Institution: University of Leicester

Unit of Assessment: UoA5 Biological Sciences

Title of case study: Tools for analysing human Y-chromosome diversity: impact of DNA testing on the development of genetic genealogy and male-specific forensic analysis

1. Summary of the impact

This case study describes the societal and cultural impact of the development of DNA-based tools for distinguishing between different lineages of the human Y chromosome, which is male-determining and passed down from father to son. The availability of highly discriminating DNA markers has had two main impacts: (i) illumination of the link between the Y chromosome and patrilineal surnames, triggering the development of genetic genealogy, the investigation by the public of historical family relationships through DNA testing; and (ii) application of Y-DNA markers in forensic casework, with particular utility in rape cases where male and female DNAs are mixed.

2. Underpinning research

The Y chromosome determines male sex early in development, and is passed down from father to son largely without undergoing genetic recombination (the "reshuffling" process that affects the rest of the genome). Analysing Y-chromosomal DNA can therefore define highly informative paternal lineages, with applications in genealogical studies, population genetics and forensics.

Mark Jobling (1992-current) came to the Department of Genetics, University of Leicester, to develop a hypervariable Y-specific minisatellite [3.1] (the marker type used in Prof Sir Alec Jeffreys' DNA fingerprinting method) under an MRC Training Fellowship [grant 1] with Jeffreys as co-sponsor. This marker remains the only known example of a polymorphic minisatellite on the Y chromosome. Jobling and his MRC-funded PhD student (1992-96), Neale Fretwell, also contributed to an international collaboration to develop and validate a different kind of Y-specific marker, a set of 13 highly discriminating short tandem repeats (STRs; also known as microsatellites), the class of markers now used in forensic DNA profiling. This collaboration led to a paper [3.2] in the *International Journal of Legal Medicine* (with 501 citations, the greatest for any published in that journal), and to an accompanying influential review (cited 235 times [3.3]), setting out the potential role of Y-chromosome markers in forensic analysis. At the same time, Jobling and Fretwell themselves developed two novel STRs (known as DYS425 and DYS426) [3.4].

In 1998 these different marker types were applied in a landmark historical genealogy study – the Thomas Jefferson paternity case. Jefferson drafted the Declaration of Independence in 1776, became third U.S. President in 1801, and today is revered, featuring on the nickel, and among the four monumental heads on Mount Rushmore. But despite his avowal that "all men are created equal", Jefferson owned a total of over 600 slaves at his Monticello estate in Virginia. It had been persistently alleged that, following the death of his wife Martha, Jefferson had a relationship with one of these slaves, Sally Hemings, and that she bore up to six children by him. This remained the preserve of historical argument until Eugene Foster, a Virginia family historian, realised that analysis of Y-chromosomal DNA might throw new light on the story.

Foster traced male-line descendants of two of Hemings' sons, who gave DNA samples that were analysed in Oxford (binary polymorphisms – Zerjal, Tyler-Smith), Leiden (STRs – Mieremet, De Knijff), and Leicester (minisatellite – Jobling, then a Wellcome Trust Research Career Development Fellow [grant 2] and his WT-funded RA [1995-99], Paul Taylor). The conclusions were published in the interdisciplinary scientific journal *Nature*, in a 1998 paper [3.5] that has been cited 112 times. They showed that the male-line descendant of Eston Hemings Jefferson, Sally's last child, carried a Y chromosome exactly matching that of the descendants of Thomas's paternal uncle, Field Jefferson. This genetic evidence was therefore consistent with Thomas fathering Eston, and this explanation was accepted by a research committee appointed by the Thomas Jefferson Foundation. The implications of the Hemings-Jefferson relationship are now used to enrich the interpretation of Monticello (a National Historic Landmark and part of a UNESCO World

Heritage Site attracting ~450,000 visitors p.a.).

With the release of the complete Y chromosome sequence in 2003, it became possible to conduct systematic searches for novel STRs. This was done as an international collaboration [3.6] led by Manfred Kayser (Leipzig) in which 166 Y-STRs were identified and analysed together with groups in Leicester (Jobling with Andy Lee, Jobling's WT-funded technician [1999-2004], and Zoë Rosser his WT-funded post-doc [2001-06]), Helsinki and Oxford.

Jobling and Turi King (2000-current), then under a WT Prize PhD Studentship [grant 3] applied Y-STRs and binary markers to the analysis of the relationship between British surnames and Y chromosomes, showing the dependency of the probability of sharing Y chromosome type on the frequency of the surname, and demonstrating the feasibility of using Y profiles to predict surnames, given appropriate databases [3.7]. They later published detailed studies of 40 surnames, showing the importance of genetic drift, and low historical non-paternity rates [3.8].

3. References to the research (indicative maximum of six references)
(bold authors at University of Leicester, underlined are corresponding; positions and dates given where not already described in main text)

3.1 **Jobling MA**, **Bouzekri N**, **Taylor PG** (1998) 'Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (*DYF155S1*)'. *Hum Mol Genet*, **7**:643-53. doi: 10.1093/hmg/7.4.643; Bouzekri was Jobling's PhD student (1994-96)

3.2 Kayser M, Caglià A, Corach D, **Fretwell N**, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, **Jobling M**, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider P, M, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P Roewer L (1997) 'Evaluation of Y-chromosomal STRs: a multicenter study'. *Int J Legal Med,* **110**:125-33. doi: 10.1007/s004140050051; large international study coordinated by Kayser and Roewer (Berlin), where Jobling and Fretwell contributed Y-STR population data

3.3 **Jobling MA**, Pandya A, Tyler-Smith C (1997) 'The Y chromosome in forensic analysis and paternity testing'. *Int J Legal Med*, **110**:118-24; Tyler-Smith was Jobling's collaborator in Oxford, and Pandya was Tyler-Smith's PhD student. doi**:** 10.1007/s004140050050

3.4 **Jobling MA**, **Samara V**, Pandya A, **Fretwell N**, **Bernasconi B**, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L, Crawford MH, Tyler-Smith C (1996) 'Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males'. *Hum Mol Genet*, **5**:1767-75. doi: 10.1093/hmg/5.11.1767; study coordinated by Jobling with international collaborators supplying DNA samples. Samara (1995-96) was Jobling's MSc student and Bernasconi (1995-96) his Erasmus project student

3.5 Foster EA, **Jobling MA**, **Taylor PG**, Donnelly P, de Knijff P, Mieremet R, Zerjal T, Tyler-Smith C (1998) 'Jefferson fathered slave's last child'. *Nature*, **396**:27-8. doi:10.1038/23835

3.6 Kayser M, Kittler R, Erler A, Hedman M, **Lee AC**, Mohyuddin A, Mehdi S, **Rosser Z**, Stoneking M, **Jobling MA**, Sajantila A, Tyler-Smith C (2004) 'A comprehensive survey of human Y-chromosomal microsatellites'. *Am J Hum Genet*, **74**:1183-97. doi: 10.1086/421531

3.7 **King TE**, **Ballereau SJ**, Schürer K, **Jobling MA** (2006). 'Genetic signatures of coancestry within surnames'. *Curr Biol*, **16**:384-8. doi: 10.1016/j.cub.2005.12.048. Ballereau was Jobling's WT-funded post-doc (2004-09)

3.8 **King TE**, **Jobling MA** (2009) 'Founders, drift and infidelity: the relationship between Y chromosome diversity and patrilineal surnames'. *Mol Biol Evol*, **26**:1093-102. doi: 10.1093/molbev/msp022

Research Grants

1. **M. Jobling**: *Genetic diversity among human Y chromosomes,* £66,000; 1992-95; **MRC Training Fellowship** (Sponsors: Prof Sir Alec Jeffreys, Prof Gabriel Dover)
2. **M. Jobling**: *Novel molecular markers for human Y chromosome diversity,* £285,193; 1995-99; **Wellcome Trust Career Development Fellowship** (Sponsor: Prof Sir Alec Jeffreys)
3. **M. Jobling**: *Surnames and genetic structure: a molecular analysis using Y-chromosomal DNA*

polymorphisms, £77,268; 2000-03**; WT Prize Studentship** (PhD student: Turi King)

4. **M. Jobling**: *What's in a name? Applying patrilineal surnames in forensics, population history and genetic epidemiology,* £251,226; 2007-11; **WT Project Grant** (post-doc: Turi King)

4. Details of the impact

1) Impact on development of genetic genealogy

As the first case in which a genealogical question was tackled using the Y chromosome, "…*the Jefferson-Hemings case created enormous publicity about DNA and its possible commercialisation and use*" [5.1], thus sparking into life 'genetic genealogy', in which the public have their DNA analysed to discover genealogical links. This activity is currently pursued via several companies. The largest is FamilyTreeDNA (649,309 tests carried out @ Sept 2013); at a conservative estimate of $100 per test, this equates to ~$65M turnover. On its website [5.2] this company offers a 'Matching Thomas Jefferson' test, and acknowledges that "*The Jefferson-Hemings DNA study, published in 1998, was one of the catalysts of what later became the field of Genetic Genealogy*." FamilyTree DNA now offers highly discriminating Y-chromosomal tests based on >100 STRs, including the original set of thirteen [3.2], plus DYS425 and DYS426 [3.4], and many of those developed in the 2004 study [3.6]. Other companies offering genetic genealogy services include Oxford Ancestors, Britains DNA, and Genebase.

Genetic genealogy has become a powerful driver for 'citizen science' and the engagement of the public in genetics; this can be seen in the thriving International Society of Genetic Genealogy (www.isogg.org), a non-commercial non-profit organization run by its members, and in the publication of papers in scientific journals by non-academics using online community data [e.g. 5.3]. Jobling & King published a review in which the public's engagement in genetic genealogy was acknowledged [5.4], and King, together with surname historians David Hey and George Redmonds, published a popular book on surnames and DNA [5.5]. A more recent popular book on the origins of surnames summarises the work of King & Jobling [5.6]. The surnames/genetics link has been a fruitful topic for public engagement events, including talks to family history societies and U3A (University of the Third Age) groups by King and Jobling; King speaks annually at 'Who Do You Think You Are?' (London Olympia), the world's biggest family history event.

Jobling & King's research into the relationships between surnames and Y-chromosomal variation contributed to a further funding application [grant 4], employing King as post-doc. They used surname-based sampling of modern populations to access past population structures [5.7], and to show high levels of Norwegian admixture in parts of western Britain colonized by Vikings. This work was described in a book for the layperson [5.8], was featured in a BBC Radio 4 programme about the genetics of the British [5.9], and has formed the basis of many talks to the public.

2) Impact on forensic DNA analysis

The development and validation of Y-STR markers has had an impact in the development of male-specific DNA profiling kits, and the application of Y-DNA analysis in forensic investigations. This activity is currently increasing, demonstrating a persistent impact into the current REF period.

The initial studies of Y-chromosomal markers were partly driven by a desire to apply them in forensic casework: in the UK (for example), 80% of serious offences and 98% of sexual assaults are committed by men, so Y-DNA tests are informative, and particularly useful in mixtures of DNA from assailant and victim. It is standard practice to enrich rape case swabs for sperm DNA by use of the differential lysis method, which preferentially destroys non-sperm cells from the victim. This allows conventional DNA profiling, which targets STRs from other parts of the genome (autosomal STRs), to give a profile from the assailant. However, when this fails, or where there are mixtures of other types (such as blood-blood, or blood-saliva), a highly discriminating male-specific test is an invaluable tool. Y-specific tests also have three specialised applications: to readily determine the number of assailants in multiple-rape cases; to give assailant-specific information when sexual assault is by a close male relative, and conventional DNA profiles of victim and assailant are therefore similar; and to refine familial searches, in which database profiles are sought that might belong to the first-degree relatives of perpetrators, and thereby lead to a suspect. As well as criminal cases, Y-DNA testing is applicable in so-called 'deficiency paternity testing'. For example,

when the alleged father of a male child is unavailable, the paternal uncle's Y-DNA profile can be compared with that of the child.

Development of the Y-STRs [3.2] set the scene for their use as forensic tools; initially this was via 'home-made' kits, but the market soon attracted large commercial suppliers of DNA profiling kits, in particular Applied Biosystems with their Y-filer kit (2004), containing 17 Y-STRs, and Promega Corporation, with PowerPlex Y (12 Y-STRs; 2002), now succeeded by the more discriminating PowerPlex Y-23 (23 Y-STRs; 2012). All of these kits are based on the core STR set validated in 1997 [3.2], with the addition of others from the systematic survey [3.6]. In a recent development, one Y-STR, DYS391, is now included in the conventional profiling kit PowerPlex Fusion (2012), together with 22 autosomal STRs, to provide preliminary male-specific information. Commercial profiling kits are optimised to give results from just 0.5ng of DNA, and Powerplex Y-23, for example, yields male-specific profiles in the presence of a 16,000-fold excess of female DNA.

Forensic scientists using Y-DNA profiling have benefited from the development of large online databases that allow the frequencies of profiles in particular populations to be estimated easily. As set out by Jobling [3.3], Y-DNA profiles are less informative than autosomal profiles, since the Y-STRs are not separated by recombination. Estimating the significance of a matching Y-profile requires a large database in which its frequency in a population of interest can be estimated. The online Y Haplotype Reference Database (YHRD; www.yhrd.org [5.10]) was constructed by Lutz Roewer and Sascha Willuweit in Berlin, with contribution of a total of 1842 profiles from 5 populations from Jobling and his WT-funded post-doc Elena Bosch (2000-02), and later his ESF-funded post-doc Emma Parkin (2003-06). This database now contains 114,256 Y-STR profiles from 851 globally distributed populations. Most of these profiles are based on the Yfiler kit, but shortly will be joined by a large dataset of profiles from the latest Powerplex Y-23 kit, in work coordinated by Roewer, to which Jobling and his Leverhulme-Trust-funded post-doc, Jon Wetton (2012-current) have contributed. Forensic applications of Y-DNA are increasing [5.11], and have included use in the exonerations of US prisoners via testing of archived case materials (e.g. A.B. Butler, Wilton Dedge, Raymond Towler; www.innocenceproject.org).

5. Sources to corroborate the impact

5.1 Weil F (2013) *Family Trees: a History of Genealogy in America*. Harvard University Press, Cambridge, Mass. & London. pp. 208-13

5.2 Family Tree DNA website: http://www.familytreedna.com/landing/matching-jefferson.aspx

5.3 Rocca RA, Magoon G, Reynolds DF, Krahn T, Tilroe VO, et al. (2012) 'Discovery of Western European R1b1a2 Y chromosome variants in 1000 Genomes project data: an online community approach'. *PLoS One,* **7**:e41634. doi: 10.1371/journal.pone.0041634

5.4 **King TE**, **Jobling MA** (2009) 'What's in a name? Y chromosomes, surnames and the genetic genealogy revolution'. *Trends Genet*, **25**:351-60. doi: 10.1016/j.tig.2009.06.003

5.5 Redmonds G, **King T**, Hey D (2011) *Surnames, DNA and Family History*. Oxford: Oxford University Press

5.6 McKie D (2013) *What's in a Surname?: A Journey from Abercrombie to Zwicker*. London: Random House Books. pp.102-3

5.7 **Bowden GR**, **Balaresque P**, **King TE**, Hansen Z, **Lee AC**, Pergl-Wilson G, Hurley E, Roberts SJ, Waite P, Jesch J, Jones AL, Thomas MG, Harding SE, **Jobling MA** (2008) 'Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in northwest England'. *Mol Biol Evol,* **25**:301-9. doi: 10.1093/molbev/msm255

5.8 Harding SE, **Jobling M**, **King T** (2010) *Viking DNA*. Countryvise Ltd., Nottingham; Harding is a collaborator from the University of Nottingham

5.9 BBC Radio 4 programme featuring Jobling on Viking ancestry: *British, More or Less*, first broadcast 27/7/11: http://www.bbc.co.uk/programmes/b012r6z8

5.10 Roewer L et al. (44 authors including **Jobling M**, **Bosch E**) (2001). 'Online reference database of Y-chromosomal short tandem repeat (STR) haplotypes'. *Forensic Sci Int***,** *118*:103-11. doi: 10.1016/S0379-0738(00)00478-3

5.11 Roewer L (2009) 'Y chromosome STR typing in crime casework'. *Forensic Sci Med Pathol*, **5**:77-84. doi: 10.1007/s12024-009-9089-5