Concerns about the case report, "Forensic Investigation of a Shawl Linked to the "Jack the Ripper" Murders," by Jari Louhelainen and David Miller

Chris Phillips

[Note that it is unclear whether the subject of this study is really a shawl, or even a garment, but for convenience it will be referred to as a shawl here.]

The provenance of the shawl

The evidence relating to the provenance of the shawl is outlined in Appendix A. In summary, it is extremely unlikely that the shawl came from the scene of Catherine Eddowes's murder, because:

(1) The only evidence to support this is a tradition in the family of PC Amos Simpson, first recorded a century after the murder, which stated incorrectly that Simpson discovered the body of Catherine Eddowes, and that he then picked up the shawl and kept it. According to a later version of the story, PC Simpson accompanied the body on its way to the mortuary and was given permission to take the shawl away by a senior police officer.

(2) Simpson belonged to the Metropolitan Police, whereas the murder of Catherine Eddowes took place in the territory of the City of London Police, a separate force. Simpson belonged to N Division, which did not even share a boundary with the City. Between 1881 and 1893, he was serving in Cheshunt, Hertfordshire, nearly 14 miles away from the murder site. There is no record of his having been assigned special duties which would have taken him outside his division.

(3) The events following the murder are well documented in contemporary records. The body was found by PC Edward Watkins and was not left unattended after its discovery. The body and its surroundings were closely examined and a drawing was made of it before its removal to the mortuary. A detailed inventory was then made of the clothing and possessions, including descriptions of the blood-stains on the clothes. The inventory does not include any item resembling the shawl. It is not credible that anyone would have been allowed to take away an item of the victim's clothing in the way that is suggested.

Presentation of results

The DNA sequence comparisons in Figure 7 are presented in a schematic form, in which the sequence variants are not identified. It is therefore impossible for the significance of the comparison to be independently assessed, or for any statistical calculations to be independently verified. It is particularly desirable that this should be done, in view of the errors made in the original presentation of the work and the subsequent failure to acknowledge or correct them (see Appendix B).

The schematic presentation tends to be misleading because the variants shown include some that are found in more than 90% of the population (see Appendix C), giving the impression of a significant match where none exists. The variants also include insertions/deletions not normally used in calculating probabilities for forensic purposes (see Appendix C).

The DNA sequences obtained from the shawl are incomplete, covering only parts of the hypervariable regions (see below and Appendix C). Therefore in Figure 7, in each case there are ranges for which information is missing, but this is not indicated. Where comparison is made with a DNA sequence from a living relation, the information in these ranges is also excluded, but again without any indication. This exaggerates the apparent significant of any matches found, as it gives

the impression of a match in the whole of the hypervariable regions, rather than only in limited parts of them, of unknown extent.

On page 9 of the paper, it is stated that the DNA sequence data had to be suppressed for legal reasons: "due to the restrictions set by the Data Protection Act, detailed nucleotide-level DNA information of living individuals should not be published." But the Data Protection Act covers only information relating to identifiable individuals, and therefore would not restrict the publication of the DNA sequence of the relation of Aaron Kozminski, who remains anonymous. Moreover, it can be presumed that Karen Miller, the relation of Catherine Eddowes, gave consent for details of her DNA sequence to be revealed, because what was believed to be an extremely rare sequence variant was published in the original edition of Russell Edwards's book.

Comparison of the "suspect" DNA sequences

A match is claimed between these two sequences in the abstract, in the introduction and in the conclusion of the paper.

But the comparison illustrated in Figure 7 shows two differences between the sequences. According to the guidelines for mitochondrial DNA analysis cited on page 9 of the paper: "If samples differ at two or more nucleotide positions (excluding length heteroplasmy), they can be excluded as coming from the same source or maternal lineage."[1] There is no suggestion of length heteroplasmy in this case. Therefore the conclusion should be that the sequences are unrelated, not that they match.

It may also be noted that the comparison was made over a limited part of the sequence, which did not include the whole of the hypervariable regions, much of HVR1 apparently being excluded (see Appendix C).

The differences between the sequences were apparently dismissed on the basis that at these points the sequence "could not be determined with high confidence" (page 6). But as the differences are included in the schematic illustration of results, this presumably represents a decision made after the sequences were compared, in the knowledge that they did not match at these points. Such post hoc selection of the evidence is not acceptable.

"Victim" DNA comparison

On page 5 it is stated that, for material obtained from the stains "judged to originate from blood or internal organ imprints," sequencing was regarded as successful for six out of eight mitochondrial DNA segments. It is also stated that "in general, only less than 100 bp reads could be achieved (at Phred20 quality)." As five of the eight segments used were longer than 100 bp, this suggests that some of the segments described as being successfully sequenced were in fact only partially sequenced.

In contrast, Dr Louhelainen's report quoted in the book states that six out of seven segments were successfully sequenced, but that two segments were considered to be the result of contamination, because they matched one of the control samples (see below).[2] Presumably this also implies that they did not match the DNA sequence obtained from the victim's relation, otherwise they might just as well have originated from the victim. (In fact, Dr Louhelainen's report quoted in the book states only that there was one segment which matched the relation's DNA and none of the controls.) The problem is that if there were such a post hoc exclusion of some segments from the comparison, the

1 Section 3.1.1 of the SWGDAM Interpretation Guidelines for Mitochondrial DNA Analysis by Forensic DNA Testing Laboratories, available at http://media.wix.com/ugd/4344b0_c5e20877c02f403c9ba16770e8d41937.pdf.
2 See Appendix D, part 1.

match between the remaining segments could not be taken at face value. Any statistical analysis would have to be modified to take the exclusion process into account.

The question of contamination

Concerns have been expressed about the possibility of the shawl having been contaminated before its examination. Dr Louhelainen had previously stated that "a maximum of six persons have handled [the shawl] in the past twelve months." (He argued that skin cells would not remain in the surface of the shawl for longer than a year because it is made from silk, which would behave differently from - for example - wool.) DNA samples were therefore taken from Mr Edwards and "the laboratory personnel who have been known to handle the shawl" for use as controls.[3] However, Figure 7 shows that only two control samples (from Mr Edwards and Dr Louhelainen) were actually compared with the DNA sequences obtained from the shawl.

Two segments of DNA were indeed considered to have been the result of contamination, because they matched one of the control samples, but there is a contradiction between the information given about this in the book and that in the paper.

In the book, these two segments are mentioned in Dr Louhelainen's account of the analysis of material obtained from six stains on the shawl (described by Mr Edwards as bloodstains) using his novel extraction technique.[4] In the paperback edition, the reference to these segments was removed from Dr Louhelainen's report, though not from Mr Edwards's commentary, and replaced with a comment that "the possibility of cross-contamination is near impossible due to the extraction method used."[5]

In contrast, on page 8 of the paper, the segments attributed to contamination are said to have been obtained from other sections of the shawl, near the corners, using traditional "double swab" sampling methods. And on page 5, it is stated that none of the sequences retrieved from the stains "judged to originate from blood or internal organ imprints" matched either of the control sequences.

Analysis of nuclear DNA

On page 8, it is stated that analysis of nuclear DNA recovered from the shawl suggested that it came from a male with brown hair and brown eyes. On page 9, these results are said to be "in full accordance with one of the very few witness statements considered reliable."

Regarding the eyewitness description, in fact there is no witness statement describing a suspect with brown eyes. From the account in the book, the witness referred to is Israel Schwartz, who gave the description "comp[lexion] fair hair dark, small brown moustache," without specifying eye colour.[6] (Note that there is no evidence that Schwartz was considered a particularly important witness; he was not called to give evidence at the relevant inquest.)

This conclusion stated on page 8 is different from the one published by Edwards (2014), which was that "our man definitely did not have red or blond hair, but that his hair colour was likely to be dark (brown or black)."[7] No information about eye colour or complexion was given in the book, but in an interview with Adam Rutherford for BBC Radio 4's Inside Science programme, Dr Louhelainen said that the hair and eye colour reported in the book matched "the most reliable witness statement

3 See Appendix D, part 1.
4 See Appendix D, part 1.
5 Edwards (2015), p. 205.
6 See Appendix D, part 2, and for Schwartz's statement Evans and Skinner (2001), p. 137. In Russell Edwards's book, the description was misquoted, to give "small dark moustache."
7 See Appendix D, part 2.

which is known." He added that "from the top of my head," the hair and eye colour had both been found to be brown.[8]

From the list of characteristics in Table 3, it is not clear how the eye colour could be determined to be brown, as the only references to eye colour are for "Blue" and "Non-blue" eyes.

In the paperback edition of the book published in 2015, the description of hair colour was the same as in the first edition, and there was again no information about eye colour or complexion. However, a new summary of the scientific findings by Dr Louhelainen was added, which said that "the hair, eye and skin colour could be determined. These characteristics perfectly matched the only reliable eyewitness statement available." But no details were given of these findings.[9]

Stains considered to be blood

On page 5, it is stated that some of the stains on the shawl are "considered to be blood." Dr Louhelainen has previously described them, in a summary of the scientific findings written for the paperback edition of Russell Edwards's book, as "stains which originated from blood."[10]

It appears that the only evidence suggesting these stains are blood came from a qualitative examination of fluorescence under UV illumination, though the characteristics of the fluorescence were not described.[11]

At the end of the paper, Dr S. McColl is acknowledged for help with presumptive blood testing, but no details of the results are given. In the book, presumptive tests for blood using the KM [Kastle-Meyer] method were said to have proved inconclusive.[12]

Stains described as semen

On pages 2 and 9, other stains on the shawl are described as semen. On page 5 they were said to have "followed the behaviour of semen stains under reflective UV light." According to the book, the technique used was not reflected UV, but a qualitative examination of fluorescence under UV illumination. Semen was suggested because the fluorescence was greenish, although it was acknowledged there that there were other possibilities.[13]

The cells extracted from the area of these stains have previously been described by Dr Louhelainen as looking "identical to semen associated cells both in size and morphology."[14] However, these were identified not as sperm cells but as squamous epithelial cells, which are normally only a minor component of semen. Although sperm cells would normally be present in much greater numbers than epithelial cells in semen, Dr Miller found none, and said he would have expected to find them if they had been present. He noted that squamous epithelial cells are also found in any bodily fluid

8 From an extended version of the interview, previously available at http://www.bbc.co.uk/podcasts/series/inscience (a copy of the MP3 file can be provided). The shorter version, broadcast on Radio 4 on 11 September 2014, two days after the publication of the book, is still available at https://www.bbc.co.uk/programmes/b04gc2y0.
9 Edwards (2015), pp. 292, 298. According to Schwartz's description the complexion was fair. It is not clear how a fair complexion could have been determined consistently with the hair colour being either brown, or else not red or blond but probably "dark (brown or black)." In the list of characteristics in Table 3, fair skin is found only in conjunction with red or blond hair.
10 Edwards (2015), p. 297.
11 Edwards (2014), Chapter 8 (Kindle location 2048).
12 Edwards (2014), Chapter 8 (Kindle location 2081).
13 Edwards (2014), Chapter 8 (Kindle locations, 2047, 2265). In the summary of the scientific findings written by Dr Louhelainen for the paperback edition, it was stated that the stains "were shown in routine testing to be aged semen stains." (Edwards, 2015, p. 297)
14 Edwards (2015), p. 297.

that washes over or bathes an epithelial surface.[15]

Appendix A: The provenance of the shawl

The family tradition

The family tradition about the shawl seems to have been first recorded in 1988 when David Melville Hayes, Amos Simpson's great-great-nephew, framed two pieces cut from it and added an inscription describing them as "Two silk samples taken from Catherine Eddowes' shawl at the time of the discovery of her body by Constable Amos Simpson in 1888 ..."[16]

In fact, Catherine Eddowes's body was discovered by PC Edward Watkins of the City of London Police (see below).

Some further details of the family tradition were published by Kevin O'Donnell in 1997, based on an interview with David Melville Hayes and his mother Elsie. Although mentally alert, Elsie knew little about the story, except that it was something to do with Jack the Ripper and Eddowes, that Simpson had been on duty at the time, and that her mother had claimed he told her to look after it as it might be valuable one day. David Melville Hayes was slightly confused as to how much of the story had come from his mother and how much from his grandmother. He said that Simpson was the first person to find the body of Catherine Eddowes, and had picked up the shawl that night and kept it. Simpson was supposed to have been on some kind of special duty, which Hayes assumed to mean that he was in plain clothes on surveillance duty with two or three other officers.[17]

In 2014 Russell Edwards published a different version of the story, in which Simpson with another policeman had accompanied the body of Catherine Eddowes as it was taken to the mortuary, asked a senior officer whether he could take it away, and was allowed to do so.[18]

Amos Simpson's police career

The murder of Catherine Eddowes on 30 September 1888 took place in Mitre Square in the City of London. Then, as now, the City of London had its own police force, with an entirely separate jurisdiction from that of the Metropolitan Police.

Amos Simpson was an officer of the Metropolitan Police Force, which he joined on 24 February 1868. On 31 August 1881 he was appointed an Acting Sergeant at Cheshunt police station in Hertfordshire. Cheshunt was then in Y Division, but was transferred to N Division three years later, and as a result Simpson was also transferred, on 26 July 1884. Simpson retired on 27 March 1893, still an Acting Sergeant in N Division, when his address was given as Jews [Gews] Corner, Cheshunt. A newspaper report of his death in 1917 stated that he had served for 13 years at St Pancras, and for 12 years at Cheshunt (the latter period corresponding to 1881-1893).[19]

There are no references in Police Orders to special duties for any N Division officers in 1888 before

15 See Appendix D, part 3.
16 O'Donnell (1997), p. 211. Despite the wording of the inscription, the pieces were not cut from the shawl until 1988.
17 O'Donnell (1997), pp. 213-215. Later, in conversation with Russell Edwards, David Melville Hayes "categorically stated" that Amos Simpson had been in Mitre Square on the lookout for Fenian terrorists (Edwards, 2014, Chapter 7; Kindle location 1722).
18 Edwards (2014), Chapter 7 (Kindle location 1664).
19 National Archives, MEPO 21/22, number 10850; MEPO 7/43, 30 August 1881, and MEPO 7/46, 28 July 1884, posted by Neil Bell at forum.casebook.org, 28 and 29 September 2014; Suffolk and Essex Free Press, 18 April 1917, quoted by O'Donnell (1997), p. 219.

October, or to plain clothes duties for Amos Simpson in 1888.[20]

Cheshunt, in Hertfordshire, where Amos Simpson was serving at the time of the murder of Catherine Eddowes, is nearly 14 miles from Mitre Square. Nor did Simpson's division - N Division - share a boundary with the City of London. The closest point was in Bethnal Green, with part of H Division (Whitechapel) lying between them.

The murder of Catherine Eddowes

The events following the discovery of Catherine Eddowes's body are documented in the depositions taken at her inquest. The body was discovered by PC Edward Watkins at 1.44am. Watkins remained with the body, sending a night-watchman employed in Mitre Square, George James Morris, to get assistance. Morris returned with PCs James Harvey and Frederick Holland. Holland then fetched Dr George William Sequeira, and members of the public were sent for other constables. Inspector Edward Collard arrived soon after 2am, when he found Dr Sequeira and several police officers there, but the body was not touched until Dr Frederick Gordon Brown reached the square about a quarter of an hour later.[21] All the police officers named in the depositions belonged to the City of London Police, not the Metropolitan Police.

Before the body was removed, Dr Brown made a detailed examination, and a drawing of it was made in situ by Frederick William Foster. The body's surroundings were also searched, and the finding of several small buttons, a thimble and a small mustard tin was recorded, these items being handed to Inspector Collard. The body was then conveyed to the mortuary, and there the clothes were carefully removed and a detailed inventory was made of all the victim's possessions, including notes of the locations of bloodstains on the garments.[22] None of the items listed in the inventory resembles the shawl.

It has been suggested by Russell Edwards that the shawl may have been removed by Simpson with the permission of a senior police officer before the inventory was taken, and that it was referred to in a press report which read "Her dress was made of green chintz, the pattern consisting of Michaelmas daisies."[23] In fact, a more detailed description of the dress appeared in other newspapers: "dark green chintz dress, with Michaelmas daisies, golden lily pattern," and was immediately followed by the phrase "three flounces." This was part of a "description of the deceased issued by the police authorities with a view to identification."[24] Clearly it is identical with the "Chintz Skirt three flounces, brown button on waistband" which appears in the inventory.[25] It must be different from the shawl, which is not green, not made of chintz, not a skirt or dress, and has no flounces, no buttons and no waistband. Furthermore, the flowers on the shawl don't resemble either Michaelmas daisies or lilies.

Appendix B: The previously published claim about a match between the "victim" DNA from the shawl and DNA from a descendant of Catherine Eddowes

In the first edition of Russell Edwards's book, Dr Louhelainen is quoted as claiming a match on the basis of an extremely rare sequence variant: "One of these amplified mtDNA segments had a sequence variation which gave a match between one of the shawl samples and Karen Miller's DNA only; i.e. the DNA sequence retrieved from the shawl did not match with control reference

20 Information posted by Neil Bell at forum.casebook.org, 29 September 2014.
21 Evans and Skinner (2001), pp. 224, 225, 228, 236.
22 Evans and Skinner (2001), pp. 225-228.
23 Edwards (2014), Chapter 5 (Kindle location 1135).
24 See, for example, Pall Mall Gazette, 1 October 1888.
25 Evans and Skinner (2001), p. 226.

sequences. This DNA alteration is known as global private mutation (314.1C) and it is not very common in worldwide population, as it has frequency estimate of 0.000003506, i.e. approximately 1/290,000. This figure has been calculated using the database at Institute of Legal Medicine, GMI, based on the latest available information."[26] This was the only information given in the book about the statistical significance of either of the claimed matches between mitochondrial DNA sequences.

In fact, as confirmed by a subsequent article in the Independent newspaper,[27] the sequence variant in question is common to almost the whole population. The claim made by Dr Louhelainen appears to have been the result of a series of errors and oversights:

(1) 314.1C was an error of nomenclature for 315.1C, since the convention is for an insertion to be indicated at one end of a homopolymeric tract (the 3' end of the light strand) and not at an interior point. According to the current version of the EMPOP database maintained by the Institute of Legal Medicine at Innsbruck (the same database used by Dr Louhelainen), the frequency of this sequence variant is more than 90%.[28]

(2) The error of nomenclature is assumed to be the main reason for the mistaken conclusion that this sequence variant was rare, though it is not clear how this would have produced the estimated frequency of approximately 1/290,000. As the EMPOP database used by Dr Louhelainen then contained only around 29,000 sequences, it is suggested that an arithmetical error reduced the frequency estimate by a further factor of ten.

(3) Dr Louhelainen described the sequence variant as a "global private mutation." This is a classification used by the application HaploGrep, which is used to determine the haplogroups of mitochondrial DNA sequences, and is defined by the HaploGrep user manual as a "mutation never observed in Phylotree, probably due to inconsistent alignments, phantom mutations or point heteroplasmies."[29] But Dr Louhelainen appears to have assumed that the term signified only a rare sequence variant, rather than an error or anomaly.

(4) Despite Dr Louhelainen's statement that the sequence variant in question was not found in the control sequences, it appears that it was found in every DNA sequence examined in this study, including the two control sequences (see Appendix C). This is only to be expected, as it is found in more than 90% of all sequences. Presumably the failure to appreciate this was also a result of the error of nomenclature.

The error in the original report, although it was identified the month after publication, has never been acknowledged or corrected by Dr Louhelainen. However, the claim of a rare mutation was silently omitted from the paperback edition of Russell Edwards's book, which was published in 2015.

Appendix C: The mitochondrial sequence previously found to match "suspect" DNA from the shawl

The matching sequence

26 See Appendix D, part 1.
27 Independent, 19 October 2014, available at https://www.independent.co.uk/news/science/jack-the-ripper-id-hinges-on-a-decimal-point-as-scientists-flag-up-dna-error-in-book-that-claims-to-9804325.html. The article, by the late Steve Connor, then Science Editor of the Independent newspaper, quoted criticisms by Professors Sir Alec Jeffreys, Mannis van Oven and Walther Parson, and Dr Hansi Weissensteiner.
28 https://empop.online/.
29 http://haplogrep.uibk.ac.at/docs/UserManual-HaploGrep.pdf.

In the first edition of Russell Edwards's book, Dr Louhelainen was said to have compared sequences obtained from the shawl with "the collection of DNA databases stored at the National Center for Biotechnology Information (NCBI), based in Bethesda, USA" and to have found "one perfect hit." This had been added to the database a few weeks earlier, was for the haplotype [haplogroup] T1a1, and came from a person whose ethnicity was recorded as "Russian".[30]

This is a reference to the NCBI's GenBank database. At the time of publication, there was only one sequence in the database fitting this description. It had been submitted directly by Family Tree DNA Ltd on 24 April 2014, and last modified in the database on 7 May 2014 (that is, four months before the publication of the book). Its accession number is KJ748572.[31] It originated from a Gentile Russian family.[32]

By comparison with the Revised Cambridge Reference Sequence, the following are the sequence variants in the parts of the hypervariable regions examined in the paper (namely, in HVR1, 15998-16400, and in HVR2, 49-357):
HVR1:
16126C, 16163G, 16186T, 16188T, 16189C, 16294T
(The fourth and fifth variants may alternatively be expressed as 16188-, 16193.1C.)
HVR2:
73G, 152C, 263G, 309.1C, 315.1C[33]

Implications

Assuming that each of the blocks in Figure 7 represents a single sequence variant, and comparing with the variants found in the matching sequence:

(1) It will be seen that for the "suspect" sequence obtained from the shawl in Figure 7, while five variants in HVR2 are shown, the same as in the matching sequence, only two in HVR1 are shown, compared with six in the matching sequence. This implies that the "suspect" sequence was incomplete, lacking a large part of HVR1.

(2) The fifth sequence variant in HVR2 is the common one 315.1C, which was misinterpreted as a rare variant in the original presentation of this work. Figure 7 shows that this was found in every DNA sequence examined, including the two controls.

(3) The sequence variants in HVR2 include two (263G and 315.1C) found in more than 90% of the population, and two others (73G and 309.1C) found in more than 50%. The insertion at 309 is one of the variants which is disregarded by default during searches in the EMPOP forensic database.

Appendix D: Extracts from previous presentations of this work

[For copyright reasons, the extracts from Russell Edwards's books originally quoted here have been replaced by references to page numbers.]

(1) Comparison of "victim" DNA sequences

The first edition of Russell Edwards's book "Naming Jack the Ripper", published in 2014, included

30 See Appendix D, part 4.
31 The sequence and associated information can be found at https://www.ncbi.nlm.nih.gov/nuccore/KJ748572.1.
32 Personal communication from the genealogist who submitted the sequence.
33 https://www.familytreedna.com/public/T1/default.aspx?section=mtresults, checked against the results of the sequence checker at http://www.ianlogan.co.uk/checker/checker6.htm.

in Chapter 9 the following summary written by Dr Louhelainen. It was included because it was thought that it might be of interest to scientists reading the book:

[Pages 204-206]

In the paperback edition of the book, published in 2015, the passages indicated in blue were omitted, the second being replaced by the following:

[Pages 204, 205]

(2) Analysis of nuclear DNA

From the Conclusion of "Naming Jack the Ripper" (2014):

[Pages 293, 294]

(3) Microscopic examination of stains described as semen

Chapter 9 of "Naming Jack the Ripper" (2014) includes an email by Dr Miller summarising the results of his examination:

[Page 194]

(4) Match of the "suspect" DNA from the shawl to a full sequence in a database

From the Conclusion of "Naming Jack the Ripper" (2014):

[Pages 291, 292]

Bibliography

Russell Edwards, "Naming Jack the Ripper" (Sidgwick and Jackson, 2014).

Russell Edwards, "Naming Jack the Ripper" (Pan Books, paperback edition, 2015).

Stewart Evans and Keith Skinner, "The Ultimate Jack the Ripper Sourcebook" (Robinson, 2001).

Kevin O'Donnell, "The Jack the Ripper Whitechapel Murders" (Ten Bells, 1997).