JOURNAL OF FORENSIC SCIENCES

# Reliable and robust molecular sexing of the hen harrier (Circus cyaneus) using PCR-RFLP analysis of the CHD 1 gene



| | |
|---|---|
| Journal: | *Journal of Forensic Sciences* |
| Manuscript ID: | JOFS-11-639.R1 |
| Manuscript Type: | Technical Note |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Henderson, Anique; University of Amsterdam, Faculty of Science<br>Lee, Christine; Home, ; University of Central Lancashire, School of Forensic Sciences<br>Mistry, Vanisha; Queen Mary University of London,<br>Thomas, Martin<br>Iyengar, Arati; University of Central Lancashire, School of Forensic Sciences |
| Keywords: | forensic science, wildlife forensics, bird of prey, wildlife persecution, molecular sexing, chromo-helicase-DNA binding protein 1 |
| | |

# Reliable and robust molecular sexing of the hen harrier (*Circus cyaneus*) using PCR-RFLP analysis of the CHD 1 gene

Anique Henderson, [1,2] M.Sc., Christine Michelle Lee, [1] M.Sc., Vanisha Mistry, [1] M.Sc., Martin Derek Thomas, [1] B.Sc. (Hons), Arati Iyengar, [1] Ph.D.

[1] Department of Forensic & Investigative Sciences, University of Central Lancashire, Preston PR1 2HE, United Kingdom

[2] Faculty of Science, University of Amsterdam, 1090 GE Amsterdam, the Netherlands

Corresponding author:

Arati Iyengar

Department of Forensic & Investigative Science

University of Central Lancashire

Preston PR1 2HE, United Kingdom

Fax: +44 1772 894981

Email: aiyengar@uclan.ac.uk

*This work was supported by the School of Forensic & Investigative Sciences, University of Central Lancashire, and the Faculty of Science, University of Amsterdam.

**ABSTRACT**

The hen harrier (*Circus cyaneus*) is a bird of prey that is persecuted in the UK and there is a need for a DNA-based individual identification and sexing system for use in forensic investigations. This study reports a new set of PCR primers for the chromo-helicase-DNA binding protein 1 gene which allows sexing using PCR-RFLP. Instead of exonic primers which amplify across a large intron, this set consists of a primer within the intron, enabling reduction in amplicon sizes from 356 bp to 212 bp and 565 bp to 219 bp in W and Z chromosomes. DNA degradation and dilution experiments demonstrate that this set is significantly more robust than one that amplifies across the intron and sequencing of the intronic primer binding region across several individuals shows that it is highly conserved. Whilst our objective is to incorporate this primer set into an STR-based individualisation kit, it may in the meantime prove useful in forensic or conservation studies.

**Keywords**: forensic science, wildlife forensics, bird of prey, wildlife persecution, molecular sexing, chromo-helicase-DNA binding protein 1

The hen harrier (*Circus cyaneus,* Falconiformes, Accipitridae, Circinae) has had a history of persecution in parts of the UK because its diet includes red grouse (*Lagopus lagopus*) that is managed on grouse moors for hunting (1,2). Despite being protected under UK legislation, illegal killing of hen harriers continues and numbers remain low in many regions (3). In order to establish a reliable forensic tool for use in cases of illegal persecution, there is a need for a DNA-based individual identification and sexing system. Development and validation of a STR based multiplex is currently underway in this laboratory but we present here a new set of PCR primers which can reliably sex hen harrier samples using PCR-RFLP.

Molecular sexing of non ratite birds has most commonly been achieved using PCR amplification of the sex linked chromo-helicase-DNA binding protein 1 (CHD 1) gene, where primers within conserved exons amplify across introns that differ in size on the avian sex chromosomes (Z and W). Upon size separation, females being heterogametic (ZW), display two bands of different sizes while males being homogametic (ZZ), display only one band. The P2/P8 (4) and the 1237L/1272H (5) primer sets amplifying across a CHD 1 intron have been used successfully to sex species across several avian orders. However, Ito et al. (6) reported that amplification across this intron failed to produce size differences in six species of raptors within Accipitridae, including the marsh harrier (*Circus spilonotus*), a species closely related to the hen harrier. We consequently used primers 2550F and 2718R described by Fridolfsson & Ellegren (7) which have also been successfully used to sex diverse species from 11 avian orders and amplify products of 400-450 (W) and 600-650 bp (Z) across a different CHD 1 intron. After sequencing homologous hen harrier sequences, we redesigned a set of primers (one within a highly conserved region of the intron) that amplifies smaller W (212 bp) and Z (219 bp) fragments. Due to smaller sizes of products, this new primer set is significantly more robust for sexing hen harrier samples containing degraded and/or very low amounts of DNA such as those likely to be encountered in forensic casework. We demonstrate this by carrying out DNA degradation and dilution experiments followed by PCR. We also evidence that the intronic primer binding region is conserved in

hen harriers across a wide geographical range within the UK by carrying out sequencing of several hen harrier individuals across this region.

**Methods**

DNA was extracted using 25 mg of muscle tissue from a male and a female bird that had died of natural causes in England using the QIAamp DNA Mini kit (QIAGEN, Hilden, Germany) following manufacturer's instructions, and RNA removed by the addition of excess RNase A (5 µl of a 10mg/ml solution of ≥70 Kunitz units/mg protein, Sigma-Aldrich, St. Louis, MO, USA) into 100 µl of DNA solution followed by incubation at 37 °C for 30 min. RNA-free genomic DNA was then purified using a QIAquick column (QIAGEN) following manufacturer's instructions and quantified using a UV spectrophotometer for use in the degradation and dilution experiments described below. DNA from a number of buccal swabs taken from male and female nestlings by licensed professionals during routine monitoring exercises at locations across England, Wales, Scotland and the Isle of Man was extracted using the QIAamp DNA Micro kit optimised for small sample amounts (QIAGEN) following manufacturer's instructions and quantified using 1 % ethidium bromide stained agarose gels containing known concentrations of lambda DNA (Promega, Madison, WI, USA) . Gel based quantification was used for these samples since they were not used for the degradation and dilution experiments and were quantified for PCR purposes only.

PCRs were carried out in a 10 µl volume containing 10-20 ng DNA (except dilution series), 3.0 mM $MgCl_2$, 1X ReddyMix PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA). PCR cycling conditions were as follows: initial denaturation at 94°C for 3 min, followed by 35 cycles of 94°C for 30 s, 52°C for 60 s (30 s for 212/219 bp products) and 72°C for 90 s (30 s for 212/219 bp products), followed by a final extension at 72°C for 10 min. Agarose gel electrophoresis (2 %) and ethidium bromide staining was subsequently carried out to visualise the products. PCRs were scaled up to 25 µl when purified products were required for sequencing. Male PCR products were purified using the QIAquick PCR Purification Kit

(QIAGEN) while for female samples, the Z and W amplicons were separated on a 3 % agarose gel followed by ethidium bromide staining, excision of bands, and purification of DNA using the QIAquick Gel Extraction Kit (QIAGEN). Sequencing was carried out using an ABI3500 using BigDye Terminator v3.1 Cycle Sequencing (Applied Biosystems, Carlsbad, CA, USA).

Quantified DNA from male and female tissue samples was artificially degraded by setting up 20 µl reactions at 37 °C containing ~ 900 ng DNA, 1 unit DNase I (Applied Biosystems) and 1 X DNase I buffer. A time series was carried out for 0, 5, 10, 20, 30, 60, 90 and 180 min (no DNase added to the 0 min sample). Reactions were terminated by adding 0.1 volume DNase inactivation reagent (Applied Biosystems), incubated for 5 min at room temperature, centrifuged at 10,000 x g for 1.5 min, and the supernatant transferred to a fresh tube. Degraded samples were then diluted to 20, 10, 5, 2, 1, 0.5, 0.25 and 0.1 ng/µl and used for triplicate PCRs followed by 2 % agarose gel electrophoresis.

For RFLP analysis, restriction enzyme digestion was carried out using 8 µl PCR product, 1 X restriction enzyme buffer and 10 units Sfc I enzyme (New England Biolabs, Ipswich, MA, USA) followed by incubation at 37 °C for 90 min. The enzyme was then heat inactivated at 65 °C for 20 min prior to 2 % agarose gel electrophoresis.

**Results and Discussion**

Using the Fridolfsson & Ellegren 2550F and 2718R primers, amplification of the 450 bp W fragment was consistently observed but amplification of the larger 650 bp Z band was highly inconsistent. Hen harrier W and Z fragments were consequently sequenced (GenBank accession numbers FJ415319 and FJ415320) and species specific primers (*HHCHD1FOR*: 5′ AGCAGAGTATCTGAAGTATCG 3′, *HHCHD1REV*: 5′ TCAATTCCCCTTTTATTGATCC 3′) designed in exonic regions. Exon/intron boundaries were defined using homologous sequences from marsh harrier (GenBank accession numbers AB112946 and AB112954).

These new primers, amplifying products of 356 and 565 bp in W and Z genes respectively, resulted in consistent amplification of both fragments in good quality samples. However, in forensic samples which have degraded and/or very low amounts of DNA, amplification of these fragments is likely to be problematic since several studies have shown that there is a significant negative correlation between maximum amplicon size and the level of degradation of the sample (e.g. 8,9). Furthermore, Bantock et al. (10) have demonstrated that reducing fragment size from 300–400 bp using the P2/P8 primer set to 200–250 bp using new redesigned primers greatly enhanced success rates in sexing museum bird specimens.

Since redesigning primers in the exonic regions to reduce amplicon size was not possible due to the large size of the introns in hen harriers (W: 288 bp, Z: 497 bp), homology within intron sequences was investigated by aligning the Z and W sequences using Bioedit 7.0.5.3 (11). Sequences were found to be highly divergent except for a region of high homology in the last ~ 200 nucleotides prior to the binding site of the reverse primer (data not shown). It was possible to design a new forward primer (*HHCHD1RFLPFOR*: 5′ AGACTGGCAATTACTATATGC 3′) within this region which in combination with the *HHCHD1REV* primer amplified products of 212 and 219 bp in W and Z sequences respectively. In the *HHCHD1RFLPFOR* primer binding sequence, two substitutions were observed between the W and Z sequences (G → T at position 7 and A → C at position 13 in Z sequence, see Fig. 1), but since they are internally situated on the primer, their effects on PCR amplification are considered minimal, as evidenced in this study and previously reported by others (12). However, given the intronic location, in order to ensure no additional substitutions, we sequenced this region in 10 males (ZZ) and 10 females (ZW) from across a wide geographical range in the UK. Results revealed only these two substitutions across all individuals, suggesting a highly conserved region (Fig. 1).

To compare amplification of the larger 356 and 565 bp fragments with that of the smaller 212 and 219 bp W and Z fragments, a degradation and dilution series of male and female DNA was used for PCR amplification, and results from triplicate PCRs are shown in Table 1. In females, with the *HHCHD1FOR* and *HHCHD1REV* primer set, presence of both 356 and 565 bp fragments was scored as a '1' while absence of one or both bands was scored as a '0'. With the *HHRFLPFOR* and *HHCHD1REV* primer set, since the 212 and 219 bp fragments were indistinguishable on a 2 % agarose gel, the presence of a band was scored as a '1' and absence of a band as a '0'. We carried out Mann-Whitney *U*-tests using SPSSv17 (13) to evaluate amplification success between the two primer sets and observed that *HHRFLPFOR* and *HHCHD1REV* performed significantly better than *HHCHD1FOR* and *HHCHD1REV*. The results are summarised in Table 2. When all template amounts and all degradation time points were considered (5 – 180 min i.e. 0 min time point excluded), results were significant when just male data were included, and highly significant when just female or both male and female data were included (test 1). Similar results were obtained when template amounts < 2 ng and all degradation time points (5 – 180 min) were considered (test 2). However, when template amounts > 2 ng and all degradation time points (5 – 180 min) were analysed, results were not significant when just male data was included, but significant when just female or both male and female data were included (test 3). This difference between the sexes when higher template amounts are used is most likely because with *HHCHD1FOR* and *HHCHD1REV* primers, amplification of a single band in males is more probable than amplification of two bands in females when DNA is degraded.

Since the 7 bp size difference between the W (212 bp) and the Z (219 bp) fragments is difficult to detect using ordinary agarose gels and any high resolution gel electrophoresis will necessarily require significantly longer running times and/or additional expense, we investigated the possibility of a RFLP within this region using WebCutter 2.0 (© Max Heiman 1997). Sfc I enzyme (restriction site: C▼TRYAG) was chosen because it cuts once within the Z sequence resulting in two fragments of 52 and 167 bp, but does not cut within the W

sequence. RFLP patterns should therefore, consist of three bands of 52 bp, 167 bp, and 212 bp in females and two bands of 52 and 167 bp in males. In order to confirm that the Sfc I restriction site is conserved across hen harriers from a wide geographic range within the UK, we carried out PCR-RFLP analysis using DNA from 10 male and 10 female individuals (1 tissue and 9 buccal swab samples from both sexes) from locations across England, Wales, Scotland and the Isle of Man. RFLP patterns from a few of these individuals are shown in Fig. 2 where three bands of expected sizes in females (lanes 3 & 5) and two bands of expected sizes in males (lanes 7, 9 & 11) are visible. We obtained 100% unambiguous RFLP patterns across all individuals providing good evidence that the Sfc I restriction site is conserved across hen harriers in the UK. Our results also indicate that the two internal substitutions within the *HHRFLPFOR* primer binding sequence do not significantly affect amplification of the 212 and 219 bp W and Z products.

In conclusion, the *HHRFLPFOR* and *HHCHD1REV* primer set described here will be highly useful in sexing hen harriers when samples contain DNA of poor quality and low quantity. Our ultimate objective is to incorporate this primer set into a STR multiplex currently being developed for individualisation but in the meantime, this primer set which allows sexing using either size separation or RFLP, could prove useful for forensic or conservation purposes. This method is also likely to work in the marsh harrier because sequences at primer binding and restriction enzyme recognition sites for both species were identical (data not shown).

*Acknowledgements*

We are very grateful to Rob Ogden (Edinburgh zoo) and Steve Downing (Wildlife Consultant) for assistance with setting up this project. Tissue samples came from birds that had died of natural causes and were recovered by Steve Downing. We are also very grateful to the following people for providing buccal swabs from chicks: Alan Leitch, Andrew Sanderman, Brian Etheridge, B. Ribbands, Bill Taylor, Chris Sharpe, Dave Sowter, E.R. Meek, Geoff Sheppard, Ian M. Spence, J. A. L. Roberts, Jim Williams, Steve Downing, Steve

Murphy, E.J. Williams. Finally, we wish to thank two anonymous referees for useful comments.

**References**

1. Etheridge B, Summers RW, Green RE. The effects of illegal killing and destruction of nests by humans on the population dynamics of the hen harrier *Circus cyaneus* in Scotland. J App Ecol 1997;34:1081–1105.

2. Thirgood S, Redpath S. Hen harriers and red grouse: science, politics and human wildlife conflict. J App Ecol 2008;45:1550–1554.

3. Sim IMW, Dillon IA, Eaton MA, Etheridge B, Lindley P, Riley H et al. Status of the hen harrier *Circus cyaneus* in the UK and Isle of Man in 2004, and a comparison with the 1988/89 and 1998 surveys. Bird Study 2007;54:256-267.

4. Griffiths R, Double MC, Orr K, Dawson RJG. A DNA test to sex most birds. Mol Ecol 1998;7:1071–1075.

5. Kahn NW, St. John J, Quinn TW. Chromosome-specific intron size differences in the avian CHD gene provide an efficient method for sex identification in birds. Auk 1998;115:1074–1078.

6. Ito H, Sudo-Yamaji A, Abe M, Murase T, Tsubota T. Sex identification by alternative polymerase chain reaction methods in Falconiformes. Zool Sci 2003;20:339–344.

7. Fridolfsson A, Ellegren H. A simple and universal method for molecular sexing of non-ratite birds. J Avian Biol 1999;30:116–121.

8. Sefc KM, Payne RB, Sorenson MD. Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. Auk 2003;120:982–989.

9. Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. J For Sci 2003;48:1054–1064.

10. Bantock TM, Prys-Jones RP, Lee PLM. New and improved molecular sexing methods for museum bird specimens. Mol Ecol Res 2008;8:519–528.

11. Hall TA. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999;41:95–98.

12. Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sninsky JJ. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucl Acid Res 1990;18:999–1005.

13. Green SB, Salkind N J. Using SPSS for Window and Macintosh: Analyzing and understanding data. 5th ed. Upper Saddle River, New Jersey: Prentice Hall, 2007.

Additional information and reprint requests:

Arati Iyengar, Ph.D.

School of Forensic & Investigative Sciences

University of Central Lancashire

Preston PR1 2HE

United Kingdom

E-mail: aiyengar@uclan.ac.uk

Table 1

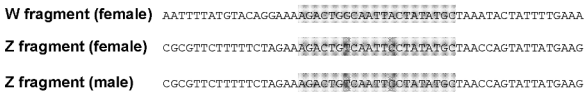
Results of the DNA degradation and dilution experiment showing amplification success in triplicate PCRs. '1' indicates successful amplification and '0' indicates lack of amplification.

| | **Amount of DNA** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 20 ng | 10 ng | 5 ng | 2 ng | 1 ng | 0.5 ng | 0.25 ng | 0.1 ng |
| **Degradation time** | | | | | | | | |
| Male, *HHCHD1FOR* + *HHCHD1REV* (565 bp) | | | | | | | | |
| 0 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 5 min | 111 | 111 | 011 | 111 | 000 | 101 | 000 | 000 |
| 10 min | 111 | 111 | 111 | 100 | 100 | 000 | 000 | 011 |
| 20 min | 111 | 111 | 111 | 111 | 111 | 011 | 010 | 011 |
| 30 min | 111 | 111 | 111 | 100 | 000 | 000 | 000 | 000 |
| 60 min | 111 | 111 | 100 | 001 | 000 | 000 | 000 | 000 |
| 90 min | 111 | 111 | 000 | 100 | 000 | 100 | 000 | 000 |
| 180 min | 100 | 010 | 000 | 000 | 000 | 000 | 000 | 000 |
| Male, *HHRFLPFOR* + *HHCHD1REV* (219bp) | | | | | | | | |
| 0 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 5 min | 111 | 111 | 111 | 111 | 111 | 111 | 110 | 110 |
| 10 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 100 |
| 20 min | 111 | 111 | 111 | 111 | 111 | 111 | 011 | 100 |
| 30 min | 111 | 111 | 111 | 111 | 110 | 100 | 101 | 000 |
| 60 min | 111 | 111 | 101 | 001 | 111 | 001 | 000 | 000 |
| 90 min | 110 | 111 | 000 | 001 | 000 | 100 | 000 | 000 |
| 180 min | 101 | 010 | 000 | 000 | 000 | 000 | 000 | 000 |
| Female, *HHCHD1FOR* + *HHCHD1REV* (356 bp, 565 bp) | | | | | | | | |
| 0 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 5 min | 111 | 111 | 111 | 111 | 000 | 111 | 000 | 000 |
| 10 min | 111 | 011 | 101 | 010 | 000 | 000 | 000 | 000 |
| 20 min | 010 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| 30 min | 000 | 100 | 000 | 000 | 000 | 000 | 000 | 000 |
| 60 min | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| 90 min | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| 180 min | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| *Female, HHRFLPFOR* + *HHCHD1REV* (212bp , 219 bp) | | | | | | | | |
| 0 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 5 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 10 min | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 20 min | 111 | 101 | 110 | 010 | 000 | 000 | 000 | 001 |
| 30 min | 110 | 101 | 101 | 001 | 000 | 100 | 000 | 000 |
| 60 min | 111 | 011 | 000 | 000 | 000 | 001 | 000 | 101 |
| 90 min | 110 | 001 | 000 | 000 | 000 | 000 | 000 | 000 |
| 180 min | 000 | 000 | 000 | 000 | 000 | 100 | 000 | 000 |

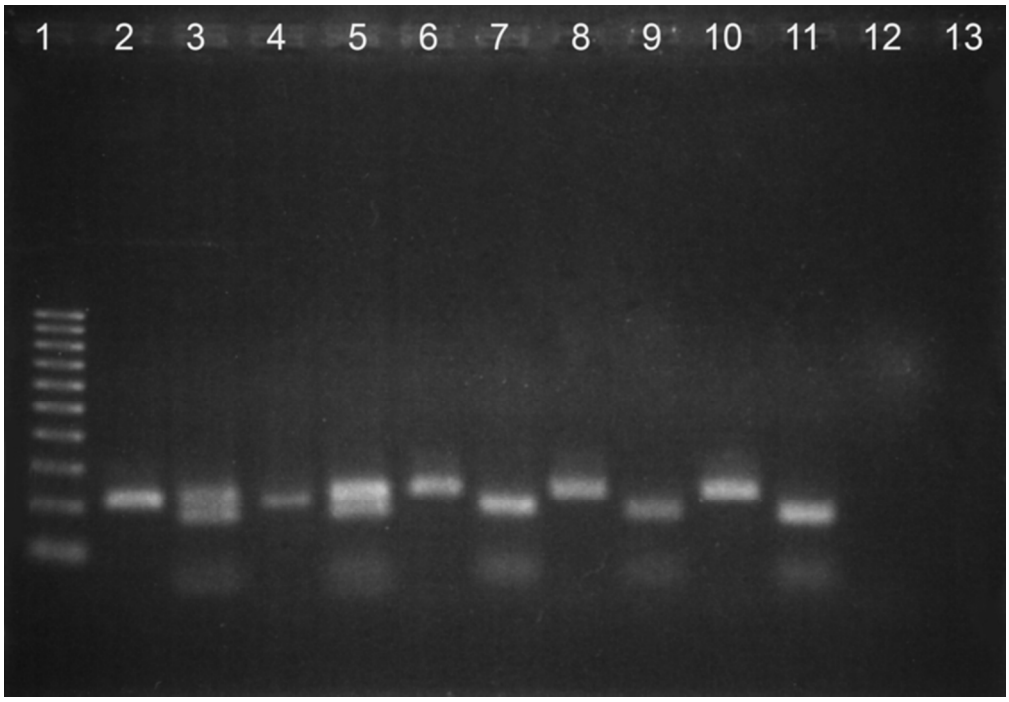
Table 2

Mann-Whitney-*U*-test 2-tailed *p* values obtained in comparisons between the primer sets for various combinations of data (significant values are in bold).

| Test | | Male | Female | Male and female |
|---|---|---|---|---|
| 1 | All template amounts (0.1 ng – 20 ng) All degradation time points (5 min – 180 min) | **0.028** | **0.000** | **0.000** |
| 2 | Template amounts < 2 ng All degradation time points (5 min – 180 min) | **0.004** | **0.000** | **0.000** |
| 3 | Template amounts > 2 ng All degradation time points (5 min – 180 min) | 0.514 | **0.010** | **0.039** |




209x297mm (300 x 300 DPI)




33x23mm (600 x 600 DPI)

Fig 1. Sequence showing the *HHRFLPFOR* primer binding site (highlighted in light grey) in W and Z fragments.

The primer binding site was identical in the W fragment from 10 females (2 from Orkney Islands, Scotland; 2 from SE Highlands, Scotland; 1 from NW Highlands, Scotland, 1 from Galloway, Scotland; 1 from Isle of Man; 2 from Wales; 1 from England), and in the Z fragment from 10 females (2 from the Orkney Islands, Scotland; 1 from SE Highlands, Scotland; 1 from NW Highlands, Scotland, 2 from Galloway, Scotland; 1 from the Isle of Man; 2 from Wales; and 1 from England) and 10 males (3 from the Orkney Islands, Scotland; 1 from SE Highlands, Scotland; 2 from NW Highlands, Scotland, 1 from Galloway, Scotland; 2 from the Isle of Man; and 1 from Wales). The two nucleotide substitutions between W and Z fragments are highlighted in dark grey.

Fig 2. PCR-RFLP patterns using Sfc I

Lane 1: 100 bp ladder; lanes 2,3: DNA from female tissue undigested, digested (location of sample: England); 4,5: DNA from female buccal swab undigested, digested (location of sample: Scotland, SE Highlands); 6,7: DNA from male tissue undigested, digested (location of sample: England); 8,9: DNA from male buccal swab undigested, digested (location of sample: Scotland, Galloway); 10,11: DNA from male buccal swab undigested, digested (location of sample: Isle of Man); 12,13: negative control.