TWO YEARS LATER: A REFLECTION ON THE IMPLEMENTATION OF STRMIX™ IN A HIGH THROUGHPUT DNA LABORATORY

Zane Kerr (corresponding author)
New South Wales Forensic & Analytical Science Service, NSW Health Pathology, PO Box 162, Lidcombe, NSW, 2141, Australia.
Email: zane.kerr@health.nsw.gov.au
Phone: +61 296460222

Introduction to the New South Wales Forensic & Analytical Science Service

The New South Wales Forensic & Analytical Science Service (NSW FASS) is a statewide service that provides independent, objective analysis to support the NSW health and justice systems. NSW FASS employs approximately 330 staff across five sites and provides testing services in a range of disciplines including forensic medicine, forensic biology and DNA, forensic toxicology, illicit drug analysis and chemical criminalistics.

NSW FASS operates a fully automated DNA laboratory processing over 35,000 crime samples per year with a turnaround time of approximately two days. The majority of exhibit examination is carried out by crime scene officers of the NSW Police Force who collect samples using specially designed field kits and then prepare sub-samples (swabs, swatches, tapelifts) which are placed into barcoded tubes for DNA analysis. These sub-samples are submitted to NSW FASS in a format that is suitable for immediate analysis in the fully automated DNA laboratory upon receipt. The DNA data is subsequently provided to the Case Management Unit for interpretation and statistical evaluation.

Impetus for Change

A significant proportion of the samples submitted to NSW FASS for analysis are “trace” or “touch” DNA samples. Increased testing sensitivity following implementation of the PrepFiler® magnetic bead extraction chemistry, PowerPlex® 21 STR kit, and ABI 3500xl genetic analyser has further enabled the recovery of results from these types of samples. Unlike samples collected from “rich” sources of DNA such as blood, semen, or saliva, trace DNA samples are more likely to produce weak and/or complex mixed DNA profiles that are difficult if not impossible to interpret using existing binary methods. The reporting of these profiles as ‘inconclusive’ represents a loss of information for both the investigative and judicial processes.

Inter-laboratory studies on mixture interpretation such as the MIX13 study carried out by the United States National Institute of Standards and Technology (NIST)[1] have demonstrated that there may be significant inconsistencies in how these difficult profiles are interpreted and reported, both within and between laboratories. Similar inter-jurisdictional exercises performed in Australasian laboratories indicated that this issue is not unique to the United States. The Australia New Zealand Policing Advisory Agency National Institute of Forensic Science (ANZPAA NIFS) initiated the Standardisation of DNA Interpretation Project and supported the formation of the Statistics Working Group (STATSWG). The charge of STATSWG was to improve inter-jurisdictional consistency in the interpretation and statistical evaluation of forensic DNA profiles with the ultimate goal of national standardisation. STATSWG recommended the use of a continuous probabilistic model for DNA interpretation. This lead to the development of the STRmix™ program by a development team consisting of scientists from Forensic Science South Australia (FSSA) and the Institute of Environmental Science and Research, Ltd., New Zealand (ESR). STRmix™ has been endorsed for use by the Australasian Biology Specialist Advisory Group (BSAG) and at the time of writing has been implemented or is in the process of being implemented by all Australasian forensic laboratories.

Introduction to STRmix™

STRmix™ is an expert system that makes use of a fully continuous probabilistic model for the interpretation and statistical evaluation of DNA profiles. STRmix™ models allelic drop-out, drop-in and stochastic peak height variation making it particularly useful for weak and/or complex DNA profiles. An established statistical sampling process known as Markov Chain Monte Carlo (MCMC) is used to "explore" the DNA data. In each round of the MCMC process, STRmix™ randomly proposes genotypes for each contributor as well as other parameters such as the amount of DNA per contributor, how degraded each contributor's DNA is, how efficiently each locus has amplified during PCR, and how efficiently each PCR test has amplified (if multiple amplification replicates of the DNA sample are being analysed). These parameters are collectively referred to as the mass parameters. A biological model based on known behaviours of DNA profiles and informed by empirical data is then used to build up a picture of what the DNA profile would be expected to look like given the proposed genotypes and mass parameters. This expected profile is compared with the observed profile. If they are similar, then the model proposed by STRmix™ provides a good explanation for the DNA data. Conversely, if there are significant differences that cannot be accounted for by stochastic peak height variation, then the proposed model offers a poor explanation for the observed data. This process is then repeated with new genotypes and mass parameters chosen. If the new model provides a better explanation for the DNA data than the previous model, STRmix™ accepts the new model and "moves" to it. In order to thoroughly explore the DNA data, STRmix™ may still accept an inferior model some of the time. This prevents the MCMC from becoming focussed on a particular model and failing to explore other possible explanations for the DNA data. This process is repeated many, many times (billions of iterations may be carried out for particularly complex profiles). In this manner, STRmix™ thoroughly explores the sample space and determines those genotype sets and mass parameters that best explain the observed data. The end result of a STRmix™ deconvolution is a genotype probability distribution (GPD) that lists all of the accepted genotype sets and their associated weights. These weights can take any value from 0 to 1 inclusive (hence, STRmix™ is often described as using a continuous model). Those genotype sets that provide a good explanation for the observed data will be assigned high weights (close to 1) whereas those that don't explain the observed data well will be given low weights (close to 0). Reference samples can then be compared to the deconvoluted profile and a likelihood ratio (LR) match statistic assigned. As STRmix™ makes better use of the quantitative data contained within the DNA profile, more informative match statistics are produced. That is, STRmix™ is more likely to provide LRs favouring inclusion for true contributors to a mixed DNA profile and LRs favouring exclusion for true non-contributors. By providing a more informative LR match statistic, STRmix™ is better able to assist the court in assessing the propositions of interest.

Validation

STRmix™ has been extensively validated. Details of the developmental validation can be found in the STRmix™ user's manual[2]. The mathematical and biological models used by STRmix™ have also been published in peer-reviewed scientific journals[3-5].

Beta testing:

NSW FASS was one of four Australasian laboratories that participated in the beta testing of STRmix™ in 2012. This involved preparing DNA samples from known contributors with varying contributor proportions (single source, two and three person mixtures). These were amplified using the Profiler Plus® STR kit and the resulting DNA profiles were deconvoluted using an early version of STRmix™ (v0.99, then known as 'DyNAmix'). The results generated by STRmix™ were then reviewed to check that they aligned with intuitive expectations and the ground truth. Generally, it was observed that the genotypes of the known contributor(s) were assigned the highest weight at each locus. The reproducibility of STRmix™ was also examined. As MCMC is a stochastic process, the results generated will vary if the analysis is repeated. To assess reproducibility, a series of profiles

was deconvoluted multiple times and a LR match statistic assigned for the known contributor(s). STRmix™ was found to be highly reproducible with most LRs typically varying by less than an order of magnitude. Increased variability in LR match statistics was observed in some unresolvable mixtures however this variability was usually within acceptable limits. Where possible, match statistics were also calculated using methods in use at NSW FASS at the time (modified random match probability and unrestricted combinatorial LR) and compared with those produced by STRmix™. Generally, the match statistics produced by STRmix™ were more informative due to better use of the quantitative data contained within the DNA profile. Throughout the beta testing process, any bugs or suggestions to improve usability were also noted. The findings of the beta testing laboratories were compiled into a report that was distributed to the BSAG and the STRmix™ development team for review.

Validation for PowerPlex® 21:

In 2012, the decision was made to expand the core STR loci used by the Australasian forensic laboratories. NSW FASS began validation of the PowerPlex® 21 kit which was implemented for person sample testing in late 2012. Following distribution of an updated version of STRmix™ (v1.05) to the Australasian laboratories, NSW FASS commenced validation of STRmix™ for use with PowerPlex® 21 profiles. This version of STRmix™ featured a number of the improvements suggested during beta testing and enabled deconvolution of mixtures containing DNA of up to four contributors. The validation of STRmix™ for use with PowerPlex® 21 profiles followed a similar approach to the beta testing that had previously been carried out. Additionally, values for the parameters and settings used by STRmix™ were determined using empirical data collected by NSW FASS and by other Australasian laboratories. In March 2013, NSW FASS implemented the PowerPlex® 21 kit for testing of crime samples along with the STRmix™ interpretation software. The intention at the time was that all mixed DNA profiles containing DNA of up to three contributors would be interpreted using STRmix™. However, due to the increased sensitivity of the PowerPlex® 21 kit, a significant number of mixtures containing DNA from four or more individuals were recovered. This necessitated further validation studies to be carried out so that STRmix™ could be used to interpret these complex mixtures.

Ongoing validation and future directions:

Ongoing validation and verification are required as each new version STRmix™ is released. The depth of testing required depends on the extent of the changes made to the software. NSW FASS have recently validated STRmix™ v2.3.06. This version provides significant improvements to the modelling, run-times and usability of the software. STRmix™ v2.3.06 also permits the user to assume more than four contributors to a mixture. Validation studies examining the ability of STRmix™ to interpret mixtures of five contributors are in progress.

Training of Staff

STRmix™ was developed to be an expert system that can be understood by forensic biologists who are not mathematicians or statisticians. Comprehension of the modelling and processes used by STRmix™ is attainable with appropriate training. Training of NSW FASS staff in the theory and use of STRmix™ began in January 2012 when two staff members attended a week-long workshop hosted by the STRmix™ development team. This workshop introduced participants to the biological modelling used by STRmix™, MCMC theory, and operation of the STRmix™ software. In July 2012, two staff members attended a three-day “train the trainer” workshop hosted by the development team. The aim of this workshop was to provide attendees with the materials and training necessary to return to their individual laboratories and carry out in-house training. Following the “train the trainer” workshop, an intensive week-long in-house training session was carried out at NSW FASS. Presentations and practical exercises were provided in order to train all of the senior reporting

biologists in the use of and theory behind STRmix™. Over the following months, regular practical exercises and discussion sessions were held to familiarise staff in the operation of the STRmix™ software including interrogation of the results output. Theory and practical exams were prepared with staff required to pass each component prior to being authorised to report on STRmix™ analyses. Presentations were also given to key stakeholders including investigators and prosecutors.

In February 2014, BSAG organised a STRmix™ advanced user's workshop that was attended by experienced STRmix™ analysts from each of the Australasian laboratories. This workshop facilitated discussion between the attending laboratories on their experiences using STRmix™ and provided an important opportunity to improve inter-jurisdictional consistency in the interpretation and reporting of complex DNA profiles. A second workshop is currently planned for June 2016.

The training of staff at NSW FASS has been ongoing with regular presentations and discussion sessions held. In particular, new releases of STRmix™ have required additional staff training as the modelling and functionality of the software has evolved. Training sessions covering the changes to STRmix™ introduced in v2.3.06 have recently been conducted and staff are undergoing updated theory and practical exams to assess their understanding of these changes.

Challenges

The implementation of a new STR typing kit and interpretational software was not without difficulty. Two of the biggest challenges are assigning the number of contributors to weak and/or complex DNA profiles and a significant increase in the workload of the Case Management Unit.

Assigning the number of contributors:

STRmix™ currently requires the user to assign the number of contributors to the DNA sample. When interpreting weak and complex mixed DNA profiles, it may be particularly difficult to assign the number of contributors with any degree of confidence. The detection of previously-unseen artefact peaks such as "post stutter" (one additional repeat unit) and "stutter of stutter" (two fewer repeat units) following implementation of the PowerPlex® 21 STR kit and ABI 3500xl genetic analyser further hinders the ability of analysts to confidently assign the number of contributors to a sample, particularly if these artefact peaks are of similar height to the minor component peaks. Increased stochastic effects such as extreme peak height imbalances and allelic drop-out have also been observed following implementation of the PowerPlex® 21 kit limiting the usefulness of peak height data when assigning the number of contributors to a mixture when the DNA template is low. We have observed that varying the number of assumed contributors can have a significant impact on the results produced by STRmix™. In our experience, the GPD and LR match statistics produced by STRmix™ for major contributors to a mixed DNA profile are generally fairly insensitive to variations in the number of assumed contributors. However, such variations can have a significant impact on the GPD and LR match statistics generated for minor contributors. Consider the PowerPlex® 21 profile shown below in Figure 1. This profile was recovered from a crime sample submitted to NSW FASS for testing. The presence of more than two alleles at several loci indicates that the DNA sample originates from more than one individual. Disregarding apparent stutter artefacts, no more than four alleles have been detected at any locus. Following review of the mixture data, it may be reasonable to proceed under the assumption that this mixture originates from two individuals. This profile was deconvoluted using STRmix™ v2.06 and assuming two contributors to the mixture. A LR was then assigned for a person of interest whose reference sample was submitted for analysis and comparison. The genotype of this individual at the twenty-one PowerPlex® 21 loci is summarised in Table 1. Visual inspection of the mixture data reveals that this individual appears to match the minor component of the recovered mixture. Notably, a LR of zero (i.e. an exclusion) was obtained at D3S1358. Inspection of the STRmix™ results shown in Figure 2 reveals that all genotypes accepted by STRmix™ for the minor contributor (Contributor 2 in Figure 2) include a 15 allele. The {16,17} genotype of the person of interest has not been accepted by STRmix™ resulting in an exclusion at

this locus. Conversely, when the mixture was deconvoluted assuming three contributors, a LR match statistic of approximately 16 billion favouring inclusion was obtained. Faced with these results, how should a comparison with the person of interest be reported? One option would be to stand by an assumption of two contributors and report that the person of interest is excluded as a contributor to the mixed DNA profile recovered. Alternatively, one might update their assumptions regarding the number contributors in light of the reference genotype and report a match under the assumption that the mixture originates from at least three contributors. A third option would be to report both results. While this provides a transparent approach, a jury may experience difficulty understanding the findings if presented with both a match and an exclusion. Where the number of contributors cannot be confidently assigned, the approach taken at NSW FASS has been to carry out further analysis of the sample in order to better support any assumptions made. This may involve re-amplifying the DNA extract or re-analysing the sample using a lower analytical threshold following capillary electrophoresis. If considerable uncertainty still exists regarding the number of contributors following further analysis then further interpretation of the DNA profile will typically not be carried out.

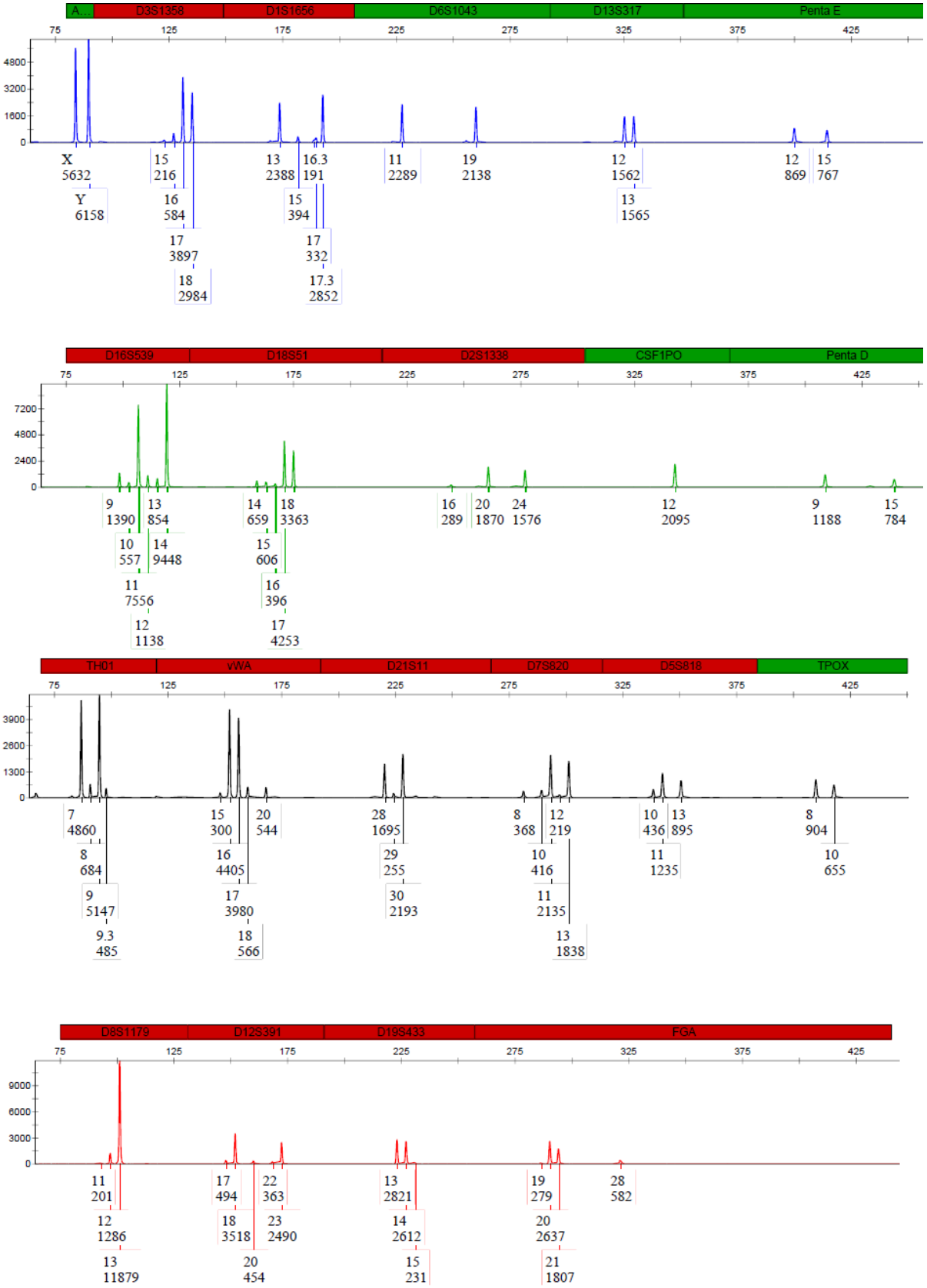

Figure 1: A PowerPlex® 21 profile recovered from a crime sample analysed at NSW FASS. There appears to be at least two contributors to the mixed DNA profile recovered.

Table 1: Genotype of a person of interest submitted for comparison with the DNA profile shown in Figure 1.

Amelogenin	X,Y
D3S1358	16,17
D1S1656	15,17
D6S1043	11,19
D13S317	8,13
Penta E	16,18
D16S539	9,12
D18S51	14,15
D2S1338	16,24
CSF1PO	10,12
Penta D	9,13
TH01	8,9.3
vWA	18,20
D21S11	31.2,33.2
D7S820	8,10
D5S818	10,10
TPOX	8,8
D8S1179	12,13
D12S391	17,20
D19S433	14,15
FGA	19,28

```
Locus D3S1358
Contributor 1
Genotype [17,18] - 100.0%
Contributor 2
Genotype [15,15] - 0.0%
Genotype [-1,15] - 0.4%
Genotype [15,16] - 43.5%
Genotype [15,17] - 41.3%
Genotype [15,18] - 14.7%
```

Figure 2: STRmix™ results produced at locus D3S1358 when the mixture shown in Figure 1 was deconvoluted using STRmix™ v2.06 and two contributors assumed. The mixture proportions proposed by STRmix™ are 0.89 (contributor 1) and 0.11 (contributor 2).

Increased workload:

The introduction of STRmix™ at NSW FASS has greatly extended our ability to interpret weak and complex results. Prior to STRmix™, many of these profiles would not have been interpreted further as they were beyond the limits of available binary interpretation methods. These results now require careful assessment to determine suitability for further interpretation using STRmix™. This pre-assessment of profiles prior to deconvolution using STRmix™ can be quite time consuming if the profile is particularly complex. Following deconvolution, the results generated by STRmix™ also require careful scrutiny to check that they align with intuitive expectations. Occasionally, STRmix™ may fail to model a profile well, even for profiles that appear fairly straightforward. Consider the PowerPlex® 21 profile shown in Figure 3. This mixed DNA profile appears to originate from two main contributors. As noted in Figure 3, there are indications throughout the profile that suggest that low levels of DNA from a third contributor may also be present. The mixture was deconvoluted using

STRmix™ v2.06 and assuming three contributors. STRmix™ proposed mixture proportions of 0.65, 0.32 and 0.03, which agree well with intuitive expectations following examination of the mixture data. However, the genotypes proposed by STRmix™ at one locus (FGA) do not align with expectations. Based on the determined mixture proportions, one would expect relatively high weight to be assigned to the {24,25} and {24,24} genotypes for contributor 3 (mixture proportion 0.65) and to the {24,24} and {25,25} genotypes for contributor 2 (mixture proportion 0.32). Given the low mixture proportion of contributor 1 and apparent masking and/or drop-out of their alleles at FGA, one would expect to obtain a fairly diffuse genotype probability distribution for this contributor. As can be seen in Figure 4, STRmix™ has given a very high weight to both of the major contributors having a genotype of {24,24}. Furthermore, STRmix™ has proposed that the 25 allele, with a peak height of over 9000 relative fluorescence units, likely originates from the minor contributor who is present at very low levels. The deconvolution was repeated several times using STRmix™ v2.06 with similar results obtained in each instance. The mixture was then deconvoluted using STRmix™ v2.3.06. This version features a number of enhancements including improved modelling. As can be seen in Figure 5, more reasonable results were obtained at FGA using STRmix™ v2.3.06. This example highlights the need to carefully review all results produced by STRmix™ in order to check that they align with intuitive expectations. It is important to realise that STRmix™ is an expert system intended for use by experienced experts.

While STRmix™ is capable of rapidly deconvoluting a profile, the review of the results produced can be quite time consuming with some complex profiles generating over 1000 pages of results. Improvements have been implemented in STRmix™ v2.3.06 which streamline the results output and greatly simplify this task. In order to combat the increased workload that has resulted following the implementation of STRmix™ at NSW FASS, additional forensic biologists have been trained in the use of STRmix™. These biologists carry out devonvolutions and LR calculations and check that the results are intuitively sensible before passing the results on to more experienced senior reporting biologists for peer-review and reporting. One other approach taken by NSW FASS to reduce the workload is to not carry out any further interpretation of complex mixtures in a case if ‘simpler’ results are available to interpret. Further interpretation of these complex profiles can be carried out if requested by investigators, prosecutors or the defence.

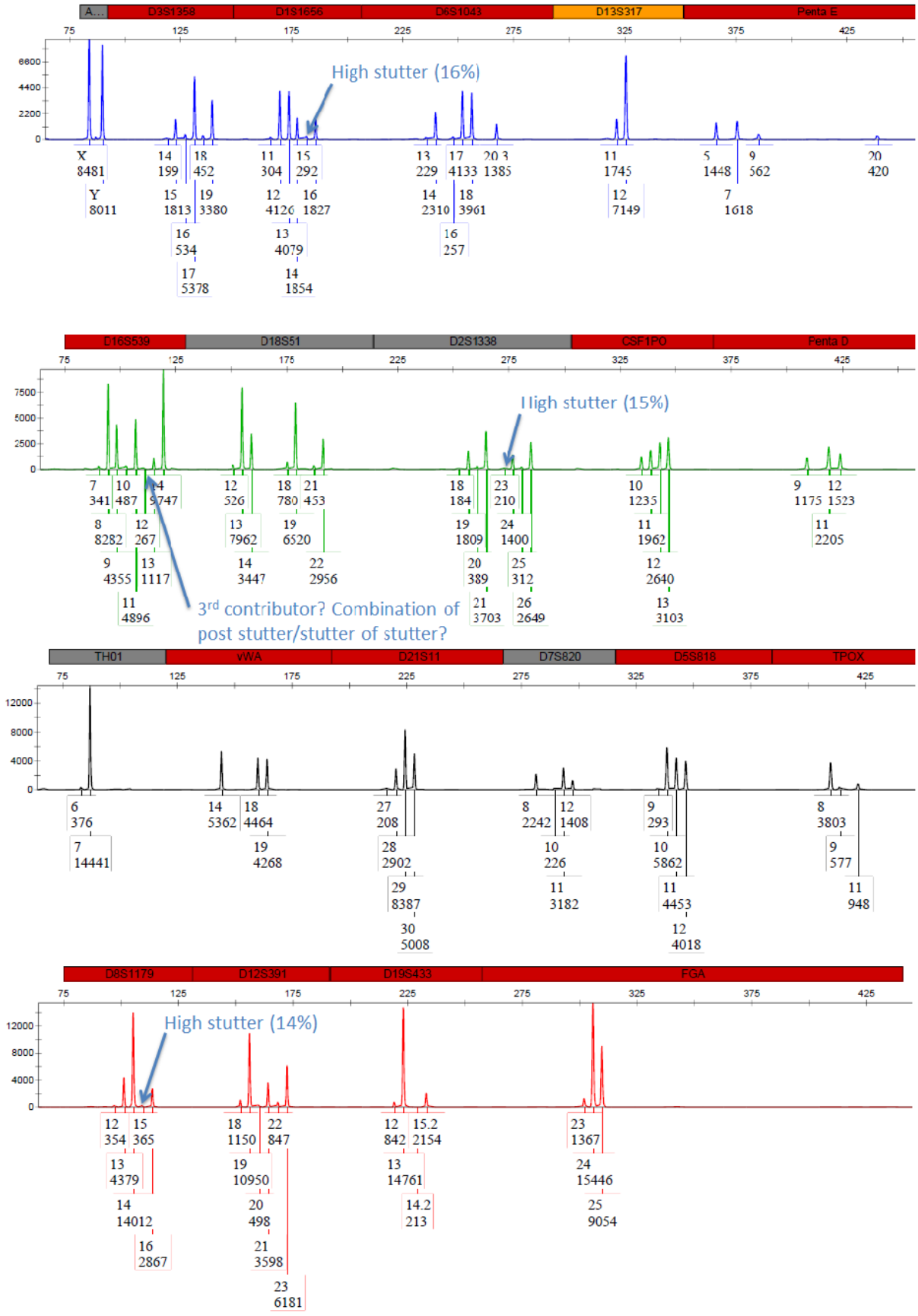

Figure 3: A PowerPlex® 21 profile recovered from a crime sample analysed at NSW FASS. The profile is a mixture that appears to originate from two main contributors. Low levels of DNA from at least one other contributor may also be present, as indicated.

```
Locus FGA
Contributor 1
Genotype [23,25] - 24.6%
Genotype [24,25] - 25.1%
Genotype [25,25] - 24.6%
Genotype [-1,25] - 25.7%
Genotype [24,24] - 0.0%
Genotype [-1,23] - 0.0%
Genotype [-1,-1] - 0.0%
Genotype [23,24] - 0.0%
Contributor 2
Genotype [24,24] - 100.0%
Genotype [25,25] - 0.0%
Genotype [24,25] - 0.0%
Contributor 3
Genotype [24,24] - 100.0%
Genotype [24,25] - 0.0%
```

Figure 4: STRmix™ results produced at locus FGA when the mixture shown in Figure 3 was deconvoluted using STRmix™ v2.06 and three contributors assumed. The mixture proportions proposed by STRmix™ are 0.03 (contributor 1), 0.32 (contributor 2) and 0.65 (contributor 3).

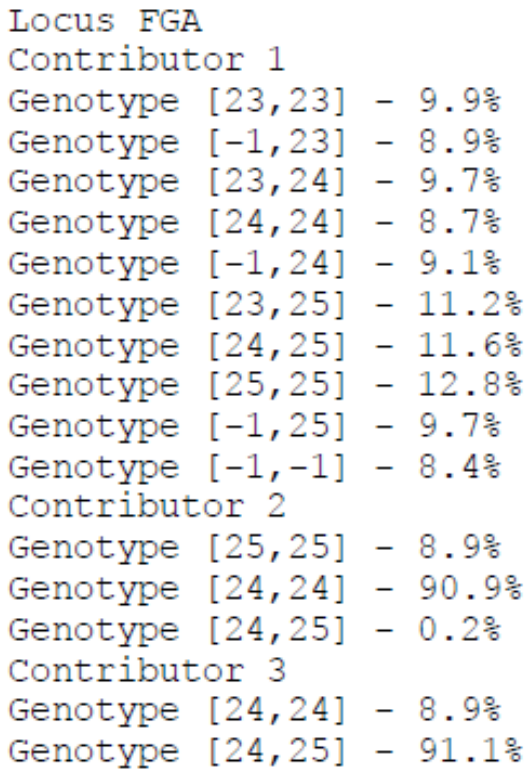

Figure 5: STRmix™ results produced at locus FGA when the mixture shown in Figure 3 was deconvoluted using STRmix™ v2.3.06 and three contributors assumed. The mixture proportions proposed by STRmix™ are 0.01 (contributor 1), 0.33 (contributor 2) and 0.66 (contributor 3).

Conclusion

Expert systems such as STRmix™ have significantly expanded the range of DNA profiles that can be interpreted. Many of these weak and complex profiles would previously have been reported as 'inconclusive' which represents a loss of information for both the investigative and judicial processes. Extensive validation of STRmix™ by NSW FASS, by other Australasian forensic laboratories, and by the STRmix™ development team has demonstrated its effectiveness, particularly in the interpretation of weak and complex DNA profiles. The biological and mathematical models used by STRmix™, while complex, are understood by forensic biologists. The implementation of any new technology is not without challenge. Difficulties in assigning the number of contributors to weak and/or complex mixed DNA profiles and a significant increase in the workload of reporting biologists have been two of the biggest challenges encountered at NSW FASS. It is important to remember that expert systems such as STRmix™ are intended for use by experienced forensic biologists.

Acknowledgments

I would like to thank Sharon Neville, Dr. John West, and the staff of the forensic biology & DNA laboratory of NSW FASS for their constructive feedback. I would also like to thank Kevin Forward and NSW Health Pathology management for providing financial support to attend the 26th International Symposium on Human Identification.

References

1. National Institute of Standards and Technology 2013, *NIST Interlaboratory Mixture Interpretation Study 2013 (MIX13)*, viewed 7 October 2015, <http://www.cstl.nist.gov/strbase/interlab/MIX13.htm>.
2. Institute of Environmental Science & Research Ltd. (ESR) 2015, *STRmix™ 2.3 User's Manual*, ESR, Auckland, New Zealand.
3. Bright, J., Taylor, D., Curran, J.M. & Buckleton, J.S. 2013, 'Developing allelic and stutter peak height models for a continuous method of DNA interpretation', *Forensic Science International: Genetics*, vol. 7, no. 2, pp. 296-304.
4. Taylor, D., Bright, J. & Buckleton, J. 2013, 'The interpretation of single source and mixed DNA profiles', *Forensic Science International: Genetics*, vol. 7, no. 5, pp. 516-528.
5. Bright, J., Taylor, D., Curran, J.M. & Buckleton, J.S. 2013, 'Degradation of forensic DNA profiles', *Australian Journal of Forensic Sciences*, vol. 45, no. 4, pp. 445-449.