Queensland Health
HealthSupport Queensland

Proposal # 231 – Verification of STRmix™ v2.8.0 Report

Angela Adamson, Cassandra James, Emma Caunt, Allan McNevin & Cathie Allen

March 2021

Queensland Government

Contents

Document Details

Contact for enquiries and proposed changes

If you have any questions regarding this document or if you have a suggestion for improvements, please contact:

Contact officer: Angela Adamson
Title: Reporting Scientist, FRIT
Phone: [redacted]
Email: [redacted]

Version history

Version	Date	Changed by	Description
1.0	8 March 2021	E Caunt, A Adamson, C James	Document Created.
2.0	30 March 2021	A Adamson A McNevin	Incorporated all feedback. Removed Figure 4 from v.1.0 – not relevant. Formatting of pages/tables

Document sign off

This document has been **approved** by:

Name	Position	Signature	Date
Cathie Allen	Managing Scientist		

The following officers have **endorsed** this document

Name	Position	Signature	Date
Justin Howes	Team Leader FRIT		

Name	Position	Signature	Date
Paula Brisotto	Team Leader ER & Q		

Name	Position	Signature	Date
Luke Ryan	Senior Scientist Analytical		

Name	Position	Signature	Date
Allan McNevin	Senior Scientist ER		

Name	Position	Signature	Date
Kirsten Scott	Senior Scientist Q & P		

Name	Position	Signature	Date
Allison Lloyd	A/Senior Scientist Intel		

Name	Position	Signature	Date
Sharon Johnstone	Senior Scientist Reporting 1		

Name	Position	Signature	Date
Kylie Rika	Senior Scientist Reporting 2		

1. INTRODUCTION

STRmix™ was introduced into Forensic DNA Analysis in December 2012 as the interpretational tool for profiles generated using the PowerPlex® 21 amplification kit (PP21). As more functionality is added to STRmix™ by the developers, new versions are made available to the user. The current version in use in Forensic DNA Analysis is version 2.7.0.

Forensic DNA analysis has recently implemented the 3500xL genetic analyser for the analysis of casework profiles which are subsequently interpreted using STRmix™ v2.7.0. Since implementation it has been noted that more complex DNA profiles, in particular four contributor mixtures, can cause an 'Out of Memory' error with STRmix™ v2.7.0 and are therefore unable to be interpreted. Those that do run to completion often trigger the 'integer overflow issue'[4]. Initial testing with STRmix™ v2.8.0 has shown that these complex profiles are able to be interpreted with this new version due to the memory improvements and correction of the 'integer overflow issue'.

STRmix™ version 2.8.0 features the following changes:

- Improvements to both low and normal memory modes to allow for processing of larger problems with lower RAM requirements
- Model improvements in allele frequency sampling to use k+1 (instead of k) within posterior allele frequency
- Change of total iteration counts from integers to longs to prevent the possibility of too many iterations causing an integer overflow

A full list of full features can be found in the STRmix™ 2.8 User's Manual located in I:\STRmix™ settings\Manuals and Release Notes.

STRmix™ v2.8.0 is backwards compatible with versions 2.6.0 to 2.6.3 and 2.7.0[2]

2. PURPOSE AND SCOPE

The purpose of this project was to:

- Verify the use of STRmix™ v2.8.0 within Forensic DNA Analysis for the interpretation of DNA profiles generated using the 3500xl Genetic Analyzer.
- Assess identified mixed profiles that have been unable to successfully run on STRmix™ v2.7.0 due to insufficient PC memory and the integer overflow issue.

3. GOVERNANCE

Project Personnel

- Project Manager: Justin Howes – Team Leader, Forensic Reporting and Intelligence Team
- Project Officer: Angela Adamson – Scientist, Reporting 1 Team.
- Project Officer: Allan McNevin - Senior Scientist, Evidence Recovery
- Project Officer: Cassandra James – Scientist, Reporting 1 Team.
- Project Officer: Emma Caunt - Scientist, Reporting 2 Team.

Decision Making Group

The Management Team are the Decision Making Group for this project and may use information obtained from parts of this project to cease or extend part or all of the experimentation at any stage. The Decision Making Group may also make modifications to this Experimental Design as required to the approved Experimental Design, however this must be documented and retained with the original approved Experimental Design.

Reporting

Updates as necessary will be provided at the management team meetings and to the Managing Scientist. Draft and final reports will be provided to the Decision Making Group for review.

4. MATERIALS, METHODS AND TECHNIQUES

4.1 Pre-existing laboratory specific settings

STRmix™ v2.8.0 will use all laboratory specific settings established for the interpretation of 3500xL profiles using STRmix™ v2.7.0. As per the release and testing report[2], Model Maker does not need to be re-run.

4.2 Genemapper Analysis

All samples were analysed with GeneMapper®ID-X v1.6 using the Promega PowerPlex®21 system analysis panels.

For the samples run on the 3500xL the conditions were as follows:

- +1 rpt and -1 rpt stutter peaks were labelled, -2 rpt stutter peaks had the labels removed.
- Amplification artefacts and CE artefacts were removed.

5. SAMPLE SELECTION

5.1 Single source DNA profiles

The single source sample amplified at a range of total DNA inputs used in *Project #219 – Validation of STRmix™ v2.7.*0 Part A was used in the Verification of STRmix™ 2.8.0 for 3500xL.

Table 1 Single source sample (Part A)

Sample #	Input template (ng)
930650041	0.5
930650041	0.338
930650041	0.182
930650041	0.065

5.2 Mixed DNA profiles

Mixed DNA samples containing two or three contributors at varying proportions and input templates used in *Project#219 – Validation of STRmix™ v2.7.0 Part A* were used in the Verification of STRmix™ 2.8.0 for 3500xL (Table 2).

Table 2 Mixture Ratios (part A)

Sample #	Contrib #	Barcode	Ratio	Input template (ng)	Number of amps used in STRmix
1	2	603317299	1:1	0.502	1/2/3
2	2	603317288	5:1	0.490	1/2/3
3	2	603317277	20:1	0.500	1/2/3
4	3	603317266	5:2:1	0.510	1/2/3
5	3	603317255	30:1:1	0.499	1/2/3
6	2	129549144	1:1	0.690	1
7	2	129549144	1:1	0.506	1
8	2	129549144	1:1	0.198	1
9	2	129549155	5:1	0.690	1
10	2	129549155	5:1	0.506	1
11	2	129549155	5:1	0.198	1
12	3	129549177	5:2:1	0.638	1
13	3	129549177	5:2:1	0.510	1
14	3	129549177	5:2:1	0.208	1
15	3	129549188	30:1:1	0.683	1
16	3	129549188	30:1:1	0.501	1
17	3	129549188	30:1:1	0.200	1

Mixed DNA samples containing two and three contributors of varying proportions used in *Project#219 – Validation of* STRmix™ *v2.7.0 Part B* were used in the Verification of STRmix™ 2.8.0 for 3500xL (Table 3).

Table 3 Mixture Ratios (Part B)

Sample	Contrib	Barcode	Ratios
1	2	706830594	20:1
2	2	706830571	10:1
3	2	706830606	5:1
4	2	706830615	2:1
5	2	706830629	1:1
6	3	706830638	30:1:1
7	3	706830647	20:10:1
8	3	706830651	10:5:1
9	3	706830660	5:2:1
10	3	706830629	1:1:1

Ten four contributor mixed DNA profiles of varying proportions used in *Project #219 – Validation of STRmix™ v2.7.0 Part C* were used as shown in Table 4.

Table 4 Mixture Ratios (Part C)

Sample	Contrib	Barcode	Ratios
1	4	706832950	20:1:1:1
2	4	706832946	10:5:2:1
3	4	706832937	10:10:1:1
4	4	706832928	4:3:2:1
5	4	706832919	1:1:1:1
6	4	706832905	20:1:1:1
7	4	706832893	10:5:2:1
8	4	706832884	10:10:1:1
9	4	706832870	4:3:2:1
10	4	706832861	1:1:1:1

6. EXPERIMENTAL DESIGN

6.1 Verification of the use of STRmix™ v2.8.0

6.1.1 Single source DNA profiles

These profiles were interpreted using Method 4.2.

Since the single source DNA profiles from *Project#219 – Validation of STRmix™ v2.7.0 Part A* were originally interpreted in STRmix™ v2.7.0 using settings determined prior to the 3500xL laser change, the profiles were re-interpreted in STRmix™ v2.7.0 using the settings determined in *Project#219 – Validation of STRmix™ v2.7.0 Part B*. This allowed a direct comparison with the results obtained from STRmix™ v2.8.0.

All profiles were deconvoluted in STRmix™ v2.8.0, and the likelihood ratios (LRs) calculated. All should have been similar to those values obtained from STRmix™ v2.7.0.

Acceptance criteria

The use of STRmix™ v2.8.0 for the interpretation of single source DNA profiles using the 3500xL genetic analyser was considered to be verified if the deconvolutions were similar to those values obtained from STRmix™ v2.7.0.

6.1.2 Mixed DNA profiles

These profiles were interpreted using Method 4.2.

Since the mixed DNA profiles from *Project#219 – Validation of STRmix™ v2.7.0 Part A* were originally interpreted in STRmix™ v2.7.0 using settings determined prior to the 3500xL laser change, the profiles were re-interpreted in STRmix™ v2.7.0 using the settings determined in *Project#219 – Validation of* STRmix™ *v2.7.0 Part B*. This allowed direct comparison with the results obtained from STRmix™ v2.8.0.

All mixed DNA profiles (Tables 2, 3 and 4) were deconvoluted in STRmix™ v2.8.0, and the likelihood ratios (LRs) calculated

Within the deconvolutions of the mixed DNA profiles, the top ten genotype combinations for each locus was compared and should be similar to those values obtained from STRmix™ v2.7.0.

The LRs should be similar to those obtained in STRmix™ v2.7.0.

Acceptance criteria

The use of STRmix™ v2.8.0 for the interpretation of mixed DNA profiles using the 3500xL genetic analyser will be considered to be verified if the deconvolutions are similar to those values obtained from STRmix™ v2.7.0.

6.2 Allele frequency sampling

LRs were calculated in STRmix™ v2.8.0 for all single source and mixed DNA profiles (Tables 1,2,3 and 4) using their STRmix™ v2.7.0 deconvolution. The LRs produced by STRmix™ v2.8.0 were compared with the LRs obtained from STRmix™ v2.7.0 to determine the effect of the change in allele frequency sampling.

6.3 Ignore loci function

Information was removed from selected reference profiles to create partial references and compared with selected DNA deconvolutions used within this project to assess the 'Ignore Loci' function within STRmix™ v2.8.0.

6.4 Additional testing to investigate 3130*xl* samples

STRmix™ v2.8.0 was set up to run 3130*xl* profiles using the settings described in Project 214[4]. Five 3130xl samples were run in STRmix™ v2.8.0 to make a simple comparison of these samples with their previous run on STRmix™ v2.7.0. The comparison consisted of one single source and one four-person mixed sample run only on the 3130xl and a two-person mixed sample and a three-person mixed sample that had both been run on the 3130xl and 3500xL combined. The LRs generated with the 3130xl and 3130xl/3500xL combined using STRmix™ v2.8.0 were compared with the previous LRs generated with STRmix™ v2.7.0 in project 219[3] and project 214[4].

7. RESULTS AND DISCUSSION

7.1 Verification of the use of STRmix™ v2.8.0

7.1.1 Single source and mixed DNA profiles

When samples were deconvoluted in both STRmix™ v2.7.0 and STRmix™ v2.8.0 the LRs obtained were close in value (refer Figure 1). Where differences were observed, these differences were less than one order of magnitude. In instances where the difference was greater than one order of magnitude this was likely a result of the variation within the deconvolution. The difference in LRs between deconvolutions in STRmix™ v2.7 and STRmix™ v2.8 is due to the inherent variation in the MCMC but also due to some of the changes between versions[2].

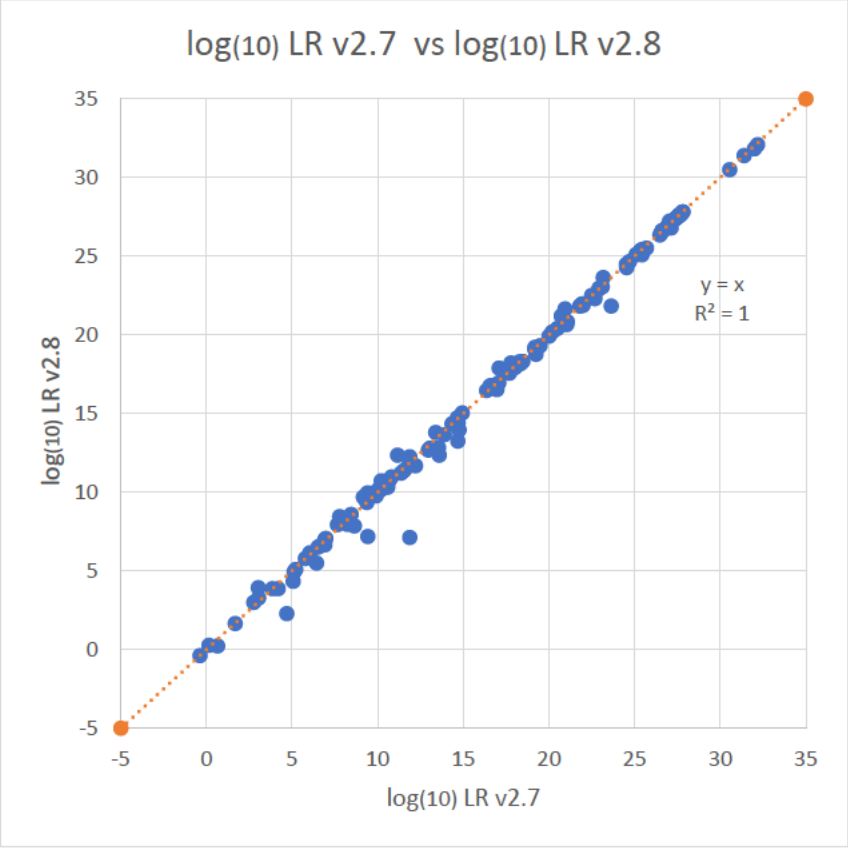

Figure 1 log(10) LR v2.7 vs log(10) LR v2.8

The top ten genotype combinations for each locus in the deconvolutions of the mixed DNA samples were compared and most were noted to be similar to those values obtained from STRmix™ v2.7.0. Any samples that showed differences in the top ten genotype combinations were due to fluctuations in the weightings caused by the variability of the MCMC.

Five of the mixed samples run in STRmix™ v2.8.0 had a high GR greater than 1.2 (ranging from 1.21 to 1.28). The samples were repeated using double accepts (20k/100k accepts) to investigate if this would decrease the GR value. The process of doubling the accepts resulted in a lower GR in all these samples however two samples still had a GR greater than 1.2.

The remaining samples were repeated with double accepts to investigate if it would be beneficial to use the double accepts settings as a default within STRmix™ v2.8.0. The GR and time were compared to the deconvolutions with normal settings (10k/50k accepts).

Out of 51 deconvolutions, the process of doubling the accepts resulted in a lower GR in the majority of samples, however there were 13 samples with a higher GR than those observed using the normal settings.

In general, the time (seconds) taken for each deconvolution with double accepts increased when compared to the deconvolution time with normal settings (see Figure 2).

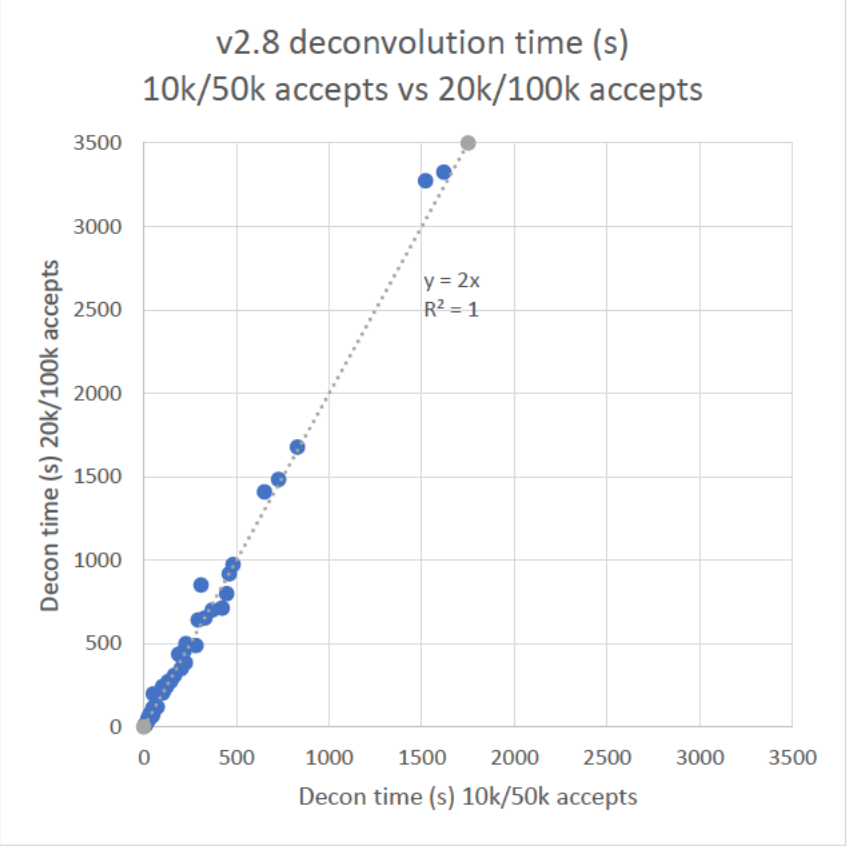

Figure 2 v2.8 deconvolution time (s) 10k/50k accepts vs 20k/100k accepts

7.2 Allele frequency sampling

The LRs were calculated in STRmix™ v2.8.0 for all single source and mixed DNA samples listed in Table 1 - Table 4 using their STRmix™ v2.7.0 deconvolution. There was no noticeable difference in the LRs produced by STRmix™ v2.8.0 using the STRmix™ v2.7.0 deconvolutions when compared with the LRs obtained from STRmix™ v2.7.0 using STRmix™ v2.7.0 deconvolutions. The results are listed below in Figure 3.

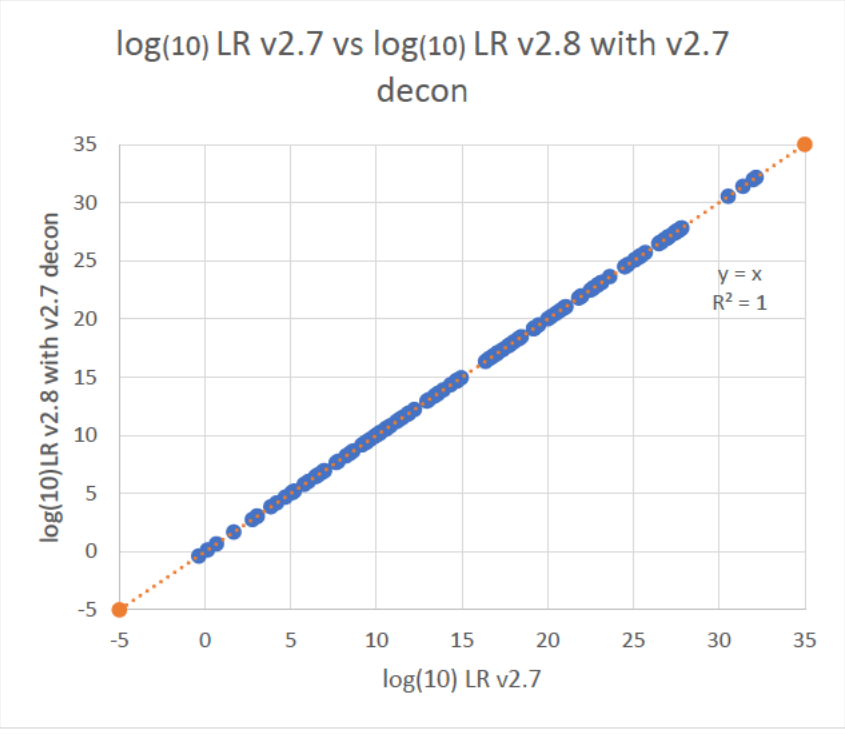

Figure 3 log(10) LR v2.7 vs log(10) LR v2.8 with v2.7 decon

Figure 3 shows that the LRs obtained from STRmix™ v2.7.0 were comparable to the LRs obtained from STRmix™ v2.8.0 using a STRmix™ v2.7.0 deconvolution. These results agree with those observed by the developers [2].

7.3 Ignore loci function

The ignore loci function exists in both STRmix™ v2.7.0 and STRmix™ v2.8.0. The use of this function when calculating LRs appears to be the same process for partial reference samples in both versions.

STRmix™ v2.7.0 however, does not allow for conditioning with a partial reference (Figure { SEQ Figure * ARABIC }).

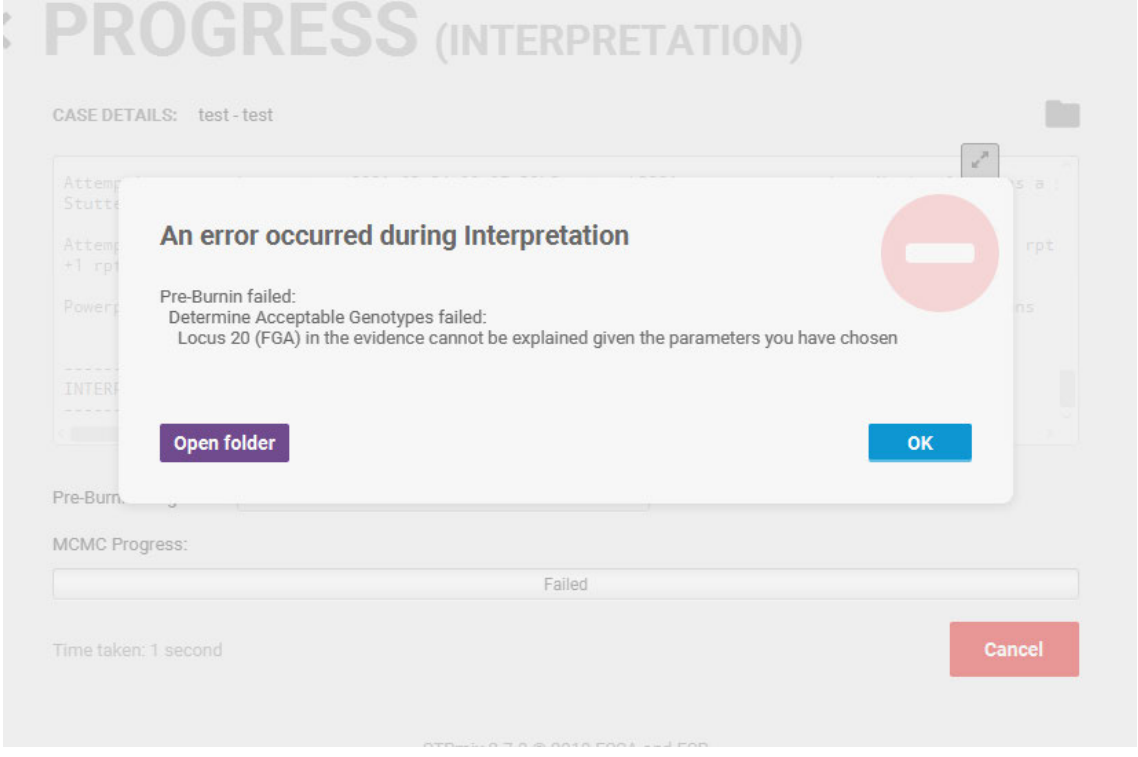

Figure { SEQ Figure * ARABIC } Error message when attempting to condition using a partial reference in STRmix™ v2.7.0

In STRmix™ v2.8.0 the change is to automatically ignore loci that are missing from a partial reference when used for conditioning. This is highlighted in in the comment section of the report (Figure 5).

EVIDENCE INPUT FILES

C01_706830571_FR1904099_EV.CSV

LOCUS	ALLELE	HEIGHT	SIZE	COMMENT
AMEL	X	5516	78	Gender Locus
	Y	4048	85	Gender Locus
D3S1358	16	941	124	Ignored - Partial Assumed Reference
	17	6115	128	Ignored - Partial Assumed Reference
	18	4458	133	Ignored - Partial Assumed Reference
D1S1656	10	267	160	
	11	4853	164	
	16.3	921	187	
	17.3	365	191	

Figure 4 Report when conditioning when using a partial reference in STRmix™ v2.8.0

7.4 Additional testing to investigate 3130*xl* samples

The single source sample and the mixed DNA samples run on the 3130*xl* were deconvoluted in STRmix™ v2.8.0, and the likelihood ratios (LRs) calculated. The LRs obtained using STRmix™ v2.8.0 were close in value to those obtained using STRmix™ v2.7.0 (refer Figure 6).

Where differences were observed, these differences were less than one order of magnitude.

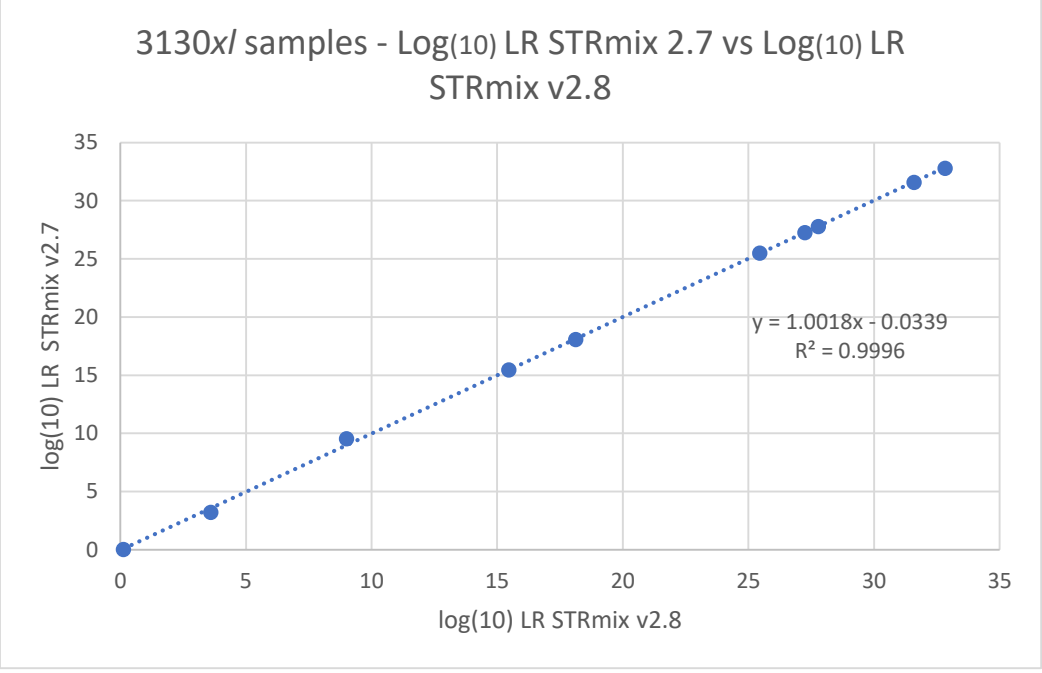

Figure 5 log(10) LR v2.7 vs log(10) LR v2.8 for 3130*xl* samples

8. CONCLUSION

Based on the findings described above, STRmix™ v2.8.0 has been verified for the interpretation of mixed DNA profiles using the 3500xL genetic analysers. Additional testing has verified the ability for STRmix™ v2.8.0 to interpret profiles run on the 3130*xl*.

The deconvolutions and LR's obtained using STRmix™ v2.8.0 were comparable to those obtained using STRmix™ v2.7.0. Any variability in the LRs is due to variations in the deconvolution.

There was no noticeable difference observed in the LR's run in STRmix™v2.7.0 and STRmix™ v2.8.0 using the deconvolutions originally run on STRmix™ v 2.7.0 indicating no observed difference from the changes to the allele frequency calculation [2].

A GR above 1.2 does not necessarily mean that a deconvolution needs to be repeated or is unusable, especially if all other diagnostics are as expected and the deconvolution agrees with the intuitive interpretation of that profile.

Doubling the accepts in deconvolutions using STRmix™ v2.8.0 has shown that generally it will lower the GR value most of the time but not all of the time. Using double accepts is not considered a requirement for use as a default setting within STRmix™ v2.8.0 however, this can be considered to investigate deconvolutions that don't converge.

STRmix™ v2.8.0 can use the conditioning function with partial references.

Considering all of the findings within this report, the use of STRmix™ v2.8.0 for the interpretation of mixed DNA profiles has been verified using the 3500xL and 3130*xl* genetic analysers.

9. REFERENCES

1 ESR, *STRmix™ V2.8 Implementation and Validation Guide*. 2020 Institute of Environmental Science and Research: New Zealand.

2 ESR, *STRmix™ V2.8 Release and Testing Report*. 2020 Institute of Environmental Science and Research: New Zealand.

3 Proposal # 219 – Verification of STRmix v2.7.0 for 3500xL, Emma Caunt, Cassandra James, Angela Adamson & Cathie Allen

4 Proposal#214 - Validation of STRmix v2.7 Emma Caunt, Hannah Pattison, Allan McNevin, Justin Howes & Cathie Allen