12	1 (2.1)	0	4 (2.9)
15-24-10-13	0	1 (4.8)	0
15-25-11-13	1 (2.1)	0	3 (2.2)
15-25-11-14	3 (6.3)	1 (4.8)	0
15-26-11-13	3 (6.3)	0	0
17-23-10-13	0	1 (4.8)	0
Total	48	21	139

Note.—Y chromosomes were genotyped at the microsatellite loci DYS19, DYS390, DYS391, and DYS393, in a single run on an Applied Biosystems 373 sequencer, with use of primers described by in de Knijff et al. (1997).

[a] Described by numbers of repeats at, consecutively, DYS19, DYS390, DYS391, and DYS393. Haplotypes at one mutational step from the core type 15-23-11-14 are denoted by an asterisk (*).

relatively little divergence on that chromosome during the past 700 years; and this makes it unlikely that the bulk of other Sykes haplotypes arose from the common 15-23-11-14 ancestor during that period. It is possible that some non–15-23-11-14 chromosomes come from other Sykes founders. However, since no other haplotype occurs at a frequency that is significantly elevated compared with that in the two control groups, this is not obvious from the data, although, since the numbers are relatively low, it cannot be excluded with high power. Since there is no record—nor reason—for widespread adoption of the Sykes surname, the third possibility must be entertained—that is, that these haplotypes are the result of the historical accumulation of nonpaternity events. If it is assumed that each non–15-23-11-14 chromosome, including the one-step derivatives, has infiltrated the Sykes genealogy as a single event, then the averaged nonpaternity rate estimate is 1.3%/generation, given the assumption that 23 generations have passed since the first common male ancestor (Sturges and Haggett 1987, p. 7). Of course, nonpaternity events involving other males with the core Sykes haplotype would not be detected. There are no reliable figures for nonpaternity rates in the modern population, and estimates from a number of studies fluctuate in a range of 1.4%–30%, although most are in the range of 2%–5% (MacIntyre and Sooman 1991, and references therein).

The frequency of *isonymy,* the marriage between individuals with the same surname, has been used as a crude method of estimation of the inbreeding coefficients within populations (Crow and Mange 1965; Crow 1980). However, it is acknowledged that the method relies on a number of assumptions, of which a monophyletic origin is one. The present study suggests that, in this case, the assumption may be valid, and similar studies with a range of surnames would indicate both the extent of monophyly and the accumulated nonpaternity rates, which could then be used to refine the isonymy method.

Although a few surnames are very common in any population, surveys in Britain and Switzerland have shown that most people have an uncommon surname (Lasker 1983; Barrai et al. 1996). "Sykes" is typical of indigenous English surnames, in being of low overall frequency but having marked local concentrations, presumably reflecting historical origins. If other surnames are shown to have an association with the Y-chromosome haplotype defined by these and other markers, there would be applications in genealogy and forensics whenever surnames follow a patrilineal pattern of inheritance. In genealogy, for example, it would be useful to investigate the link between branches with the same surname for which a common ancestor could not be established from written records. In forensics it could be used as a primary screen in cases, such as rape, in which a male leaves his Y chromosomes at the scene of the crime. The only comparable study showed a high frequency of a Y-chromosome haplotype in the Cohanim, the hereditary Jewish priesthood known to follow a strict patrilineal descent (Skorecki et al. 1997; Thomas et al. 1998).

Acknowledgments

We are grateful to the volunteers who took part in this study, to Vincent Macaulay for statistical help, to Trevor Ogden and George Redmonds for advice on surname origins, and to the Wellcome Trust for support.

References

Barrai I, Scapoli C, Beretta M, Nesti C, Mamolini E, Rodriguez-Larralde A (1996) Isonymy and the genetic structure of Switzerland. Ann Hum Biol 23:431–435

Crow JF, Mange AP (1965) Measurement of inbreeding from

the frequency of marriages between persons of the same surname. Eugen Q 12:199–203

Crow JF (1980) The estimation of inbreeding from isonymy. Hum Biol 52:1–14

de Knijff P et al. (1997) Chromosome Y microsatellites: population genetics and evolutionary aspects. Int J Legal Med 110:134–140

Kimura M, Ohta T (1978) Stepwise mutational model and distribution of allelic frequencies in a finite population. Proc Natl Acad Sci USA 75:2868–2872

Lasker GW (1983) The frequencies of surnames in England and Wales. Hum Biol 55:331–340

MacIntyre S, Sooman A (1991) Non-paternity and prenatal genetic screening. Lancet 338:869–871

Redmonds G (1992) Yorkshire surnames series. Vol 2. GR Books, Huddersfield UK

Skorecki K, Selig S, Blazer S, Bradman R, Bradman N, Warburton PJ, Ismajlowicz M, et al (1997) Y chromosomes of Jewish priests. Nature 385:32

Sturges CM, Haggett BC (1987) Inheritance of English surnames. Hawgood Computing, London

Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Golstein DB (1998) Origins of Old Testament priests. Nature 394:138–140