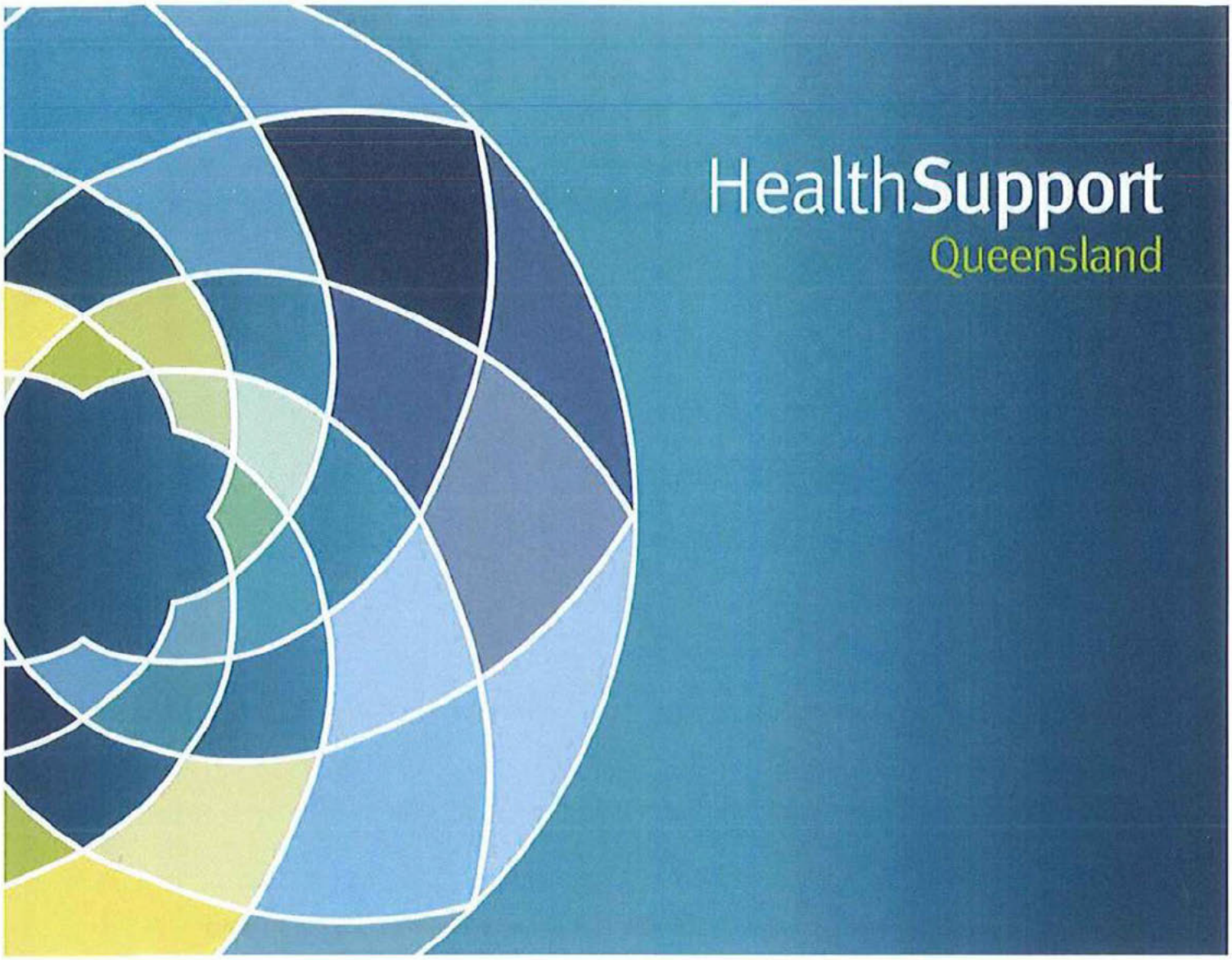

3500xL Genetic Analyzer Validation for Reference samples Amplified with Powerplex®21 using Direct Amplification

Biljana Micic, Megan Mathieson, Luke Ryan and Cathie Allen

February 2015

Great state. Great opportunity.

Queensland Government

3500xL Validation for Reference Samples Amplified with PowerPlex® 21 using Direct Amplification

Published by the State of Queensland (Queensland Health), February 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
0	December 2014	Megan Mathieson	Document Created
0.1	January 2015	Luke Ryan, Megan Mathieson, Biljana Micic	Document edited
1.0	January 2015	Luke Ryan	Formatting edit

Document sign off

This document has been **approved** by:

Name	Position	Signature	Date
Cathie Allen	Managing Scientist	[redacted]	13/02/15

The following officers have endorsed **this document**

Name	Position	Signature	Date
Paula Brisotto	Team Leader ER&Q	[redacted]	12-02-2015

Name	Position	Signature	Date
Sharon Johnstone	A/Team Leader FRIT	[redacted]	12/02/15

Name	Position	Signature	Date
Luke Ryan	Senior Scientist Analytical	[redacted]	12-02-2015

Name	Position	Signature	Date
Allan McNevin	Senior Scientist ER	[redacted]	13.02.2015

Name	Position	Signature	Date
Kirsten Scott	Senior Scientist Q&P	[redacted]	12/02/2015

Name	Position	Signature	Date
Allison Lloyd	A/Senior Scientist Intel	[redacted]	13/02/2015

Name	Position	Signature	Date
Amanda Reeves	Senior Scientist Reporting 1	[redacted]	120215

Name	Position	Signature	Date
Kylie Rika	Senior Scientist Reporting 2	[redacted]	12/02/2015

Contents

1. Abstract 7
2. Introduction 7
3. Resources 9
   - 3.1 Reagents 9
   - 3.2 Materials 9
   - 3.3 Equipment 10
4. Methods 11
   - 4.1 Sample Selection 11
   - 4.2 Direct Amplification 11
   - 4.3 DNA Fragment Capillary Electrophoresis Setup and Conditions 11
   - 4.4 Profile Interpretation 1 12
   - 4.5 Profile Interpretation 2 12
   - 4.6 Profile Interpretation 3 13
   - 4.7 Profile Interpretation 4 13
   - 4.8 Profile Interpretation 5 14
   - 4.9 Profile Interpretation 6 14
5. Experimental Design 15
   - 5.1 Samples/Plate Preparation 15
   - 5.2 Baseline, Limit of Detection and Limit of Reporting 16
   - 5.3 Stutter thresholds 17
   - 5.4 Peak Height Ratio 18
   - 5.5 Homozygote peak threshold 18
   - 5.6 Concordance and Sensitivity 18
   - 5.7 Drop In 19
   - 5.8 Artefacts 19
   - 5.9 Off Scale Allele Peaks 19
   - 5.10 Repeatability and Reproducibility 20
   - 5.11 Carry Over and Cross Talk 21
6. Results and Discussion 21
   - 6.1 Baseline, Limit of Detection and Limit of Reporting 21
   - 6.2 Stutter 25
   - 6.3 Peak Height Ratio 29
   - 6.4 Homozygote Peak Threshold 30
   - 6.5 Concordance and Sensitivity 31
   - 6.6 Drop-In 34
   - 6.7 Artefacts 34
   - 6.8 Off Scale Allele Peaks 44

6.9 Repeatability and Reproducibility 44
6.10 Carry Over and Cross Talk 48
7. Conclusion 48
8. Recommendations 48
Abbreviations 49
References 50

Figures

Figure 1 Amplification plate layout for 3130*xl* 15
Figure 2 Amplification plate layout for 3500xL 16
Figure 3 Example of the 172bp ART in the CC5 dye 23
Figure 4 Example of the 121bp ART in the CC5 dye 23
Figure 5 Example of the 93bp ART in the CC5 dye 24
Figure 6 Example of stutter peaks 25
Figure 7 Average Peak Height Ratios per locus 30
Figure 8 Comparison of average peaks between the 3500xL and the 3130*xl* 32
Figure 9 Comparison of the maximum peaks detected on the 3500xL to the 3130*xl* 33
Figure 10 Comparison of minimum peak heights detected on the 3500xL and 3130*xl* 33
Figure 11 ART 1 37
Figure 12 ART 2 38
Figure 13 ART 3 38
Figure 14 ART 4 39
Figure 15 ART 5 40
Figure 16 ART 6 and 7 41
Figure 17 BB 10 and BB 13 43
Figure 18 Repeatability for the 3500xL 46
Figure 19 Repeatability for the 3130*xl* 46
Figure 20 Reproducibility of the 3500xL 47
Figure 21 Reproducibility for the 3130*xl*. 47

Tables

Table 1 - PCR cycling conditions for PowerPlex®21 System 11
Table 2 – 3130*xl* CE protocol conditions. 12
Table 3 – 3500xL Series CE protocol conditions 12
Table 4 - 3500xL Baseline Summary of each Dye 22
Table 5 - 3500xL Baseline Summary of all Dyes 22
Table 6 - Artefacts in Orange Dye 23
Table 7- Summary of Stutter Data 27
Table 8 - Comparison of Stutter data points used and Current thresholds between 3130 and 3500xL. 28

Table 9 - Data of the PHR and calculated AI for each locus and overall 29
Table 10 - The average difference between peak heights..... 32
Table 11- Artefacts observed in Fluorescein (blue) dye 36
Table 12- Artefacts observed in JOE (green) dye 39
Table 13- Artefacts observed in TMR-ET (yellow) dye 40
Table 14- Artefacts observed in CXR-ET (red) dye..... 42

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 with 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 the 3500xL Genetic Analyzer is fit for purpose for the analysis of direct amplification reference samples through capillary electrophoresis within Forensic DNA Analysis. A limit of detection of 50 RFU, limit of reporting of 175 RFU, homozygote threshold of 650 RFU and allelic imbalance of 55% will be adopted for the analysis of reference samples processed through direct amplification. The average peak heights on the 3500xL are 6 to 7 times higher than on the 3130*xl*.

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 with 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 Analyzers. 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 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 reference samples amplified by direct amplification using Promega 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 homozygote thresholds
- Concordance and sensitivity
- Repeatability and reproducibility
- Artefacts
- Off scale peaks
- Drop in

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)
- HID 5-DYE Installation Standard (Life Technologies Applied Biosystems, Foster City, CA, USA)

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.)
- Combi-tips (Eppendorf AG, 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)

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)B
- 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 3130xl 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

4. Methods

4.1 Sample Selection

Reference FTA™ samples that have been submitted by the Queensland Police Service for routine testing were used to generate the required data sets.

4.2 Direct Amplification

FTA™ samples underwent direct amplification as per QIS 24823 *FTA™ Processing and Work Instructions.*

Table 1 lists the PCR cycling conditions utilised in this investigation.

Table 1 - PCR cycling conditions for PowerPlex®21 System

PowerPlex®21 Kit	Standard
GeneAmp 9700 mode	FTA PP21
	26 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

4.3 DNA Fragment Capillary Electrophoresis Setup and Conditions

The plates for DNA fragment analysis on the 3130*xl* were prepared as per QIS 19978 *Capillary electrophoresis set up.* The PCR fragments separated by capillary electrophoresis (CE) on the 3130*xl* were performed according to QIS 15998 *Procedure for the Use and Maintenance of the AB 3130xl Genetic Analysers.* Table 2 outlines the 3130*xl* Genetic Analyser running conditions ([2]).

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 set up*. The PCR fragments separated by capillary electrophoresis (CE) 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 – 3500xL Series CE protocol conditions

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

4.4 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_v1.1.1, with additional conditions as follows:

- Samples were analysed at 1RFU
- All true alleles, -2, -1, and +1 repeat stutter of true alleles and spectral pull-up were removed
- As defined by Promega artefact peaks in the N-2bp and/or N+2bp position at D1S1656, D6S1043, D13S317, vWA, D21S11, D7S820, D5S818, D12S391 and D19S433 loci and in the N-1bp position at Amelogenin were removed (2). Any other observed artefacts were removed
- Any peaks determined to be due to carry-over were also removed. Carry-over is defined as the physical transfer of DNA from one injection to the next.

4.5 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_v1.1.1, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOD 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
- 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.

4.6 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_v1.1.1, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated LOD
- Homozygous loci were excluded
- Loci where a true allele falls in 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 are also to be removed.

4.7 Profile Interpretation 4

Profile Interpretation 4 was used to determine:

- Concordance & sensitivity
- Off scale allelic peaks
- Repeatability and reproducibility

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using the analysis panel PowerPlex_21_IDX_v1.5, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated stutter, LOR, LOD, homozygote and allelic imbalance thresholds

- 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 removed.

4.8 Profile Interpretation 5

Profile Interpretation 5 was used for the following experiments:

- Drop in

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using the analysis panel PowerPlex_21_IDX_v1.1.1, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated stutter, LOD, homozygote and allelic imbalance thresholds
- All known alleles were left on
- All -1, -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.

4.9 Profile Interpretation 6

Profile Interpretation 6 was used to determine:

- Artefacts

All DNA profiles were analysed with GeneMapper®ID-X v1.4 using the analysis panel PowerPlex_21_IDX_v1.5, with additional conditions as follows:

- Samples were analysed using the 3500xL calculated stutter, LOD, homozygote and allelic imbalance thresholds
- All true alleles and artefacts were left on
- All -2, -1 and +1 repeat stutter of true alleles and spectral pull-up were removed.
- Any peaks determined to be carry-over peaks were removed.

5. Experimental Design

5.1 Samples/Plate Preparation

5.1.1 Direct Amplification Data Set

A total of 536 reference samples from routine FTA direct amplification sample processing were processed on the 3500xL. These samples were used to create the direct amplification data set that was used to determine the following:

- Baseline, limit of detection and limit of reporting
- Stutter thresholds
- Peak height ratio and allelic imbalance threshold
- Homozygote peak threshold
- Concordance & sensitivity
- Drop in
- Artefacts
- Off scale allele peaks
- Repeatability and reproducibility
- CE carry-over and cross talk

Samples were processed according to Methods 4.2 and 4.3.

5.1.2 Repeatability and Reproducibility Data Set

One plate was prepared using the amplified product of three different reference samples from the direct amplification data set which previously gave full DNA profiles on both the 3130*xl* and the 3500xL, with no off scale peaks. Each of the three reference samples were run in quadruplicate. The plate was prepared for the 3130*xl* according the plate map in Figure 1 and for the 3500xL according to the plate map in Figure 2.

Samples were processed according to Methods 4.2 and 4.3. (B=Blank)

A	Pos	Ladder1	B2	Ladder2							
B	Neg	B1	B3	B10							
C	1-1	1-2	B4	B11							
D	1-3	1-4	B5	B12							
E	2-1	2-2	B6	B13							
F	2-3	2-4	B7	B14							
G	3-1	3-2	B8	B15							
H	3-3	3-4	B9	B16							

Figure 1 Amplification plate layout for 3130*xl*

A	Pos	Ladder1	B2									
B	Neg	B1	B3									
C	1-1	1-2	B4									
D	1-3	1-4	B5									
E	2-1	2-2	B6									
F	2-3	2-4	B7									
G	3-1	3-2	B8									
H	3-3	3-4	B9									

Figure 2 Amplification plate layout for 3500xL

5.2 Baseline, Limit of Detection and Limit of Reporting

From the direct amplification data set 100 samples with varying profiles were selected (extremely excess samples were not selected).

Samples were analysed and profiles interpreted as per Methods 4.4. 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 reviewed the analysed results from the first analysis to ensure the interpretation was in accordance with Methods 4.4.

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

From the direct amplification data set 536 samples with varying profiles were selected (extremely excess samples which could not be interpreted were excluded). Samples were analysed and profiles interpreted as per Methods 4.5. Locus specific stutter thresholds for -2, -1 and +1 repeat stutter were calculated.

The stutter ratio (SR) for each locus and stutter type 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

From the direct amplification data set 534 samples with varying profiles were selected (extremely excess samples were excluded). Samples were analysed and profiles interpreted as per Methods 4.6.

Peak height ratio for heterozygous loci will be calculated by dividing the lower peak height by the higher peak height, as per the equation 5([5]):

Equation 5

$$PHR = LPH / HPH$$

(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 will be calculated as per equation 6 ([6],[7]):

Equation 6:

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

(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 7 ([8]):

Equation 7

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

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

5.6 Concordance and Sensitivity

From the direct amplification data set 100 samples with full profiles when processed on the 3130*xl* and 3500xL were selected (extremely excess samples were excluded).

Samples were analysed and profiles interpreted as per Methods 4.7. The allele designations obtained from the 3500xL were compared to the allele designations previously obtained from the 3130*xl* and the peak heights obtained from the 3500xL were compared to the peak heights obtained on the 3130*xl*.

In Microsoft Excel, for each sample the average peak height was calculated using the AVERAGE function (Arithmetic mean); the maximum peak height was calculated using the MAXIMUM function and the minimum peak height was calculated using the MINIMUM function.

5.7 Drop In

From the direct amplification data set 100 samples and eleven negative amplification controls were selected (extremely excess samples were excluded).

Samples were analysed and profiles interpreted as per Methods section 4.8. A drop in peak was determined to be present based on a comparison of the profile obtained when processed on the 3500xL compared to the expected DNA profile for each sample obtained previously on the 3130*xl*.

5.8 Artefacts

From the direct amplification data set 536 samples were selected (extremely excess samples were excluded).

Samples were analysed and interpreted as per Methods section 4.9. Known artefacts acknowledged by Promega([2]), and those previously observed by Forensic DNA Analysis were included in this analysis. All artefacts were catalogued according to their known location. Any artefacts not previously identified were also noted and recorded.

The prevalence (i.e. count) and severity (i.e. average and maximum peak heights) was recorded for each artefact noted from the 3500xL and compared to previous observations for the 3130*xl*.

The prevalence of the N-2 / N+2 and any other artefacts observed were noted for each loci.

The average and maximum peak height for each artefact was calculated using the AVERAGE and MAXIMUM function in Microsoft Excel.

5.9 Off Scale Allele Peaks

From the direct amplification data set 84 extremely excess samples were selected.

Samples were analysed and interpreted as per Methods section 4.7. Samples were determined to be excess, based on an assessment of the baseline and peak heights. The prevalence of off scale allelic peaks on samples run on the 3500xL was compared to the prevalence of off scale allelic peaks on samples run on the 3130*xl*.

5.10 Repeatability and Reproducibility

Samples were analysed and interpreted as per Methods 4.7. The repeatability and reproducibility plate was used as per Section 5.1.2 and the run comparisons were done based on allele call and peak height.

Repeatability was assessed on the 3500xL and 3130*xl* by:

- Preparing three samples four times in the one plate (Plate 1) by the same operator (Operator 1)
- Plate 1 was processed on the 3500xL twice by Operator 1 on Day 1.
- Plate 1 was processed on the 3130xl twice by Operator 1 on Day 1.

Repeatability on the 3500xL was measured based on consistency of peak heights and concordance of allele designations when comparing the two runs of Plate 1 run by Operator 1 on Day 1.

Repeatability on the 3130*xl* was measured based on consistency of peak heights and concordance of allele designations when comparing the two runs of Plate 1 run by Operator 1 on Day 1.

Reproducibility was assessed on the 3500xL and 3130*xl* by:

- The plate which was used for the Repeatability experiment was also used for the Reproducibility experiment.
- The plate being run under the same conditions once on each instrument by Operator 2
- Plate 1 was processed on the 3500xL once by Operator 2 on Day 2
- Plate 1 was processed on the 3130xl once by Operator 2 on Day 2

Reproducibility on the 3500xL was measured based on the consistency of peak heights and concordance of allele designations when comparing the results from Plate 1 on Day 1 by Operator 1 to the results from Plate 1 on Day 2 by Operator 2.

Reproducibility on the 3130*xl* was measured based on the consistency of peak heights and concordance of allele designations when comparing the results from Plate 1 on Day 1 by Operator 1 to the results from Plate 1 on Day 2 by Operator 2.

The peak height variation of run to run was calculated using the Student's *t*-test function (2, 3) in Microsoft Excel to determine if there was any statistical difference between runs.

5.11 Carry Over and Cross Talk

Any samples which required additional processing due to capillary electrophoresis carry-over or cross talk were recorded during the analysis of samples for the direct amplification data set.

6. Results and Discussion

6.1 Baseline, Limit of Detection and Limit of Reporting

For this validation 100 samples containing DNA 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 reporting scientist reviewed the baseline plate after it had been read by the project officer and confirmed that it was read in accordance with Methods 4.

The LOR is the threshold in which a peak can be confidently distinguished from the baseline (background fluorescence)([18]). The LOR was calculated as per equation 2 in section 5.2.

Table 4 shows the results of the baseline calculations for each of the individual dyes. The highest average peak height (19.61 RFU) and the highest average standard deviation (22.35 RFU) were in the CC5 (orange) dye. This standard deviation is two and a half times larger than the next highest standard deviation from the JOE (green) dye (8.66 RFU). This then resulted in the CC5 dye yielding the highest LOD (86.67 RFU) and the highest LOR (243.15 RFU).

During the analysis of the baseline calculations it was noted that there were artefacts at 172 base pairs (bp) (figure 3), 121bp (figure 4) and 93bp (figure 5), which elevated the average peak height and standard deviation for the CC5 dye. The baseline calculations with the 172bp artefact removed results in the CC5 dye still with the highest average peak height (18.72 RFU), the highest standard deviation (19.36 RFU), highest LOD (76.80 RFU) and the highest LOR (212.32 RFU). With the 172bp, 121bp and 93bp artefacts removed the results, even though slightly less, still has the CC5 dye with the highest average peak height (18.38 RFU), the highest standard deviation (18.49 RFU), highest LOD (73.85 RFU) and highest LOR (203.27 RFU). The CC5 dye (orange) is the dye used for the size standard which defines the sizes of known fragments ([8]). The artefacts seen in the CC5 (orange dye) size standard were not seen in any of the other dye colours. These artefacts do not interfere with the designated size standard peaks and it does not interfere with the sizing of peaks in the other dyes.

Table 4 - 3500xL Baseline Summary of each Dye

	Min RFU	Max RFU	Average RFU	SD	3SD	10SD	LOD (Ave+3SD)	LOR (Ave+10SD)
Fluoresin (Blue)	1	94	5.56	4.09	12.27	40.91	17.84	46.47
JOE (Green)	1	215	9.72	8.66	25.97	86.57	35.69	96.2
TMR (Yellow)	1	104	11.44	6.06	18.18	60.59	29.62	72.03
CXR (Red)	1	217	12.65	6.25	18.76	62.52	31.409	75.17
CC5 (orange)	1	214	19.61	22.35	67.06	223.54	86.67	243.15
CC5 (Orange) (-172 pk removed)	1	153	18.72	19.36	58.08	193.60	76.80	212.32
Orange (-172, 93, 121 pks removed)	1	120	18.38	18.49	55.47	184.89	73.85	203.27

Table 5 shows the results of the baseline calculations when the results are averaged across all of the dyes. The average peak height was 11.32 RFU, the average standard deviation was 11.14 RFU, LOD was 44.75 RFU and the LOR 122.75 RFU. When the 172bp artefact is removed from the calculations the average peak height is 11.19 RFU, standard deviation is 10.21 RFU, LOD is 41.81 RFU and LOR is 113.26 RFU. When the 172bp, 121bp and 93bp artefacts are removed from the calculations the average peak height is 11.14 RFU, standard deviation is 9.93 RFU, LOD is 40.94 RFU and LOR is 110.47 RFU.

Table 5 - 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	217	11.32	11.14	33.43	111.42	44.75	122.75
All Dyes (-172)	1	217	11.19	10.21	30.62	102.07	41.81	113.26
All Dyes (- 172, 121, 93)	1	217	11.14	9.93	29.80	99.34	40.94	110.47

Table 6 shows the number of times the artefacts in the orange dye were seen and the size range in base pairs where the artefacts sized. The 121bp artefact appears to 'shoulder' the size standard 120bp peak. It was also noted some of the other size standard peaks also had shouldering.

Table 6 - Artefacts in Orange Dye

Artefact (~bp)	Min Range (bp)	Max Range (bp)	Occurrence	Min (RFU)	Max (RFU)	Average (RFU)
172	172.27	172.45	105/105	88	214	156
121	121.51	121.64	105/105	33	136	79
94	94	94.97	84/105	8	116	50
93	93.78	93.98	71/105	36	153	91

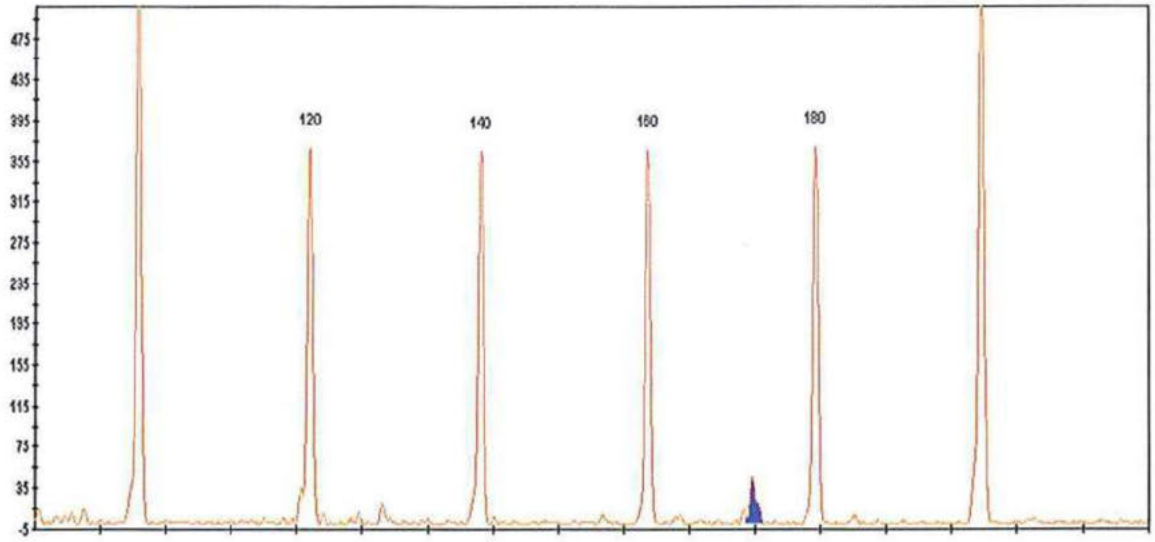

Figure 3 Example of the 172bp ART in the CC5 dye

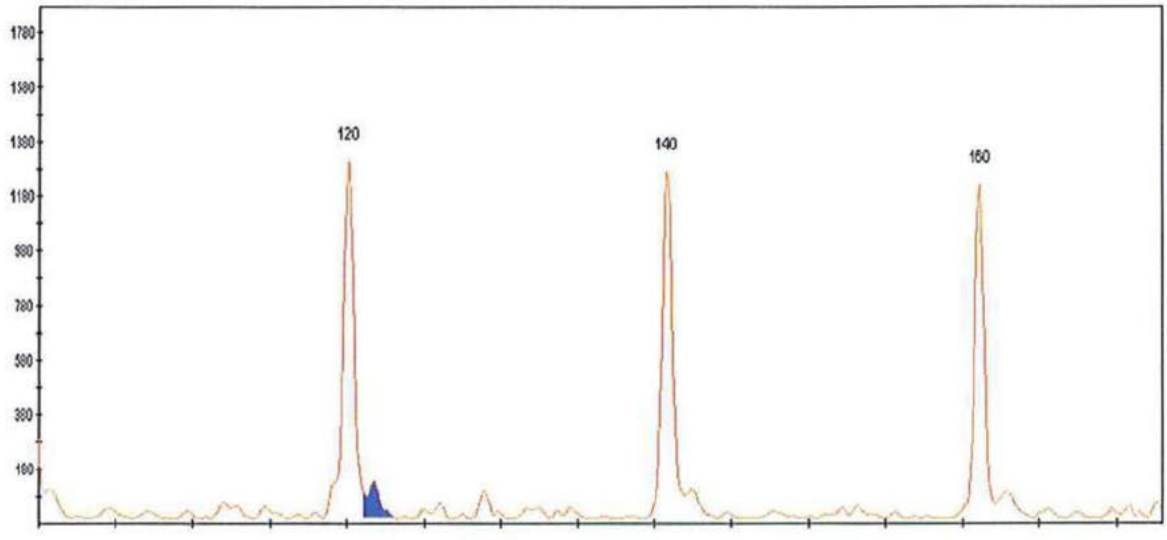

Figure 4 Example of the 121bp ART in the CC5 dye

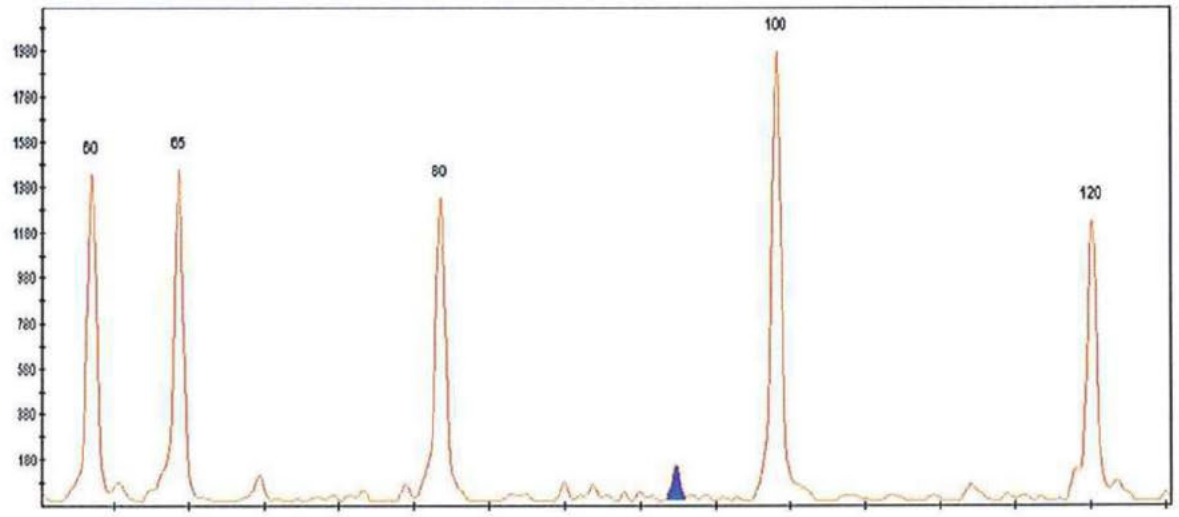

Figure 5 Example of the 93bp ART in the CC5 dye

The most conservative calculation for LOD and LOR included all dyes and artefacts (i.e. LOD 44.75 RFU and LOR 122.75 RFU). Based on these results, there is no evidence that the LOR should be set above the manufacturer recommended default LOR of 175 RFU ([1,2]). Therefore the LOR will be set to the manufacturer recommended default of 175 RFU.

The LOD will be set to 50 RFU (using a conservative rounding from the calculated LOD of 44.75 RFU).

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

The published expectations for difference in RFU scale between 3130*xl* and 3500xL is approximately 3-4 times ([1]). In this project, the scale difference observed between the 3130*xl* and 3500xL was approximately 6-7 times (see Section 6.5 Concordance), however the variation seen for calculated variables was approximately 5 times (see Section 6.4 Homozygote Peak Threshold).

The 3500xL LOD (50 RFU) is approximately 3 times greater than the 3130*xl* LOD (16 RFU).

The 3500xL LOR (175 RFU) is approximately 4 times greater than the 3130*xl* LOD (40 RFU).

The LOD and LOR scale variation is consistent with the published expectations and with the variation seen in the calculated Homozygote Peak Threshold. Accordingly the 3500xL passes this component of the acceptance criteria. (It should be noted that the Homozygote Peak Threshold calculation includes the LOR, as per Equation 7, and therefore these two variables are not independent).

The LOD and LOR scale variation is less than the variation calculated for this project (i.e. approximately 7 times, see Section 6.5 Concordance). Given that the LOD and LOR are calculated using baseline peak height data, and the calculations in Section 6.4 Concordance use allelic peak height data, the direct comparison of these two different data sets may not be valid and as such will not be used to assess whether the 3500xL has passed or failed this experiment.

Overall, the 3500xL has passed this experiment.

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]) and 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.

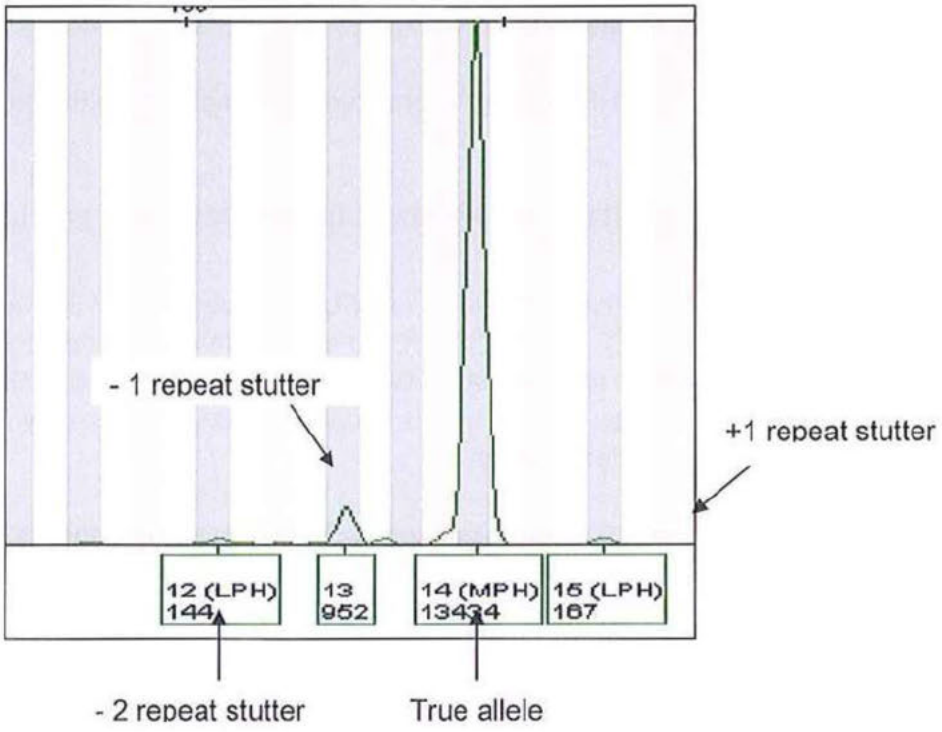

Figure 6 Example of stutter peaks

Table 7 shows the number of times stutter was observed, the average stutter ratio, standard deviation, stutter threshold for -2 repeat, -1 repeat and +1 repeat stutter for each locus. -2 repeat stutter was observed and stutter thresholds calculated for D3S1358, D1S1656, D6S1043, D16S539, D18S51, D2S1338, CSF1PO, vWA,

D21S11, D8S1179, D12S391, D19S433 and FGA. The remaining loci where -2 repeat thresholds have not been calculated is either due to no stutter observed or only observed once. -1 repeat stutter was observed and stutter thresholds calculated for all loci. +1 repeat stutter was observed and stutter thresholds calculated for D3S1358, D1S1656, D6S1043, D13S317, D16S539, D18S51, D2S1338, CSF1PO, vWA, D21S11, D7S820, D5S818, D8S1179, D12S391, D19S433 and FGA.

Table 8 shows the stutter thresholds for -2 repeat, -1 repeat and +1 repeat stutter calculated for the 3500xL, the -1 repeat stutter thresholds from the Promega PowerPlex®21 Genemapper stutter file and the current thresholds used for the 3130*xl*.

-2 repeat stutter has not previously been calculated for samples processed through direct amplification.

For -1 repeat stutter, where the calculated thresholds are lower than the Promega PowerPlex®21 Genemapper stutter file threshold, the Promega PowerPlex®21 Genemapper stutter file threshold is implemented as the analysis threshold. It is only where the calculated stutter threshold is greater than the stutter file threshold that the calculated stutter threshold would be implemented as the analysis threshold. In Table 8, for -1 repeat stutter, the recommended analysis thresholds are highlighted.

For -1 repeat stutter, 18 out of the 20 loci had 3500 stutter thresholds that were lower than current 3130*xl* thresholds. Only D12S391, D19S433 had stutter thresholds which were higher for the 3500xL than the stutter file and the 3130*xl*.

For +1 repeat stutter, thresholds were calculated for 16 of the 20 loci. Of these 6 loci had thresholds which were higher than current 3130*xl* thresholds and 4 loci did not have a 3130*xl* threshold.

Table 7- Summary of Stutter Data

Locus	-2 STUTTER				-1 STUTTER					+1 STUTTER				
	No. Observed	Average	STD Deviation	Threshold % (Ave+3STD)	No. Observed	Average	STD Deviation	Threshold % (Ave+3STD)	Current Threshold %	No. Observed	Average	STD Deviation	Threshold % (Ave+3STD)	Current Threshold %
D3S1358	21	0.0074	0.0022	1.4	270	0.0777	0.0152	12.3	14.0	19	0.0070	0.0040	1.9	1.8
D1S1656	36	0.0143	0.0097	4.3	555	0.0775	0.0230	14.6	15.5	51	0.0108	0.0041	2.3	3.7
D6S1043	6	0.0069	0.0021	1.3	455	0.0601	0.0122	9.7	14.0	54	0.0090	0.0025	1.6	1.7
D13S317	1	0.0041	-	-	195	0.0451	0.0142	8.8	11.0	29	0.0097	0.0021	1.6	0.8
Penta E	0	-	-	-	115	0.0364	0.0144	8.0	10.0	1	0.0073	-	-	6.2
D16S539	23	0.0061	0.0031	1.5	289	0.0571	0.0171	10.8	12.0	117	0.0083	0.0036	1.9	2.2
D18S51	55	0.0173	0.0150	6.2	525	0.0777	0.0270	15.9	16.0	81	0.0117	0.0125	4.9	2.3
D2S1338	8	0.0085	0.0023	1.5	440	0.0751	0.0157	12.2	16.0	3	0.0111	0.0035	2.2	-
CSF1PO	3	0.0033	0.0003	0.4	123	0.0518	0.0117	8.7	11.0	11	0.0137	0.0067	3.4	1.9
Penta D	0	-	-	-	57	0.0196	0.0075	4.2	9.0	0	-	-	-	-
THO1	0	-	-	-	188	0.0223	0.0095	5.1	6.0	1	0.0546	0.0518	-	1.0
vWA	9	0.0143	0.0067	3.4	281	0.0687	0.0200	12.9	14.0	6	0.0197	0.0133	6.0	5.1
D21S11	11	0.0067	0.0037	1.8	498	0.0706	0.0166	12.0	15.2	77	0.0112	0.0050	2.6	4.8
D7S820	1	0.0040	-	-	179	0.0480	0.0166	9.8	11.0	6	0.0067	0.0025	1.4	7.5
D5S818	1	0.0034	-	-	155	0.0496	0.0140	9.2	10.0	28	0.0083	0.0025	1.6	-
TPOX	0	-	-	-	54	0.0269	0.0124	6.4	7.0	0	-	-	-	-
D8S1179	33	0.0058	0.0025	1.3	337	0.0631	0.0151	10.8	12.0	61	0.0072	0.0035	1.8	3.8
D12S391	118	0.0121	0.0085	3.8	584	0.0887	0.0288	17.5	17.0	26	0.0070	0.0046	2.1	1.8
D19S433	20	0.0059	0.0034	1.6	349	0.0649	0.0179	11.9	11.0	8	0.0095	0.0045	2.3	-
FGA	10	0.0098	0.0039	2.1	368	0.0633	0.0171	11.5	12.0	36	0.0067	0.0039	1.8	-

Table 8 - Comparison of Stutter data points used and Current thresholds between 3130 and 3500xL.

	-2 Repeat Stutter		-1 Repeat Stutter			+1 repeat stutter	
Locus	Threshold % (Ave+3STD)	Current Threshold %	Threshold % (Ave+3STD)	Promega Stutter File	Current Threshold %	Threshold % (Ave+3STD)	Current Threshold %
D3S1358	1.4	-	12.3	14.0	14.0	1.9	1.8
D1S1656	4.3	-	14.6	15.0	15.5	2.3	3.7
D6S1043	1.3	-	9.7	14.0	14.0	1.6	1.7
D13S317	-	-	8.8	11.0	11.0	1.6	0.8
Penta E	-	-	8.0	10.0	10.0	-	6.2
D16S539	1.5	-	10.8	12.0	12.0	1.9	2.2
D18S51	6.2	-	15.9	16.0	16.0	4.9	2.3
D2S1338	1.5	-	12.2	16.0	16.0	2.2	-
CSF1PO	0.4	-	8.7	11.0	11.0	3.4	1.9
Penta D	-	-	4.2	9.0	9.0	-	-
THO1	-	-	5.1	6.0	6.0	-	1.0
vWA	3.4	-	12.9	14.0	14.0	6.0	5.1
D21S11	1.8	-	12.0	13.0	15.2	2.6	4.8
D7S820	-	-	9.8	11.0	11.0	1.4	7.5
D5S818	-	-	9.2	10.0	10.0	1.6	-
TPOX	-	-	6.4	7.0	7.0	-	-
D8S1179	1.3	-	10.8	12.0	12.0	1.8	3.8
D12S391	3.8	-	17.5	17.0	17.0	2.1	1.8
D19S433	1.6	-	11.9	11.0	11.0	2.3	-
FGA	2.1	-	11.5	12.0	12.0	1.8	-

Accepted Threshold Stutter Thresholds =

6.2.2 Acceptance Criteria – Stutter Thresholds

There are no -2 repeat stutter thresholds for direct amplification on the 3130*xl* therefore the results for the 3500xL cannot be compared or included in an assessment of whether the 3500xL passed the acceptance criteria for this experiment.

For -1 repeat stutter, there were only two loci (D12S391 and D19S433) which had thresholds that were greater than the stutter file thresholds and the current 3130*xl* thresholds. In both cases the figures were similar. The 3500xL has passed this component of this experiment.

For +1 repeat stutter the 3500xL thresholds did vary from current 3130*xl* thresholds. This has resulted from the use of different data sets, of different sizes, and the fact that the number of observations for +1 repeat stutter is much lower than for -1 repeat stutter. Given these analysis issues, the +1 repeat stutter thresholds must be accepted as calculated and cannot be used to assess whether the 3500xL has passed or failed this experiment.

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 543 samples from the direct amplification data set were used to calculate the peak height ratio. Table 9 summarises the results of the average PHR and 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 80% and a standard deviation of 0.14, while the overall average PHR is 86% with a standard deviation of 0.11.

Figure 7 graphically shows the average PHR and the allelic imbalance across all loci. The lowest observed allelic imbalance was at Penta E with 37%. Using Equation 6 in section 5.4 the overall Allelic Imbalance threshold (AI_{TH}) calculates to 54%. This is 6% lower than the threshold previously calculated for the 3130*xl* ([11]).

Table 9 - Data of the PHR and calculated AI for each locus and overall

	Average	STDEV	3xSTDEV	AI_{TH}	n*
AMEL	0.84	0.12	0.35	0.49	396
D3S1358	0.85	0.14	0.43	0.43	93
D1S1656	0.90	0.08	0.23	0.67	321
D6S1043	0.87	0.11	0.32	0.55	243
D13S317	0.88	0.08	0.23	0.65	149
Penta E	0.80	0.14	0.43	0.37	293
D16S539	0.88	0.11	0.34	0.54	95
D18S51	0.84	0.12	0.35	0.49	248
D2S1338	0.86	0.11	0.32	0.54	303
CSF1PO	0.86	0.11	0.34	0.53	32
Penta D	0.85	0.11	0.33	0.52	148
TH01	0.91	0.08	0.23	0.69	262
vWA	0.87	0.09	0.26	0.61	125
D21S11	0.90	0.08	0.23	0.67	257
D7S820	0.89	0.08	0.25	0.64	132
D5S818	0.88	0.07	0.22	0.65	53
TPOX	0.86	0.11	0.33	0.53	135
D8S1179	0.88	0.09	0.26	0.62	148
D12S391	0.84	0.10	0.30	0.54	280
D19S433	0.89	0.09	0.27	0.62	144
FGA	0.87	0.10	0.30	0.57	217
All samples	0.86	0.11	0.32	0.54	4074

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

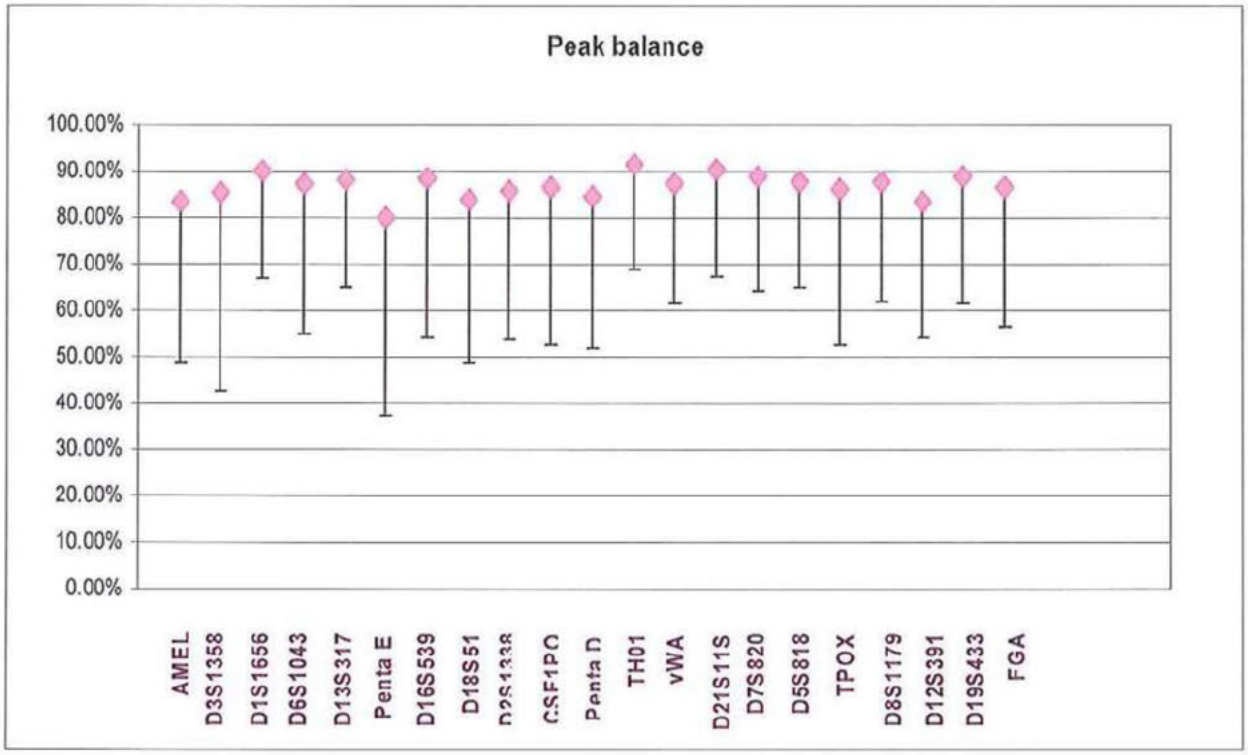

Figure 7 Average Peak Height Ratios per locus

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

6.3.2 Acceptance Criteria – Allelic Imbalance Threshold

The 3500xL AI_{TH} for direct amplification was calculated as 54%. The current AI_{TH} for direct amplification on the 3130*xl* is 60%. The 3500xL passes this experiment as these thresholds are comparable.

The AI_{TH} for direct amplification will be rounded to 55% for implementation.

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 7 in section 5.5 was used to calculate the homozygote threshold.

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

$= 175 \times (1/ 0.544415) \times 2$

$= 642.892$

The homozygote threshold will be rounded to 650 RFU for implementation.

6.4.1 Acceptance Criteria – Homozygote Peak Threshold

The current homozygote peak threshold for direct amplification on the 3130*xl* is 130 RFU. The calculated homozygote peak threshold for direct amplification on the 3500xL was 643 RFU, which

is approximately 5 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 Concordance and Sensitivity

A total of 100 samples (number of alleles = 4204) run on the 3500xL were found to be concordant to the same 100 samples run on the 3130xl (number of alleles = 4201). Three extra peaks were noted on the 3500xL run however when the profiles from the 3130xl were reviewed the extra peaks were present below threshold and each of the extra peaks were in the +1 repeat stutter position. The extra peaks were at D6S1044, D21S11 and D8S1179.

The average, maximum and minimum peak height for 100 samples were analysed and outlined in table 10. The average difference in peak height across all loci is 7.17 times higher on the 3500xL than the 3130xl with a standard deviation of 1.71.

Figure 8 shows the comparison of the average peak heights observed on the 3500xL and on the 3130xl. On the 3500xL and 3130xl the highest average peak heights were observed at D12 and lowest average peak heights observed at Penta E.

Figure 9 shows the comparison of the maximum peak heights seen on the 3500xL and the 3130xl. On the 3500xL and 3130xl the highest maximum peak heights were observed at D19 and lowest maximum peak heights observed at Penta E.

Figure 10 shows the comparison of the minimum peak heights observed on the 3500xL and the 3130xl. On the 3500xL and 3130xl the highest minimum peak heights were observed at D19 and lowest minimum peak heights observed at TPOX.

Figures 8, 9 and 10 show a consistent scale difference between the 3130xl and 3500xL of approximately 7 times for average, maximum and minimum peak height.

Table 10 - The average difference between peak heights

	Average PH (RFU) (3500xL)	Max PH (RFU) (3500xL)	Min PH (RFU) (3500xL)	Average PH (RFU) (3130*xl*)	Max PH (RFU0 (3130*xl*)	Min PH (RFU) (3130*xl*)	AVE DIFF	STDEV
AMEL	4542.37	15985.00	626.50	548.50	2334.00	78.00	8.75	2.68
D3S1358	5385.78	14101.00	641.00	726.33	2762.00	77.00	7.93	2.42
D1S1656	4507.73	11621.00	567.00	665.43	2168.00	82.50	7.35	2.22
D6S1043	4375.67	18051.00	787.50	643.11	3444.00	79.00	7.45	2.26
D13S317	3390.92	13376.00	439.50	464.53	1942.00	73.50	7.83	2.28
Penta E	1949.48	6247.00	340.00	264.63	1066.50	47.50	7.96	2.37
D16S539	9151.03	28315.00	945.00	1269.12	3865.00	173.00	7.55	1.65
D18S51	6903.05	17921.00	753.50	1044.72	3856.00	164.00	6.81	1.43
D2S1338	3252.69	12371.00	503.50	492.27	1768.00	95.50	6.93	1.44
CSF1PO	3365.40	15025.00	666.00	475.47	2225.00	101.50	7.40	1.49
Penta D	2640.01	8667.00	395.50	383.09	1364.00	57.50	7.17	1.42
TH01	4705.30	17131.00	826.00	733.40	3180.00	92.00	6.69	1.63
vWA	4751.45	13429.00	524.00	863.04	2655.00	144.50	5.69	1.40
D21S11	4962.10	13948.00	599.50	914.05	3832.00	164.00	5.67	1.31
D7S820	3483.52	11853.00	546.50	612.75	2209.00	138.00	5.85	1.32
D5S818	3864.34	13783.00	574.00	637.03	2402.00	138.50	6.21	1.38
TPOX	1963.51	11287.00	212.00	335.85	2051.00	45.00	6.16	1.36
D8S1179	6844.75	19064.50	731.50	827.26	3172.00	128.50	8.62	2.09
D12S391	10466.89	29532.00	490.50	1506.95	4960.00	102.00	7.28	1.71
D19S433	8202.87	31351.00	990.50	1174.99	5305.00	169.00	7.32	1.66
FGA	6127.33	27246.00	718.00	811.06	3844.00	134.50	7.84	1.73
Overall							7.17	1.71

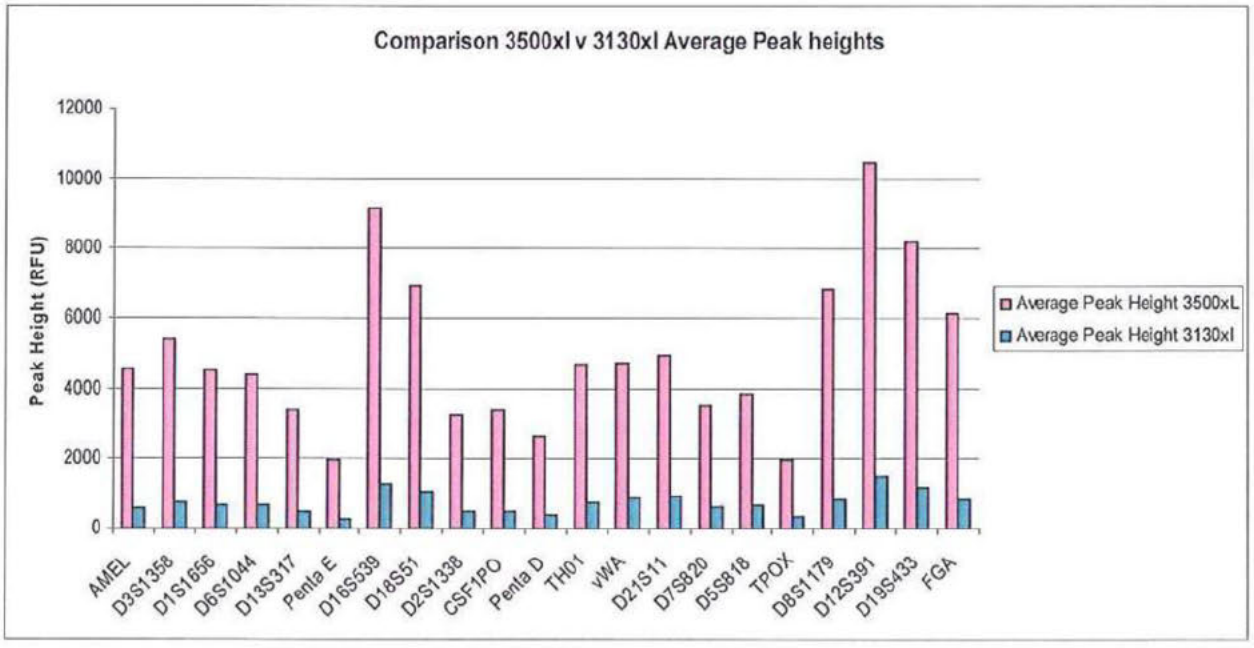

Figure 8 **Comparison of average peaks between the 3500xL and the 3130*xl***

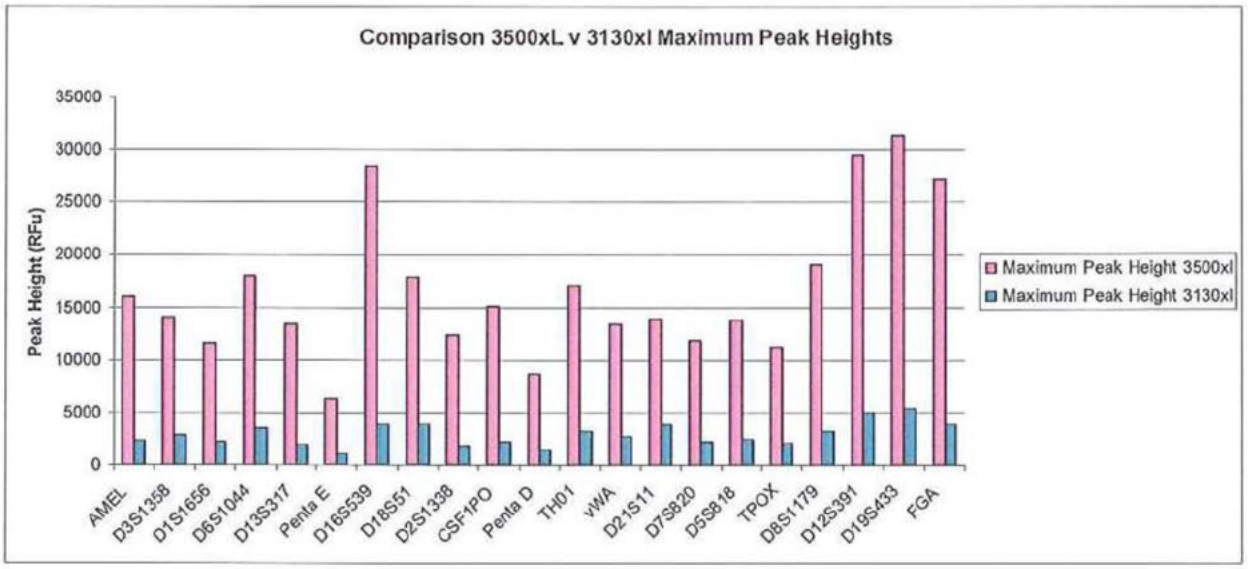

Figure 9 **Comparison of the maximum peaks detected on the 3500xL to the 3130*xl***

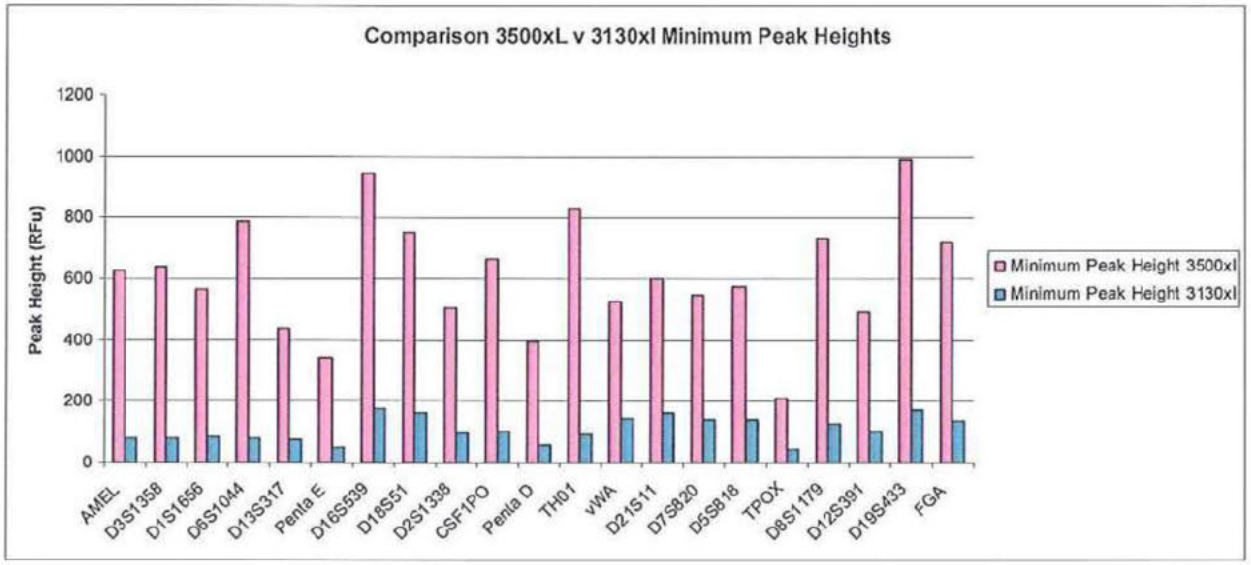

Figure 10 **Comparison of minimum peak heights detected on the 3500xL and 3130*xl***

6.5.2 Acceptance Criteria – Concordance and Sensitivity

For the 100 samples in this experiment, all allele designations were concordant. Three extra peaks were noted on the 3500xL run; however these were present on the 3130*xl* in the +1 repeat stutter position, below the LOR. The 3500xL passed this component of the acceptance criteria.

The combined average peak height for all loci on the 3500xL was approximately 7 times the average peak height for the 3130*xl*. This is greater than the published expectation of approximately 3-4 times ([1]). The calculated scale variation was also greater than the variation observed for other variables calculated in this project (i.e. LOD, LOR and Homozygote Peak Threshold).

The LOD and LOR are calculated using baseline peak height data whereas this experiment compared allelic peak height data (N.B. Homozygote Peak Threshold is calculated using the LOR as per Equation 7 and is therefore linked to the variation seen in the LOR). The direct comparison of peak height variation between baseline and allelic peak height data sets will not be assessed as a part of the acceptance criteria for this experiment.

6.6 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]).

No drop-in events were observed in the 100 samples and 11 negative amplification controls analysed on the 3500xL or on the 3130*xl*.

6.6.1 Acceptance Criteria – Drop in

No drop in peaks were observed in samples included in this experiment. Given that drop in is a result of amplification and not capillary electrophoresis, and the success of the concordance study (i.e. whereby amplification product which is visualised and sized on the 3130*xl* is also expected to be visualised and sized on the 3500xL), it is reasonable to conclude that if drop in were detected on the 3130*xl* that it would also be detected on the 3500xL. Therefore, the 3500xL has passed this experiment.

6.7 Artefacts

A total of 536 samples (533 samples and 3 positive controls) from the direct amplification data set were analysed at limit of detection (LOD), 50 RFU. Any samples that were extremely excess, NSD (no sizable data) or NAD (no analysed data) were excluded. Peaks that were considered to be raised (bad) baseline, stutter (-2, -1 and +1) and pull up, were removed.

In the analysed data set, a few samples were observed to have peak shadow artefacts on the 3500xL which were not observed on the original 3130*xl* results. These artefacts were not included in the artefact count (Tables 11-14) as they cannot be categorised into a set location (size range) as the peaks vary between individual samples (for more information on peak shadowing refer to pg 14-15 of QIS: 31389 *STR fragment analysis of PowerPlex®21 profiles using GeneMapper® ID-X software*).

Peaks that fell into an artefact position (eg. N-2) which could be attributed to pull up were removed. These peaks were excluded from the artefact observed count due to the inability to determine if these peaks are a true artefact or present due to pull up (PU) or a combination of both. It is possible that by either including or excluding these peaks, the number of times an artefact is observed may be either over or under represented in the data set. The artefact/pull up combination would also give the peak an elevated peak height and thus should not be used to calculate an average and maximum peak height, as the data may be skewed to a higher value.

Promega has noted the presence of artefact peaks observed at some of the PowerPlex® 21 System loci ([2]). These include N-2 and N+2 artefacts at the following loci: D1S1656, D6S1043, D13S317, vWA, D21S11, D7S820, D5S818, D12S391 and D19S433. Promega has also noted an

N-1 artefact sometimes present at Amelogenin. In this validation and in routine plates (processed on 3130*xl*) N-1, N-2 and N+2 artefacts have been observed in loci not specifically noted by Promega. Also an N-3 and N+3 artefact that has not been noted by Promega has been observed a number of times at a few different loci. The observed N-/+1, N-/+2, N-/+3 artefacts did not have their size (bp) and peak height (average and maximum RFU) recorded in Tables 11-14 as these values are dependant on the main allele location and peak height and thus would not provide useful information.

In tables 11-14 artefacts that consisted of one or two peaks at a locus (either labelled with an OL or allele designation) that was not commonly seen, or could possibly be a part of the baseline, was designated "BB" (bad baseline). Artefacts that appeared to be more commonly seen (in the same location) that sometimes ranged from one to a series of multiple peaks was given the name "ART" (artefact).

Artefacts that have both an average and maximum peak height <LOR (175 RFU) are not expected to be labelled routinely. Artefacts that have an average peak height <LOR and a maximum peak height of >LOR are expected on average not be labelled routinely, but can be expected to be labelled on occasion. Lastly, peaks with both an average and maximum peak height >LOR are expected to be labelled routinely.

Table 11- Artefacts observed in Fluorescein (blue) dye

Loci	Artefact [allele designations]	Artefact Description	Number of Observations	Total Number of Observations	Min Size (bp)	Max Size (bp)	Average Peak Height (RFU)	Max Peak Height (RFU)
AMEL	N-1 (excluded*)	*These peaks were excluded from the artefact observed count due to the inability to determine if these peaks are a true artefact or pull up. However are recorded separately due to the high number of occurrences.	80	14	-	-	-	-
	N-1	Artefact identified by Promega	9		-	-	-	-
	N-2 (excluded*)	*	1		-	-	-	-
	N-2		4		-	-	-	-
	N-3 (excluded*)	*	34		-	-	-	-
	N+2		1		-	-	-	-
D3S1358	N-2		11	73	-	-	-	-
	N+2		2		-	-	-	-
	8		9		93.18	93.95	132	265
	9 (excluded*)	*	101		-		-	-
	9		5		97.84	98.98	244	325
	10		11		101.94	102.24	97	267
	11		7		106.18	106.5	158	299
	12		2		110.44	110.57	54	56
	ART 1 [OLs, 16, 17, 17.1, 18, 19, 20.1]	Ranges from 1 peak to multiple peaks, usually after the main alleles towards the end of loci, can resemble bad baseline	25		127.11	148.15	82	182
	BB 1[OL]	LHS of main alleles (sample is PA with BB)	1		120.77	-	-	177
D1S1656	N-2	Known artefact identified by Promega	226	229	-	-	-	-
	BB 2[OL]	LHS of main alleles (samples have raised baseline/BB)	2		165.16	166.55	159	253
	BB 3[OL]	LHS of main alleles, very close to the D3/D1 marker. Query possible cross over of BB/artefacts from the end of D3 loci.	1		150.4	-	-	69
D6S1043	N-2	Known artefact identified by Promega	4	16	-	-	-	-
	N+2	Known artefact identified by Promega	1		-	-	-	-
	ART 2 [OLs, 8, 10]	LHS of main alleles, range from 1 to multiple peaks labelled.	4		215.56	225.84	82	145
	ART 3 [OLs, 13, 14, 15]	Usually multiple peaks in between or after the 2 main alleles.	7		232.59	246	82	126
D13S317	N-2	Known artefact identified by Promega	66	66	-	-	-	-
Penta E	N-2		2	2	-	-	-	-

Peak height less than the LOR = Peak height greater than LOR =

Table 11 shows the observed artefacts in loci labelled with the Fluorescein (blue) dye. At Amelogenin, 14 artefacts were observed which include: N-1 (9 times); N-2 (4 times); N+2 (once).

As previously mentioned, peaks were excluded from the artefact observed count if they coincided with pull up. Due to the location of D16S539 and TH01, a number of peaks at Amelogenin and D3S1358 (artefact in position 9) were excluded from the artefact count due to possible pull up from these two loci. Due to the high number of these possible artefact/pull up occurrences, it was decided to record these peaks separately. At Amelogenin the following possible artefact/pull up peaks were observed: N-1 (80 times); N-2 (once); and N-3 (34 times).

At D3S1358, 73 artefacts were observed which include; N-2 (11 times); N+2 (twice); 8 (9 times); 9 (5 times); 10 (11 times); 11 (7 times); 12 (twice); ART 1 (25 times) and BB 1 (once). Possible artefact/pull up peaks in the 9 bin position were observed 101 times. An artefact categorised as ART 1 (in Table 11), ranges from one to multiple peaks, sometimes resembling raised baseline. See Figure 11 for an example of ART 1. BB1 is a single peak observed once in an individual sample. This particular sample exhibits preferential amplification (PA) with bad baseline and the artefact observed (BB1) could be bad baseline, but was not removed due to the peak height being >LOR and higher than the surrounding baseline.

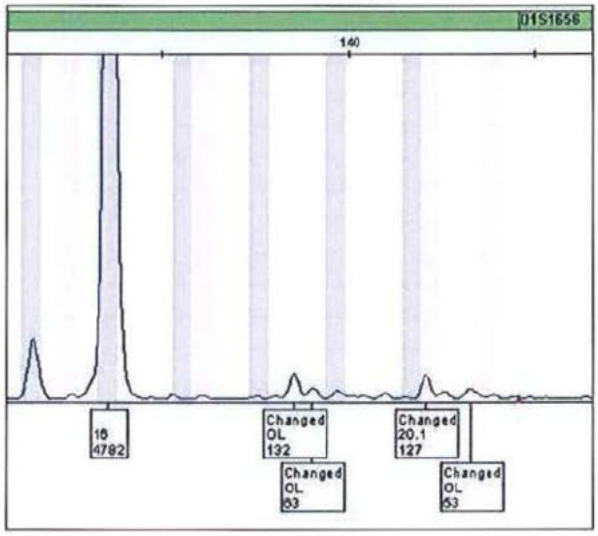

Figure 11 ART 1

Artefacts 8, 10, 11 and ART 1 all have an average peak height that is <LOR and a maximum peak height >LOR. The 9 artefact (that does not coincide with PU) has both an average and maximum peak height >LOR. This indicates that the artefact (when present) will be labelled in routine plates; however these peaks also have a low occurrence (observed only 5 times). Thus it is not expected to be seen often. The artefact in the 12 position has both an average and maximum peak height <LOR.

A total of 229 artefacts were observed at D1S1656; with the N-2 artefact observed 226 times. Two peaks (BB2) were noted from two different samples in similar locations; however both these samples have been noted to also have bad baseline. Also a single peak (BB3) located very close to the D3/D1 marker was observed once in an individual sample. This peak has been noted separately in Table 11 as this particular sample also has ART 1 in D3 that looks as though the ART 1 (or bad baseline) from D3 has crossed the marker into D1. The average peak height of BB2 and peak height of BB3 are <LOR, however the maximum peak height of BB2 is >LOR.

Sixteen artefacts were observed at D6S1043 which include; N-2 (4 times); N+2 (once); ART 2 (4 times); and ART 3 (7 times). ART 2 can be described as ranging from 1 to multiple peaks at the left hand side (LHS) of the main alleles (see Figure 12). ART 3 artefact was observed as (usually) multiple peaks either in between or to the RHS of the main alleles (see Figure 13). Both ART 2 and ART 3 have average and maximum peak heights <LOR.

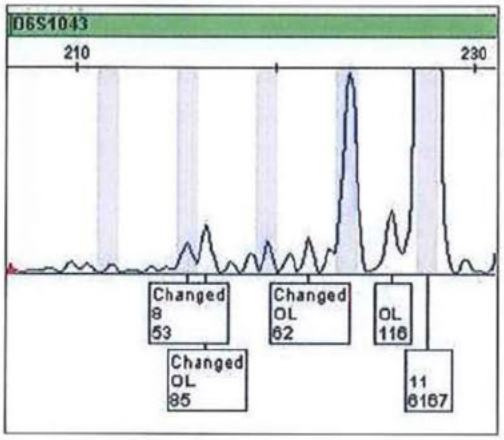

Figure 12 ART 2

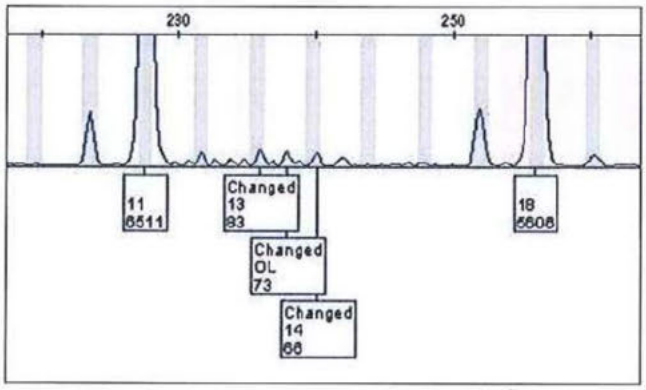

Figure 13 ART 3

The N-2 artefact was observed 66 times at D13S317 and twice at Penta E; these were the only artefacts observed at these loci.

Table 12- Artefacts observed in JOE (green) dye

Loci	Artefact [allele designations]	Artefact Description	Number of Observations	Total Number of Observations	Min Size (bp)	Max Size (bp)	Average Peak Height (RFU)	Max Peak Height (RFU)
D16S539	N-2		5	6	-	-	-	-
	N-3		1		-	-	-	-
D18S51	N-2		11	38	-	-	-	-
	22.2	Single peak (some 22.2 peaks have been removed as PU and excluded from the observed count).	6		193.46	193.6	77	129
	ART 4 [OLs, 11, 11.2, 12]	Regularly seen as a series of multiple peaks close together, sometimes with the baseline raised (similarly to a blob).	21		146.96	152.08	134	312
D2S1338	N-2		1	22	-	-	-	-
	ART 5 [OL]	Regularly seen very close to (almost behind) the LHS red loci marker (triangle).	19		214.36	214.51	66	90
	BB 4[OL]	Single peak, LHS of main alleles.	2		247.63	247.76	74	93
CSF1PO	-		-	-	-	-	-	-
Penta D	-		-	-	-	-	-	-

Peak height less than the LOR = Peak height greater than LOR =

Table 12 shows the observed artefacts in loci labelled with the JOE (green) dye. Six artefacts were observed at D16S539; N-2 (5 times); and N-3 (once). Loci D18S51 had 38 artefacts observed which include; N-2 (11 times); 22.2 (6 times); and ART 4 (21 times). The 22.2 is a single peak artefact that has both an average and maximum peak height <LOR. ART 4 is regularly seen as multiple close together peaks regularly labelled with OLs, 11, 11.2 and 12 (146.96-152.08 bp). See Figure 14 for an example of ART 4. ART 4 has an average peak height <LOD and a maximum peak height >LOD.

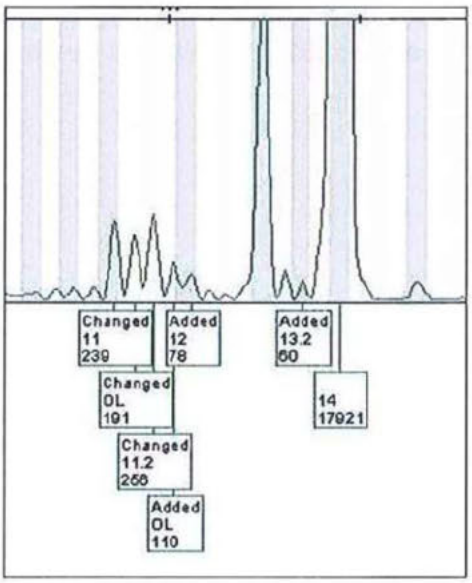

Figure 14 ART 4

At D2S1338, 22 artefacts were observed which include; N-2 (once); ART 5 (19 times); and BB4 (twice). ART 5 is a peak regularly seen very close to the D2S1338 marker triangle (see Figure 15). BB 4 consists of two peaks that were noted from two different samples in similar locations. Both ART 5 and BB 4 have an average and maximum peak height < LOD.

No artefacts were observed at CSF1PO and Penta D.

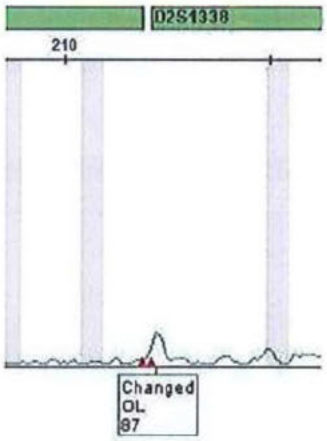

Figure 15 ART 5

Table 13- Artefacts observed in TMR-ET (yellow) dye

Loci	Artefact [allele designations]	Artefact Description	Number of Observations	Total Number of Observations	Min Size (bp)	Max Size (bp)	Average Peak Height (RFU)	Max Peak Height (RFU)
THO1	N-1		1	11	-	-	-	-
	N-2		9		-	-	-	-
	N-3		1		-	-	-	-
vWA	N-2	Known artefact identified by Promega	2	56	-	-	-	-
	N-3		50		-	-	-	-
	N+2	Known artefact identified by Promega	2		-	-	-	-
	N+3		2		-	-	-	-
D21S11	N-2	Known artefact identified by Promega	3	24	-	-	-	-
	ART 6 [OLs, 34.2, 35, 36, 37]	First (LHS) of 2 blob like artefacts regularly seen close together	11		246.28	256.42	87	199
	ART 7 [OLs, 37.2, 38, 38.2, 39]	Second (RHS) of 2 blob like artefacts regularly seen close together	8		258.6	265.56	72	106
	BB 5[24.2]	Single peak, LHS of main alleles (present in 3130 run)	1		205.96	-	-	55
	BB 6[25.3]	Single peak, LHS of main alleles	1		211.61	-	-	70
D7S820	N-2	Known artefact identified by Promega	7	8	-	-	-	-
	BB 7[OL]	RHS of main alleles (sample is PA with BB, also present in 3130 run)	1		300	-	-	82
D5S818	N-2	Known artefact identified by Promega	1	1	-	-	-	-
TPOX	-		-	-	-	-	-	-

Peak height less than the LOR = Peak height greater than LOR =

Table 13 shows the observed artefacts in loci labelled with the TMR-ET (yellow) dye. At TH01, 11 artefacts were noted; an N-1 (once); N-2 (9 times); and N-3 (once). At vWA, 56 artefacts were observed which include; N-2 (twice); N+2 (twice); N+3 (twice); and N-3 (50 times). A total of 24 artefacts were noted at D21S11; N-2 (3 times); ART 6 (11 times); ART 7 (8 times); BB 5 (once); and BB 6 (once). ART 6 and ART 7 are two blob like artefacts that are often seen together or can be seen individually (see Figure 16). The two peaks BB 5[24.2] and BB 6[25.3] were noted from two different samples both similarly found on the left hand side of the main alleles. ART 7, BB 5 and BB 6 all have an average and maximum peak height < LOD. ART 6 had an average peak height <LOD and maximum peak height >LOD.

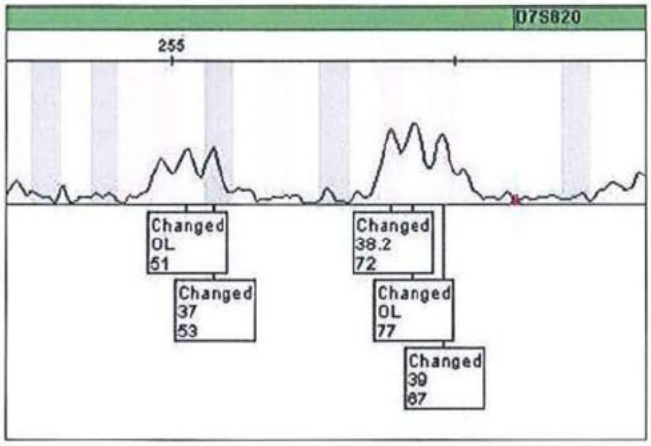

Figure 16 ART 6 and 7

Eight artefacts were noted at D7S820; N-2 (7 times) and BB 7 (once). BB 7 is a single peak observed once in a sample that also exhibits PA with bad baseline and has a peak height <LOR.

Only a single N-2 artefact was observed at D5S818 and no artefacts observed at TPOX.

Table 14- Artefacts observed in CXR-ET (red) dye

Loci	Artefact [allele designations]	Artefact Description	Numnber of Observations	Total Number of Observations	Min Size (bp)	Max Size (bp)	Average Peak Height (RFU)	Max Peak Height (RFU)
D8S1179	N-1		1	34	-	-	-	-
	N-2		25		-	-	-	-
	N-3		4		-	-	-	-
	BB 8[OLs, 19]	Single peak on RHS of main peaks, (19 labelled peak is in original 3130 run)	4		122.6	130.87	195	293
D12S391	N-1		1	158	-	-	-	-
	N-2	Known artefact identified by Promega	9		-	-	-	-
	N-3		84		-	-	-	-
	N+2	Known artefact identified by Promega	3		-	-	-	-
	N+3		1		-	-	-	-
	BB 9[17.3]**	**Consists of either 1 or 2 peaks	11		150.6	150.75	74	79
	N+3[17.3]**	***This peak coincides with both an observed (labelled) artefact and is in a N+3 artefact position.	7		-	-	-	-
	BB 10 [18.3]**		16		154.63	154.8	78	134
	N+3[18.3]**	***	5		-	-	-	-
	N-1[18.3]**	This peak coincides with both an observed (labelled) artefact and is in a N-1 artefact position.	1		-	-	-	-
	BB 11[19.3]**		1		158.81	-	-	50
	N+3[19.3]**	***	1		-	-	-	-
	BB 12[20.3]**		4		162.7	162.93	84	96
	BB 13[24.3]**		3		178.8	178.94		
	BB 14[OLs]**		11		134.59	186.93	82	99
D19S433	N-2	Known artefact identified by Promega	31	37	-	-	-	-
	BB 15[5.2,OL]	Single peak, LHS of main alleles	2		193.9	194.03	78	91
	BB 16[OLs, 11, 12]	Single peak, LHS of main alleles	4		213.89	219.01	117	264
FGA	-		-	-	-	-	-	-

Peak height less than the LOR = Peak height greater than LOR =

Table 14 shows the observed artefacts in loci labelled with the CXR-ET (red) dye. A total of 34 artefacts were noted at D8S1179; N-1 (once), N-2 (25 times), N-3 (4 times) and BB 8 (4 times). BB 8 are all single peaks on the RHS of the main alleles from four different samples with the average and maximum peak height >LOD.

A total of 158 artefacts were observed at D12S391 which include; N-1 (once), N-2 (9 times), N-3 (84 times), N+2 (3 times), N+3 (once). The remaining 60 artefacts all consist of 1 or 2 peaks that all fall within the size range (134.59-186.93 bp) noted for BB 14. See Figure 17 for an example of these artefacts. Some of these peaks are labelled with allele designations which include 17.3, 18.3, 19.3, 20.3 and 24.3 (BB 9 to BB 13). A number of these peaks also coincide with the N+3 artefact position, thus it was difficult to determine if these peaks were either a single peak artefact or an N+3 artefact. For this reason these artefacts were recorded separately and not included together with BB 14. For example, a peak at 17.3 was observed a total of 18 times, however 7 of those the peak also coincided with a N+3 artefact position. Thus it was recorded that 17.3 was observed 11 times and N+3[17.3] was observed 7 times. All artefacts observed at D12S391 have an average and maximum peak height < LOR.

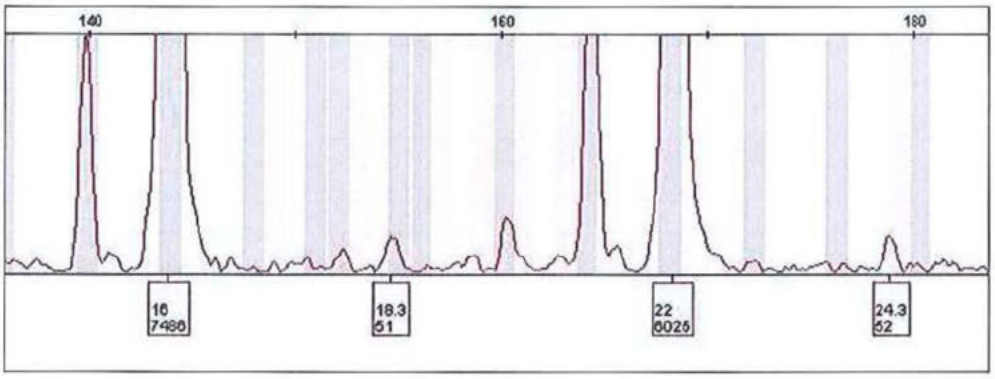

Figure 17 BB 10 and BB 13

A total of 37 artefacts were noted at D19S433; N-2 (31 times); BB 15 (twice) and BB 16 (4 times). BB 15 and BB 16 are all single peaks on the LHS of the main alleles, that are grouped together based on there location (closeness in size bp). BB 15 has both average and maximum peak heights <LOD. BB 16 has and average peak height <LOD and a maximum peak height >LOD.
No artefacts were observed at FGA.

In summary, loci with artefact (excluding N-/+1, N-/+2, N-/+3) peak heights >LOR include: D3S1358; D1S1656; D18S51; D21S11; D8S1179; D19S433. This indicates that it is possible for these artefacts to be labelled routinely. However from these loci, artefacts at D1S1656, D21S11, D8S1179, D19S433 and D3S1358 [10; BB1; ART 1] only have one individual peak that is >LOR. Thus although possible, it is unlikely to see these artefacts labelled routinely. The remaining artefacts at D3S1358 have the following number of peaks >LOR; 8 (2); 9 (3); and 11 (4). Again, due to the low number of peaks >LOR, it is possible, but not likely to see these artefacts labelled routinely. ART 4 at D18S51 has 14 peaks >LOR, from 9 different samples. This artefact appears to be the artefact most likely to be labelled routinely. This artefact is also regularly observed in routine plates on the 3130xl.

6.7.2 Acceptance Criteria – Artefacts

The prevalence of amplification related artefacts (i.e. N-1, N-2, N-3, N+2, N+3, ARTs) was not observed to be higher on the 3500xL than the 3130xl. This is expected for amplification related artefacts given the findings of the Concordance experiment (see Section 6.5), whereby amplification product which has been visualised and sized on the 3130xl is also visualised and sized on the 3500xL.

The prevalence of capillary electrophoresis related artefacts (i.e. blobs, bad baseline) was not observed to be higher for the 3500xL than the 3130xl. It should be noted that capillary electrophoresis related artefacts are not reproducible from run to run on the same instrument, and therefore comparing the prevalence from the 3500xL to the 3130xl using this project's limited sized data set may not be informative.

The 3500xL has passed this experiment as the prevalence of artefacts was observed to be equivalent to the 3130xl and there were no new/unexpected artefacts observed.

6.8 Off Scale Allele Peaks

All samples run on the 3500xL with a flag of OS (over size) determined by GeneMapper ID-X v1.4 were analysed and compared to the profile for the same sample when run on the 3130*xl*. This validation showed all samples that flagged as OS on the 3130xl were still flagged as OS when run on the 3500xL. It was also noted that samples that appeared to be preferentially amplified (PA) on the 3130xl were more likely to be flagged OS on the 3500xL.

Of the eighty-four samples flagged OS on the 3500xL forty-six samples were not flagged OS when run on the 3130xl. Nineteen samples were labelled with the maximum peak height (MPH) flag when run on the 3130xl but flagged OS on the 3500xL. Eleven samples displaying a PA profile but without a MPH or OS flag when run on the 3130xl were flagged OS when run on the 3500xL.

Eight samples with an "OK" profile when run on the 3130xl were flagged OS when run on the 3500xL.

The highest peak height seen across the eighty-four samples on the 3500xL was 35464 RFU. The highest peak height seen across the eighty-four samples on the 3130xl was 9192 RFU.

6.8.1 Acceptance Criteria – Off Scale Allele Peaks

The OS flag is not a sole indicator of profile quality, but identifies samples which may be excess and require manual interpretation by a plate reader. The plate reader then assesses a range of factors including baseline, spectral pull up and peak morphology to make an overall assessment of profile quality. Comparing the number of OS calls between the 3130xl and 3500xL does not take into consideration these additional factors, all of which are instrument dependant and are expected to vary between the 3130xl and 3500xL. Therefore the 3500xL will not be assessed against the assessment criteria for this experiment.

6.9 Repeatability and Reproducibility

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

Repeatability

Figure 18 shows the results of the repeatability testing on the 3500xL. The results show the majority of run to run variation of peak heights range from 12% to -14%. One of the samples showed a percentage change ranging from -22% to -26% with a significant difference (p =

0.000167) 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 run 2.

Figure 19 shows the results of the repeatability testing for the 3130xl. The results show the majority of run to run variation of peak heights range from 3% to -14%. One of the samples showed a percentage change from 11% to 18% with a significant difference ($p = 0.0367$) in peak heights between run 1 and run 2. Peak heights on run 2 were higher than run 1 which could be due to run to run variation. One other sample had a significant difference of ($p = 0.04793$) with the peak heights on run 1 higher than run 2.

There was only minor difference in the performance of the two instruments for the repeatability experiment, however the 3500xL did perform better than the 3130*xl* (i.e. the 3500xL had only one result which was significantly different, whereas the 3130*xl* had two results which were significantly different).

Reproducibility

Figure 20 shows the results of the reproducibility testing on the 3500xL. The results show that when samples are run on day 1 then run again on day 2 by a different operator the variation of peak heights range from 15% to -35%. Seven samples had no significant difference between peak heights and six samples had a significant difference in peak heights.

Figure 21 shows the results of the reproducibility testing on 3130*xl*. The results show when samples are run on day 1 then run again on day 2 the variation of peak heights range from -3% to -44%. Two samples had no significant difference between peak heights and eleven samples had a significant difference in peak heights.

The reduction in peak heights observed in the reproducibility experiment is expected when one preparation of a plate is processed multiple times (in this experiment a total of six times). This reduction in peak height is caused by some of the DNA present in a sample being consumed each time the sample is processed on the 3500xL or 3130*xl*. The objective of the reproducibility experiment was to compare the performance of the 3500xL and the 3130*xl*. The 3500xL performed better in this experiment than the 3130*xl* (i.e. the 3500xL had seven samples which were significantly different, whereas the 3130*xl* had eleven samples which were significantly different). Given the number of times the plate was processed, direct comparison of 3500xL and 3130*xl* results showed that whilst there was a slight decrease in peak heights this was considered to be not unexpected and acceptable.

The results from this experiment can also be used to verify current routine practices whereby a plate preparation is not processed more than three times.

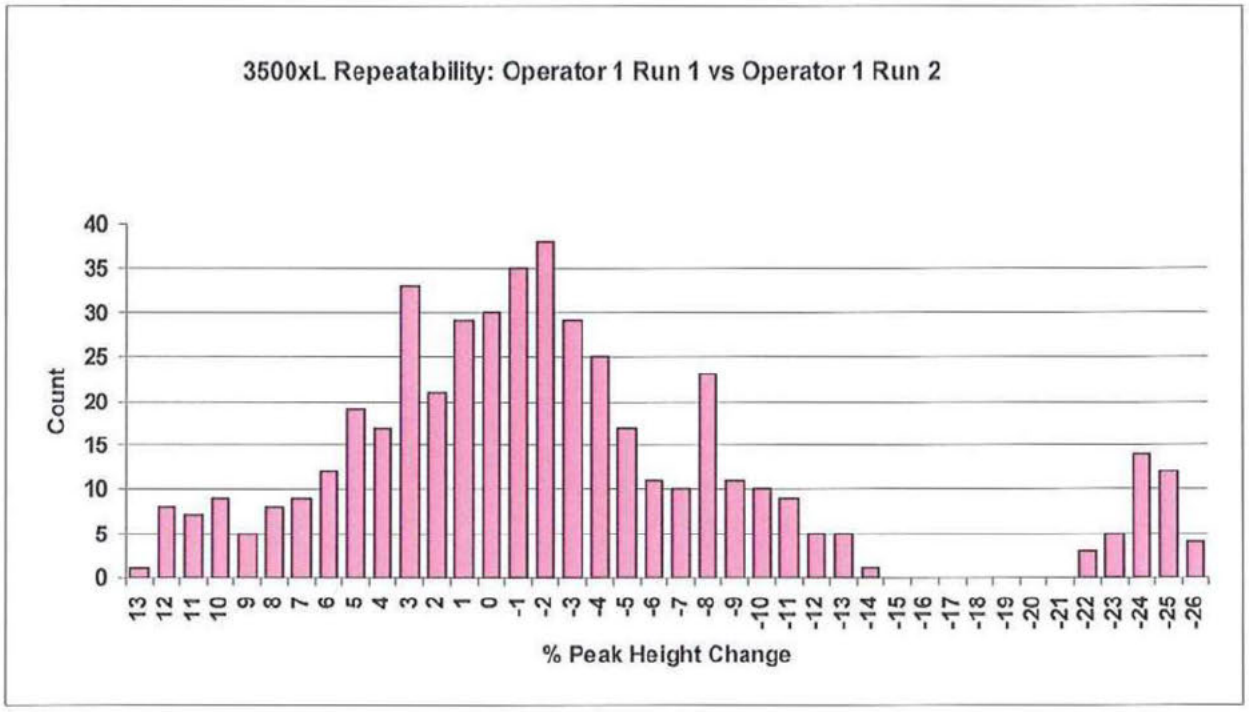

Figure 18 Repeatability for the 3500xL

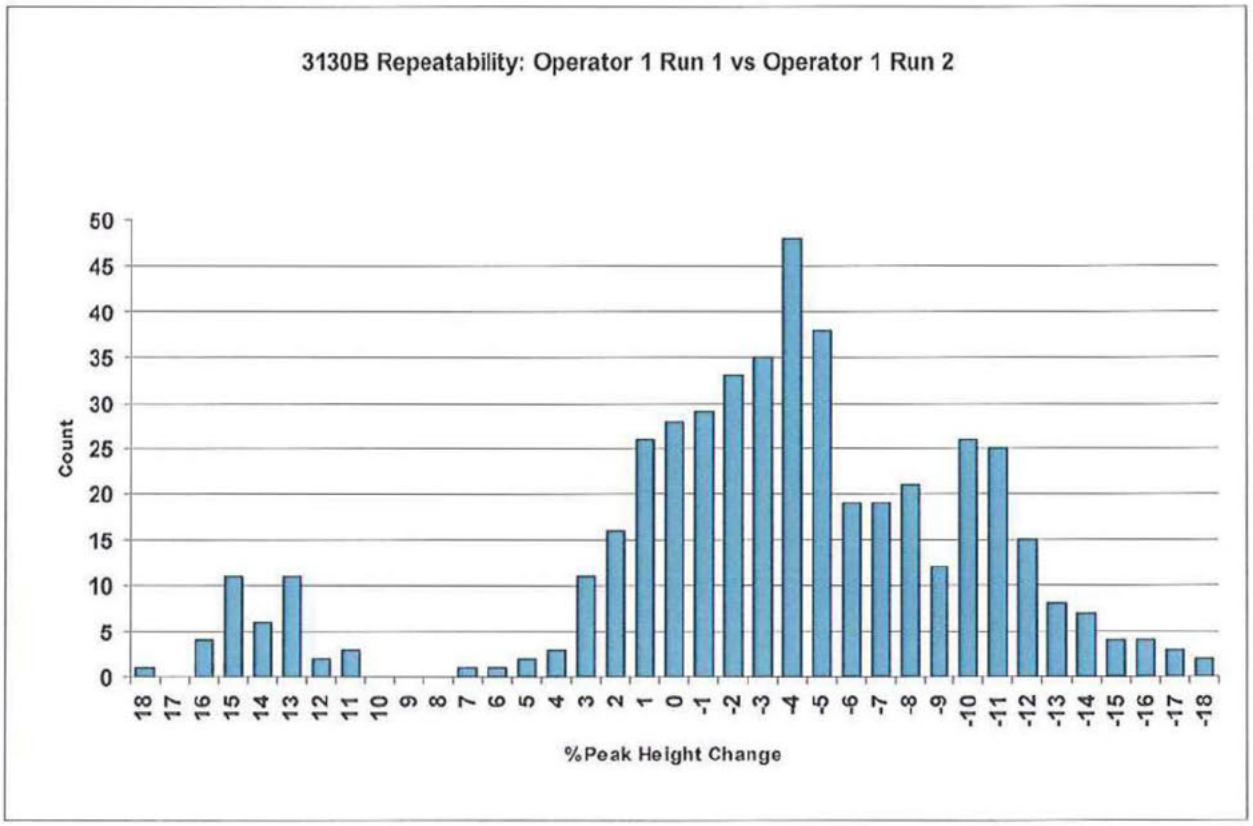

Figure 19 Repeatability for the 3130*xl*

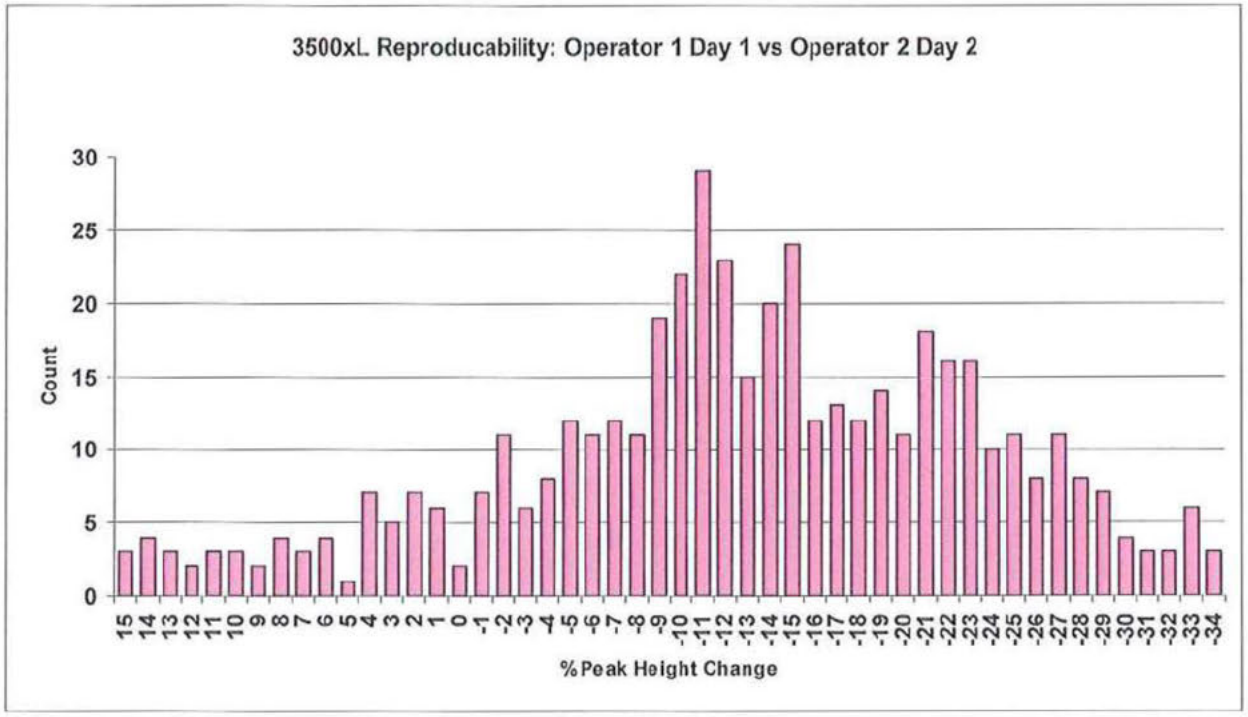

Figure 20 Reproducibility of the 3500xL

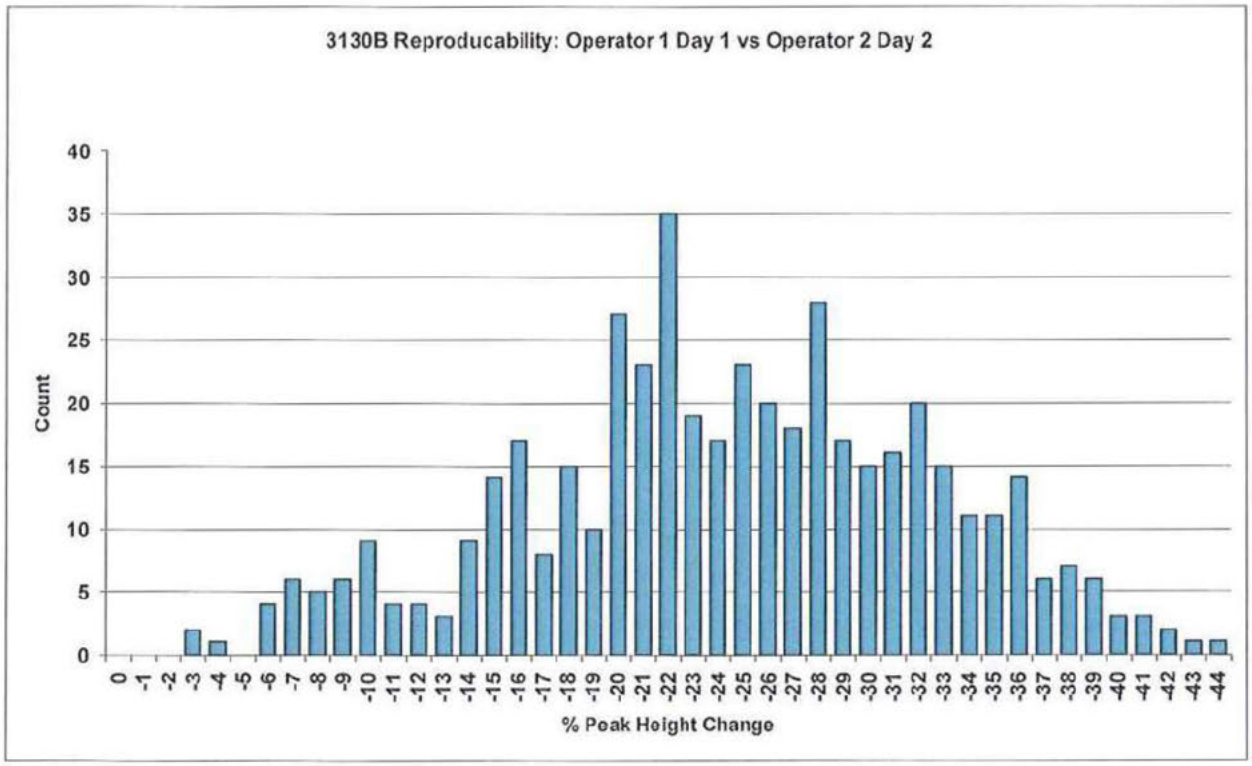

Figure 21 Reproducibility for the 3130*xl*.

6.9.2 Acceptance Criteria – Repeatability and Reproducibility

The 3500xL has passed this experiment because allele designations for repeatability and reproducibility were completely concordant.

6.10 Carry Over and Cross Talk

During the analysis and interpretation of the samples within the direct amplification data set no samples were re-analysed due to capillary electrophoresis carry-over or cross talk.

6.10.1 Acceptance Criteria – Carry Over and Cross Talk

The 3500xL has passed this experiment as no carry over or cross talk was observed.

7. Conclusion

The 3500xL has not been failed on any acceptance criteria defined for this project. Comparisons to the 3130*xl* have shown that the 3500xL performs as well or better than the 3130*xl*. 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. The results from this validation support that the 3500xL capillary electrophoresis instrument is suitable to use for the analysis of reference sample processed through direct amplification using PowerPlex®21.

8. Recommendations

1. Implement the use of the 3500xL for the analysis of reference samples processed through direct amplification.
2. Set LOD to 50 RFU
3. Set LOR to 175 RFU
4. Set AI to 55%
5. Set homozygote threshold to 650 RFU
6. Set stutter thresholds as per table 8
7. Conduct a post implementation review to determine if thresholds set in this report are appropriate when more 3500xL data is available. The post implementation review should also include determination of thresholds which were not possible in this project due to sample size (i.e. some -2 and +1 repeat stutter thresholds and N-2 and N+2 artefact thresholds).

Abbreviations

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

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.