HealthSupport
Queensland

Forensic and Scientific Services

4P Mixture Discussion Paper

Rhys Parry, Emma Caunt, Allison Lloyd 01/10/2020

Introduction

The introduction of routine analysis of four-contributor (4P) mixtures to DNA Analysis has increased the complexity of DNA mixture interpretation. Papers such as Paoletti *et al.* (2005) and Dembinski *et al*. (2018) have outlined the probabilities of mixed DNA profiles being inaccurately interpreted based on the true number of contributors versus the observed number of contributors using synthetic mixtures. These two papers and the increased complexity of interpretation over conventional two- (2P) and three-contributor (3P) mixtures has caused considerable debate about the complexities of higher-contributor interpretation. However, it will be shown that these papers do not reflect the real-world analysis of multiple contributor mixed DNA profiles and, therefore, should not be relied upon as a source of uncertainty.

There will always be a level of subjectivity involved, and this is increased as the complexity of a mixture increases especially where numerous low-level peaks are involved. Difficulties lie in formal guidelines, as they can never cover all situations. Ultimately, the scientist should be able to use logic and probabilistic rationalisation to justify the decision made. This discussion paper aims to cover some of the most salient issues pertaining to 4P mixture analysis and interpretation.

Determining the Number of Contributors

For the purposes of this paper, a "major" part of the DNA profile is any part that can be reasonably separated out from the remaining profile, either visually or by deconvolution. In a 4P mixture, this could conceivably consist of 1, 2, or 3P contributions. As the number of contributors to the major profile increases, the more likely it is that potential masking will go undetected. This is because the probability of larger peaks masking low-level DNA contributions is greater. Additionally, the stutter peaks produced by these larger contributions add extra information in the stochastic range of the profile, which is where low-level contributions of DNA are most likely to be observed. Therefore, there is a greater potential for peaks to be masked in a DNA profile with 2P or 3P in the "major" than there is in a mixture with only a 1P "major".

Four-contributor mixtures fall on a spectrum from "single clear major with 3P minor" to "4P low-level" mixtures where, in the latter, there is no clear separation between any of the contributors. The former type tends to be relatively easy to interpret for the reasons outlined below. In the latter type however, interpretation will rely on the scientist's ability to reasonably make assumptions about the quality of the profile. For example, if all potential contributions are roughly equal and well above stochastic threshold, it should be a relatively trivial matter to decide on an appropriate number of contributors. This is because any uncertain low-level contributions will either have a meaningless effect on the deconvolution or will be obvious as stand-alone peaks in the minor part of the profile. Conversely, if all contributors are predominantly in the stochastic range, then it is

unlikely that a reasonable conclusion can be drawn as to the number of contributors. Most of the complexity of 4P mixture interpretation is from profiles that fall between these two extremes.

The interpretation of a 4P mixed DNA profile is no different in principle to the interpretation of a 3P profile. The main risk in determining the number of contributors in a potential 3P/4P/5P mixture is the risk that the number of contributors is not apparent due to masking of low level allelic peaks. It is obvious that the probability of masking increases as the true number of distinct contributors increases. It should be noted however, that this risk is not new. This risk exists for 2P and 3P mixtures as well; the only difference being that the probability of masking is increased in 4P mixtures. The principal risks are:

- False Inclusion[1]: False inclusion is a possibility when adding a contributor (n+1), however this is true for all mixtures. It must be noted that most loci will not have indications of 4P, so very few individuals are likely to match in the minor contribution across the entire profile without relying on dropout peaks in order to fit. Therefore, anyone matching the minor contribution would have to be very well represented in order to generate a statistic in support of contribution, and this in itself would imply that there is a reasonable possibility of them having contributed DNA. Ultimately, adding an extra contributor will move all LRs towards 1 and so the risk of false inclusion is low.
- False Exclusion[2]: False exclusion is possible in any mixture but is more likely with less contributors. The risk of false exclusion for a 2P mixture that is actually 3P is relatively high. However, as extra possible contributors are added the risk diminishes. A false exclusion in a 4P mixture would be exceptionally rare.

Accordingly, the same rules for determining the number of contributors in 2P and 3P mixtures also apply to 4P mixtures. However, there are additional considerations of which to be mindful. As demonstrated in the 2P/3P reproducibility analysis (Morgan & Caunt, 2014), adding an additional contributor is acceptable and is generally reflective of the true number of contributors where reproducibility is low. By extension, this is, with high probability, true of 3P/4P mixtures as well. As the complexity of a mixture increases, the probability of the drop-out and stochastic effects on minor peaks increases, and in turn these effects will tend to obscure the true number of contributors. However, as per Bright *et al.* (2014b), which is discussed below, the actual number of contributors is relatively immaterial in situations where there is a strong major and an unknown number of minor contributors. They concluded that the lowest number of contributors required to explain the profile is the most effective assumption for determining the number of contributors. It must be pointed out that there is a fundamental issue that requires this assumption to be employed differently when dealing with 2P/3P versus 3P/4P or 4P/5P mixtures. In a 2P mixture, incorrectly assuming 2P when the mixture is 3P can lead to issues with incorrect remaining profiles, false exclusions where none should exist, and overstating the LR for potential inclusions, especially when the inclusion is based on a partial profile. Hence, for a 2P mixture, it may be better to assume an additional contributor if confidence is not high as to the exact number of contributors. The main risk of assuming an additional contributor is false inclusion, but this is not likely to lead to a meaningful statistic unless the POI is very well represented, in which case, there is a reasonable possibility that the POI is a true contributor. Furthermore, the addition of an extra contributor will impel the LR towards 1, which will not have a significant impact on a "major" contributor's LR. Similarly, this

[1] False inclusion is defined as a LR > 1 for a true non-contributor

[2] False exclusion is defined as a LR = 0 for a true contributor

principle applies to 3P/4P mixtures and though the risk of false inclusion increases, it is strongly counterbalanced by the fact that any false inclusion is much more unlikely to generate a supportive LR.

One of the main concerns expressed by case managers centres around the uncertainty that a 4P mixture could actually be a 5P and that there is no way of knowing for certain that it isn't. As previously, stated Bright *et al*. (2014b) conclude that the minimum number of contributors needed to explain the mixture is acceptable and has negligible effect on the obtained LR. Therefore, adding an extra contributor is unnecessary in the absence of evidence, unlike with a 2P/3P scenario, and as outlined above, the risk of a supportive false inclusion is minimal.

Some loci are more discriminatory than others (see Appendix). The most discriminatory loci are D18, D12 and D6, followed by D1, D2 and Penta E. Thus, these loci are more likely, though not guaranteed, to be more informative with respect to the actual number of contributors to the mixture. Clear indications of 4P at loci other than these, when these loci indicate <4P, could be a strong indicator that significant masking of alleles is occurring, and greater consideration should be given as to whether the profile might be complex.

Points to consider:

- An increased number of contributors in the minor contribution means a decreased risk with regard to the number of contributors, and conversely, an increased number of contributors to the major indicates an increased risk
- For 2P mixtures, assuming n+1 contributors is preferable unless confidence is high that this is not warranted. This is not necessary with 3P or 4P mixtures.
- For 3P/4P or 4P/5P mixtures, it is recommended to use the minimum observed number of contributors as the basis for interpretation. STRmix™ relies on the lowest number of contributors to reasonably explain the profile. If there is no obvious reason to consider the profile too complex for interpretation, then it is recommended that the profile be interpreted.
- Designating a mixture 4P is acceptable where there are 1P or 2P contributors to a strong major and there is no *clear* indication of >3P or >2P in the minor respectively.
- DNA profiles consisting of 4P, where all contributions are well above the stochastic range (>500rfu), whether equal contributions or not, can be interpreted. Even though masking is more likely to go unobserved, its potential presence is not likely to have a meaningful impact on the deconvolution of the profile. When interpreting a 4P mixtures where all contributions are equal, consideration should be given as to how meaningful the result may be.

When to Consider "Complex"

There are several situations where it is not possible to be certain of the number of contributors and this uncertainty is likely to have problematic effects on the ability of STRmix™ to deconvolute the mixture effectively. The examples given below are not meant to be taken as rules; merely guiding principles. These include scenarios where:

- There are three amplifications of the DNA profile and valid uncertainty still exists around the number of contributors. Valid concerns might include aberrant deconvolution results; or a consistent single peak above the highly stochastic region (<200rfu) that does not seem to

have a matching partner and the introduction of such a partner would mean the locus was >4P.

- The ratio between contributors is approximately 1:1:1:1 and almost all alleles are in the <500rfu region. Note: this ratio should not be taken too literally as a ratio of 1:2:2:1 (or similar combination) would be similarly complex due to the possibility of stochastic allelic imbalances.
- Where all contributions, including any major, fall in the stochastic range (<200rfu) across the majority of the DNA profile.
- The major DNA profile is markedly degraded, but the minor contributions are not. Given that Penta E is one of the most discriminatory loci, an inability to separate mixture proportions clearly here is a good indication that the profile might be problematic. While the major may be obvious at lower molecular weight loci, the ability of STRmix™ to accurately model the overall profile may be compromised.
- Clear indications of 4P at loci other than the most discriminatory (see section above). As an extreme example, indications of 4P at TPOX is probably indicative of far greater than four contributors across the profile as a whole.
- There are multiple (>3) loci with clear indications of >6 alleles. It should be noted that there is a distinct difference between 7-allele loci and 8-allele loci from a statistical perspective. The former allows a number of "Q" designations for potentially missing alleles whereas this is not necessarily the case with 8-allele loci. For example, if the eight alleles are distinct, unique designations, STRmix™ will have increased confidence in modelling a restricted number of allele combinations. By contrast, if there are only seven distinct alleles or the eighth possible allele is based on a stutter peak that falls above threshold, STRmix™ will model a much wider array of possible allele combinations. Additionally, there is less risk of masking at 7-allele loci than at 8-allele loci.
- STRmix™ is overly confident (p≥0.99) of allelic pairings in the minor despite them being in the highly stochastic range (<200rfu) and not particularly consistent between runs. The possible combinations in a 4P mixture can be assessed in STRmix™ in the same manner as a 3P or 2P analysis; it just takes longer as there are more allelic combinations to consider. There is no need to try to check that STRmix™ has all the individual allelic pair percentages correct, just that it has considered the combination.
- There are 3P in a distinct "major" contribution and 1P in a stochastic "minor". The main issue with this scenario is that STRmix™ will assign a p=1.00 to any labelled allele in the minor, which means there is an increased risk of false attribution, despite masking being possible. The number of alleles in the "minor" contribution may assist in determining whether to call such a profile complex.

STRmix™ is not currently validated within the lab for the interpretation of 5P mixtures. Thus, if there is direct evidence of >4P in the profile, it must be considered complex, irrespective of any of the considerations listed above. However, if a profile is indicative of a 4P mixture, and there is no evidence to suggest a fifth contributor, then considering the profile as complex because a fifth contributor may be present (but is being masked) is not appropriate as the same could be said for any mixture.

Reworking:

As with any mixed DNA profile, reamplification may assist with obtaining greater clarity as to how many peaks are present, especially in the minor contributions. For 3P mixtures, additional reworks

are not undertaken unless there is a reasonable level of uncertainty. Similarly, there is no need to repeatedly confirm a 4P mixture unless there are other aspects of the DNA profile that indicate that it might not truly be 4P. There are situations where one amplification will be sufficient to infer the number of contributors, however two amplifications are advised in most circumstances particularly where one or more contributors are low level. If there is no direct indication of >4P following two amplifications, then the DNA profile is most likely acceptable for further analysis as a 4P mixture. Care should be taken with routinely using three amplifications as this can lead to STRmix™ having increased certainty of allelic pairings. Whilst this may be correct, it can cause uncertainty in the mind of the scientist as to whether such weightings are appropriate for a profile with so many contributors. Accordingly, use three amplifications with caution.

One of the most difficult configurations to assess is the 1P strong major contribution, with a 1P sub-major contribution and a 2P minor. This type of profile requires increased consideration due to potential issues with combined peaks. There are no exact ratios of major:sub-major:minor but the scientist's experience with how peaks combine and how stochastics affects peaks will be the most informative approach. Reworking is strongly advised in this situation to ensure that the sub-major profile is consistent between runs. If the rerun is not consistent or the separation between major, sub-major, minor is uncertain after examining the STRmix™ deconvolution, then it is reasonable to consider that the profile may be complex.

Things to looks for in a STRmix™ Deconvolution:

As with the assessment of any 2P or 3P mixture, the STRmix™ deconvolution of a 4P mixture requires checking for the robustness of the result. Checking all the possible combinations is daunting and immensely time consuming but is not required. However, there are some areas where checks must be performed. It is incumbent upon the case-manager to make whatever checks they feel necessary based on the profile they are analysing. These include, but are not limited to:

- Do the degradation curves fit with the guidelines given above for potentially accepting a 4P mixture for analysis?
- Does any deconvoluted major profile intuitively fit with the observed profile?
- Is dropout (Q) modelled for all minor peaks where expected?
- Has double dropout (Q,Q) been modelled for at least the lowest contribution at all loci where there are less than 7 distinct alleles?
- Do the diagnostics meet expectation?

It should be reiterated that a probabilistic result of 0.99 does not mean p=1.00. The 0.99 value is just the NCIDD upload threshold. Just because something is deconvoluted 99% of the time does not mean it is the only answer, or the "correct" answer. It is just the most likely answer based on the profile obtained, allowing for some of the common profile anomalies, such as degradation, amplification inefficiencies, and stochasticity. Remember, STRmix™ is a statistical model and models are only ever an approximation of reality. The exact probabilities given by STRmix™ should not be a major impasse as long as all other reasonable combinations for a locus have been modelled and are an intuitive fit to the observed profile.

Relevant Literature

Bright et al. (2014b)

> *"Abstract: The effect of uncertainty in the number of contributors to a profile is a matter of some contention in forensic DNA interpretation. Interpretation methods are moving towards continuous models. Within this paper the effect of misspecification of the number of contributors to a profile caused by one artefactual peak, either a large back stutter or a forward stutter, was investigated using a continuous model. The misassignment of the number of contributors to a profile either has no significant effect or decreases the LR for the true contributors. It often increases the chance of an adventitious link."*

Bright et al. (2014a), which was published prior to the above study, showed that adding an extra contributor where one was not present did not significantly affect the LR for a clear major contribution. They concluded that there may be a significant decrease in the LR assigned to minor or trace contributions.

> *"The assumption of additional contributors to a profile beyond that suggested by allele count alone tended to lower the LRC for the true minor and major contributors and increase the number of low grade adventitious links, where 1 < LRC < 1000. A match against the database is unlikely for a trace contributor that has very few alleles either present above the analytical threshold and present in non-masked allele positions."*

It should be noted that the profiles from which these conclusions were drawn were clear mixtures with all peaks being >50rfu and the profile ratios being in the order of 3:1 or 4:1. The addition of an extra contributor would be unreasonable based on the examples given in the paper. These types of EPGs do not generally present an issue within DNA Analysis where most of the uncertainty is closer to the LOD and mixture ratios are generally in to 10-20:1 ratio.

Subsequently, in Bright et al (2014b), the authors consider a small, and unspecified, number of synthetic mixtures where the template ratios could be controlled, which they compared to 1000 synthetic reference samples. Similar to their 2014a paper, they found that the addition of an extra contributor did not significantly affect the LR generated by known contributors. The number of adventitious matches was generally higher when an extra contributor was added, but not always. It should be noted that in the examples given, there were adventitious matches to most profiles irrespective of whether the contributor count was correct or n+1. Where numerous (A,Q) and (Q,Q) designations are required to explain such profiles, the number of adventitious matches to uncertain minor contributions with an LR>1 is relatively unlikely.

They conclude:

> *"The assumption of an increased number of contributors may significantly decrease the LR assigned to known minor and trace contributors to the mixture. This was observed in two key situations in this study. First, if there is no indication of an additional contributor, then the additional contributor's alleles must all be shared with the minor, effectively diffusing the weights across many more genotype combinations. Second, the treatment of a larger than threshold stutter as an additional contributor has almost no effect on the LR of the known contributors until that stutter peak is sufficiently large that it interacts with*

the minor contributors. When the additional height above expected stutter is similar in height to the alleles from a minor or trace component the LR of that minor is decreased."

In interpreting DNA profiles, the current practice is to not add an additional contributor unless there is some evidence to do so. As such, the first conclusion of this paper is not particularly applicable. Even if this is based on a high stutter or similar artefact, it is unlikely to have a marked impact on the LR for actual contributors. The statement "...until that stutter peak is sufficiently large that it interacts with the minor contributors." Suggests that the minor peaks in the DNA profiles are larger than any stutter generated. This is not often the case in the types of DNA profile that cause the most uncertainty. As the minor peaks decrease in size, the weightings applied will also decrease and the number of allowable A,Q and Q,Q designations may also increase. As a consequence, the undesirable effects noted by Bright et al. are not going to be as significant in most profiles of concern. Hence, if there is evidence to do so, there is no issue with adding an additional contributor to an uncertain 3P mixture that may be a 4P mixture.

Paoletti et al. (2005)

"Abstract: Samples containing DNA from two or more individuals can be difficult to interpret. Even ascertaining the number of contributors can be challenging and associated uncertainties can have dramatic effects on the interpretation of testing results. Using an FBI genotypes dataset, containing complete genotype information from the 13 Combined DNA Index System (CODIS) loci for 959 individuals, all possible mixtures of three individuals were exhaustively and empirically computed. Allele sharing between pairs of individuals in the original dataset, a randomized dataset and datasets of generated cousins and siblings was evaluated as were the number of loci that were necessary to reliably deduce the number of contributors present in simulated mixtures of four or less contributors. The relatively small number of alleles detectable at most CODIS loci and the fact that some alleles are likely to be shared between individuals within a population can make the maximum number of different alleles observed at any tested loci an unreliable indicator of the maximum number of contributors to a mixed DNA sample. This analysis does not use other data available from the electropherograms (such as peak height or peak area) to estimate the number of contributors to each mixture. As a result, the study represents a worst-case analysis of mixture characterization. Within this dataset, approximately 3% of three-person mixtures would be mischaracterized as two-person mixtures and more than 70% of four-person mixtures would be mischaracterized as two- or three-person mixtures using only the maximum number of alleles observed at any tested locus."

This paper was published in 2005, prior to the introduction of continuous interpretation and probabilistic modelling. The researchers only modelled mixtures by allelic designation and did not consider any information that might be adduced by the relative mixture proportions, an omission noted by the authors themselves. This also means that homozygotic contributions would be indicated by only a single allelic designation rather than a double height peak which would be more likely to be evident in a normal DNA profile. So, for example, a locus with four peaks would be only ever be considered as indicating 2P in the Paoletti *et al.* study, whereas if one of those peaks was considerably higher than the other three, indicating the possible presence of a homozygote, it would indicate the potential for a third contributor at that locus in a normal DNA profile.

The authors point out that the ability of a DNA profiling system to accurately reflect the true number of contributors increases as the number of loci increases. The additional loci required to reach an

acceptable level of mischaracterization of contributor number (<5%) is much of the focus of this paper. This study utilized the 13 CODIS loci that were available at that time and so does not examine a number of the most discriminatory loci that are available in PP21, such as Penta E, D1, D2 and D12. Accordingly, the ability of the 13 CODIS loci to indicate the true number of contributors is generally going to be much less informative than that of a PP21 profile.

In retrospect, this paper was an interesting theoretical paper at the time of its publication, but it is no longer particularly relevant given the major changes that have occurred in DNA profile interpretation, modelling, and the subsequent expansions of profile systems to 21 or more loci.

Dembinski et al. (2018)

> *"Abstract: DNA mixtures are more frequently encountered in casework due to increased kit sensitivity, protocols with increased cycle number, and requests for low copy number DNA samples to be tested. Generally, the first step in mixture interpretation is determining the number of contributors, with the most common approach of maximum allele count. Although there are previous studies regarding the accuracy of this approach, none have evaluated the accuracy with the newly expanded U.S. core STR loci. In this work, 4,976,355 theoretical mixture combinations were generated with the PowerPlex® Fusion 6C system which includes 23 autosomal STR loci and three Y-STR loci. The number of contributors could be correctly assumed for 100% two-person and 99.99% three-person mixtures, whereas, four-, five-, and six-person mixtures were correctly assumed in 89.7%, 57.3%, and 7.8% of mixtures, respectively. Y-STR analysis showed the 3 Y-STR markers are only accurate for two-person male mixtures (96.7%). This work demonstrates that maximum allele count using the expanded U.S. core loci is not much improved from previous smaller panels, reiterating that this method is not as accurate beyond three contributors."*

The kit used in this study uses three more loci than PP21 and so would potentially be expected to be more discriminatory. It does not include information about stutter, allelic drop-out and amplification effects, which are relevant to contemporary real-world DNA profile analysis. In ~910 thousand known 4P mixture combinations, ~90% had at least one locus that indicated >7 alleles and in ~960 thousand 5P mixtures there was a locus with at least nine alleles in ~43% of the mixtures.

Similar to Paoletti *et al*. it does not attempt to look at mixture proportions. So the same critique given above applies to this paper, in that the EPG will generally allow for improved contributor discrimination due to differing mixture proportions where contributions are above stochastic levels. However, there is potential that numerous contributions at sub-stochastic level peaks could impair discrimination, but the authors make the point that the profile needs to be examined as a whole.

> *"When considering total allele count across a four-person mixture compared to a three-person mixture, there is a decrease in loci that exhibit four alleles and increase in those exhibiting five and six alleles."*

Thus, the proportion of loci indicating the true number of contributors is generally shifted towards the true number of contributors. Similarly, a true 5P mixture will tend to have more loci exhibiting seven or eight alleles. However, they note that this relationship becomes less marked as the number of contributors increases above five.

Appendix

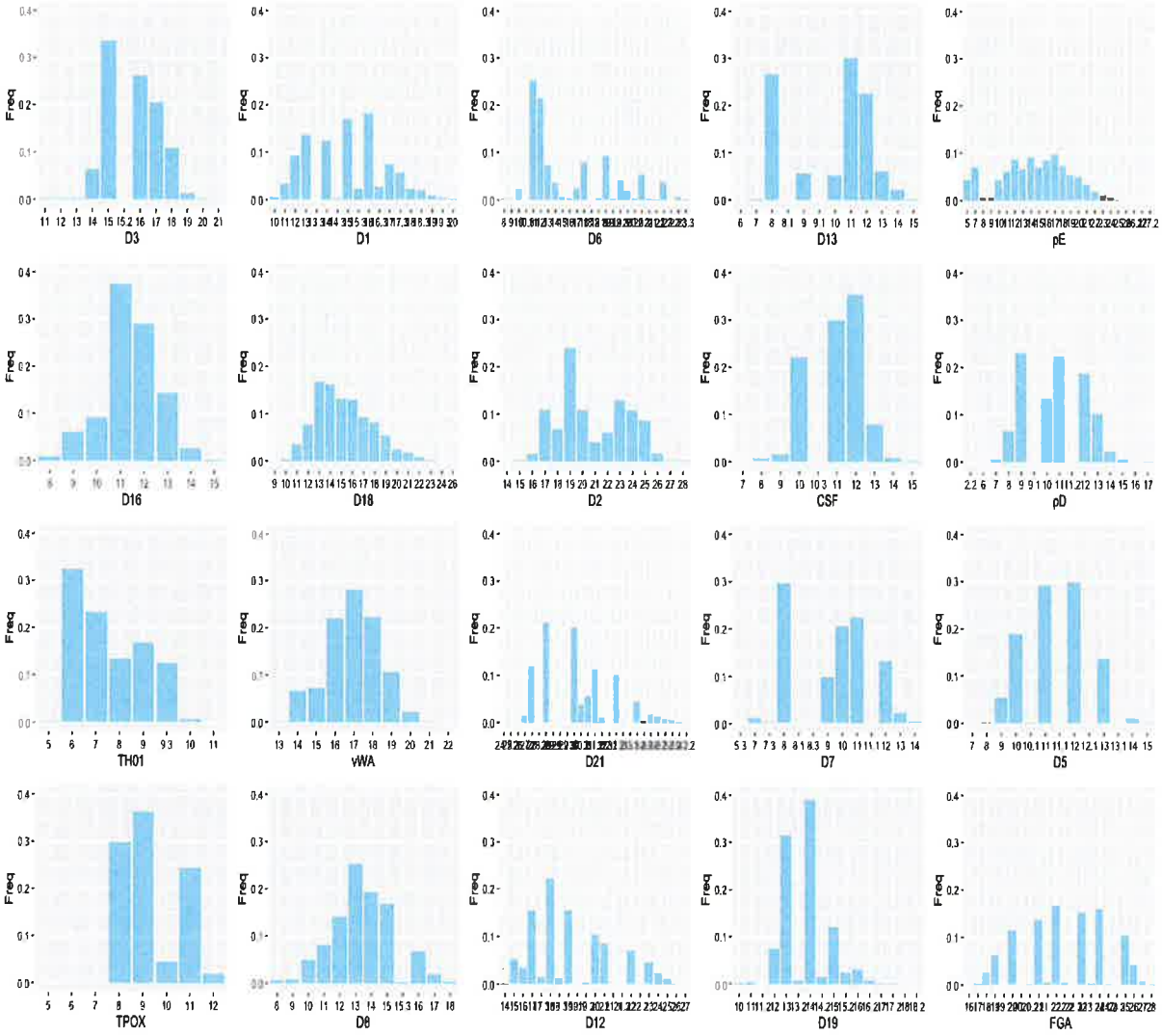

Figure 1. Bar plots of PP21 loci by frequency of allele.

The above figure shows the relative frequencies for individual alleles by locus. The following points are relevant to 4P mixture interpretation:

- The higher the bar the more likely that allele is to be represented in any given DNA profile. For example, a 6 STR at TH01 has a relative frequency of approximately 32.5 %. This means that roughly 32.5% of the Australian population will have this allele designation at TH01.
- A locus with fewer possible allele designations will be less discriminatory than a locus with a larger array of possible designations. However, this fact is counter-balanced by the relative frequency of the alleles. For example, D19 has a large number of possible designations but over 80% of the population is represented by just three alleles (13,14, and 15); the rest being infrequent (0.03-5% frequency) or rare. Thus, D19 is not considered a highly discriminatory locus despite having a larger array of possible alleles.
- The most discriminatory loci are those that have a relatively equal frequency distribution across a wide array of possible allelic designations. For example, the large range of possible alleles at Penta E, D1, D2, D12 and D18 are more evenly distributed within the population.

This makes these loci much more discriminatory and therefore more likely to give a better indication of the number of contributors to a mixture.

References

Bright J, Taylor D, Curran J, Buckleton J (2014a) Searching mixed DNA profiles directly against profile databases, Forensic Sci. Int. Genet. 9 (2014) 102–110.

Bright J, Curran JM, Buckleton JS (2014b) The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation. Forens Sci Int:Genetics 12:208–214.

Dembinski GM, Sobieralski C, and Picarda CJ (2018) Estimation of the number of contributors of theoretical mixture profiles based on allele counting: Does increasing the number of loci increase success rate of estimates? Forens Sci Int: Genetics. 33:24–32.

Morgan R and Caunt E (2014) Assessment of the Number of Contributors for Mixed PowerPlex® 21 DNA Profiles within Forensic DNA Analysis. Internally Published Guideline. FSS DNA Analysis.

Paoletti DR, Doom TE, Krane CM, Raymer ML, and Krane DE (2005) Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures. J Forensic Sci. 50(6):1-6.