11-1-2017

Case Report: Coincidental Inclusion in a 17-Locus Y-STR Mixture, Wrongful Conviction and Exoneration

Greg Hampikian
Boise State University

Gianluca Peri
Boise State University

Shih-Shiang Lo
Taiwan Association for Innocence

Mon-Hwa Chin
National Chiao Tung University School of Law

Kuo-Lan Liu
Criminal Investigation Bureau

Publication Information

Hampikian, Greg; Peri, Gianluca; Lo, Shih-Shiang; Chin, Mon-Hwa; and Liu, Kuo-Lan. (2017). "Case Report: Coincidental Inclusion in a 17-Locus Y-STR Mixture, Wrongful Conviction and Exoneration". *Forensic Science International: Genetics, 31*, 1-4. https://doi.org/10.1016/j.fsigen.2017.08.004

Case Report: Coincidental Inclusion in a 17-locus Y-STR Mixture, Wrongful Conviction and Exoneration

Author names and affiliations:

Greg Hampikian [a], Gianluca Peri [a], Shih-Shiang Lo [b], Mong-Hwa Chin [c], Kuo-Lan Liu [d].

[a] Boise State University, 1910 University Dr., Boise, ID 83725, USA

[b] Taiwan Association for Innocence, 10F, No.75, Sec.3 Roosevelt Rd., Taipei City 10646, Taiwan

[c] National Chiao Tung University School of Law, 1001 University Rd., Hsinchu City 30010, Taiwan

[d] Criminal Investigation Bureau, No.5, Lane. 553, Chung Hsiao, East Road Section 4, Xinyi area, Taipei, Taiwan 11072, R.O.C.

Abstract:

We report the case of a suspect (Suspect-3) who was convicted (and later exonerated) of participating in the multiple-attacker rape of two women. The forensic evidence against him was his inclusion in a 17-marker Y-STR mixture isolated from semen on one victim's clothing. The DNA inclusion produced a match statistic with a combined probability of inclusion of 1 in 741, and a Likelihood Ratio of 3296. While the defense team was told that Suspect-3 was included in the semen DNA mixture, they were not told that all of the Y-STR alleles could also be explained by just the other two accused attackers' haplotypes. Suspect-3 was subsequently freed after the Taiwan Association for Innocence requested re-examination of the incriminating mixed DNA sample. The Criminal Investigation Bureau was then able to exclude him using an extended set of Y-STR markers (23 loci), leading to his exoneration.

Keywords: Coincidental Inclusion, Y-STR, DNA Mixture, Complex DNA, Exoneration, Innocence.

Case Timeline:

Suspect-3 and two co-defendants were accused of raping two women on May 25th, 2009. The victims arrived at Suspect-3's rental apartment that day at 3 in the morning. Suspect-3 maintained that he left his apartment to pick up his wife from work right after the victims arrived, and that he was not involved in the crime. The testimony of Suspect-3's wife confirmed that he came to her workplace to pick her up, and the record from her workplace supported her testimony that she punched-out at 4:13. Also, a co-defendant testified that Suspect-3 left around 4, but he was not able to identify the exact time. None of the witnesses accused Suspect-3 of taking part in the crime.

Still, the court convicted Suspect-3 of rape. The key evidence that led to Suspect-3's conviction was a 17-locus Y-STR DNA test from the mixed DNA sample collected from semen on the underpants of one of the victims. That test led to a report concluding that Suspect-3 "or men who share the same paternal line cannot be excluded". Suspect-3 was convicted and sentenced to prison in 2012.

In 2013, Suspect-3 contacted the Taiwan Association for Innocence (TAI). The lawyers of TAI sought assistance from DNA experts (including author GH) who advised them to request the electropherogram of the Y-DNA STR test, and to consider further testing to resolve any possible coincidental inclusion. In June 2013, TAI filed a motion for retrial on his behalf.

The Taichung High Court conducted a hearing on September 23rd, 2013. The court summoned the analysts who did the tests from the Criminal Investigation Bureau (CIB) to clarify what they meant when they wrote "cannot be excluded" in the original expert opinions. The experts testified that there were two possible interpretations of the mixed DNA sample. It could have come from all three defendants (including Suspect-3), or from just two defendants (not Suspect-3), because all of Suspect-3's alleles are found in the combination of the other two defendants (Table 1.). In fact, Suspect-3 is the only one of the three suspects whose alleles can be completely removed from the interpretation, leaving all the mixture's alleles explained by the remaining two suspects.

During testimony, an analyst from the CIB also revealed that the Bureau recently adopted a new 23-locus Y-STR DNA test kit, and that the new method could clarify whether Suspect-3's DNA was included in the sample. In December 2013, the result of the 23-locus test was reported, and Suspect-3 was excluded from the sample (Table 2.). The Taichung High Court granted a retrial a week later, and reversed his guilty verdict on March 26th, 2014.

Y-STR mixture analysis:

The analysis of Y-chromosome Short Tandem Repeats (Y-STRs) is now common in forensic laboratories. These Y-STRs are particularly useful in cases where an evidentiary mixed DNA sample has a very low concentration of male DNA, and a high concentration of female DNA (i.e. sexual assaults) [1-3]. Y-STR typing can also be used to resolve complex relationship testing [4].

Despite the methodological similarities between standard Y-STR and autosomal STR testing, the DNA markers on the Y chromosome require a significantly different method for their statistical analysis. Y-STRs are not considered to be subject to recombination. In fact, all markers on the Y chromosome form a haplotype that is inherited as a whole from father to son. For this reason, it can be assumed that all male subjects in the same male line should share the same Y-STR haplotype, unless a mutation has occurred.

The low evidentiary strength of Y-STR typing is a well-known issue in forensic DNA. Some researchers have tried to address this problem by suggesting an expansion of the Y-STR locus set with more informative markers [5]. The first comprehensive guidelines aimed for the use of Y-chromosome STRs in a forensic setting were published by the DNA Commission of the International Society of Forensic Genetics (ISFG) in 2001 [6], which was followed by an updated set of recommendations [7].

While matching a single contributor male Y-STR evidence profile to a reference sample is usually straightforward, caution must be taken when considering mixtures. In this case report we show how the use of a 17-marker system led to a coincidental inclusion in a Y-STR mixture. The report concerns a man's conviction based on the fact that the evidentiary DNA from semen initially included his Y-STR haplotype. Evidence gathered from this multiple-attacker rape initially resulted in a Probability of Inclusion of 1 every 741 randomly selected unrelated males in the Han Chinese database, and a Likelihood Ratio of 3296 times more likely that Suspect-3 contributed to the mixture than the hypotheses that he was not a contributor, favoring the prosecutor hypothesis that Suspect-3 contributed to the mixture.

Statistical evaluation:

We have used a database of 4451 Han Chinese Y-STR haplotypes [8] to evaluate the probative value of the initial analysis on this DNA mixture. When considering only the 17 Y-STR markers, as had been done initially in this case, Suspect-3 cannot be excluded from the mixture (see Table 1); however all alleles can also be explained by the two other suspects alone (S1 and S2).

Y-STR Marker	**DNA Mixture**	**Suspect-3**	**S1**	**S2**
DYS456	15	15	15	15
DYS389 I	12	12	12	12
DYS390	23,25	23	25	23
DYS389 II	28,29	28	28	29
DYS458	18,20	20	20	18
DYS19	14	14	14	14
DYS385 a/b	13,15,18	15,18	13,18	15,18
DYS393	12	12	12	12
DYS391	10	10	10	10
DYS439	12,13	12	13	12
DYS635	20	20	20	20
DYS392	14	14	14	14
Y GATA H4	12,13	12	12	13
DYS437	15	15	15	15
DYS438	11	11	11	11
DYS448	20	20	20	20

Table 1. The Y-STR haplotypes of the three suspects (Suspect-3, S1 and S2), and the DNA mixture (from semen stain) profiled using the 17-marker system.

Of the 4451 haplotypes recorded in the population database [8], only 6 subjects could not be excluded from the semen stain DNA mixture in this case. As a result, we can infer that we would expect a Probability of Inclusion (PI) for this DNA mixture of 0.00135 (6/4451). The Probability of Exclusion (PE) can be estimated as 0.99865 (PE = 1−PI). According to the PI calculated, we would expect to find a randomly selected subject that could be included in the mixture, once in every 741 males in the Han Chinese population (1/0.00135 = 741).

Before trial Suspect-3's defense team was told only that the DNA evidence matched him, with no explanation of the potential coincidental match. However, the risk of coincidental match is obvious in this case when one examines the other suspects' DNA profiles. In order to provide statistical weight for the various hypotheses, we have performed individual statistical analysis for each of the three suspects, using the Mixture Analysis tool from YHRD [9, 15].

First, we calculated the Likelihood Ratio of two mutually exclusive hypotheses: 1) Suspect-3 is a contributor to the mixture with one other unknown contributor, 2) Suspect-3 is not a contributor and two unknown males have originated the mixture. This was performed through the YHRD mixture analysis tool using the Yfiler set of markers (17 Y-STRs). The tool estimated the likelihood ratio at 3409 in favor of hypothesis 1, that Susepct-3 is a contributor. This was based on the YHRD dataset of 136,443 Yfiler haplotypes (17loci) from global populations (Release 54, June 6, 2017). According to the ENFSI Guideline for Evaluative Reporting in Forensic Science [10] the obtained Likelihood Ratio (LR = 3296), falling between the range of 1,000 and 10,000, would indicate that this piece of evidence provides "strong support for the first proposition rather than the alternative", meaning that there is strong support that Suspect-3 is a contributor of the mixture. The same Likelihood Ratio was obtained for the other two co-defendants, when their profiles were each substituted for Suspect-3's in the above hypotheses. These statistical results are of high probative value, especially for Y-STRs. It is obvious from these calculations, that using inclusion statistics for a Y-STR mixture can lead to a misleading inclusion, when reasonable alternative hypothesis are not considered. In this case, the probability of the evidence given the hypothesis that all three defendants contributed DNA, is the same as the probability of the evidence given that only two of the co-defendants (S1 and S2) contributed DNA (without Suspect-3). Both probabilities are 1.0, that is, the 17-marker DNA evidence can be fully explained by all three suspects contributing DNA or just S1 and S2. However, the original trial lawyers for Suspect-3 were only told that the evidence was consistent with all three of the suspects having contributed DNA.

To further illustrate the problem of a coincidental match in Y-STR mixtures, artificial mixture of database haplotypes can be constructed [12]. We have used a homemade script to simulate 1000 random 2-person mixtures using the subjects of the Han Chinese database [8]. For each mixture, we calculated its Power of Inclusion. The average Power of Inclusion obtained was 0.00327, equal to one random male in every 306 subjects (Interval of Confidence 95%: 75-4451).

As we have previously mentioned, for Suspect-3's post-conviction the number of Y-STR markers was extended from the initial 17-locus system to a 23-locus system. The extended haplotypes obtained for the evidentiary stain and the three suspects are reported in Table 2.

Y-STR Marker	DNA Mixture	Suspect-3	S1	S2
DYS576	19,21	**17**	19	21
DYS389 I	12	12	12	12
DYS448	20	20	20	20
DYS389 II	28,29	28	28	29
DYS19	14	14	14	14
DYS391	10	10	10	10
DYS481	23	23	23	23
DYS549	12,14	**13**	14	12
DYS533	11,12	12	11	12
DYS438	11	11	11	11
DYS437	15	15	15	15
DYS570	17	17	17	17
DYS635	20	20	20	20
DYS390	23,25	23	25	23
DYS439	12,13	12	13	12
DYS392	14	14	14	14
DYS643	11,12	12	11	12
DYS393	12	12	12	12
DYS458	18,20	20	20	18
DYS385 a/b	13,15,18	15,18	13,18	15,18
DYS456	15	15	15	15
Y GATA H4	12,13	12	12	13

Table 2. The Y-STR haplotypes of the three suspects (Suspect-3, S1 and S2), and the DNA mixture (from the semen stain) profiled using the 23-marker system. Suspect-3's alleles that are not present in the mixture have been highlighted.

From the newly obtained genetic data, it can be seen that two Y-STR markers (DYS576 and DYS549) exclude Suspect-3 from the mixture. Based on this new information Suspect-3 was exonerated.

Conclusions:

This case demonstrates how a coincidental Y-STR DNA mixture inclusion, with match statistics favoring a prosecution hypothesis, can lead to an erroneous conviction. The report shows the importance of considering all relevant hypotheses to explain DNA mixtures, and the importance of communicating these possible interpretations to interested parties including the triers-of-fact. It also demonstrates the power of using more Y-STR markers to discriminate between alternative mixture hypotheses. To our knowledge this is the first published report of an overturned wrongful conviction based on coincidental Y-STR mixture. It is beyond the scope of this paper to estimate the frequency of such coincidental matches that have resulted in wrongful convictions.

To date there is no consensus on Y-STR mixture analysis methods: the International Society for Forensic Genetics (ISFG) has not published guidelines for interpretation; and the Scientific Working Group on DNA Analysis and Methods (SWGDAM) “has not yet reached consensus” [11] on how to interpret Y-STR mixtures, suggesting that they may only be used for exclusion. Lacking authoritative guidance for Y-STR mixture statistics, we have opted to interpret this particular mixture by using the Probability of Inclusion [12] used by other authors, in order to show how coincidental inclusions may arise in casework.

While exclusion from such mixtures is a trivial matter (if drop-out is not considered), Y-STR mixtures are known to be problematic when it comes to inclusion. In our analysis of the Han Chinese database [8] the rarest 17-locus haplotype produces a single contributor Match Probability of 0.00022 (one match in the total of 4451 haplotypes), while the most common haplotype (26 matching haplotypes of 4451) produces a match probability of 0.00584 (one in every 171). We note that the estimated Probability of Inclusion for the evidence considered in this case was somewhat higher than the average PI for a random 2-person mixture from the database. The original 17-locus match presented here would be expected to include as potential contributors potentially 900,000 Han Chinese men worldwide. Of course, few of these men would have access to the crime scene

As part of every Y-STR match report, the CIB laboratory now informs the court that the 23-locus Y-STR random match probability for the Taiwan Han database is between 0.01 and 0.001. Since Taiwan has 23.52 million people, and approximately half of the population is male, between 10,000 and 100,000 men in Taiwan share each Y-STR haplotype. Thus, it is important to communicate to lay-people that a match report itself cannot establish guilt or innocence, which is a matter for the triers-of-fact to determine. There are often other factors to consider when weighing DNA evidence in sexual assault cases, such as questions of consent or possible coincidental matches to consensual partners who have not provided elimination samples. Finally, the possibility that an assailant leaves no trace of DNA is a question that cannot be settled by science, as one cannot generally prove a negative. A well-known aphorism in forensic science states, “the absence of evidence, is not the evidence of absence,” but for the innocent person absence (exclusion) is the only evidence their DNA can provide. DNA exclusion can be crucial evidence of innocence, and was sufficient in Suspect-3’s case to reverse his conviction.

In conclusion, while coincidental inclusion in DNA mixtures is not new to our community [12, 13, 14], peer-reviewed literature concerning this key issue remains scarce. We note that SWGDAM guidelines recommend Y-STR mixtures that do not present a clear major contributor should only be used for exclusion.

References:

[1] Purps, J., Geppert, M., Nagy, M. and Roewer, L. (2015) Validation of a combined autosomal/Y-chromosomal STR approach for analyzing typical biological stains in sexual assault cases. *Forensic Science International: Genetics*. 19, 238-242.

[2] McDonald, A., Jones, E., Lewis, J. and O'Rourke, P. (2015) Y-STR analysis of digital and/or penile penetration cases with no detected spermatozoa. *Forensic Science International: Genetics*. 15, 84-89.

[3] Sibille, I., Duverneuil, C., de la Grandmaison, G. L., Guerrouache, K., Teissière, F., Durigon, M. and de Mazancourt, P. (2002) Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. *Forensic Science International*. 125, 212-216.

[4] González-Andrade, F., Sánchez, D., Penacino, G., Jarreta, B.M. (2009) Two fathers for the same child: A deficient paternity case of false inclusion with autosomic STRs. *Forensic Science International: Genetics*. 3, 138-140.

[5] Roewer, L. (2009) Y chromosome STR typing in crime casework. *Forensic Science, Medicine, and Pathology*. 5(2), 77-84.

[6] Gill, P., Brenner, C., Brinkmann, B., Budowle, B., Carracedo, A., Jobling, M.A., de Knijff, P., Kayser, M., Mayr, W.R., Morling, N., Olaisen, B., Pascali, V., Prinz, M., Roewer, L., Schneider, P.M., Sajantila, A. and Tyler-Smith, C. (2001) DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. *Forensic Science International*. 124, 5-10.

[7] Gusmão, L., Butler, J.M., Carracedo, A., Gill, P., Kayser, M., Mayr, W.R., Morling, N., Prinz, M., Roewer, L., Tyler-Smith, C. and Schneider, P.M. (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Science International*. 157, 187-197.

[8] Wu, W., Pan, L., Hao, H., Zheng, X., Lin, J. and Lu, D. (2011) Population genetics of 17 Y-STR loci in a large Chinese Han population from Zhejiang Province, Eastern China. *Forensic Science International: Genetics*. 5, e11-e13.

[9] Y-SR Mixture Calculation. Available Online at: https://yhrd.org/mixture [Accessed May 1st 2017].

[10] European Network of Forensic Science Institutes (2015) ENFSI Guideline for evaluative reporting in forensic science, Approved version 3.0., page 17. Available Online at: http://enfsi.eu/news/enfsi-guideline-evaluative-reporting-forensic-science [Accessed May 1st]

[11] Scientific Working Group on DNA Analysis Methods (2014) SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories. Available Online at: http://www.swgdam.org/publications [Accessed Nov. 15th 2016].

[12] Ge, J., Budowle, B. and Chakraborty, R. (2010) Interpreting Y chromosome STR haplotype mixture. *Legal Medicine*. 12, 137-143.

[13] Dror, I.E. and Hampikian, G. (2011) Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice*. 51, 204-208.

[14] Miller, V.R. and Bentley, J.R. (2009) Frequency of apparent false inclusions in complex Y-STR mixtures. Poster at the 20th International Symposium on Human Identification in Las Vegas(NV). Abstract available Online at: https://www.promega.com/-/media/files/resources/conference-proceedings/ishi-20/poster-abstracts/poster3.pdf?la=en [Accessed Nov. 15th 2016].

[15] Willuweit, S. and Roewer, L. (2007) Y chromosome haplotype reference database (YHRD): Update. *Forensic Science International: Genetics*. 1, 83-87.