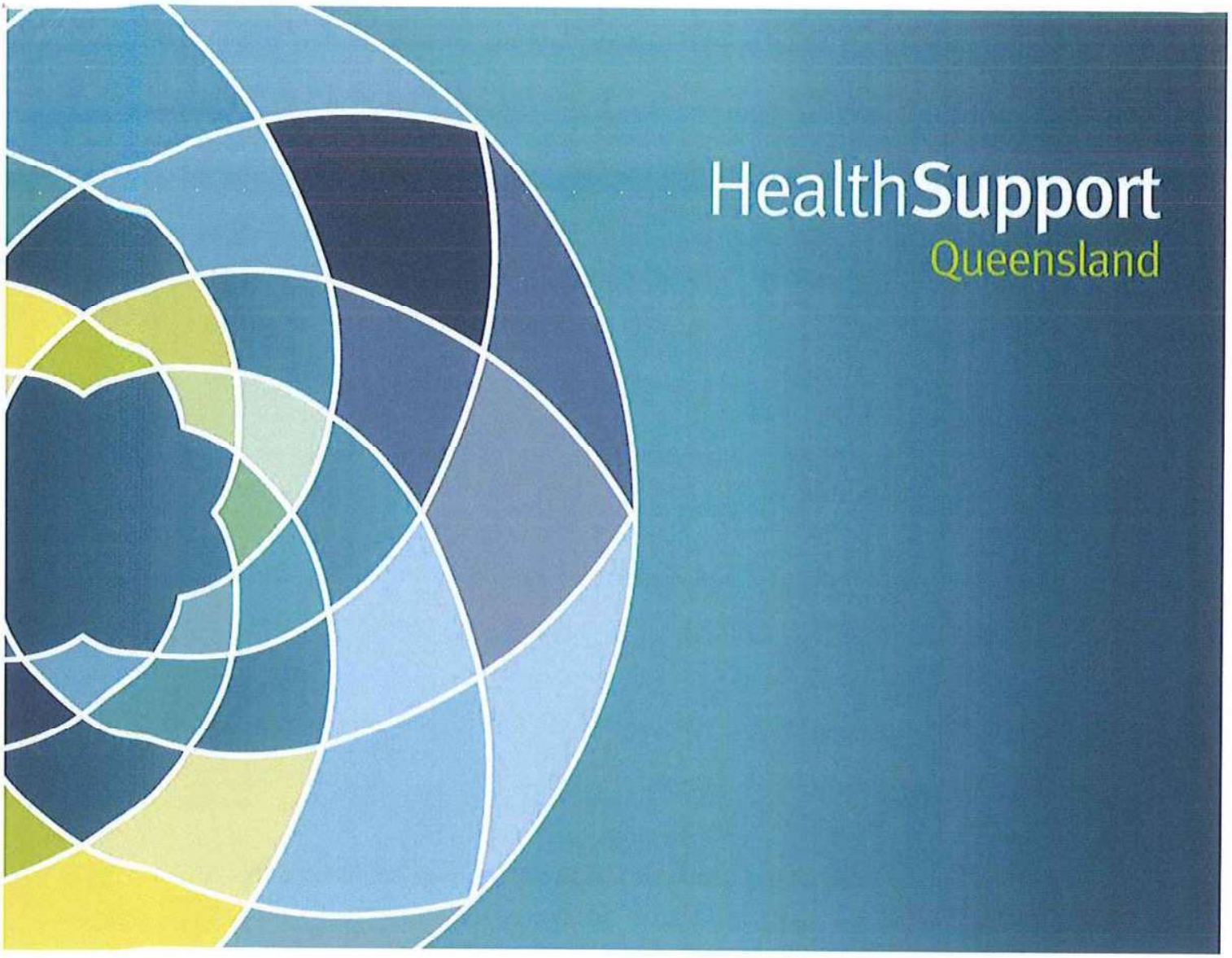

3500xL Genetic Analyzer Validation for Casework Samples Amplified with Powerplex®21

Megan Mathieson, Biljana Micic, Luke Ryan and Cathie Allen

September 2015

Queensland Government

3500xL Genetic Analyzer Validation for Casework Samples Amplified with PowerPlex®21

Published by the State of Queensland (Queensland Health), September 2015

For more information contact:
Forensic DNA Analysis, Department of Health, GPO Box 48, Brisbane QLD 4001.

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: Luke Ryan
Title: Senior Scientist – Analytical
Phone: [redacted]
Email: [redacted]

Version history

Version	Date	Changed by	Description
1.0	June 2015	Biljana Micic, Megan Mathieson, Luke Ryan	Document Created

Document sign off

This document has been **approved** by:

Name	Position	Signature	Date
Cathie Allen	Managing Scientist	[redacted]	01/10/2015

The following officers have endorsed **this document**

Name	Position	Signature	Date
Luke Ryan	A/Team Leader ER&Q	[redacted]	29-09-2015

Name	Position	Signature	Date
Justin Howes	Team Leader FRIT	[redacted]	29.09.15

Name	Position	Signature	Date
Megan Mathieson	A/Senior Scientist Analytical	[redacted]	29-09-2015

Name	Position	Signature	Date
Allan McNevin	Senior Scientist ER	[redacted]	29.09.2015

Name	Position	Signature	Date
Kerry-Anne Lancaster	A/Senior Scientist Q&P	[redacted]	29.09.2015

Name	Position	Signature	Date
Sharon Johnstone	Senior Scientist Intel	[redacted]	29/09/15

Name	Position	Signature	Date
Amanda Reeves	Senior Scientist Reporting 1	[redacted]	29/09/15

Name	Position	Signature	Date
Kylie Rika	Senior Scientist Reporting 2	[redacted]	01/10/2015

Contents

Figures

Tables

1 Abstract

Forensic DNA Analysis currently uses the 3130*xl* Genetic Analysers (Life Technologies, Applied Biosystems, Foster City, CA, US) for capillary electrophoresis. These instruments have been superseded by the 3500 Series Genetic Analyser and the servicing and sourcing of replacement parts has become difficult and time consuming.

Currently Forensic DNA Analysis uses the Promega PowerPlex®21 System amplification kit for all reference and the majority of casework samples. This kit has had inherent problems with artefacts and untidy baseline when processed using the 3130*xl* Genetic Analyzers. The 3500 Series may alleviate this by the improved hardware, software and running conditions.

This validation has shown that the 3500xL performs as well as or better than the 3130xl for all experiments excluding the Mixture Studies. Saturation and mixture studies were not completed during this validation due to the quality of the mixture samples and poor spectral separation for PowerPlex®21 produced on the 3500xL. The results from this validation show the 3500xL Genetic Analyzer instruments are unsuitable to use for the analysis of extracted casework samples amplified using PowerPlex®21 at this time.

2 Introduction

Forensic DNA Analysis currently uses the 3130*xl* Genetic Analysers (Life Technologies, Applied Biosystems, Foster City, CA, US) for capillary electrophoresis. These instruments have been superseded by the 3500 Series Genetic Analyser and the servicing and sourcing of replacement parts has become difficult and time consuming.

The updated 3500 Series of Genetic Analysers has improved hardware, pre-packaged consumables with RFID tags and HID specific software to produce increased reliability, performance and ease of use. The 3500 Series instruments operate on a RFU (relative fluorescence units) scale, which is approximately four times the scale of the 31xx Series instruments [1].

Currently Forensic DNA Analysis uses the Promega PowerPlex®21 System amplification kit for all reference and the majority of casework samples. This kit has had inherent problems with artefacts and untidy baseline when processed using the 3130*xl* Genetic Analysers. The 3500 Series may alleviate this by the improved hardware, software and running conditions.

The potential benefits for using the 3500 Series of Genetic Analyzers include:

- solid-state laser utilises standard power supply and requires no heat removal, reducing energy consumption
- reduced signal variation from instrument to instrument, run to run and capillary to capillary
- internal standard normalisation is designed to reduce peak height variation across instruments

- a redesigned oven with an advance thermal system design to improve temperature controls for more consistent data migration and reduced run times
- a new compact polymer pump design to reduce instrument setup time and potential for bubble formation
- new pre-packaged consumable design with RFID technology
- simplified run setup and intuitive software
- increase number of capillaries to 24 (3500xL only), increases the number of samples per injection and reduces (higher throughput) run times

The purpose of this document is to describe the validation performed for casework sample amplified with the PowerPlex®21 System and analysed on the 3500xL.

This validation includes the calculation and comparison of:

- Baseline, limit of detection (LOD) and limit of reporting (LOR) thresholds
- Stutter thresholds
- Peak height ratio and Allelic Imbalance thresholds
- Homozygote thresholds
- Mixture study and saturation
- Drop in
- Repeatability and reproducibility

This validation does not include the GeneMapper®ID-X normalisation function, as it is not compatible with the PowerPlex®21 System.

3 Resources

The following resources were used for this validation.

3.1 Reagents

- 5% v/v Bleach White N Bright (Ecolab, NSW, AU)
- 5% v/v Trigene Advance (CEVA DEIVET Pty. Ltd. Seven Hills, NSW, AU)
- Ethanol (Recochem Incorporated, Wynnum, QLD, AU)
- Nanopure water (Forensic DNA Analysis, Brisbane, QLD, AU)
- Hi-Di™ Formamide (Life Technologies Applied Biosystems, Foster City, CA, US)
- Running Buffer (Gel Company, San Francisco, CA, US)
- 3130 POP-4™ Polymer (Life Technologies Applied Biosystems, Foster City, CA, US)
- Promega PowerPlex®21 system (Promega Corp., Madison, WI, US)
- Promega PowerPlex®21 Allelic Ladder (Promega Corp., Madison, WI, US)
- Promega CC5 Internal Lane Standard (Promega Corp., Madison, WI, US)
- Promega PowerPlex 5 Dye Matrix Standard (Promega Corp., Madison, WI, US)
- 2800M Control DNA, 10ng/µL (Promega Corp., Madison, WI, US)
- Water amplification grade (Promega Corp., Madison, WI, US)

- Anode buffer container (ABC) (Life Technologies Applied Biosystems, Foster City, CA, US)
- Cathode buffer container (CBC) (Life Technologies Applied Biosystems, Foster City, CA, US)
- POP-4™ Polymer 3500 Series (Life Technologies Applied Biosystems, Foster City, CA, US)
- Conditioning reagent (Life Technologies Applied Biosystems, Foster City, CA, US)
- Sarcosyl (Sigma-Aldrich® Corporation, St Louis, MO, US)
- Proteinase K (20mg/mL) (Affymetrix USB, Cleveland, Ohio, USA)
- Dithiothreitol (Affymetrix USB, Cleveland, Ohio, USA)
- TNE Buffer (Forensic DNA Analysis, Brisbane, QLD, AU)
- Quantifiler® Human DNA Quantification kits (Life Technologies Applied Biosystems, Foster City, CA, US)
- Promega Genomic Male DNA (Promega Corp., Madison, WI, US)
- DNA IQ™ Casework Pro Kit for Maxwell® 16 (Promega Corp., Madison, WI, US)
- Amphyl (Rickitt Benckiser Inc. Parsippany, NJ, US)

3.2 Materials

- 96-well PCR half skirt micro-plates (Axygen Scientific Inc., Union City, CA, US)
- 3500xL (24 capillary) arrays (Life Technologies Applied Biosystems, Foster City, CA, US)
- 3130*xl* (16 capillary) arrays (Life Technologies Applied Biosystems, Foster City, CA, US)
- Tape pads adhesive film (Qiagen Pty. Ltd., Doncaster, VIC, AU)
- Sterile 2mL screw-cap tubes (Axygen Scientific Inc., Union City, CA, US)
- Sterile 1.5mL screw-cap tubes (Axygen Scientific Inc., Union City, CA, US)
- ART filtered 1000, 300, and 20P pipette tips (Molecular BioProducts Inc., San Diego, CA, US)
- F1-Clip Tip pipette tips 10µL (ThermoFisher Scientific Inc.)
- Combitips advanced® 0.5mL (Eppendorf Biopur, Hamburg, DE)
- Rediwipes (Cello Paper Pty. Ltd., Fairfield, NSW, AU)
- Septa Cathode buffer container 3500 Series (Life Technologies Applied Biosystems, Foster City, CA, US)
- Reservoir septa (Life Technologies Applied Biosystems, Foster City, CA, US)
- Axygen plate septa (Axygen Inc. Union City, CA, US)
- 96 Well Optical Plates (Life Technologies Applied Biosystems, Foster City, CA, US)
- Optical Adhesive Covers (Life Technologies Applied Biosystems, Foster City, CA, US)

3.3 Equipment

- BSD Duet 600 Series II (BSD Robotics, AU)
- LaboGene Scanspeed 1248 Centrifuge (Labgear, Lynge, Denmark)
- Hot-block (Ratek Instruments Pty. Ltd., Boronia, VIC, AU)

- Biological safety cabinets class II (Westinghouse Pty. Ltd., Newport, AU)
- Refrigerators and freezers (Westinghouse Pty. Ltd., AU)
- FTA® collection kits (Whatman)
- GeneMapper®ID-X ver.1.4 (Life Technologies Applied Biosystems, Foster City, CA, USA)
- GeneAmp PCR system 9700 (Life Technologies Applied Biosystems, Foster City, CA, USA)
- AB 3130*xl* Genetic Analyzer (B) (Life Technologies Applied Biosystems, Foster City, CA, USA)
- Vortex (Ratek Instruments Pty Ltd, Melbourne, VIC, AU)
- Micro centrifuge (Tomy, Tokyo, JP)
- Pipettes (Eppendorf, Hamburg, DE and Thermo Fisher Scientific (Finnpipette), Waltham, MA, US)
- AB 3500xL Genetic Analyser (Life Technologies Applied Biosystems, Foster City, CA, USA)
- Promega Maxwell® 16 MDx Instrument (Promega Corp., Madison, WI, USA)
- Milli-Q® Integral 3 (A10) System with Q-POD™ (Millipore™, Billerica, MA, USA)
- AB 7500 Real Time PCR System (Life Technologies Applied Biosystems, Foster City, CA, US)
- STORstar instrument (Process Analysis & Automation, Hampshire, GB)
- MultiPROBE II PLUS HT EX with Gripper Integration Platform (PerkinElmer, Downers Grove, IL, US)
- Thermomixer (Eppendorf AG, Hamburg, DE)
- MixMate (Eppendorf AG, Hamburg, DE)

4 Methods

4.1 Sample Selection

The extracts of 528 reference FTA™ samples submitted by the Queensland Police Service for routine testing after 1 January 2014 and gave full DNA profiles were used as the data set for this project.

4.2 DNA Extraction

Reference FTA™ samples were extracted using the DNA IQ™ Casework Pro Kit for Maxwell®16 according to QIS 29344 DNA IQ™ Extraction using the Maxwell®16.

4.3 DNA Quantification

All quantifications were prepared by manual methods or using the MultiPROBE II plus HT EX platform according to QIS 19977 Quantification of Extracted DNA using the Quantifiler™ Human DNA Quantitation Kit.

4.4 DNA Amplification

All amplifications were full volume amplifications prepared by manual methods or using the MultiPROBE II plus HT EX platform according to QIS 31511 Amplification of Extracted DNA using the PowerPlex®21 System. Table 1 lists the PCR cycling conditions utilised in this validation.

Table 1 PCR cycling conditions for PowerPlex®21 System

PowerPlex®21 Kit	Standard
GeneAmp 9700 mode	Max
	30 cycles
Activation	96°C for 1 minute
Cycling	94°C for 10 seconds
	59°C for 1 minute
	72°C for 30 seconds
Extension	60°C for 10 minutes
	4°C Soak

4.5 DNA Fragment Capillary Electrophoresis

Plates for DNA fragment analysis on the 3130*xl* were prepared as per QIS 19978 Capillary electrophoresis setup and PCR fragments separated by capillary electrophoresis (CE) according to QIS 15998 Procedure for the Operation and Maintenance of the AB 3130*xl* Genetic Analysers. Table 2 outlines the 3130*xl* Genetic Analyser running conditions.

Table 2 3130*xl* CE protocol conditions

Injection time	Injection voltage	Run time
5s	**3kV**	**1500s**

The plates for DNA fragment analysis on the 3500xL were prepared as per QIS 19978 Capillary electrophoresis Setup. The PCR fragments separated by capillary electrophoresis on the 3500xL were performed according to manufacturer specifications. Table 3 outlines the 3500xL Genetic Analyser running conditions as per manufacturer specifications [2].

Table 3 3500 Series CE protocol conditions

Injection time	Injection voltage	Run time
24s	1.2kV	1210s

4.6 Profile Interpretation 1

Profile Interpretation 1 was used to determine the:

- Baseline, limit of detection and limit of reporting

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using analysis panel PowerPlex_21_IDX_v1.7. The data set was analysed twice using two different analysis methods; one analysed at 1RFU and then again at 20 RFU. The 1RFU data set had the following conditions:

- All peaks were left on

The 20 RFU data set had the following conditions:

- All true alleles, -2, -1, and +1 repeat stutter of true alleles and spectral pull-up were left on
- As defined by Promega artefact peaks in the N-2 bp and/or N+2 bp position at D1S1656, D6S1043, D13S317, vWA, D21S11, D7S820, D5S818, D12S391 and D19S433 loci and in the N-1 bp position at Amelogenin were left on [(1)]
- All known artefacts were left on (or later removed from result file): Fluorescein (blue) 66-69 bp; JOE (green) 60-62 bp and 82-83 bp; TMR-ET (yellow) 60-67 bp; and CXR-ET (red) 58-65 bp and 76-77 bp
- Any peaks determined to be due to carry-over were also left on
- All other labelled peaks were removed

4.7 Profile Interpretation 2

Profile Interpretation 2 was used to determine:

- Stutter thresholds

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using analysis panel PowerPlex_21_IDX_v1.7, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOD (75 RFU) and with stutter thresholds of zero
- Any true alleles that coincided with a stutter position (-2, -1 and +1 repeat positions) were excluded from analysis
- Stutter (-2, -1 and +1 repeat positions) peaks that overlapped were excluded from analysis. For example, if a peak in a -2 stutter position also aligned with +1 stutter position of an adjacent allele, this peak was then excluded
- Stutter (-2, -1 and +1 repeat positions) peaks that may have had an elevated peak height due to pull up were excluded from analysis

- All true alleles, -2, -1 and +1 repeat stutter of true alleles, which did not follow the point above were left on
- All artefacts and spectral pull-up were removed
- Any peaks determined to be carry-over peaks were removed

4.8 Profile Interpretation 3

Profile Interpretation 3 was used to determine:

- Peak height ratio (Intra-locus balance)

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using the analysis panel outlined by PowerPlex_21_IDX_v1.7, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOR (215 RFU)
- Homozygous loci were excluded
- Loci where a true allele fell into a stutter position (-2, -1 and +1 repeat positions) were excluded from analysis
- All true alleles were left on
- All -2, -1 and +1 repeat stutter of true alleles, all artefacts and spectral pull-up were removed
- Any peaks determined to be carry-over peaks were also removed

4.9 Profile Interpretation 4

Profile Interpretation 4 was used to determine:

- Repeatability and Reproducibility, Mixture Study, Variance determination

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using the analysis panel PowerPlex_21_IDX_v1.7, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOD (75 RFU), LOR (215 RFU)
- All true alleles were left on
- All -1 repeat stutter of true alleles were left on
- All -2 and +1 repeat stutter of true alleles, all artefacts and spectral pull-up were removed
- Any peaks determined to be carry-over peaks were removed

4.10 Profile Interpretation 5

Profile Interpretation 5 was used to determine:

- Drop in

The 54 extraction negative controls from the baseline data set were analysed with GeneMapper®ID-X v1.4 using the analysis panel PowerPlex_21_IDX_v1.7, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOD (75 RFU)
- All true alleles were left on
- All -1 repeat stutter of true alleles were left on

- All -2 and +1 repeat stutter of true alleles, all artefacts and spectral pull-up were removed

4.11 Preparation of a Dilution Series

For each sample selected for the baseline, a dilution series was prepared using amplification grade water provided with the Promega PowerPlex®21 System. A spreadsheet used previously in the PowerPlex®21 System validation study [18] was utilised to calculate the dilution series to obtain the specified concentrations.

5 Experimental Design

5.1 Samples and Plate Preparation

5.1.1 Casework Sample Set

A total of 528 reference sample extracts amplified under casework conditions were processed on the 3500xL. These samples were used to create the casework sample set to determine the following:

- Stutter thresholds
- Peak height ratio and allelic imbalance threshold
- Homozygote peak threshold

Samples were extracted, quantified, amplified, processed through capillary electrophoresis and analysed as per Methods 4.2, 4.3, 4.4 and 4.5.

5.1.2 Baseline, Limit of Detection and Limit of Reporting Data Set

Thirteen samples from the extracted reference sample data set were selected and re-quantified (Method 4.3). Any samples that had quantification results that varied more than 30% were excluded from this data set. This resulted in 3 samples being excluded, leaving 10 samples in this data set. Based on the average quantification result a dilution series was carried out as per Method 4.11. A dilution series from 25 pg to 0.5 ng was performed. Table 4 lists the total amount of DNA template added.

Table 4 Dilution series

#	Total DNA (ng)
1	0.500
2	0.447
3	0.394
4	0.342
5	0.289
6	0.236
7	0.183
8	0.131
9	0.078
10	0.025

Each of these dilutions was tested with a total amplification volume of 25µL. See Figures 1 and 2 below for the amplification plate layouts which also included 54 negative extraction controls.

Each amplification plate included the kit positive (2800M Control DNA) and an amplification negative control (amplification grade water). The 54 extraction negative controls were spread across the two amplification plates.

All amplification reaction set-ups were performed using a MultiPROBE II plus HT EX with Gripper integration platform. Amplification cycling conditions and DNA fragment analyses were conducted as per Methods 4.4 and 4.5.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Pos	Ladder 1	B4	Ladder 2	B9	Ladder 3	B14	Ladder 4	B19	Ladder 5	B24	Ladder 6
B	Neg	1-7	B5	2-7	B10	3-7	B15	4-7	B20	5-7	B25	6-7
C	1-1	1-8	2-1	2-8	3-1	3-8	4-1	4-8	5-1	5-8	6-1	6-8
D	1-2	1-9	2-2	2-9	3-2	3-9	4-2	4-9	5-2	5-9	6-2	6-9
E	1-3	1-10	2-3	2-10	3-3	3-10	4-3	4-10	5-3	5-10	6-3	6-10
F	1-4	B1	2-4	B6	3-4	B11	4-4	B16	5-4	B21	6-4	B26
G	1-5	B2	2-5	B7	3-5	B12	4-5	B17	5-5	B22	6-5	B27
H	1-6	B3	2-6	B8	3-6	B13	4-6	B18	5-6	B23	6-6	B28

Figure 1 Amplification plate layout for plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A	Pos	Ladder 1	B32	Ladder 2	B37	Ladder 3	B42	Ladder 4	B47			
B	Neg	7-7	B33	8-7	B38	9-7	B43	10-7	B48			
C	7-1	7-8	8-1	8-8	9-1	9-8	10-1	10-8	B49			
D	7-2	7-9	8-2	8-9	9-2	9-9	10-2	10-9	B50			
E	7-3	7-10	8-3	8-10	9-3	9-10	10-3	10-10	B51			
F	7-4	B29	8-4	B34	9-4	B39	10-4	B44	B52			
G	7-5	B30	8-5	B35	9-5	B40	10-5	B45	B53			
H	7-6	B31	8-6	B36	9-6	B41	10-6	B46	B54			

Figure 2 Amplification plate layout for plate 2

5.1.3 Repeatability and Reproducibility Sample Set

Five samples were selected which previously gave full DNA profiles with no off scale peaks. Each of the five samples was prepared five times in the one capillary electrophoresis plate as per Figure 3. The capillary electrophoresis plate also included the kit positive (2800M Control DNA) and an amplification negative control (amplification grade water).

	1	2	3	4	5	6	7	8	9	10	11	12
A	Pos	Ladder 1	3-4	5-2	Ladder 2							
B	Neg	2-2	3-5	5-3								
C	1-1	2-3	4-1	5-4								
D	1-2	2-4	4-2	5-5								
E	1-3	2-5	4-3									
F	1-4	3-1	4-4									
G	1-5	3-2	4-5									
H	2-1	3-3	5-1									

Figure 3 Repeatability and reproducibility plate layout

Four CE plates; two for the repeatability and two for the reproducibility experiments were prepared and processed as per Method 4.5.

For the repeatability experiment:

- Preparation 1 of the repeatability plate was processed on the 3500xL twice by the same operator on the same day.
- Preparation 2 of the repeatability plate was processed on the 3130*xl* twice by the same operator on the same day.

For the reproducibility experiment:

- Preparation 1 of the reproducibility plate was processed on the 3500xL once by operator 1 on day 1 and a second time by operator 2 on day 2.
- Preparation 2 of the reproducibility plate was processed on the 3130*xl* once by operator 1 on day 1 and a second time by operator 2 on day 2.

5.1.4 Saturation Sample Set

Five samples with sufficient quantification results were selected (S1, S2, S3, S4 and S5) from the casework sample data set. Samples were quantified twice as per Method 4.3. Amplification reactions were prepared as per Method 4.4 with each of the five samples amplified at total template inputs of 1.0 ng, 1.25 ng, 1.5 ng, 1.75 ng and 2.0 ng. These samples were then prepared (together with the mixture study samples) as per Figure 4, and then run on the 3500xL once as per Method 4.5.

However, results showed the samples amplified at template inputs between 1.0 - 2.0 ng were too excess and did not provide usable data. Thus the experiment was redone with lower DNA template inputs ranging from 0.247 ng/µL to 0.369 ng/ µL. The five samples were then re-amplified as per Method 4.4 with the new total template inputs of 0.2 ng, 0.4 ng, 0.6 ng, 0.8 ng and 1.0 ng. The samples were then prepared as per Figure 4 and then run on the 3500xL once as per Method 4.5. DNA profile interpretation was performed as per Method 4.9.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Pos	Ladder 1	S3 0.8ng	S5 0.4ng	Ladder 2							
B	Neg	S2 0.4ng	S3 1.0ng	S5 0.6ng								
C	S1 0.2ng	S2 0.6ng	S4 0.2ng	S5 0.8ng								
D	S1 0.4ng	S2 0.8ng	S4 0.4ng	S5 1.0ng								
E	S1 0.6ng	S2 1.0ng	S4 0.6ng									
F	S1 0.8ng	S3 0.2ng	S4 0.8ng									
G	S1 1.0ng	S3 0.4ng	S4 1.0ng									
H	S2 0.2ng	S3 0.6ng	S5 0.2ng									

Figure 4 Saturation plate layout

5.1.5 Mixture Study Data Set

The mixture study data set contained samples from the following three experiments: Number of contributor's assessment; two person mixtures; and three person mixtures. The samples from all three experiments were amplified together in the one plate as per Method 4.4 and prepared as per Figure 5.

From the one amplification plate, two CE plates; one for the 3500xL and one for the 3130*xl* were prepared and processed as per Method 4.5. Each plate was run on the corresponding CE instrument three times. With each subsequent run (i.e. the second and third analyses) processed immediately after the previous run so to minimise any loss of sample and/or run quality over time. Each of these runs was uniquely identified.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Pos	Ladder 1	P1P2-1	S1 1.75ng	Ladder 2	S4 1.5ng						
B	Neg	579901322	Blank	S1 2.0ng	S3 1.25ng	S4 1.75ng						
C	57990 1261	579901333	P1P2P3 -1	S2 1.0ng	S3 1.5ng	S4 2.0ng						
D	57990 1270	57990 1344	P1P2P3 -2	S2 1.25ng	S3 1.75ng	S5 1.0ng						
E	57990 1284	57990 1355	Blank	S2 1.5ng	S3 2.0ng	S5 1.25ng						
F	57990 1293	Blank	S1 1.0ng	S2 1.75ng	S4 1.0ng	S5 1.5ng						
G	57990 1300	P1P2-3	S1 1.25ng	S2 2.0ng	S4 1.25ng	S5 1.75ng						
H	57990 1311	P1P2-2	S1 1.5ng	S3 1.0ng	S3 1.5ng	S5 2.0ng						

Figure 5 Mixture study and saturation plate layout

(□ - Number of Contributors Assessment; □ - Two Person Mixtures; □ - Three Person Mixtures; □- Saturation)

For the number of contributors assessment, ten samples with mixed DNA profile which were prepared for Project Proposal #149[21], were used. Table 5 lists the samples used for this experiment, including the number of contributors and mixture ratios.

Table 5 Samples for number of contributors assessment

Barcode	Number of contributors	Mixture Ratio
579901261	2	20:1
579901270	2	30:1
579901284	2	50:1
579901293	3	1:1:1
579901300	3	5:2:1
579901311	3	30:1:1
579901322	3	30:1:1
579901333	3	30:1:1
579901344	4	30:1:1:1
579901355	4	20:10:1:1

For the two person mixture assessment, two samples (P1 and P2) that had a low number of shared alleles were selected from the casework sample set. Both samples were quantified twice as per Method 4.3. Mixtures were created as per the ratios in Table 6.

Table 6 Two person mixture ratios

Sample	Male : Female Ratio
P1P2-1	20:1
P1P2-2	5:1
P1P2-3	1:1

For the three person mixture assessment, three samples (P1, P2 and P3) that had a low number of shared alleles were selected from the casework sample set. All three samples were quantified twice as per Methods 4.3. Mixtures were created as per the ratios in Table 7.

Table 7 Three person mixture ratios

Sample	Male : Female Ratio
P1P2P3-1	5:1:1
P1P2P3-2	20:10:1

5.2 Baseline, Limit of Detection and Limit of Reporting

The baseline, limit of detection and limit of reporting data set which consisted of 100 samples (ten samples amplified ten times at varied concentrations) and 54 negative controls was used.

Samples were analysed and profiles interpreted as per Method 4.6. The samples were first analysed by a project team member who is competent to train in GeneMapper. A second experienced reader, who is competent in PowerPlex®21 case management, then reviewed the analysed results from the first analysis to ensure the interpretation was in accordance with Method 4.6.

The average peak height RFU (μ_{PK}) for each dye channel was calculated using the AVERAGE function (Arithmetic mean) in Microsoft Excel. The standard deviation (σ_{PK}) was calculated using the STDEV function in Microsoft Excel.

The baseline was assessed to determine the limit of detection and the limit of reporting using the following equations.

The limit of detection [3, 4] was calculated using equation 1:

Equation 1

$$LOD = \mu_{PK} + 3\sigma_{PK}$$

(LOD = Limit of detection, μ_{PK} = average peak height, σ_{PK} = standard deviation of peak height)

The limit of reporting [3, 4] was calculated using equation 2:

Equation 2

$$LOR = \mu_{PK} + 10\sigma_{PK}$$

(LOR = Limit of reporting, μ_{PK} = average peak height, σ_{PK} = standard deviation of peak height)

5.3 Stutter Thresholds

The casework data set consisting of 501 samples was used. Samples were analysed and profiles interpreted as per Method 4.7. Locus specific stutter thresholds for -2, -1 and +1 repeat stutter were calculated.

The stutter ratio (SR) for each locus for -1 repeat stutter was calculated using equation 3:

Equation 3

$$SR = E_S/E_A$$

(SR = stutter ratio; E_S = stutter height; E_A = allele height)

The stutter threshold (ST) [4] for each locus and stutter type was calculated using equation 4:

Equation 4

$$ST = \mu_{SR} + 3\,\sigma_{SR}$$

(ST = stutter threshold, μ_{SR} = average stutter ratio, σ_{SR} = standard deviation of stutter ratio)

The average stutter ratio (μ_{SR}) for each locus and stutter type was calculated using the AVERAGE function (Arithmetic mean) in Microsoft Excel. The stutter ratio standard deviation (σ_{SR}) for each locus and stutter type was calculated using the STDEV function in Microsoft Excel.

To determine thresholds for -2, and +1 repeat stutter peaks the stutter ratio (SR) was calculated for each locus using Equations 5 & 6 (modified versions of Equation 3).

Equation 5

$$SR = E_{S-2}/E_A$$

(SR = Stutter ratio, E_{S-2} = -2 repeat stutter height, E_A = Allele height)

Equation 6

$$SR = E_{S+1}/E_A$$

(SR = Stutter ratio, E_{S+1} = +1 repeat stutter height, E_A = Allele height)

The -2 and +1 repeat stutter thresholds (ST) for each locus was calculated as per Equation 4 from the main allele peak.

5.4 Peak Height Ratio and Allelic Imbalance Threshold

The casework data set consisting of 521 samples was used. Samples were analysed and profiles interpreted as per Method 4.8.

Peak height ratio for heterozygous loci was calculated by dividing the lower peak height by the higher peak height, as per the equation 7[5]:

Equation 7

$$PHR = LPH / HPH$$

(PHR = peak height ratio; LPH = lower peak height; HPH = higher peak height)

The overall average PHR (μ_{PHR}) across all loci was calculated using the AVERAGE function (Arithmetic mean) in Microsoft Excel and the PHR standard deviation (σ_{PHR}) was calculated using the STDEV function in Microsoft Excel.

The allelic imbalance (AI) threshold was calculated as per equation 8 [6, 7]:

Equation 8:

$$AI_{TH} = \mu_{PHR} - 3\,\sigma_{PHR}$$

(AI_{TH} = Allelic Imbalance threshold; μ_{PHR} = overall average PHR; σ_{PHR} = standard deviation of the PHR)

5.5 Homozygote Peak Threshold

The homozygote peak threshold was calculated using equation 9 [8]:

Equation 9

$$Th_{Hom} = LOR \times (1 / AI_{TH}) \times 2$$

(Th_{Hom} = homozygote peak threshold; LOR = limit of reporting as determined in Section 5.2; AI_{TH} = allelic imbalance threshold)

5.6 Drop In

The 54 extraction negative controls from the baseline data set were used and were analysed as per Method 4.10. Any peaks considered to be drop in were noted. All previous amplifications of the negative controls were checked to determine if possible drop in peaks were present previously. A reporting case scientist then reviewed the peaks to determine if the peaks present were due to drop in.

5.7 Repeatability and Reproducibility

The repeatability and reproducibility data set was used and profile interpretation performed as per Method 4.9. Run comparisons were performed based on allele call and peak height.

Repeatability was measured for the 3500xL and 3130*xl* based on the consistency of peak heights and allele designations between the two runs on the same day.

Reproducibility was measured for the 3500xL and 3130*xl* based on the consistency of peak heights and allele designations between the two runs, by different operators on different days.

5.8 Saturation

The saturation data set was used and profile interpretation performed as per Method 4.9.

5.9 Mixture Study

For all mixture study experiments, the mixture study data set was used and profile interpretation was performed as per Method 4.9.

5.9.1 Mixture Study: Number of Contributors Assessment

The results analysis for this section was to involve an experienced STRmix™ reporting scientist performing an intuitive assessment to determine the number of contributors for each of the 10 samples, on each of the three repeat runs. These assessments would then be compared for each sample on the 3500xL and also to previous number of contributor assessments for these samples using 3130*xl* results. Due to the quality of the mixed DNA profiles obtained, this analysis was not performed – see results section for detail.

5.9.2 Mixture Study: Two and Three Person Mixtures

The results analysis for this section was to involve calculation of a mixture ratio averaged across all loci and locus specific mixture ratios for each repeat of each sample.

The locus specific and average mixture ratios obtained from each CE instrument (3500xL and 3130*xl*) was compared to assess consistency of each CE instrument from run to run.

Locus specific and average mixture ratios for the 3130*xl* and 3500xL were then to be compared to assess the consistency between the two instruments.

All 3500xL and 3130*xl* results were then to be processed through STRmix™ to determine modelled mixture ratios. An experienced STRmix™ reporting scientist was to perform an intuitive assessment of each profile run on the 3130*xl* and 3500xL and compare to the STRmix™ output. Due to the quality the mixed DNA profiles obtained, this analysis was not performed – see results section for detail.

6 Results and Discussion

6.1 Baseline, Limit of Detection and Limit of Reporting

For this validation 100 samples containing DNA and 54 negative controls were used to assess the baseline and calculate the limit of detection (LOD) and the limit of reporting (LOR). The use of thresholds for reporting is essentially a risk assessment [9], if the thresholds are set too low then labelling of artefacts and noise may occur, if set too high then real peaks will not be labelled and information will be lost [5, 10, 18].

The LOD is the lowest signal that can be distinguished from the baseline (background fluorescence) and may vary between capillary electrophoresis instruments. The LOD was calculated as per equation 1 in Section 5.2. A PowerPlex®21 case manager reviewed the baseline plate after it had been read by the project officer and confirmed that it was read in accordance with Method 4.6. The LOR is the threshold in which an allelic peak can be confidently distinguished from the baseline (background fluorescence) [18]. The LOR was calculated as per equation 2 in Section 5.2.

Table 8 shows the results of the baseline calculations for each of the individual dyes. The highest average peak height (53.92 RFU) and the highest average standard deviation (48.66 RFU) were in the CC5 (orange) dye. This standard deviation is over three times larger than the next highest standard deviation from the TMR-ET (yellow) dye (14.65 RFU). This then resulted in the CC5 dye yielding the highest LOD (199.89 RFU) and the highest LOR (540.48 RFU).

Table 8 3500xL baseline summary of each dye

	Min RFU	Max RFU	Average	SD	3 SD	10 SD	LOD (Ave+3SD)	LOR (Ave+10SD)
Fluorescein (Blue)	1	335	6.83	6.89	20.66	68.85	27.48	75.68
JOE (Green)	1	198	10.30	7.82	23.47	78.24	33.77	88.54
TMR-ET (Yellow)	1	651	13.96	14.65	43.96	146.54	57.92	160.50
CXR-ET (Red)	1	396	15.24	7.94	23.82	79.42	39.07	94.66
CC5 (Orange)	1	300	53.92	48.66	145.97	486.56	199.89	540.48
CC5 (Excl 172, 93/94,121 pks)	1	231	49.68	43.04	129.13	430.43	178.81	480.11

As noted previously in Proposal #145 3500xL validation for direct amplification[19] and extracted reference samples[20], there were a number of artefacts with larger peak heights observed at 93-94 bp, 121 bp and 172 bp, which elevated the average peak height and standard deviation for the CC5 dye. These artefacts were also present in the CC5 for the casework data set along with additional artefacts at 101 bp, 141 bp and 161 bp. The 101 bp, 141 bp and 161 bp artefacts appear to 'shoulder' an adjacent size standard peak, like the 121 bp artefact [19]. It was also noted some of the other size

standard peaks also had shouldering. The CC5 dye is the dye used for the size standard which defines the sizes of known fragments [8]. The artefacts seen in the CC5 size standard were not seen in any of the other dye colours. These artefacts do not interfere with the designated size standard peaks, nor do they interfere with the sizing of peaks in the other dyes.

The highest LOD (199.89 RFU), LOR (540.48) and standard deviation (48.66) were in the CC5 (orange) dye. The second highest LOD (57.92 RFU), LOR (160.50 RFU) and standard deviation (14.65) were in the TMR-ET (yellow) dye. The largest maximum peak height (651 RFU) was also observed in the TMR-ET (yellow) dye. This peak along with a number of other large peaks were observed as artefacts and pull up in TH01. These artefacts elevated the average peak height and standard deviation for the TMR-ET dye.

Table 9 shows the results of the baseline calculations when the results are averaged across all of the dyes. The average peak height is 14.70 RFU, the average standard deviation is 19.70 RFU, LOD is 73.79 RFU and the LOR is 211.67 RFU. When the CC5 dye is removed from the calculations, all the values decrease to average peak height of 12.85 RFU, standard deviation of 14.20 RFU, LOD of 55.45 RFU and LOR is 154.84 RFU.

Table 9 3500xL baseline summary of all dyes

	Min RFU	Max RFU	Average RFU	SD	3 SD	10 SD	LOD (Ave+3SD)	LOR (Ave+10SD)
All Dyes	1	651	14.70	19.70	59.09	196.97	73.79	211.67
All Dyes (Excl CC5)	1	651	12.85	14.20	42.60	142.00	55.45	154.84

The most conservative calculation for LOD and LOR included all dyes and artefacts (LOD of 73.79 RFU and LOR of 211.67 RFU). By using conservative rounding from the calculated LOD and LOR thresholds, the LOD will be set to 75 RFU and LOR will be set to 215 RFU.

6.1.1 Acceptance Criteria – Baseline, Limit of Detection and Limit of Reporting

The 3500xL LOD (75 RFU) is approximately 5 times greater than the 3130*xl* LOD (15 RFU). The 3500xL LOR (215 RFU) is approximately 5 times greater than the 3130*xl* LOR (40 RFU). These thresholds are:

- greater than the 3-4 times scale difference which were reported by the manufacturer[1] and which were observed for the direct amplification validation[19] LOD and LOR.
- consistent with the scale variation observed in the Concordance experiment (Section 6.5) in the direct amplification validation[19].

LOD and LOR thresholds are instrument, kit and reaction type (i.e. half or full volume) specific, therefore comparison to other internally and externally published data using different instruments, kits and reaction types should be used to identify results which are grossly different and/or unexpected. The results from this experiment were not grossly different or unexpected. Further, given the thresholds have been calculated for a

specific instrument and application, using appropriate data sets and analysis methods, they should be accepted as calculated and implemented.

6.2 Stutter

Stutter peaks are Polymerase Chain Reaction (PCR) artefacts commonly observed in all STR analysis [2,12]. They are usually observed as a peak one repeat unit smaller in size than the true allele peak [12,13] which will be called -1 repeat stutter. The stutter mechanism has been attributed to slippage of the DNA strand during replication [12, 18].

For this validation -2, -1 and +1 repeat stutter thresholds were calculated. -2 repeat stutter is observed as a peak two repeat units less in size than the true allele [13,14]. +1 repeat stutter is observed as a peak one repeat unit larger in size than the true allele [14,15]. Figure 6 shows examples of -2, -1 and +1 repeat stutter.

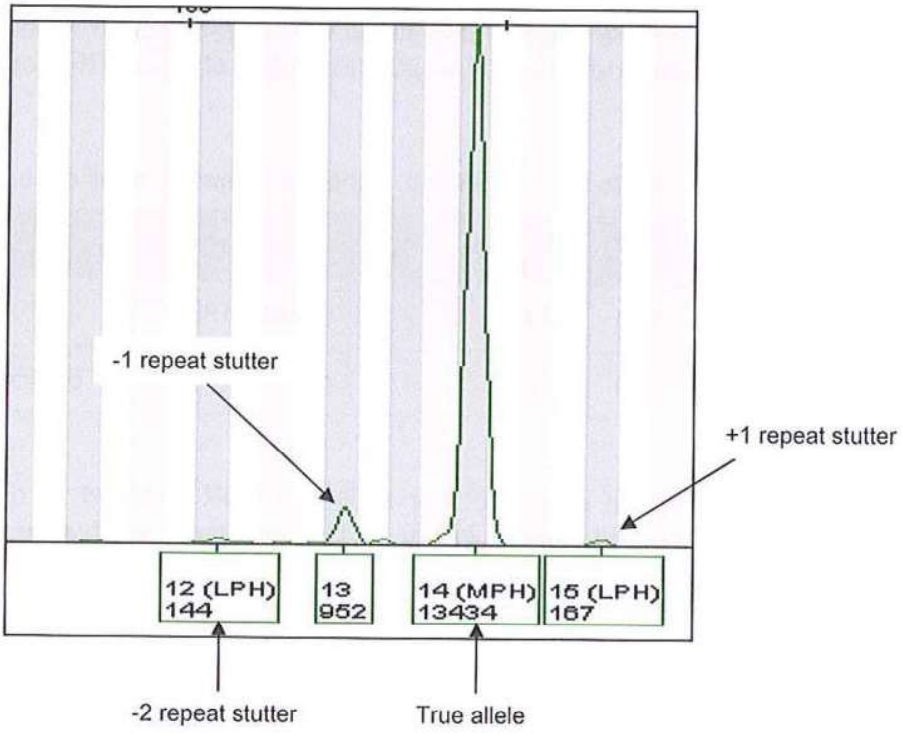

Figure 6 Example of stutter peaks

Table 10 shows the number of times stutter was observed, the average stutter ratio, standard deviation, and stutter threshold for -2 repeat, -1 repeat and +1 repeat stutter for each locus.

For the -2 repeat stutter, thresholds were calculated for 19 out of 20 loci. From these, 15 out of 20 had lower thresholds, 4 out of 20 had greater thresholds, and 1 out of 20 had the same thresholds as the current 3130*xl* thresholds. Due to no observations of -2 repeat stutter at Penta D stutter thresholds were unable to be calculated. The -2 repeat stutter thresholds calculated for the 3500xL are the recommended analysis thresholds for implementation as there are no recommended -2 repeat stutter thresholds from Promega.

For -1 repeat stutter, where the calculated thresholds are lower than the Promega PowerPlex®21 GeneMapper stutter threshold, the Promega PowerPlex®21 GeneMapper stutter threshold will be implemented as the analysis threshold. Where the calculated stutter threshold is greater than the Promega PowerPlex®21 stutter threshold the calculated stutter threshold will be implemented as the analysis threshold. For the -1 repeat stutter, 7 out of 20 loci had higher stutter thresholds, 12 had lower stutter thresholds and one was the same as the Promega stutter thresholds.

For the +1 repeat stutter, thresholds were calculated for all loci. Of these 17 of the 20 had lower thresholds, 3 out of 20 had greater thresholds than the current 3130*xl* thresholds. The +1 repeat stutter thresholds calculated for the 3500xL are the recommended analysis thresholds for implementation as there are no recommended +1 repeat stutter thresholds from Promega.

Table 11 shows the stutter thresholds for -2 repeat, -1 repeat and +1 repeat stutter calculated for the 3500xL, the -1 repeat stutter thresholds for PowerPlex®21 from Promega and the current thresholds used for the 3130*xl*.

Table 10 Summary of Stutter Data

	-2 STUTTER					-1 STUTTER					+1 STUTTER				
Locus	No. Observed	Average	Standard Deviation	New 3500xL Casework Threshold %	3130x/ Current Casework Threshold %	No. Observed	Average	Standard Deviation	New 3500xL Casework Threshold %	3130x/ Current Casework Threshold %	No. Observed	Average	Standard Deviation	New 3500xL Casework Threshold %	3130x/ Current Casework Threshold %
D3S1358	90	0.0082	0.0039	2.0	2.3	263	0.0864	0.0141	12.9	14.2	127	0.0080	0.0064	2.7	4.3
D1S1656	118	0.0113	0.0102	4.2	4.3	627	0.0827	0.0219	14.9	17.2	300	0.0128	0.0060	3.1	6.7
D6S1043	62	0.0078	0.0032	1.8	2.7	574	0.0690	0.0135	10.9	12	305	0.0127	0.0061	3.1	7.4
D13S317	20	0.0065	0.0038	1.8	3.4	382	0.0483	0.0206	11.0	11.8	156	0.0141	0.0070	3.5	7.4
Penta E	2	0.0155	0.0000	1.6	2.1	510	0.0351	0.0155	8.2	10.7	24	0.0157	0.0087	4.2	1.7
D16S539	108	0.0066	0.0063	2.5	2.1	287	0.0659	0.0173	11.8	12.1	294	0.0109	0.0048	2.5	4.3
D18S51	126	0.0087	0.0054	2.5	2.8	497	0.0809	0.0226	14.9	16	272	0.0109	0.0093	3.9	4.9
D2S1338	178	0.0082	0.0037	1.9	2.9	627	0.0840	0.0183	13.9	14.9	45	0.0159	0.0113	5.0	6.5
CSF1PO	28	0.0050	0.0013	0.9	3.7	201	0.0592	0.0142	10.2	13.7	161	0.0107	0.0053	2.7	4.4
Penta D	0	0.0000	0.0000	0.0	0	326	0.0183	0.0067	3.8	8.2	8	0.0143	0.0094	4.3	8.8
TH01	104	0.0071	0.0040	1.9	2.6	499	0.0222	0.0094	5.0	8.7	67	0.0162	0.0116	5.1	2.1
vWA	18	0.0107	0.0088	3.7	3.5	280	0.0659	0.0274	14.8	15.2	39	0.0189	0.0135	5.9	5.6
D21S11	95	0.0188	0.0204	8.0	4	510	0.0831	0.0173	13.5	14.1	272	0.0157	0.0102	4.6	7.1
D7S820	33	0.0140	0.0102	4.5	3.4	354	0.0500	0.0208	11.3	11.4	81	0.0100	0.0061	2.8	5.8
D5S818	3	0.0132	0.0024	2.1	2.3	216	0.0534	0.0168	10.4	12	144	0.0117	0.0060	3.0	5.6
TPOX	3	0.0037	0.0011	0.7	2.7	435	0.0273	0.0104	5.9	9	18	0.0163	0.0095	4.5	6.3
D8S1179	122	0.0062	0.0021	1.3	2.6	394	0.0708	0.0149	11.5	13.2	257	0.0086	0.0046	2.2	5.5
D12S391	219	0.0121	0.0050	2.7	2.8	619	0.0949	0.0303	18.6	18.8	129	0.0109	0.0065	3.0	5.3
D19S433	121	0.0067	0.0041	1.9	3.2	377	0.0693	0.0154	11.5	12.8	35	0.0097	0.0076	3.3	7.1
FGA	105	0.0090	0.0035	1.9	3.1	457	0.0697	0.0180	12.4	13.8	168	0.0083	0.0046	2.2	5.9

Table 11 Comparison of stutter thresholds between 3130*xl* and 3500xL

	-2 STUTTER		-1 STUTTER			+1 STUTTER	
Locus	New 3500xL Casework Threshold %	Current Casework Threshold %	New 3500xL Casework Threshold %	Promega Stutter File	Current Casework Threshold %	New 3500xL Casework Threshold %	Current Casework Threshold %
D3S1358	2.0	2.3	12.9	14.0	14.2	2.7	4.3
D1S1656	4.2	4.3	14.9	15.0	17.2	3.1	6.7
D6S1043	1.8	2.7	10.9	14.0	12.0	3.1	7.4
D13S317	1.8	3.4	11.0	11.0	11.8	3.5	7.4
Penta E	1.6	2.1	8.2	10.0	10.7	4.2	1.7
D16S539	2.5	2.1	11.8	12.0	12.1	2.5	4.3
D18S51	2.5	2.8	14.9	16.0	16.0	3.9	4.9
D2S1338	1.9	2.9	13.9	16.0	14.9	5.0	6.5
CSF1PO	0.9	3.7	10.2	11.0	13.7	2.7	4.4
Penta D	0.0	0.0	3.8	9.0	8.2	4.3	8.8
TH01	1.9	2.6	5.0	6.0	8.7	5.1	2.1
vWA	3.7	3.5	14.8	14.0	15.2	5.9	5.6
D21S11	8.0	4.0	13.5	13.0	14.1	4.6	7.1
D7S820	4.5	3.4	11.3	11.0	11.4	2.8	5.8
D5S818	2.1	2.3	10.4	10.0	12.0	3.0	5.6
TPOX	0.7	2.7	5.9	7.0	9.0	4.5	6.3
D8S1179	1.3	2.6	11.5	12.0	13.2	2.2	5.5
D12S391	2.7	2.8	18.6	17.0	18.8	3.0	5.3
D19S433	1.9	3.2	11.5	11.0	12.8	3.3	7.1
FGA	1.9	3.1	12.4	12.0	13.8	2.2	5.9

- Accepted Stutter Thresholds

6.2.1 Acceptance Criteria – Stutter Thresholds

Stutter is an amplification artefact and is not caused by capillary electrophoresis. Variation in stutter thresholds between the 3130*xl* and 3500xL are likely the result of amplification variation and/or differences in the size and composition of the data sets [20]. Further, as stutter thresholds are a ratio of stutter peak height to allele peak height, they are not impacted by the peak height scale variation between the 3130*xl* and 3500xL.

The 3500xL stutter thresholds were generally comparable to the 3130*xl* thresholds and those in the Powerplex®21 stutter file. Variation between the 3130*xl* and 3500xL is likely due to amplification variation and/or the data set as outlined above rather than a result of 3130*xl* / 3500xL instrument factors. Therefore the stutter thresholds calculated for the 3500xL must be accepted and implemented, with a recommendation that these are reviewed post implementation using a larger data set for increased robustness.

6.3 Peak Height Ratio

Peak height ratio (PHR) is the ratio between the two peaks in a heterozygous pair [18]. Under optimal conditions the amplification of a pair of alleles should result in equal peak heights however, input DNA, inhibitors and quality of DNA will affect the amplification [16,17].

A total of 521 samples from the casework sample data set were used to calculate the peak height ratio. Table 12 summarises the results of the average PHR and allelic imbalance threshold (AI_{TH}) data calculated for all loci in the PowerPlex®21 system. All loci displayed high peak balances within each locus. The lowest observed PHR was at Penta E with 48 and a standard deviation of 0.1180, while the overall average PHR is 53% with a standard deviation of 0.1058.

Table 12 Data of the PHR and calculated AI

	Average	SD	3 SD	AI_{TH}	n*
AMEL	0.8612	0.0972	0.2917	0.5695	392
D3S1358	0.8594	0.1061	0.3184	0.5411	81
D1S1656	0.8679	0.0949	0.2848	0.5831	307
D6S1043	0.8569	0.1053	0.3160	0.5408	256
D13S317	0.8374	0.1173	0.3518	0.4855	160
Penta E	0.8316	0.1180	0.3541	0.4776	344
D16S539	0.8719	0.1027	0.3082	0.5637	98
D18S51	0.8434	0.1038	0.3113	0.5321	231
D2S1338	0.8385	0.1117	0.3351	0.5034	284
CSF1PO	0.8202	0.1074	0.3223	0.4979	32
Penta D	0.8526	0.1005	0.3015	0.5511	174
TH01	0.8773	0.0806	0.2417	0.6356	226
vWA	0.8421	0.1056	0.3169	0.5252	135
D21S11	0.8546	0.1005	0.3014	0.5531	238
D7S820	0.8500	0.1066	0.3197	0.5354	136
D5S818	0.8305	0.1021	0.3063	0.5421	52
TPOX	0.8439	0.1177	0.3532	0.4907	155
D8S1179	0.8554	0.1002	0.3005	0.5549	161
D12S391	0.8438	0.1110	0.3330	0.5107	295
D19S433	0.8605	0.1115	0.3346	0.5259	137
FGA	0.8508	0.1099	0.3296	0.5212	231
All Loci	0.8514	0.1058	0.3175	0.5339	4125

*n = number of times peak height ratio was calculated for a locus

Figure 7 graphically shows the average PHR and the allelic imbalance across all loci. Using Equation 8 in Section 5.4 the overall allelic imbalance threshold (AI_{TH}) calculates to 53%. This is 8% higher than the threshold previously calculated for the 3130*xl* [11].

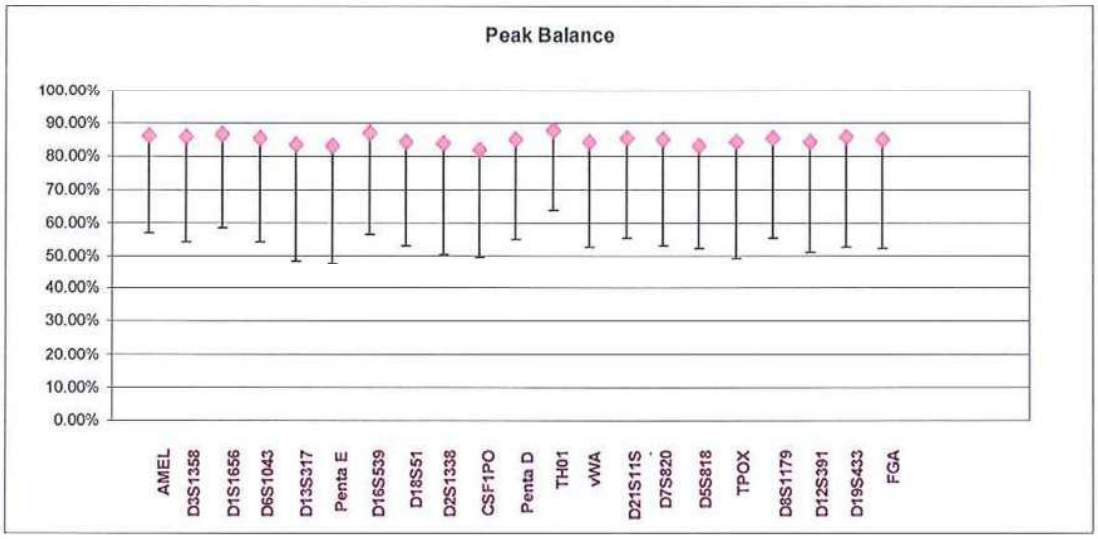

Figure 7 Average Peak Height Ratios per locus

Note: error bars represent the mean PHR minus three times standard deviation

6.3.1 Acceptance Criteria – Allelic Imbalance Threshold

The ratio of peaks in a heterozygous pair (PHR) is primarily influenced by amplification and not capillary electrophoresis. Variation in the AI_{TH} between the 3130*xl* and 3500xL therefore is likely due to amplification variation and/or differences in the data sets used, rather than instrument factors. Further, as the AI_{TH} is a ratio of peak heights in a heterozygous pair, it is not impacted by the peak height scale variation between the 3130*xl* and 3500xL.

The 3500xL AI_{TH} for casework samples was calculated as 53%. The current AI_{TH} for casework samples on the 3130*xl* is 45%, which is comparable to the 3500xL threshold and therefore the 3500xL passes this experiment.

The AI_{TH} for casework samples will be accepted and rounded to 55% for implementation purposes.

6.4 Homozygote Peak Threshold

The homozygote threshold is the threshold above which you can be confident that a heterozygote locus will not be incorrectly called as a homozygote locus [18].

For this validation equation 9 in Section 5.5 was used to calculate the homozygote threshold.

Th_{Hom} = LOR x (1 / AI_{TH}) x 2

= 215 x (1/ 0.533897) x 2

= 805.3987 RFU

It is recommended the homozygote threshold be rounded up to 805 RFU for implementation purposes.

6.4.1 Acceptance Criteria – Homozygote Peak Threshold

The current homozygote peak threshold for casework samples on the 3130*xl* is 200 RFU. The calculated homozygote peak threshold for casework samples on the 3500xL was 805 RFU, which is approximately 4 times greater than the 3130*xl* threshold. The 3500xL passes this experiment as this result is consistent with published expectations of approximately 3-4 times [1].

6.5 Drop In

Allelic drop in is due to spurious amplification products from unknown DNA, which makes allele drop in a random event [12,13]. The phenomenon of allelic drop in is usually not reproducible and can be detected through testing samples multiple times [13].

Any drop in peaks will be noted and later used to determine drop-in rates and enable calculation of STRmix™ drop in parameters. A total of 54 extraction negative controls from the baseline, LOD, LOR sample set were analysed at the calculated LOD (75 RFU) and checked for possible drop in peaks. A total of four possible drop in peaks from 3 different samples were noted, these are noted in Table 13. Two of these peaks were below the LOD and cannot be used in the calculation of STRmix™ drop in parameters. The other two drop in peaks (187 and 145 RFU) were both greater than the LOD (75 RFU) and less than the LOR (215 RFU). None of the noted drop in peaks were present in any previous amplification of those extracts.

Table 13 Drop in peaks

Sample	Peak Height (RFU)	Loci	Allele Designation
Sample1	~60 (<LOD)	D3S1358	18
Sample1	~65 (<LOD)	D16S539	13
Sample2	187	D21S11	29
Sample3	145	TH01	8

6.5.1 Acceptance Criteria – Drop In

The data obtained will be used to generate the drop in settings for STRmix™.

6.6 Repeatability and Reproducibility

Complete and concordant profiles were obtained from all the samples on all runs for repeatability and reproducibility testing on both the 3130*xl* and 3500xL. The peak height data from each run was compared by calculating the percentage change and performing a Student's t–test.

Repeatability is an assessment of the ability of the 3500xL to produce the same results when one sample set is processed a number of times by one user, under the same conditions.

Figure 8 shows the results of the repeatability testing on the 3130*xl*. The results show that the majority of the run to run variation of peak heights range from 6% to -22%.

Two samples showed a significant difference (p = 0.0036 and 0.0419) in peak heights between run 1 and run 2. Peak heights on run 1 were higher than run 2 which may be due to run to run variation. For all other samples there was no significant difference ($p \geq 0.05$) in peak heights between run 1 and 2.

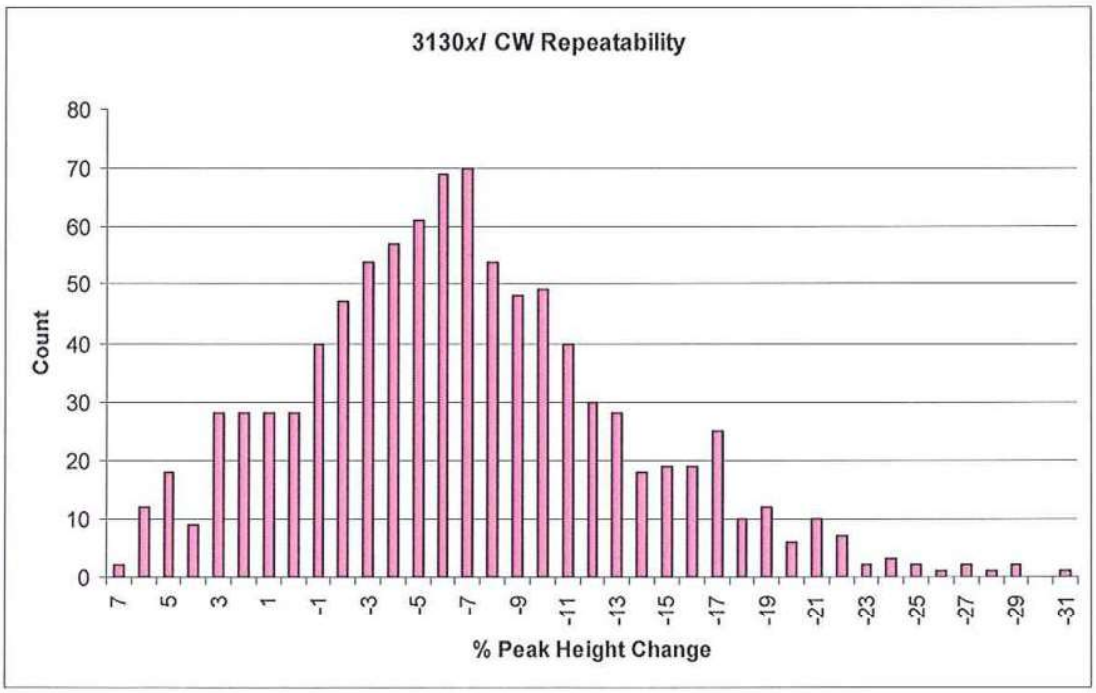

Figure 8 3130*xl* CW Repeatability

Figure 9 shows the results of the repeatability testing on the 3500xL. The results show the majority of the run to run variation of peak heights range from -3% to -18%. For all samples, there were no significant differences ($p \geq 0.05$) in peak heights between run 1 and 2.

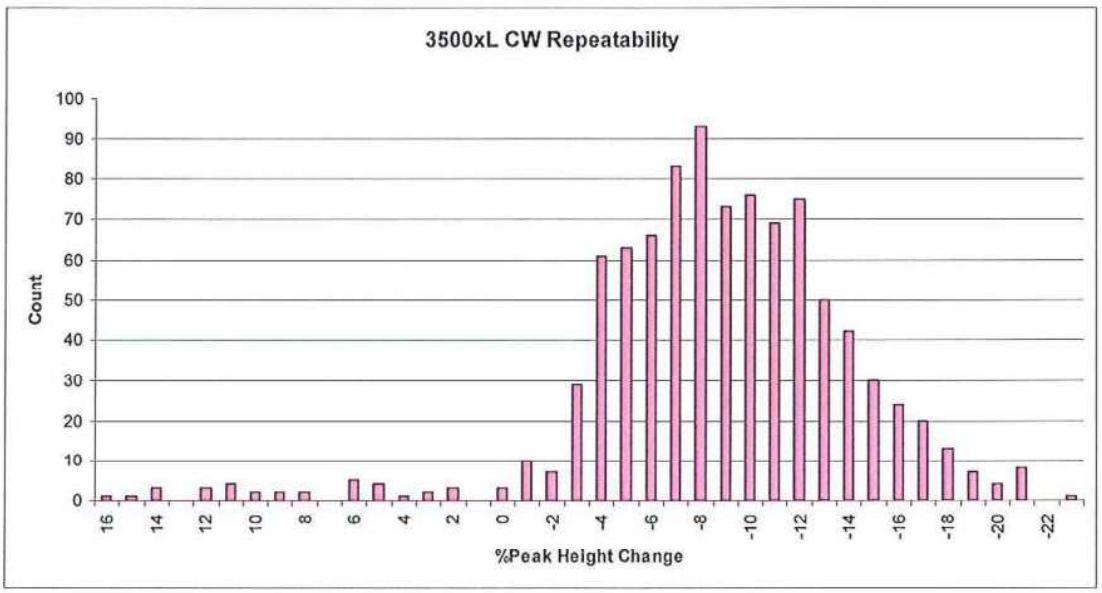

Figure 9 3500xL CW Repeatability

Reproducibility is an assessment of the ability of the 3500xL to produce the same results when one sample set is processed by different operators under different conditions.

Figure 10 shows the results of the reproducibility testing on the 3130*xl*. The results show that the majority of the run to run variation of peak heights range from 0% to -30%. Five samples showed a significant difference (p = 0.0484, 0.0090, 0.0153, 0.0054 and 0.0031) in peak heights between run 1 and run 2. Peak heights on run 1 were higher than run 2 which may be due to run to run variation. For all other samples there was no significant difference ($p \geq 0.05$) between run 1 and 2.

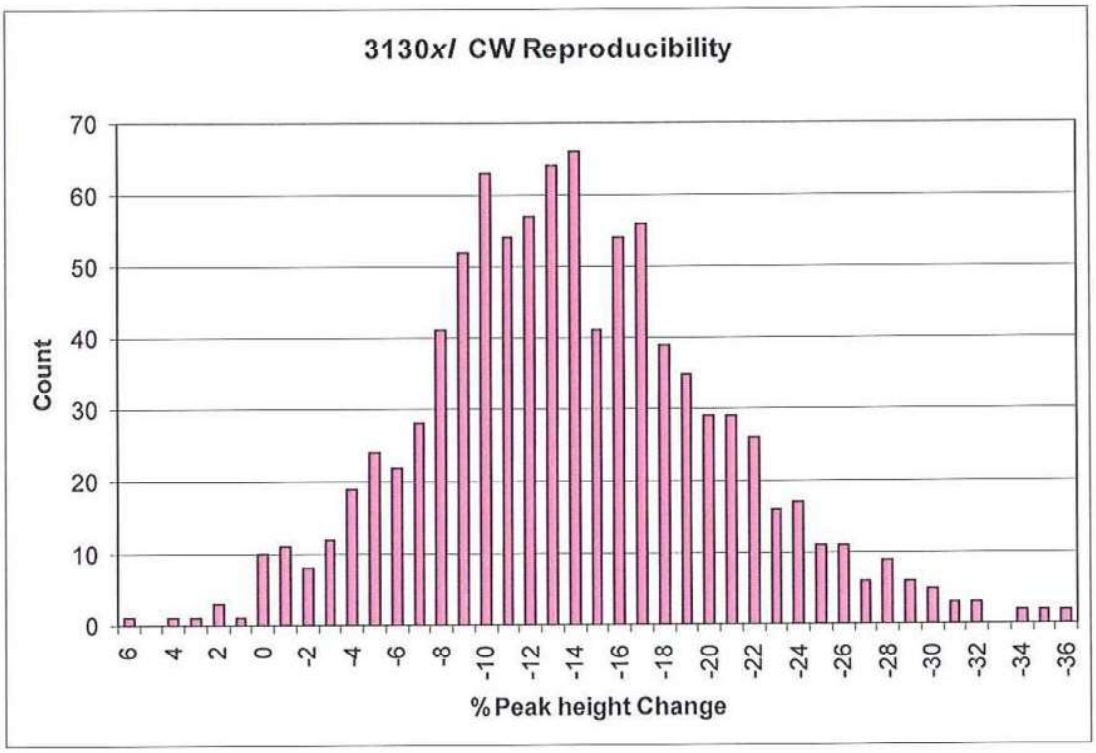

Figure 10 3130*xl* CW Reproducibility

Figure 11 shows the results of the reproducibility testing on the 3500xL. The results show the majority of the run to run variation of peak heights range from -24% to -50%. All samples showed there was significant difference ($p \geq 0.05$) in peak heights between run 1 and run 2. Peak heights on run 1 were higher than run 2. Between run 1 and run 2 a new pouch of POP4 polymer (same lot number) was loaded to the 3500xL at the beginning of day 2. An electrical discharge error also occurred prior to the processing of run 2. The process was aborted and re-started. The electrical discharge error is usually due to bubbles in the tubing containing the polymer. This could explain the variation between run 1 and run 2. Due to this the reproducibility was repeated. The samples used for the original reproducibility plates were consumed and new samples were selected. Figure 12 shows the results of the repeated reproducibility. The results show the majority of the run to run variation of peak heights range from 1% to -38%. Nine samples showed there was a significant difference ($p \geq 0.05$) and 16 samples showed there was no significant difference ($p \leq 0.05$) in peak heights between run 1 and run 2. For the samples that showed a significant difference the peak heights on run 1 were higher than run 2.

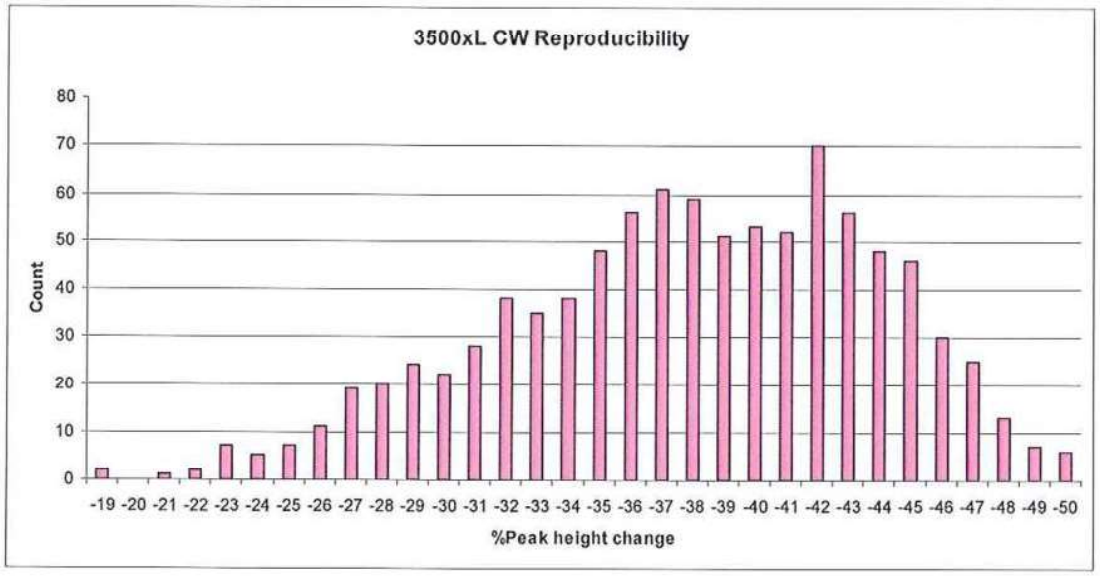

Figure 11 3500xL CW Reproducibility

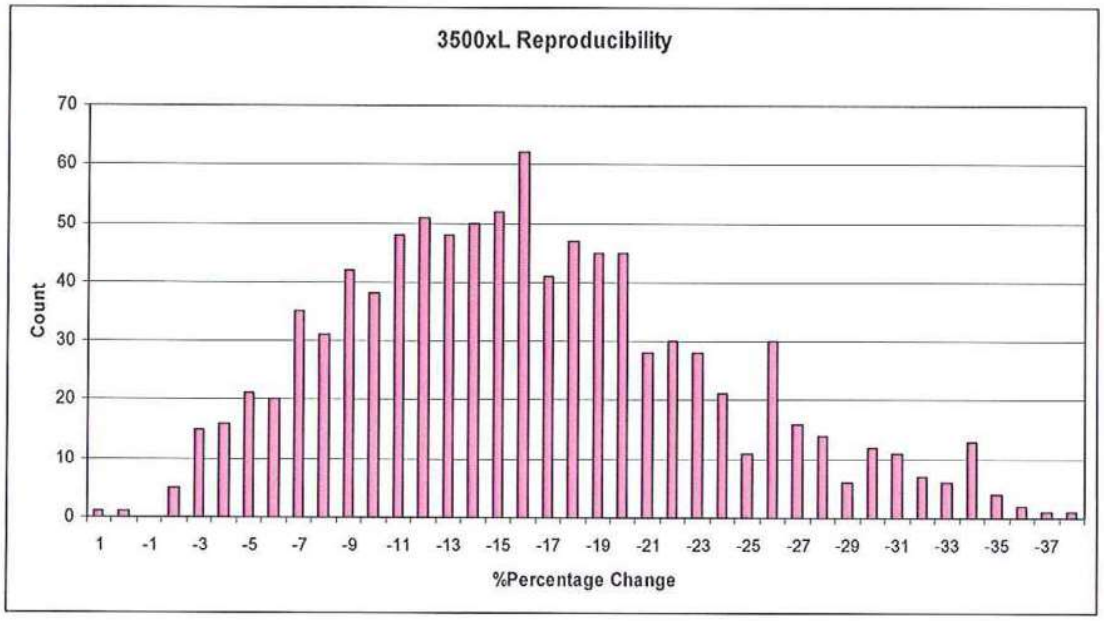

Figure 12 3500xL CW Reproducibility repeat

6.6.1 Acceptance Criteria – Repeatability and Reproducibility

All allele designations for samples on the repeatability and reproducibility plates were completely concordant.

Repeatability – the 3500xL showed no significant difference in peak height between runs. Whereas the 3130*xl* showed 2 samples were significantly different in peak height between runs.

Reproducibility – the 3500xL showed 9 samples having a significant difference in peak height between runs. Whereas the 3130*xl* showed 5 samples were significantly different.

Based on these results, the 3500xL performed better than the 3130*xl* in terms of repeatability, and worse in terms of reproducibility. Given the results of Section 6.8 Number of Contributors do not support the implementation of the analysis of casework samples amplified with PowerPlex®21, the results of this experiment have not been assessed as to whether the 3500xL has passed or failed.

6.7 Saturation

Five samples were selected from the casework data set with quantification values between 0.247 ng/µL to 0.369 ng/µL and with replicate quant values with less than a 30% change in quantification results. The samples were initially amplified with DNA templates of 1 ng, 1.25 ng, 1.5 ng, 1.75 ng and 2 ng. The results showed all samples were extremely excess and unusable at these DNA template inputs. The five same samples were amplified again at DNA templates of 0.2 ng, 0.4 ng, 0.6 ng, 0.8 ng and 1 ng. Further analysis of this data was required to determine the STRmix™ saturation setting. Given the quality of the mixed DNA profiles obtained in Section 6.8 Number of Contributors, further investigation into saturation was not performed.

6.8 Mixture Study: Number of Contributors

A review and an intuitive assessment of the ten samples with mixed DNA profiles was performed.

6.8.1 Acceptance Criteria – Mixture Study: Number of Contributors

Upon review of the mixed DNA profiles, excessive pull-up was noted. In some instances the pull-up was as high as approximately 400 rfu. The effect of this was that it was not possible to reliably determine the number of contributors to the mixtures analysed in this experiment. With a proposed LOR of 215 rfu, it was determined that this level of pull-up is not acceptable for casework samples as it would inhibit the ability of the scientist to distinguish between allelic and pull-up peaks. The pull-up peaks also caused excessive raised baseline which has the potential to mask allelic peaks.

In an attempt to decrease the amount of pull-up seen, the spectral calibration records were reviewed. A spectral calibration had been performed approximately 2 weeks prior to the mixture samples being run on the 3500xL. It was decided to run the samples again with a timeframe of only 1 week since a spectral calibration.

The more recent spectral calibration had the affect of reducing the height of the pull-up seen, however it was not reduced enough to enable the number of contributors to be reliably determined.

Given these findings it was decided that the 3500xL is not suitable for implementation for the analysis of casework samples amplified with PowerPlex®21 at this time.

6.9 Mixture Study: Two and Three Person Mixtures

For the two person mixture assessment, one male profile and one female profile with a low number of shared alleles were selected from the casework sample data set.

For the three person mixture assessment, one male and two female profiles with a low number of shared alleles were selected from the casework sample data set.

6.9.1 Acceptance Criteria – Mixture Study: Two and Three Person Mixtures

The DNA profiles obtained for this experiment were not assessed due to the findings from Section 6.8 Number of Contributors.

7 Conclusion

Comparisons to the 3130*xl* have shown that the 3500xL performs as well or better than the 3130*xl* for all experiments excluding the Mixture Studies. Analysis thresholds and variables which have been calculated are specific to the 3500xL and therefore comparisons to the 3130*xl* are qualitative assessments only, intended to compare the performance of the two instruments. Saturation and mixture studies were not completed during this validation due to the quality of the mixture samples due to poor spectral separation for PowerPlex®21 produced on the 3500xL. The results from this validation show the 3500xL Genetic Analyzer instruments are unsuitable to use for the analysis of extracted casework samples amplified using PowerPlex®21 at this time.

8 Recommendations

1. Do not implement the 3500xL Genetic Analyzer instruments for the analysis of extracted casework samples amplified using PowerPlex®21.
2. If the spectral pull up issues identified in the Mixture Studies can be resolved, the analysis thresholds and variables determined in this study can be implemented.
3. If the spectral pull up issues identified in the Mixture Studies can be resolved, the repeatability/reproducibility, saturation and mixture studies must be repeated prior to implementation.

9 Abbreviations

ART	Artefact
bp	Base pair
LOD	Limit of detection
LOR	Limit of reporting
OS	Off scale data
PA	Preferential Amplification
pk / pks	Peak / peaks
RFU	Relative fluorescence units
SD	Standard deviation

10 References

1. Applied Biosystems 3500/3500xL Genetic Analyzer User Bulletin June 2011
2. Promega PowerPlex® 21 System Technical Manual; 2014 7/14 Contract No.:TMD034
3. Gilder JR, Doom TE, Inman K, Krane DE. Run-Specific Limits of Detection and Quantitation for STR-based DNA Testing. Journal of Forensic Sciences. 2007;52 (1):97-101.
4. STATSWG, STATSWG recommendations for the interpretation of DNA. 2011.
5. SWGDAM, SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. 2010, Scientific Working Group on DNA Analysis Methods (SWGDAM).
6. Gilder, J., et al., Magnitude-dependent variation in peak height balance at heterozygous STR loci. International Journal of Legal Medicine, 2011. 125(1): p. 87-94.
7. Leclair, B., et al., Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. J Forensic Sci, 2004. 49(5): p. 968-80.
8. Weber, C., McNevin, A., Muharam, I., & Ientile, V., Peak Height and Allelic Imbalance Thresholds. 2008, Forensic and Scientific Services. p. 15.
9. Butler J, editor. Data Interpretation & Statistical Analysis. Topics and Techniques for Forensic Analysis; 2012; New York City,NY,US. NIST.
10. Rakay CA, Bregu J, Grgicak CM. Maximizing allele detection: Effects of analytical threshold and DNA levels on rates of allele and locus drop-out. Forensic Science International: Genetics.
11. Nurthen, T., Mathieson, M., Scott, K. and Allen, C. PowerPlex®21 - Direct Amplification of Reference FTA® samples validation, 2012.
12. Butler, J.M., Fundamentals of Forensic DNA Typing. 2010: Academic Press.

13. Walsh, P.S., N.J. Fildes, and R. Reynolds, Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res, 1996. 24(14): p. 2807-12

14. Gibb AJ, H.A., Simmons MC, Brown RM, Characterisation of forward stutter in the AmpFISTR SGM Plus PCR. Science and Justice, 2009. 49(1): p. 24-31.

15. Westen, A.A., et al., Assessment of the stochastic threshold, back- and forward stutter filters and low template techniques for NGM. Forensic Science International: Genetics, 2012. 6(6): p. 708-715.

16. Ensenberger, M.G., et al., Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic Sci Int Genet, 2010. 4(4): p. 257-64.

17. Alaeddini, R., S.J. Walsh, and A. Abbas, Forensic implications of genetic analyses from degraded DNA—A review. Forensic Science International: Genetics, 2010. 4(3): p. 148-157.

18. Nurthen, T., Mathieson, M and Allen, C., PowerPlex®21 - Amplification of Extracted DNA Validation. 2013.

19. Micic, B., Mathieson, M., Ryan, L., and Allen, C., 3500xL Genetic Analyzer Validation for Reference Samples Amplified with PowerPlex®21 using Direct Amplification. 2015.

20. Micic, B., Mathieson, M., Ryan, L., and Allen, C., 3500xL Genetic Analyzer Validation for Extracted Reference Samples Amplified with PowerPlex®21. 2015.

21. Caunt, E., Morgan, R., Howes, J., and Allen, C., Project Proposal #149 – Development of guidelines for the determination of number of contributors to a PowerPlex®21 profile. 2015