Review of DNA analysis undertaken in the Blackburn case

23 November 2022

PREPARED FOR: **The Commission of Inquiry into Forensic DNA Testing in Queensland**

AUTHOR: **Ms Johanna Veth, Senior Scientist, Forensic Biology, Institute of Environmental Science and Research Limited (ESR)**

REVIEWED BY: **Dr Bruce Budowle**

DISCLAIMER AND DECLARATION OF CONFLICTS OF INTEREST

The Institute of Environmental Science and Research Limited (ESR) owns STRmix Limited (a fully owned subsidiary company) which licenses STRmix™ software to laboratories around the world, including the forensic DNA Laboratory at Queensland Health Forensic and Scientific Services.

The author has used all reasonable endeavours to ensure that the information contained in this client report is accurate. However, the author and ESR do not give any express or implied warranty as to the completeness of the information contained in this client report or that it will be suitable for any purposes other than those specifically contemplated during the Project or agreed in writing by ESR and the Client. The conclusions in this report are opinions based on the information provided to the author at a point in time, and may change if additional information is provided.

Contents

1. EXECUTIVE SUMMARY 1

2. Introduction 3

3. Review of the Blackburn casefile 4

3.1 TESTING FOR BLOOD BY QPS STAFF 4

3.2 POOR DNA RECOVERY FROM SAMPLES DESCRIBED AS BLOODSTAINED 7

3.3 THE INTERPRETATION OF MIXED DNA PROFILES 11

3.4 INTERPRETABLE PROFILES DESCRIBED AS 'COMPLEX' 16

3.5 SAMPLES REPORTED AS 'DNA INSUFFICIENT' OR 'NO DNA DETECTED' 17

4. Review of OQI34043 20

4.1 TIMELINE 20

4.2 WAS THE RESPONSE ADEQUATE? 22

4.3 COULD THE PROBLEM HAVE BEEN DETECTED EARLIER? 22

4.4 WAS THE PROBLEM CONTAINED ONCE IT WAS IDENTIFIED? 23

4.5 WAS A ROOT CAUSE DETERMINED? 23

4.6 WHAT CHANGES DID THE LABORATORY IMPLEMENT? 24

4.7 IMPACT ON THE BLACKBURN CASE 24

5. Review of OQI34817 26

5.1 TIMELINE 26

5.2 WAS THE RESPONSE ADEQUATE IN RELATION TO CASEWORK? 27

5.3 IMPACT ON BLACKBURN CASE 27

5.4 WHAT WERE THE IMPLICATIONS FOR THE VALIDATION OF POWERPLEX® 21 AND STRMIX™ 28

5.5 WAS THE RESPONSE ADEQUATE FOR THE AT AND DROP-IN CAP? 30

5.6 ADDITIONAL CONSIDERATIONS: CONTAMINATION AND CARRYOVER 33

6. Opportunities for retesting 35

1. EXECUTIVE SUMMARY

1. During the processing of the Blackburn case samples the Queensland Health Forensic and Scientific Services (QHFSS) laboratory staff were dealing with substantial changes to operational processes which include implementation of a new multiplex short tandem repeat (STR) kit called PowerPlex® 21 and a new way of mixture interpretation relying on probabilistic genotyping with software known as STRmix™.

2. Reporting results as soon as possible remained a priority despite there being significant issues which include:

   a. *At least* 2 large quality investigations

   b. An apparent increase in drop-in, contamination and ongoing capillary electrophoresis (CE) carryover concerns (one type of contamination)

   c. Difficulties adjusting to PowerPlex® 21 and incorporating STRmix™ into the workflow.

3. There is evidence to suggest that the extractions carried out on the MultiProbe® II platform were not recovering DNA optimally. Several Blackburn case samples were processed using this method. It is strongly recommended that the laboratory investigate this further as there are implications for other cases where samples were processed using this method.

4. The poor profiling results obtained from Ms Blackburn's bloodstained shirt may be related to the shirt itself rather than, as has been proposed, a result of a defective batch of Proteinase K.

5. The mis-characterisation by Queensland Police Service (QPS) staff of bloodstaining for some samples in the Blackburn case resulted in possibly unwarranted criticism when QHFSS failed to detect DNA in the samples.

6. Misguided policies related to interpretation of DNA profiles or designed to avoid generating potentially complex DNA profiles resulted in:

   a. Very low-level results in mixed DNA profiles being used for comparison purposes

   b. Apparent single source but incomplete (but interpretable) profiles not being interpreted

   c. Samples stopped after quantification and reported as "No DNA detected"

7. Opportunities to reassess drop-in and adjust settings were not taken. This issue is concerning given the apparent increase in drop-in as seen in the reassessment of the baseline analysis of the 3130xl B CE instrument running with a 5-second injection time and from the laboratory's own monitoring of drop-in and contamination.

8. Ongoing issues with CE carryover[1] were reported in various documents. This issue should have been resolvable. In conjunction with the increased drop-in and contamination events, the use of low-level results for comparison purposes was problematic.

9. Problems that indicate the laboratory was not focussed on quality assurance. These include:

[1] Carryover is defined as the physical transfer of DNA from one capillary injection into the next injection into the same capillary. In most cases, if the CE instrument is functioning correctly, carryover occurs below the background noise of the instrument and is therefore not detectable.

a. Ongoing contamination from different sources (CE carryover, drop-in and gross contamination events)

b. DNA recovery using the MultiProbe® II extraction platform seemed to be lower than when using the Maxwell® platform.

c. An incorrect injection time on a CE instrument was not detected for several months.

d. Key information is missing from the opportunity for quality improvement (OQI) documentation including some of the troubleshooting undertaken. Some requisite investigative steps do not appear to have been undertaken at all.

e. A new reagent was validated and subsequently implemented into casework without laboratory staff being aware the reagent had already expired or what the correct storage conditions were.

f. The laboratory used a defective reagent despite it having been quarantined as a result of a quality investigation.

g. New lots of critical reagents were used in laboratory processes without first undergoing quality checks.

h. Inadequate disclosure of quality incidents that had affected the outcome of DNA profiling.

10. Some documentation was missing from the Blackburn casefile which made review of the laboratory processing and interpretation of profiling results challenging. Where required for this review, the missing information was requested via Commission staff and subsequently supplied by the laboratory. If the contents of the Blackburn file are indicative of QHFSS casefiles in general then the missing documentation makes casefile review difficult, whether it be by another scientist, defence analyst or other recipient. The missing information includes but is not limited to:

a. Lack of detail in relation to what samples were affected by quality incidents and any decision-making as to rework and reporting of results

b. Batch and batch quality information such as the performance of positive and negative controls

c. Quantification results and quality flags for case samples and extraction controls.

11. Large cases like the Blackburn case require more oversight than is possible by the method of reporting results electronically as soon as they become available.

2. INTRODUCTION

12. This review of key documents has been conducted as part of the Commission of Inquiry into Forensic DNA testing undertaken by the Queensland Health Forensic and Scientific Services (QHFSS). A large number of documents, which include the DNA profiling casefile for the Shandee Blackburn homicide investigation, reports, emails, spreadsheets and meeting minutes, have been made available. These documents, which for the most part span a period between mid-2012 and late 2014, relate to the following:

a. Documents related to two quality investigations

   i. OQI34043 Positive Extraction Controls with low DNA yields

   ii. OQI34817 Incorrect conditions used for Capillary Electrophoresis.

b. A review of the Blackburn DNA Analysis prepared by Dr Kirsty Wright

c. Documentation related to the validation and implementation of the following:

   i. Capillary Electrophoresis 3130 instruments

   ii. PowerPlex® 21 DNA multiplex kit

   iii. STRmix v1.05 and v2.0 DNA interpretation software.

13. The Commission of Inquiry requested that these matters be addressed:

a. Whether the work done by QHFSS in the Blackburn case was done in accordance with best practice

b. Whether the way in which the matters referred to in the two OQI documents were dealt with, or resolved, was in accordance with best practice

c. Whether the matters referred to in the two OQI documents had, or could have had, any effect on the analyses performed by QHFSS in the Blackburn matter

d. The soundness of the opinions of Dr Wright

e. Identify any opportunities for retesting of samples or exhibits in the Blackburn matter

f. Whether the validation of the Capillary Electrophoresis 3130 instruments, PowerPlex® 21 or STRmix were performed in accordance with best practice and, if not, whether any failures had, or could have had, an effect upon the results of DNA testing in the Blackburn case.

14. The first 5 items are addressed in this report. The 6th item, the review of the validations, is not included in this report although any particular issues affecting the Blackburn case will be addressed in this document. To avoid repetition, Dr Wright's opinions are referred to, where applicable, directly in the sections relating to the review of the Blackburn casefile and the two OQIs.

15. The conclusions reached in this report are based solely on the documentation made available to the author at the time of this report. If further information is made available, the conclusions may need to be revised.

3. REVIEW OF THE BLACKBURN CASEFILE

16. In the review of the Blackburn casefile and associated documentation there is evidence of a busy laboratory struggling to manage a great deal of change to their processes. In December 2012 the laboratory had transitioned to the PowerPlex® 21 DNA profiling kit and adopted a new way of interpreting mixed DNA profiles and calculating likelihood ratios using STRmix™ v1.05.

17. During the processing of samples from the Blackburn case in early 2013 the laboratory was also dealing with significant quality issues which include, but perhaps are not limited to, a poorly performing batch of the Proteinase K reagent, a malfunctioning industrial dishwasher, an incorrect injection setting on one of the 3130xl capillary electrophoresis instruments, an increase in detectable drop-in and instances of gross contamination, and an ongoing and seemingly unresolved issue of carryover in one or both capillary electrophoresis instruments. Not all of these issues directly affected the laboratory processing or interpretation of profiling results in the Blackburn case but that they were occurring at the same time as the Blackburn case sheds some light on the difficulties the laboratory staff members were facing.

18. While the focus of this report is reviewing the analysis and reporting undertaken in the Blackburn case, some of the issues raised are likely systemic in nature, occurring as a result of policies that are at odds with the underlying science, a culture that prioritises turnaround times over quality assurance, and the structural relationship between QPS and QHFSS.

3.1 TESTING FOR BLOOD BY QPS STAFF

19. The laboratory has attracted much public criticism for failing to detect DNA in a number of samples taken from a vehicle which were described as bloodstained. The examination of the vehicle, the testing for blood, and the subsequent sampling were undertaken by QPS staff. Therefore, the following comments are more applicable to the bioscreening (testing for blood) undertaken by QPS, rather than the scientific work undertaken by the QHFSS DNA laboratory.

20. Twelve samples taken from vehicle 706LHN were described as swabs of blood when they were submitted to QHFSS for DNA profiling. QHFSS reported that DNA was not detected in any of these samples, and they were not processed further by QHFSS as per the workflow in place at the laboratory at the time.

21. The Blackburn casefile (pages 2157-2168) contains screenshots from AUSLAB of the bioscreening notes of these samples. The table below summarises the bioscreening notes entered into AUSLAB.

Sample	Description of staining	Combur result	Luminol result	Polilight result
V14 – handbrake well	"potentially transfer stain (non-visible)"	"Combur neg (very slow)"	neg	neg
V15 – clutch pedal	"potentially transfer stain (non-visible)"	"Combur neg (very slow)"	neg	neg
V16 – brake pedal	"potentially transfer stain (non-visible)"	"Combur neg (very slow)"	neg	neg
V17 – accelerator pedal	"potentially transfer stain (non-visible)"	"Combur neg (very slow)"	neg	neg
V31 – rear interior driver's side door handle	"potentially transfer stain (non-visible)"	"Combur pos (very slow)"	neg	neg
V32 - rear interior driver's side window wind	"potentially transfer stain (non-visible)"	"Combur pos (very slow)"	neg	neg
V33 - rear interior driver's side handle to door	"potentially transfer stain (non-visible)"	"Combur pos (very slow)"	neg	neg
V34 - rear interior driver's side door trim	"potentially transfer stain"	"Combur pos (very slow)"	neg	neg
V48 – steering wheel	"potentially transfer stain (non-visible)"	neg	pos	pos
V49 – ignition	"potentially transfer stain (non-visible)"	neg	pos	neg
V50 – rear of driver's seatback	"potentially transfer stain (non-visible)"	neg	pos	pos
V51 – front passenger footwell	"potentially transfer stain (non-visible)"	neg	pos	neg

Table 1: Results of bioscreening for vehicle samples as captured in AUSLAB

22. The use of the phrase "potentially transfer stain (non-visible)" is particularly confusing – if the stain is non-visible how can it be described as "potentially transfer stain"? It also is unclear why some of the "very slow" Combur results are described as "pos" in the AUSLAB notes whereas others are described as "neg". The examination and bioscreening of the vehicle were undertaken by a QPS Scientific Officer. In the handwritten examination notes the Scientific Officer records that all of the above samples are Combur negative[2]. The Scientific Officer also stated:

    "In my opinion slow positive could be result of use over last few weeks or false positive to grime and dust."[3]

23. Forensic laboratories and crime scene analysts that undertake bioscreening for blood generally have some criteria to categorise the results obtained documented in their standard operating procedures (SOP).

24. For instance, a laboratory may adopt criteria that depend on the visual appearance of the staining in combination with the results of a presumptive test. Such criteria may be that in

[2] Pages 209-224 of Document QPS.0001.0065.0039
[3] Pages 208-209 of Document QPS.0001.0065.0039

order for a sample to be described as potentially bloodstained the staining must first have the visual appearance of bloodstaining and provide a positive result from a presumptive test. An example of this can be found on page 38 of the witness statement prepared by the QPS Scientific Officer[4], where the section titled "Identification of Blood" includes the following text:

"…*the reference to 'blood' in this statement should be interpreted as 'a substance that has the appearance of blood, is presumptive positive for blood and in the context of the scene reliably is blood'."*

25. A very slow or weak Combur reaction from a non-visible stain may be reported as possible blood. Finally, a luminol positive result may be described as probable or possible blood depending on the intensity of the reaction, or may simply be described as a luminol positive stain. Other laboratories may differ in how they categorise their bioscreening results from non-visible stains. Moreover, it is standard to include a caveat explaining that the presumptive tests used cannot confirm the presence of blood as other substances can provide a positive result in these tests.

26. Presumptive tests for blood are useful as they act as screening tests, identifying areas that may contain human DNA for further processing. There is always uncertainty as to what the outcome of DNA processing might be. While the presumptive tests for blood do not detect if *human* blood is present or if any degradation of the sample has occurred, the criteria for reporting presumptive test results are important as they provide some expectations as to the potential success of the DNA profiling undertaken. For example, good quality DNA profiling results are generally expected from non-degraded bloodstains that are human in origin, whereas samples described as "possible bloodstains" may not be biological in origin at all.

27. Given that the above samples from the vehicle apparently did not have the visual appearance of bloodstaining and the results obtained from the presumptive testing were somewhat inconclusive, that the samples were described as "blood swabs" was misleading and not in accordance with the text in the Scientific Officer's statement.

28. There were records of communications[6] dated 23 January 2014 and 11 February 2014 between the reporting scientist and QPS staff inquiring about the "no DNA detected" findings for the vehicle swabs but there was no discussion about the actual findings of the bioscreening undertaken. On 23 January 2014 the reporting scientist explained that a positive Combur result could be generated by rust or non-human blood, which is correct. However, this appears to have been a more general discussion about Combur testing, rather than based on a review of the actual bioscreening results from the testing undertaken in the vehicle.

29. Around 2017 AUSLAB screenshots of the bioscreening were added to the casefile. On 7 February 2017 there were communications with an investigating officer in which the reporting scientist explained that based on the presumptive testing and the failure to detect DNA, the samples were likely to not be biological in origin, and the luminol positive results could be a reaction to metal salts. In an addendum statement[7] issued on 23 February 2017 the reporting scientist explained that the presumptive testing was undertaken by QPS staff and that he could not provide information about this testing, which is entirely appropriate.

30. In her report[9], Dr Wright expressed some concern that no DNA was detected in the 12 samples from the vehicle that were described as bloodstained. It is understood that Dr Wright did not have access to the original QPS examination and bioscreening notes for

[4] Page 36 of Document QPS.0001.0065.0039
[6] Page 1931 of the Blackburn casefile
[7] Page 2044 of the Blackburn casefile
[9] Section 2.2 of Dr Wright's Review of Blackburn DNA Analysis report

these samples when she prepared her report. The descriptions of the visual appearances and the results from the presumptive testing may not be indicative of the presence of bloodstaining from which one would expect to obtain good quality results.

31. The above samples should not have been described as bloodstained given they did not have the visual appearance of blood. However, if QHFSS staff understood these samples to be bloodstained the absence of DNA in these samples should have been investigated further at the time the samples were quantitated in 2013. There is an expectation that DNA would be detected in bloodstained samples if the blood is human in origin and assuming no significant environmental insults. If no DNA is detected in bloodstains, it would be normal practice to consider possible explanations – at the very least there ought to have been a discussion about whether the blood could be non-human in origin and this inquiry would likely have revealed the true nature of the bioscreening results. Had QHFSS personnel demonstrated some curiosity about the failure to DNA in these samples when they were processed in 2013 or during communications with QPS staff in early 2014 they may have prevented some of the public criticism that later ensued in relation to these samples or may have been better prepared to address criticisms in this regard.

32. In summary, there are two primary areas of concern. The first is the way in which QPS staff interpreted the results of bioscreening for blood. This interpretation matters because if a sample is described as bloodstained then there is an expectation as to the potential success of the DNA profiling results that will be obtained assuming the blood is human in origin and not degraded. Because these samples were widely assumed to be bloodstained, criticism that the laboratory processes failed to obtain DNA profiling results from these samples may be unwarranted.

33. The second area of concern relates to the lack of oversight of the case overall. In order to report results to QPS as quickly as possible, the profiling results of each sample are reported electronically as soon they become available. This system may work well for small cases where there are only a few samples. However, larger cases require more oversight; preferably a case manager who reviews all the results as they become available and in the context of the case as a whole, and who is responsible for the reporting of the case. Prior to the issuing of reports there should be a full technical review of the laboratory work and the interpretation of the results undertaken by a qualified reporting scientist and a further administrative review of the report. If these steps had occurred in the Blackburn case then the lack of DNA recovery from the vehicle samples described as bloodstained and the poor DNA profiling results obtained from Ms Blackburn's shirt could have been investigated and addressed in a timely fashion. Subsequent external criticisms may have been avoidable.

3.2 POOR DNA RECOVERY FROM SAMPLES DESCRIBED AS BLOODSTAINED

34. There are further samples in the case which were described as bloodstained and from which either the profiling results were not as expected from bloodstains or which were reported as "no DNA detected". These include the 'S' series of samples from gutters (for example S14 and S15), 3 samples from a white 'Effekt' T-shirt (ML series) and 4 samples from a knife (reference FBC 585528200).

35. Information provided suggests the 'S' series of samples were taken several hours[10] after Ms Blackburn's body was located and it is possible that if human blood associated with her death was present then it may have started to degrade, particularly if it was a hot and sunny

[10] Time stamps of photographs of S14 and S16 areas of bloodstaining indicates the samples were taken after 8.20PM on 9 February 2013. See document: In-situ photographs of samples S14 and S16.pdf. It has been assumed that the S15 sample was taken in the same timeframe as S14 and S16.

day. S14 was reported as “No DNA detected” and poor results consistent with originating from Ms Blackburn were obtained from S15. In contrast, a swab described as “A - Heavy wet and congealing - Blood stain pattern on footpath - Boddington Street, Mackay" was collected much earlier in the day[11]. Very good DNA profiling results, consistent with originating from Ms Blackburn, were obtained from this sample.

36. The examination notes for the T-shirt described Combur positive results from “dirty stains”[12] although in the QPS Scientific Officer’s formal witness statement the results of the Combur testing are described as “slow weak”[13]. As other substances such as bleach, rust, plant material and paint can cause a positive reaction in a Combur test it is possible that the dirty staining observed on this T-shirt was a substance other than blood. No DNA was detected in the 3 “blood soaked fabric” samples that were collected from this T-shirt.

37. The examination notes for the 4 knife samples have not been provided. However, as 3 of the sample names contained the terms “blood/rust” it is possible that blood may not be present at all in these samples. All 4 of these knife samples were reported as “no DNA detected.”

38. There are also a number of bloodstained samples taken from Ms Blackburn’s shirt from which low level DNA profiling results, corresponding to Ms Blackburn’s reference DNA profile, were obtained. For these samples it is reasonable to assume that they were bloodstained, given the nature of Ms Blackburn’s injuries. Some of these samples were processed in March 2013 when a sub-optimal batch of Proteinase K was in use in the laboratory. However, a review of the associated extraction batch paperwork indicates that the faulty Proteinase K was not used in the processing of these samples. Furthermore, lower than expected results were also obtained from a second set of bloodstained samples from the same shirt, which appear to have been tested more than a year later in May 2014.

39. One plausible explanation is that whatever was causing the inhibition of the DNA profiling reaction could be related to the shirt fabric. It is possible that there was something inherent to the shirt that caused the poor performance from the bloodstains, compared to the relatively good results obtained from the trace samples from the same shirt. DNA extractions can be inhibited by the presence of substances such as dyes. A key difference between the bloodstains and the trace samples was that the bloodstained samples where pieces of fabric taken from the shirt whereas the trace samples were tape lifts. The different sampling strategies could result in dyes in the fabric would have been present in the bloodstain extractions but not in the tapelift extractions, which may be an explanation for the poor performance of the bloodstained samples.

40. Could bacteria have been the cause of the poor DNA recovery in samples from the Blackburn case? This question was raised by QPS[17] specifically in relation to a trace of foreign DNA under Ms Blackburn’s lefthand fingernail samples but it is worth considering for other samples in the case. As Ms Blackburn’s body was located and transported to the mortuary very quickly after her death environmental factors at the scene are less likely to support bacterial growth. No information as to the packaging and storage of the shirt from between it being recovered from Ms Blackburn’s body and its examination by QPS staff has been provided, but if the shirt had been packaged in plastic while the bloodstaining was still wet then conceivably bacteria or mould could have developed. It should be noted however, that the QPS forensic biologist who examined the shirt stated that when she received the

[11] Time stamp of photograph area ‘A’ of bloodstaining indicate the sample was taken around 5:07AM on 9 February 2013. See document: In-situ photograph of sample 585592064.pdf.

[12] Page 184 of Document QPS.0001.0065.0039

[13] Page 28 of the QPS Scientific Officer’s formal written statement dated 03 December 2014

[17] Document QPS.0001.0607.0461, page 461 from operation Lima Zimzala Running Log and Document 1982-109777 FSS.0001.0082.1777

shirt for examination on 19 April 2013 it was packaged correctly in a paper bag[18], not plastic. It seems unlikely that the shirt would have been packaged in plastic at the mortuary and they transferred to paper packaging later.

41. Another possible explanation for the poor recovery of DNA from some of the Blackburn samples was discovered through a review of the extraction positive control quantification data[19]. Further investigation of the positive control data revealed an anomaly between the quantification results obtained from extractions undertaken on the MultiProbe® II platform compared to those obtained from batches processed on the Maxwell® platform[20]. The MultiProbe® II platform extraction positive controls had much lower quantification results than the Maxwell® extracted positive controls. This difference suggests that DNA was not being recovered effectively from the MultiProbe® II batches. Due to time constraints, this matter could not be investigated further and there may be other reasons for the lower positive control quantification results. However, this is a compelling indication that there was something about the the MultiProbe® II extraction method that was resulting in a lower recovery of DNA when compared to the Maxwell® method.

42. This finding is significant for the Blackburn case as several samples were extracted in the MultiProbe® II batches[21], including the first set of bloodstained samples from Ms Blackburn's shirt, some of the bloodstained "S" series samples from the scene, the vehicle samples that were described as bloodstained and the ML series from the white 'Effekt' T-shirt. In comparison, trace samples processed using the Maxwell® platform generally produced quite good results, demonstrating effective DNA recovery.

43. The following table lists the samples in the Blackburn case that were processed using the MultiProbe® II extraction method.

---

[18] Document DPP.0003.001.2256 at paragraph 4.

[19] We thank Dr Wright for providing an analysis of several months of extraction positive control quantification data.

[20] Spreadsheet QP1300165446_all_samples_Ext_Pos_Quant, AUSLAB CS samples tab

[21] Document 1982-1438 which contains a list of Blackburn samples and associated batches

Samples processed in MultiProbe® II extraction batches

Barcode	Description
320117416	PM Samples - from deceased Shandee BLACKBURN
320117425	L hand f/nail clippings (320117381)
320117390	R hand f/nail clippings
572984230	Item F10 - a trace DNA swab - collected from the rear left side of the mobile phone (item F)
572984202	Item F5 - a trace DNA swab - collected from bottom front of the mobile phone (item F)
572984197	Item F6 - a trace DNA swab - collected from front upper and sides of the mobile phone (item F)
572984241	Item F9 - a trace DNA swab - collected from the rear right side of the mobile phone (item F)
572439163	Item L12a - bloodstained fabric - ~5mmx5mm excised from distal back left sleeve of shirt
572439181	Item L13a - bloodstained fabric - ~5mmx5mm excised from proximal back left sleeve of shirt
572439205	Item L14a - bloodstained fabric - ~5mmx5mm excised from back right armpit area of shirt
572439024	Item L1a - bloodstained fabric - ~5mmx5mm excised from front right upper chest area of shirt
572984114	Item L24: Cutting of Blood Soaked Fabric - from LHS rear of knee of pants in 585580649
572439041	Item L2a - bloodstained fabric - ~5mmx5mm excised from front right button hole area of shirt
572439068	Item L3a - bloodstained fabric - ~5mmx5mm excised from front RHS below collar area of shirt
572984461	Item L48: Swab of Blood - from front upper sole of RHS shoe in 585580626
572439085	Item L4a - bloodstained fabric - ~5mmx5mm excised from front LHS adjacent to buttons of shirt
572984493	Item L51: Swab of Blood - from front upper label of LHS shoe in 585580626
572984504	Item L52: Swab of Blood - from medial arch area on lower sole of LHS shoe in 585580626
572439109	Item L5a - bloodstained fabric - ~5mmx5mm excised from front LHS lower chest area of shirt
572439127	Item L6a - bloodstained fabric - ~5mmx5mm excised from front LHS chest area of shirt
572439145	Item L9a - bloodstained fabric - ~5mmx5mm excised from front left sleeve of shirt
585528094	Item ML2: Cutting of Blood Soaked Fabric - from rear of T-Shirt contained in 585580812
585528112	Item ML4: Cutting of Blood Soaked Fabric - from front of T-Shirt contained in 585580812
585528123	Item ML5: Cutting of Blood Soaked Fabric - from rear of T-Shirt contained in 585580812
572572940	Item S14: Swab of Blood - from gutter on Boddington Street
572572956	Item S15: Swab of Blood - from gutter on Boddington Street just west of Item S14
572572967	Item S16: Swab of Blood - from upper gutter verge on Boddington Street even further west of Item S14
572573279	Item V14: Swab of Blood - from handbrake well of Qld Rego 706LHN
572573285	Item V15: Swab of Blood - from clutch pedal of Qld Rego 706LHN
572573296	Item V16: Swab of Blood - from brake pedal of Qld Rego 706LHN
572573309	Item V17: Swab of Blood - from accelerator pedal of Qld Rego 706LHN
572573448	Item V31: Swab of Blood - from rear interior driver's side door handle of Qld Rego 706LHN
572573452	Item V32: Swab of Blood - from rear interior driver's side window wind of Qld Rego 706LHN
572573461	Item V33: Swab of Blood - from rear interior driver's side handle to door of Qld Rego 706LHN
572573470	Item V34: Swab of Blood - from rear interior driver's side door trim of Qld Rego 706LHN
572573612	Item V48: Swab of Blood - from steering wheel of Qld Rego 706LHN
572573623	Item V49: Swab of Blood - from ignition of Qld Rego 706LHN
572573634	Item V50: Swab of Blood - from rear of driver's seatback of Qld Rego 706LHN
572573640	Item V51: Swab of Blood - from front passenger's side footwell of Qld Rego 706LHN
601775354	Swab from combur +ive area - collected from the left side of the blade (FBC 601775336)
601775345	Swab from combur +ive area - collected from the right side of the blade (FBC 601775336)
585528222	Trace DNA swab - collected from the missing stud hole of the knife handle (FBC 585528200 F765052)

585528361	Blood swab - collected from the right side of the blade and hilt of the knife.
585528255	blood/rust swab - collected from the knife hilt (FBC 585528200 F765052)
585528233	blood/rust swab - collected from the hole at the end of the knife handle (FBC 585528200 F765052)
585528244	blood/rust swab - collected from the top edge of the knife blade (FBC 585528200 F765052)

Table 2: Samples processed in MultiProbe® II batches

44. If the MultiProbe® II extraction method was performing sub-optimally then this is particularly problematic for samples that likely had low DNA template to begin with. These include the samples from Ms Blackburn's fingernails, the samples from the vehicle that had been described as bloodstained and the samples from a knife. Furthermore, there are obvious consequences for samples from other cases that were processed in these batches.

45. It is strongly recommended that the laboratory review the results from their MultiProbe® II extraction batches to determine if there indeed was a problem with this method or with something specific to the way these particular extractions were conducted, resulting in poor DNA recovery[22]. Investigations are required to determine how many batches were affected and how this has affected the reporting of results from those batches.

46. In this section a number of explanations have been proposed to explain why there was poor DNA recovery from several samples in the Blackburn case. These include factors such as degradation (specifically in relation to the S series of samples), inhibition (in relation to the bloodstained samples from Ms Blackburn's shirt) and indications that the MultiProbe® II extraction method was performing sub-optimally when compared to the Maxwell® method. It is possible that a combination of these factors were in play. It is imperative that the performance of the MultiProbe® II extractions is investigated further given the wider implications.

3.3 THE INTERPRETATION OF MIXED DNA PROFILES

47. There are several samples taken from Ms Blackburn's clothing where the majority of the recovered DNA originated from Ms Blackburn herself but also present were very low-level DNA profiling results from at least one other person. Despite these additional results being very few in number (in some instances there is only one additional low-level result) QHFSS had a policy of interpreting these mixtures through STRmix™, and then comparing the results of all the components of the STRmix™ deconvolution, major and (trace) minor, to <u>all</u> of the reference DNA profiles using the STRmix™ database search functionality. QHFSS then reported every likelihood ratio (LR) that was generated in Intelligence Reports.

48. Before continuing, it is useful to explain the purpose of the STRmix™ database search functionality and a key difference between the LRs calculated by this functionality, as reported in the QHFSS Intelligence Reports, and the evidentiary LRs subsequently reported in formal witness statements. The database search function can be used to provide investigative information by allowing a deconvoluted DNA profile to be compared to a database of reference samples or staff profiles if contamination is suspected. The LRs for any matches generated by the database search function are intended for investigative purposes only as they do not include corrections for sub-population or sampling effects,

[22] A number of factors must be explored. For example, was the method itself sub-optimal? Was there an issue with a particular reagent that affected a limited number of batches processed during this time period? Could the issue be localised to a particular technician?

unlike the evidentiary LRs calculated for formal witness statements. For evidentiary LRs QHFSS also utilise a verbal scale in order to express in words the level of support the DNA evidence has for a particular proposition.

49. LRs calculated for a full set of matching PowerPlex® 21 results are usually in excess of one hundred billion. For trace minor components in a mixture, LRs will converge to 1, where 1 is uninformative or neutral. In some instances, the LR supports inclusion and in others it supports exclusion. The earlier Intelligence Reports issued in the Blackburn case reported all inclusionary and exclusionary LRs, but later reports only listed the inclusionary LRs, along with a corresponding level of support, which for LRs close to 1 is defined by QHFSS as "low support". This verbal equivalent differs from the verbal scale QHFSS use for evidentiary LRs and no explanation or scale was provided in Intelligence Reports.

50. If very few additional low-level DNA profiling results are present in a mixed DNA profile, several people may be reported as possible sources of the DNA, each with an LR close to 1. This scenario is the case for sample L45 which is described as a trace tapelift from the rear LHS upper leg area of pants. Sample L45 is a useful example because the reported LRs, one in particular reported in relation to Mr Daniel, attracted much attention. However, several samples in the Blackburn case generated profiles with similar mixture proportions where the majority of the DNA originated from Ms Blackburn and very few additional trace results indicated the presence of DNA from at least one other person.

51. The majority of the DNA recovered from L45[23] is consistent with coming from Ms Blackburn, which is not unexpected as the sample was taken from her clothing. There were two additional results that were very low level compared to the results from Ms Blackburn, and it was not possible from the testing undertaken to determine whether this low-level DNA had originated from a male or a female. Furthermore, one cannot have confidence that the low-level DNA had originated from just one person. With no replication of these low-level results, one cannot even be sure if either of these two results were actually DNA inherent to the sample, or an instance of drop-in or carryover, or some other artefact.

52. For the STRmix™ interpretation it was assumed that the total number of contributors to this mixture was 2 including Ms Blackburn. Based on these assumptions the STRmix™ deconvolution[24] estimated that the major contributor (Ms Blackburn) had contributed 96% of the DNA to the mixture with the remaining 4% originating from the minor contributor.

53. When the STRmix™ deconvolution results from sample L45 were compared to the 76 reference profiles, there were 6 inclusionary LRs. As expected, 1 of these LRs was in relation to Ms Blackburn and was greater than one hundred billion. The remaining 5 LRs were between 2 and 13. Two people were reported as having an LR of 13, one of whom was Mr Daniel, and another was a female.

54. It is important to note that by comparing the low-level component from this mixed DNA profile to so many reference DNA profiles it is expected that there would be matches that have occurred entirely by chance. Indeed, if this low-level component was compared to a database of profiles from people completely unconnected to the Blackburn case, further inclusionary LRs close to 1 would be expected.

55. The concern with reporting that each of these 5 people could be the source of this low-level DNA is that there may be a tendency to think that 1 of the people must be the source. This is would be an incorrect assumption given the limitations of the evidence and it is entirely possible that none of the reference candidates is the source of the DNA.

---

[23] See electropherogram at page 0778 of the Blackburn casefile
[24] See STRmix™ deconvolution report at page 0773 of the Blackburn casefile

56. It is clear from communications in the casefile (for example, pages 2022 and 2032) and in particular the singling out of Mr Daniel as a possible contributor, that the LRs being reported were not well-understood by investigators or other recipients of the reports. As late as December 2016 questions were still being asked about the meaning of the LR reported in relation to Mr Daniel despite the reporting scientist's repeated attempts to explain the (lack of) significance of this LR. Crucially, no context was given in any of the communications, reports or witness statements as to the quality of the DNA profiling results being interpreted – that in this sample the only evidence of a second contributor was the presence of 2 discrete low-level results.

57. The confusion continued when the reporting scientist was asked to undertake an evidentiary statistical evaluation of the profiling results from L45 in relation to Mr Daniel and using the Aboriginal population database. This evaluation resulted in a weaker LR of 6. The reporting scientist had forewarned investigators of this eventuality both in Intelligence Reports where there was some standard text explaining the limitations of the LRs reported, and in communications with QPS and DPP staff as can be seen at page 2033 of the Blackburn casefile.

58. Communications[25] between the reporting scientist and the QPS Scientific Officer, dated 16 February 2016, suggest that the Scientific Officer was being called to court to explain the significance of the LR from the L45 sample generated with regard to Mr Daniel, as reported in the reporting scientist’s Intelligence Report dated 26 September 2014. This is worrying as QPS staff should not be testifying to work undertaken by QHFSS staff. The conclusion of this discussion seems to be an agreement that the Scientific Officer would only speak to the contents of the Intelligence Report and if the Court required further detail as to the significance of the LR then the reporting scientist would need to be called, which is entirely appropriate.

59. It is possible that the Blackburn case was one of the earliest to incorporate the reporting of LRs generated by STRmix™ database searches. After a review of the original STRmix™ v1.05 validation documentation and the related SOPs that came into effect late 2012, it remains unclear why the laboratory chose to report LRs for results that are essentially uninformative and perhaps the laboratory was following guidance from other Australian laboratories. How to report LRs close to 1 continues to be a topic for debate internationally. In 2018 the Scientific Working Group on DNA Analysis Methods (SWGDAM) released a report that included a section on reporting LR values that are close to 1 which states “…likelihood ratios appropriately express the strength of the evidence and should be reported no matter how low or high the numerical value.”[26] However, it is incumbent upon the laboratory to have robust mechanisms in place. A key step prior to undertaking any statistical evaluation is an assessment of the DNA profiling results obtained. For low-level results analysts must have confidence that the results are from DNA inherent to the sample rather than from other sources such as carryover, and perhaps apply some guidelines as to the minimum number of results required. Furthermore, in the reporting of LRs associated with low-level results it is important to include an explanation of the quality of the results on which the LR is based and any assumptions made as to the number of contributors or any other limitations of the data or the LR calculated.

60. From their validation documentation the laboratory was aware that very low-level results could provide uninformative LRs. For instance, in the reproducibility study of their STRmix™ v1.05 validation[27] the laboratory identified that when a two-person mixture with a contributor

[25] Pages 2032-2033 of the Blackburn casefile.
[26] Page 4 of the Recommendations of the SWGDAM Ad Hoc Working Group on Genotyping Results Reported as Likelihood Ratios.
[27] Page 17 of Verification of the DNA Profile Analysis module of STRmix™ using Promega PowerPlex 21 system – 12.2012.

ratio of 50:1 was repeatedly deconvoluted and compared to the DNA profiles of the 2 known contributors, the LRs generated for the known trace minor contributor were close to 1, either weakly supporting inclusion or exclusion. The report correctly explained that this is not unexpected and is largely due to the loss of allelic information from that contributor. It follows that had the mixture been compared to a database of non-contributors that some LRs weakly supporting inclusion would also likely have occurred by chance.

61. It is important to note that this is not a case of STRmix™ getting the wrong answer. The program is simply using the DNA profiling information that is available to inform the interpretation. When an LR close to 1 is generated, it quite correctly reflects the weakness or insufficiency of the DNA results being interpreted. To put it another way, as peak heights decrease or if few DNA results are present one generally expects to see a convergence of the LR to 1.

62. The QHFSS STRmix™ v1.05 validation report concluded that for very low template contributors interpretation difficulties arise. The validation report also recommended that "The input template is considered before deciding whether a profile will be suitable for reliable interpretation." However, the laboratory seemed to apply this guideline to the total input template, rather than to contributor-specific template amounts or proportions.

63. Document '31523 - Interpretation and Statistical Analysis of DNA Profiles Using the STRmix Expert System' (v1) SOP, which came into effect 14 December 2012, stated that the primary function of the database search functionality is to allow the comparison of deconvoluted case profiles to multiple reference DNA profiles in cases where more than one reference sample has been submitted. The SOP further stated "This process needs to be performed against every casework sample" but there is no advice as to determining whether the quality of the profiling results for the different components of the mixed DNA profile is suitable for comparison purposes. For example, in a scenario where the majority of the DNA is from one person and there are a few additional low-level results, then it may be concluded that the majority of the DNA (the major component) is suitable for comparison purposes, but the remaining low-level DNA is unsuitable for comparison purposes. There is some subjectivity inherent to the determination as to whether or not low-level profiling results are suitable for comparison purposes. Laboratories can utilise guidelines to assist in this determination, such as specifying a minimum number of low-level results at a minimum number of loci.

64. A more serious issue is that during the timeframe that the Blackburn samples were being processed, laboratory staff members were concerned about increased drop-in and gross contamination events, and ongoing carryover issues that they seemed unable to resolve. Drop-in and in-laboratory contamination[28] are each a reflection of the general hygiene of the laboratory and the meticulousness of the laboratory processes undertaken. Any increases in drop-in rates or in-laboratory contamination events indicate that anti-contamination procedures need to be addressed. Continued detectable carryover suggest a malfunctioning CE instrument and this ought to have been resolved prior to continuing with casework. These issues are discussed in more depth later in this report. However, that the laboratory continued its policy to report LRs based on very low-level data is difficult to understand in light of the observations of increased drop-in, contamination and carryover.

65. There may have been a belief that reporting inclusionary LRs that were close to 1 would not have serious consequences. In an email chain from June and July 2014[29] there was a discussion about a situation where 1 extra peak would change a single source DNA profile to a mixed DNA profile "but we have always accepted that this is low risk because you are not going to get a strong stat as a result of this (because of the limited information from that

[28] In-laboratory contamination is a contamination event that occurs during the processing of samples
[29] Document '28. Corro – Determination of number of contributors – 14.07.2014.pdf'

contribution).” However, if the statistic and the poor quality of the results on which it is based are not well understood by investigators or other audiences, then the consequences for the person or persons named in the resulting reports could be serious. Conversely, erroneous exclusions based on uninformative DNA profiling results also could have serious implications during an investigative phase.

66. It must be noted that there are indications that some staff members were uneasy with the policy to report LRs based on very low-level results. The clearest indication of this concern is in an email chain from April 2014[30] discussing the implementation of STRmix™ v2 and optimisation of PowerPlex® 21. With regard to the purpose of re-optimising the PowerPoint 21 kit, one staff member stated, “Reproducibility is a challenge with any kit…but with continuous interpretation and being forced to interpret ‘minors’, we have to investigate ways to minimise artefacts…” A year earlier in a document dated 30 April 2013[31] which discussed the reassessment of drop-in parameters and raised concerns about increased drop-in rates, contamination events and carryover, there was a proposal to include all instances of drop-in <u>and</u> contamination into the drop-in cap and rate analysis thereby inflating the drop-in value. This course of action would “… mean that most "minor" profiles that we are looking at presently will be considered as potential drop-in thereby lowering any LR values considerably for profiles that do match, or even potentially excluding some genotypes. This potentially will "solve" some of our current issues around low-level uncertainty as the problem will to a marked degree "just disappear."”[32]

67. While LRs of very low-level mixture components should not have been reported at all, they continued to be reported in the Blackburn case, including in the formal written statements issued by the reporting scientist in 2016 and 2017. The earlier intelligence reports that included the LRs based on trace mixture components could and should have been retracted, especially in light of the background levels of drop-in, contamination and carryover. The general policy to report LRs based on very little profiling information should also have been reassessed.

68. Dr Wright expressed concerns about the way in which low-level results obtained from samples in the Blackburn case had been interpreted. She quite rightly notes that one of the findings of the STRmix™ validation was that very low levels of DNA, especially in mixed DNA profiles, cannot be reliably interpreted. Dr Wright proffers that the incorporation of the 3-second injection time during the validation of the PowerPlex® 21 DNA profiling multiplex kit resulted in an artificially low analytical threshold (AT) and drop-in cap. In terms of the Blackburn case, she concludes that trace results were included in interpretations when they should have been excluded, ultimately resulting in several database search matches generating inclusionary LRs close to 1. Given the manner in which the baseline experimental data were averaged across both instruments and all dye channels, the incorporation of the 3 second injection data may not have had a significant effect on the setting of the AT. However, in relation to setting the drop-in cap, without having a clear understanding of what was actually being observed in the rerun of the negative control data at 5 seconds, it is difficult to assess what effect the incorporation of the 3-second data had in any subsequent STRmix™ interpretations. The possible effects should have been investigated further and if required, the STRmix™ drop-in cap and parameters adjusted accordingly.

69. Dr Wright also presented an example of where a peak below the AT was used in order to exclude a person identified as having an inclusionary LR through the database search

---

[30] Document ‘40 Corro - topic to discuss - 14.04.2014.pdf’

[31] Document ‘Drop-in re-assessment discussion 30-04-2013.doc’

[32] This proposal does not appear to have been adopted as other documentation suggests that as late as December 2014 the drop-in cap still had not been adjusted from the original 40rfu setting.

functionality[33]. Below threshold peaks with good morphology but clear of other baseline noise may well be from DNA, rather than noise. There are situations where below threshold peaks could be used for exculpatory reasons. If for example, a case profile has low level profiling results at most of the DNA sites tested and there are no indications that the DNA is from more than one person. In this example good morphology peaks that are just below AT may be useful in excluding people whose DNA profiles otherwise match the case profile based solely on above AT data. However, in the particular example that Dr Wright illustrates in her report, the peak used for exculpatory purposes is not sufficiently clear of other baseline peaks to be certain with any confidence that it is DNA from the second contributor. Furthermore, when evaluating peaks below threshold extra caution must be taken when drop-in and carryover peaks have been observed in negative control data.

70. In summary, many of the trace profiling results observed in several of the mixed DNA profiles obtained from Blackburn case samples were unsuitable for comparison purposes. The QHFSS validation of STRmix™ v1.05 demonstrated that known contributors who have contributed very little DNA to a mixture may generate LRs close to 1, and sometimes the LR will support exclusion. This behaviour is expected because when very little DNA profiling information is present for a contributor(s), the LR will trend toward 1. Furthermore, when so few profiling results are available, it cannot be assumed that they all originate from one person, nor can the possibilities of drop-in or carryover be ignored, especially when the laboratory is observing increasing drop-in rates and carryover, as was the case at QHFSS at the time. These factors should have influenced QHFSS into taking a cautionary approach with regard to the interpretation of low-level peaks.

3.4 INTERPRETABLE PROFILES DESCRIBED AS 'COMPLEX'

71. There are some samples that have been described as "complex" and have not been interpreted. As Dr Wright states in her report at section 5.2, some of these profiles could easily have been interpreted. That they were not interpreted deprived the investigation of potentially useful information.

Samples from vehicle 706LHN

- Item V13, described as a trace DNA tapelift from front driver's seat belt
- Item V24, described as a swab of saliva from 'Mt Franklin' water bottle
- Item V41, described as a swab of saliva from 'Coke' bottle.

72. Incomplete and low-level DNA profiling results were obtained from each of these samples[34]. However, there is sufficient profiling information to conclude that if the assumption is made that in each of these samples the DNA has originated from just one person, then Mr Peros cannot be excluded as being the source of the DNA recovered from each of these samples. This conclusion should have been reported, even though the finding may not be unexpected given the samples were taken from items in his own vehicle.

Sample described as a trace DNA tapelift from Ms Blackburn's left forearm

73. The original qualification results for this sample was 0.006ng/µL and the sample was reported as "No DNA detected." It was later reworked at the request of QPS[35]. The

[33] In Dr Wright's report, Figure 17 on page 51.
[34] Electropherograms at pages 0152-0159 (V13), 0143-0150 (V24), 0134-141 (V41) of Blackburn casefile
[35] Page 1800 of the Blackburn casefile

subsequent DNA profile was reported as "Complex"[36]. The profiling results obtained were very low level and incomplete, which is not unexpected from skin samples as many cells in the outer layer of skin are keratinised which means they may not have DNA-containing nuclei.

74. All but one of the DNA profiling results recovered from this sample corresponded with Ms Blackburn's reference DNA profile, which is not unexpected given that the sample was taken from her skin. The single DNA profiling result that did not correspond with Ms Blackburn's reference profile is unsuitable for meaningful comparison purposes.

75. Dr Wright posits that the above 4 samples could have been affected by the defective Proteinase K and so the laboratory chose to report these results using what Dr Wright describes as a "neutral reporting option."[38] However, where a DNA profile appeared to be single source but incomplete and showing signs of stochastic effects such as peak imbalance, the laboratory may have had a policy of reporting these as "unsuitable for meaningful interpretation" (or "complex"). If this is the case then this policy runs the risk of not providing potentially inculpatory, exculpatory or otherwise useful information to the investigation. It is also very strange that the same policy was not applied to the very trace DNA components detected in mixed DNA profiles discussed in the previous section.

<u>Sample described as trace DNA tapelift from handle of a knife (sample 585528352)</u>

76. The profiling results obtained from this sample were reported as "complex." A mixed DNA profile originating from at least three people was obtained[39]. The majority of the DNA is from a male, and this major DNA profile is interpretable without the need for further laboratory processing. The remaining DNA profiling results are low level and due to the uncertainty as to the number of contributors are unsuitable for comparison purposes.

3.5 SAMPLES REPORTED AS 'DNA INSUFFICIENT' OR 'NO DNA DETECTED'

77. Several samples submitted for DNA profiling were stopped after quantification and reported as "DNA insufficient for further processing" or "no DNA detected". The quantification results for the Blackburn samples were not documented in the case file but were supplied for this review following a request from the author.

78. It is understood from the documentation provided that <u>prior</u> to the implementation of the PowerPlex® 21 profiling kit in late 2012, the laboratory had an agreed policy with QPS where if a <u>volume</u> crime sample returned a quantification value less than 0.01ng/µL then the sample would not be processed further and it would be reported as "DNA insufficient for further processing" or "no DNA detected."

79. As long as the threshold has been properly validated this is an acceptable practice for volume crime samples because usually the primary aim of processing these samples is to achieve a DNA profile suitable for comparison to a local or national DNA database. Processing samples with low amounts of DNA may not provide DNA profiles that meet the criteria for comparison to DNA databases.

[36] Electropherogram at pages 1111-1114 of the Blackburn casefile
[38] Page 45 of Dr Wright's report.
[39] Electropherogram at pages 0436-0443 of Blackburn casefile

80. However, for more serious crime incomplete or low-level DNA profiling results that are unsuitable for comparison to DNA databases may still be informative to a criminal investigation. This is because reference DNA profiles from the victims or other people believed to be associated with the victim or the crime scene can be directly compared to profiling results obtained from crime samples. For example, if the incomplete profiling results obtained from a victim's fingernail samples correspond with the victim's reference DNA profile but there are indications of DNA from a second person then this allows for some decision-making around using a more sensitive DNA profiling technique. Even if no foreign DNA is detected in the sample, then this information is more informative than simply reporting "DNA insufficient for further processing" or "no DNA detected."

81. There are two documents that provide some background as to why the "no DNA detected" policy was extended to include major crime samples. The first is the PowerPlex® 21 validation document[40] which indicated that profiling results were achievable from samples containing DNA less than 0.01ng/µL but that these profiling results exhibited stochastic effects. Stochastic effects are common in low-level DNA profiles and include phenomena such as peak imbalance and drop-out (absence) of some or many DNA profiling results. Determining the minimum number of contributors can also be difficult.

82. Recommendation 3 of the PowerPlex® 21 validation document[41] stated that samples with concentrations below 0.01ng/ µL would not be routinely processed. This recommendation was incorporated into the Procedure for Case Management SOP[42], stating that any sample with a quantification value less than 0.01ng/µL would not be amplified or processed further using a DNA concentration step, and these would be reported as "DNA insufficient for further processing." Samples with an "undetermined" quantification result would be reported as '"no DNA detected."

83. Although the presence of stochastic effects can complicate profile interpretation, low level DNA profiles may still be interpretable. Therefore, the recommendation to not routinely profile any major case sample with a DNA concentration below 0.01ng/µL is surprising. Setting a threshold designed to essentially avoid having to interpret DNA profiles demonstrating stochastic effects runs the risk of not detecting potentially probative or otherwise informative profiling results.

84. The second document is an email dated 06 March 2013[43] that informed QPS that the validations of PowerPlex® 21 and STRmix™ and the resulting changes to workflow had placed a significant burden on the laboratory during the previous year. The email was very focussed on turnaround times and warned that due to the extra information contained in PowerPlex® 21 DNA profiles, these may take longer to interpret and report.

85. The email further stated that due to a workflow change, "No DNA detected" samples would now be reported more quickly than previously and that:

"Previously, samples where no DNA had been detected were supplied for volume crime, however this workflow has now been implemented for both Major and Volume crime."

While this text only refers to "no DNA detected" samples it seems as though it must also apply to "DNA insufficient for further processing" samples too, given the recommendation in the PowerPlex® 21 validation and the directive in the SOP.

86. This text seems to suggest that the extension of the policy to major crime samples was designed to off-set the longer turnaround times anticipated from interpreting and reporting

---

[40] Document 2. PowerPlex 21 – Amplification of Extracted DNA Validation – 14.12.2012

[41] Document 2. PowerPlex 21- Amplification of Extracted DNA Validation – 14.12.2012.pdf at page 64.

[42] Document 7. SOP -17117V16 – Procedure for Case Management – 11.12.2012.pdf at page 7.

[43] Document 25. Corro – Update – 06.03.2013

the results obtained from samples processed through PowerPlex® 21. This change was communicated as an advantage because it would likely produce faster turnaround times for a significant number of samples that would now be automatically stopped after the quantification step.

87. It is unclear if the extension of this policy to major crime samples was previously agreed with QPS. However, this change in policy seems to have been driven by the observation that low-template DNA profiles were more difficult and time consuming to interpret. This change effectively avoided having to interpret potentially complex results by not generating profiling results in the first place, thus significantly reducing the turnaround times for these samples. However, given that the PowerPlex® 21 validation indicated that profiling results were achievable from samples with DNA concentrations less than the 0.01 ng/µL cut-off, and that the STRmix™ software was designed to assist with interpreting complex and low-level DNA profiling results, it could be argued that not profiling the low-template samples was a disservice to QPS. The prioritisation of faster turnaround times over the potential information that may have been gained from processing these samples is questionable.

88. In the Blackburn case several samples were reported as “DNA insufficient for further processing” or “no DNA detected”. At the request of QPS, some of these samples subsequently underwent a DNA concentration process and were profiled. These samples include the trace DNA tapelift from Ms Blackburn’s left forearm and the 3 samples from vehicle 706LHN, the results of which were discussed in the previous section of this report. While in these instances the profiling results obtained were not probative, the conclusions would likely be more informative to the investigators than the initial “no DNA detected” reporting. That interpretable results were achieved from case samples initially reported as “DNA insufficient for further processing” or “no DNA detected” demonstrates that the policy of halting DNA testing of low template samples from major crimes is a missed opportunity to obtain informative or potentially probative DNA results.

4. REVIEW OF OQI34043

OQI34043 Positive Extraction Controls with low DNA yields (also known as the "Dishwasher OQI")

89. This quality incident was reviewed in order to determine whether the issue was investigated and resolved appropriately, and also determine as far as possible whether or not samples in the Blackburn matter were directly affected.

4.1 TIMELINE

90. The following is a timeline of the events related to this quality investigation, which is based on various documents provided.

91. 20 March 2013: An entire batch of reference samples (REFQUA20130320-01) had very low quantification values. The quantification was immediately repeated, and similarly poor results were obtained indicating that something had negatively affected the recovery and extraction of DNA from these samples and the extraction of the positive control.

92. 21 March 2013: Following some immediate investigations that identified other batches similarly affected, it was determined that the issue was limited to batches that had been extracted on the Maxwell® instruments. At 3.44pm an email (document 9b) was sent with high importance stating that no further Maxwell® extractions be undertaken until further investigations could be completed.

93. 25 March 2013: An email (document 9e) sent at 9.29am advised that the Maxwell® kit reagents could be ruled out as a possible cause of the issue and that further troubleshooting of other reagents was in progress. An email (document 9d) sent at 5.13pm identified that it was a specific USB-Affymetrix Proteinase K lot (lot 1251021-A) that had caused the issue. The email also advised that routine extractions could be re-started the following day and that the process of organising re-extractions of the affected samples would be started.

94. Further documentation (spreadsheet 7e and document 9g) indicates that aliquots of other Proteinase K lots (manufactured by Sigma Pro) were tested as part of the troubleshooting and these lots performed as expected, thus isolating the issue to the specific USB-Affymetrix Proteinase K lot 1251021_A. The 'A' in the lot identifier related to the particular vial the aliquots were made from and a further comment in document 9g noted "a different aliquot from a different bottle but of the same manufacturer gave OK results." Testing of aliquots made from vial or bottle A of the USB-Affymetrix Proteinase K yielded a pH much higher than expected (pH of 14), well beyond the normal working range of Proteinase K (pH of 7-8).

95. There is also an email chain (document 7c) from early April 2013 between a QHFSS laboratory staff member and staff at In Vitro Technologies, the suppliers of USB-Affymetrix products in Australia. In this email chain there were discussions around the correct storage of Proteinase K once it has been prepared for use in the laboratory, correct concentrations of the product and how the reagent should be prepared both for immediate use and long-term use.

96. Within this email chain, on 11 April 2013, the QHFSS staff member explained that three vials of the USB-Affymetrix Proteinase K had originally been received, all with lot number 1251021. Of these three vials, one had been used for the original validation of the reagent, and it had been determined that this reagent worked satisfactorily. This comment suggests that this particular brand of Proteinase K had only been in use in the laboratory a short time, perhaps providing some context for the earlier questions in the email chain about correct

storage and preparation[44]. Other information in the email chain suggests that In Vitro Technologies had sent these vials to the laboratory on 4 March 2013.

97. The 11 April 2013 email further detailed that the remaining two vials, called A and B, were subsequently prepared into aliquots for routine use although it is unclear whether these aliquots were prepared at the same or different times. Only the aliquots from vial A had the very high pH indicating that there was an issue specific either to the original vial A product or specific to the preparation or storage of the aliquots prepared from vial A.

98. Later that same day there was a reply from In Vitro Technologies stating that the three vials that had been supplied had actually been expired stock and should never have been sent to the laboratory. The response from the laboratory suggests that the expiry dates of reagents were not being checked and recorded upon receipt into the laboratory. Further questions also suggest that aliquots made from the reagent may have been stored at 4°C rather than -20°C and that the information related to expiry dates and storage that accompanies the bottles of reagents had not been reviewed by the laboratory staff member responsible for the validation of the reagent. That the USB-Affymetrix Proteinase K was validated and subsequently incorporated into laboratory processes without expiry dates having been checked and recorded, and without a clear understanding of appropriate storage is not in line with best practice.

99. An email from In Vitro Technologies dated 18 April 2013 stated that USB-Affymetrix had been contacted. USB-Affymetrix were unable to explain the pH issue, and no other laboratories had reported this issue. The manufacturer was not able to re-test that particular lot as they no longer had any in stock. They suggested "something else must be going on" and that the high pH is unlikely to be a result of the product being expired or from incorrect storage.

100. According to the OQI document, during the Proteinase K investigation there was another investigation underway related to the malfunctioning of an industrial dishwasher that was used to clean the laboratory glassware. The faulty dishwasher was mentioned in the Proteinase K OQI because it was proposed that a measuring cylinder used to prepare the aliquots of the affected Proteinase K may have contained residue of the caustic cleaning agent used in the dishwasher, resulting in the very high pH. The Proteinase K OQI included the information that the repair of the dishwasher was underway and an alternative process for washing laboratory glassware was in place. From the documentation provided, it is unclear as to whether there was a separate OQI raised in relation to the malfunctioning dishwasher that details the nature of the problem, how long it had been occurring and what measures were taken to evaluate and mitigate any downstream effects. If no such investigation was undertaken, then this is cause for concern and it is impossible to conclude that the dishwasher was the root cause of the defective Proteinase K.

101. 04 April 2013: An intelligence report (contained in document 9a) was sent to the DNA Management Section at Queensland Police detailing the initial suboptimal processing of 186 casework and 26 reference samples, describing the reprocessing of these samples and warning that for the case samples "…the results obtained may not be a true representation of the DNA that may have been in the original sample, but the best we are able to obtain given the circumstances." The report lists the sample barcodes of the 186 casework samples. Reference samples were not included in the list as reprocessing of these would produce unaffected results.

102. The OQI summary document also revealed that once the particular lot of Proteinase K had been identified as causing the issue, staff had been notified and that the reagent was not to be used. However, the reagent was again used in 6 subsequent extraction batches dated

[44] Meeting Minutes dated 01.08.2012 indicated that the laboratory was trying to "verify cheaper version of Proteinase K".

02 April and 04 April 2013, affecting a further 77 casework samples and a batch of reference samples.

103. 09 April 2013: A further intelligence report (contained in document 9a) was issued to the DNA Management Section at Queensland Police identifying and listing an additional 77 case samples that had been affected by the suboptimal Proteinase K.

4.2 WAS THE RESPONSE ADEQUATE?

104. In considering all the information provided in relation to this quality incident, once the laboratory was aware of the issue their response was immediate and appropriate. Although the exact timings of events are not available, there was approximately a day between the issue being first observed and the request to cease all extractions using the Maxwell® instruments. This response seems reasonable. During this time the quantification of the batch identified as having an issue had to be repeated to determine if the issue was isolated to the quantification process. Once re-quantification confirmed that the issue was more likely to have occurred at the extraction stage the quality investigator had to review all other batches to see if they were similarly affected and then determine a common denominator, which in this case was that all affected batches had been extracted on a Maxwell® instrument. This process immediately led to the cessation of all work on the Maxwell® instruments so that further, more refined troubleshooting could be undertaken. Once the cause of the issue was identified, the problem and all affected case samples were communicated to QPS. The affected samples were subsequently reworked in what appears to be a timely and appropriate manner.

105. However, there are some gaps in the investigation, which are described below, that may have helped identify a root cause(s). It is possible that these steps were undertaken but were not documented in the OQI document. The OQI document also does not include some pertinent information that was included in emails or other documentation provided. It is also unclear if there was a separate investigation for the malfunctioning dishwasher. Given the dishwasher was proposed as a possible root cause for the defective Proteinase K, the dishwasher OQI, if it existed, should have been cross referenced in the Proteinase K OQI. These omissions and the lack of detail as to which samples and cases were affected and when, are not entirely appropriate.

4.3 COULD THE PROBLEM HAVE BEEN DETECTED EARLIER?

106. It is difficult from the documentation provided to determine how many batches were affected prior to the alarm being raised after the review of the quantification results from batch REFQUA20130320-01. In a spreadsheet (7b) that contains a tab called 'Affected batches' there are 16 casework extraction batches and 2 reference extraction batches, 1 of which presumably became REFQUA20130320-01, with dates within their extraction batch identifiers between 15 March and 25 March 2013. If it is assumed that the dates within the extraction batch identifiers relate to the date the DNA extraction process was undertaken and that normally the quantification process would be undertaken within a day or so after extraction, then perhaps the problem may have been detectable as early as 18 or 19 March 2013 when the quantification of the first affected casework batch likely occurred. However, this issue is more easily identifiable when reviewing the quantification results of a batch of reference samples, where good amounts of DNA are expected from all the samples as well as the extraction positive control. For a batch of casework samples, especially trace samples where little or no DNA may be recovered, only the extraction positive control may indicate an issue with the DNA extraction process. In a busy laboratory this single indicator

may be overlooked at the quantification stage but would hopefully be explored further once the DNA profiling results were analysed, although this latter stage may not occur until some days after the quantification stage. Furthermore, if different people are reviewing the quantification results of different casework batches a pattern of low quantification results from the extraction positive controls across the batches may not be detected at this stage of the DNA profiling process.

4.4 WAS THE PROBLEM CONTAINED ONCE IT WAS IDENTIFIED?

107. It is difficult to explain why some of the extraction batch identifiers contain dates that are between 22 March and 25 March 2013 given that the email stating the cessation of extractions on the Maxwell® instruments was sent during the afternoon of 21 March 2013. In theory, there should have been no DNA extractions using the Maxwell® instruments from 22 March until the end of day 25 March 2013, when the email was sent authorising the resumption of extractions the following day. It is possible that the assumption that the dates in the batch identifiers relate to the dates of processing is incorrect.

108. The aliquots of the poorly performing Proteinase K were not adequately quarantined during the course of the investigation. This improper action resulted in a further 6 batches (including 77 case samples) being extracted using the defective Proteinase K in early April 2013.

4.5 WAS A ROOT CAUSE DETERMINED?

109. Although the reason for the poor extractions was determined to be the excessively high pH of the aliquots prepared from a particular vial of Proteinase K, the actual cause of the high pH has not been established with any certainty. A representative of the supplier of the product, In Vitro Technologies, admitted that the vials of Proteinase K they had supplied on 4 March 2013 had expired. However, expiration does not appear to be the cause of the poorly performing aliquots as other aliquots prepared from 2 of the 3 vials supplied appear to have worked adequately. None of the information related to the Proteinase K being expired, the performance of the different vials, the preparation of the aliquots, or how the aliquots may have been stored incorrectly was included in the OQI document. Ideally the report of a quality incident should include all troubleshooting steps including any discussions about the issue with suppliers or manufacturers.

110. The malfunctioning dishwasher was also proposed as a possible root cause if glassware containing a caustic detergent had been used to prepare the aliquots of the Proteinase K. While this may be possible, it seems unlikely that only this one reagent would have been affected. Information related specifically to the investigation into the malfunctioning dishwasher and its downstream effects, assuming such an investigation was undertaken, was not provided. If an investigation into the cause and downstream effects of the malfunctioning dishwasher was not undertaken then this is a serious lapse in quality control.

111. A more likely explanation may be that the aliquots from vial A were simply prepared incorrectly. However, without scrutinising the reagent preparation records from this time, assuming these records existed, this conclusion cannot be reached with any certainty. Neither the OQI document nor any of the other documents provided details of any investigation as to the actual preparation and storage of the suboptimal aliquots from vial A compared to the aliquots prepared from vial B and from the vial used in the original validation for this reagent. Any differences as to whom, how or when these aliquots were

prepared and how they were subsequently stored should have been a fundamental part of the quality investigation.

4.6 WHAT CHANGES DID THE LABORATORY IMPLEMENT AS A RESULT OF THIS PROBLEM?

112. As a result of this quality incident, the laboratory implemented a number of preventive actions aimed to minimise the possibility of using a sub-optimal reagent in future extractions. These steps include a process to check the quantification results of positive controls, a better process for quarantining kits and reagents implicated in quality investigations and quality control testing of the performance and pH of aliquots of new lots of Proteinase K (and Dithiothreitol) prior to use in the laboratory.

113. However, it must be stated that it is highly unusual that quality control processes for testing the performance of new lots of critical reagents were not already in place in 2013 as it was standard practice for forensic DNA laboratories internationally. For example, the Quality Assurance Standards for Forensic DNA Testing Laboratories (QAS) which took effect 1 July 2009 state:

Standard 9.3: The laboratory shall identify critical reagents and evaluate them prior to use in casework. These critical reagents shall include but are not limited to the following:

9.3.1 Test kits or systems for performing quantitative PCR and genetic typing

9.3.2 Thermostable DNA polymerase, primer sets and allelic ladders used for genetic analysis that are not tested as test kit components under Standard 9.3.1.

114. Although Standard 9.3 does not mention Proteinase K specifically, the performance of this reagent when used in a DNA extraction procedure is critical to its success.

4.7 IMPACT ON THE BLACKBURN CASE

115. Case samples from the Blackburn case do not appear in the list of samples and cases that the laboratory has reported as being affected by the defective Proteinase K. The OQI report is included in the casefile as reference samples from the case may have been affected. However, this has no overall effect on the case as reference samples can easily be resampled and tested.

116. There are a number of samples described as bloodstains taken from the clothing of Ms Blackburn that were either reported as ‘no DNA detected’ or had lower than expected quantification values and profiling results. While it could appear as though these samples in particular were also affected by the poorly performing Proteinase K, a review of the associated extraction batch paperwork indicates that Sigma brand Proteinase K was used, not the sub-optimal USB-Affymetrix brand.

117. In summary, there is no evidence that case samples from the Blackburn case were affected by this particular quality incident.

118. Dr Wright raised concerns that the profiling results obtained from many of the reference DNA samples submitted in the Blackburn case were weaker than expected. Reference DNA samples should contain good amounts of DNA and, in theory, produce consistently good profiling results. Dr Wright clarifies that in the Blackburn case the reference samples that generated weak profiles span more than a year and that initially, reference DNA samples undergo a different process than case samples, which does not contain Proteinase K. She

proposes, as an explanation for the poor results, that perhaps the effects of the malfunctioning dishwasher were more widespread than disclosed by the laboratory.

119. However, the results from reference samples, which are stored on FTA[46] cards, can be variable, especially when processed using a direct amplification[47] method as is the case at QHFSS. While direct amplification methods are ideal for the high throughput processing of good quality samples from known people, such as reference DNA samples, it is not possible to control the amount of DNA that is amplified, sometimes resulting in too little or too much DNA being processed. This may require a new sample to be taken from the FTA card and processed, either through the direct amplification method again or through the standard extraction, quantification and amplification steps that are used to process casework samples. Another confounding issue is that the DNA on the FTA card may not be evenly distributed, and a sample taken from the card may not contain sufficient, or any, DNA. For these reasons the results obtained from reference samples submitted in this case do not raise concerns.

[46] FTA cards contain chemicals that burst cells, denature proteins and protect DNA. A typical process for taking a reference DNA sample is that the inside cheek area of the mouth is swabbed and then the material on the swab is transferred onto a FTA card for storage.
[47] In a direct amplification method there are no separate extraction or quantification steps, as there are for case samples.

5. REVIEW OF OQI34817

OQI34817 Incorrect conditions used for Capillary Electrophoresis

120. This quality incident was reviewed to determine whether the investigation was conducted and resolved appropriately, and also determine as far as possible whether or not samples in the Blackburn matter were directly affected.

5.1 TIMELINE

121. On 8 July 2013 (document 13a) it was discovered that the Genetic Analyzer 3130xl B capillary electrophoresis (CE) instrument had an incorrect injection time of 3 seconds in its run module settings. It was also determined that the 3130xl B instrument had likely always been incorrect, including during the period of time when data for the PowerPlex® 21 validation were being collated. Therefore, some data used to validate the PowerPlex® 21 amplification kit and any casework samples subsequently run on this particular genetic analyzer were affected. The running of casework data was halted immediately after the 3-second injection time issue was identified on the 3130xl B instrument.

122. In order to determine the magnitude of effect of this difference in injection times, a set of samples previously run at 3 seconds was rerun on the same instrument using the corrected 5-second injection time. This comparison was undertaken on 10 July 2013 (document 13b). It was determined that peak heights were on average 1.75 times higher in the 5-second data, and this was linear across a range of peak heights. There were also an additional 70 peaks that were above the reporting threshold in the 5-second data. These differences were expected.

123. As additional profiling results were obtained from the 5-second data all casework batches run on the 3130xl B instrument since the PowerPlex® 21 implementation had to be identified, prioritised and rerun at the correct 5-second setting. This process appears to have taken some time, with the first reruns starting on 16 July 2013. An email dated 15 August 2013 (document 13u) indicated there were still 12 batches left to rerun.

124. An email dated 16 July 2013 (document 13h) stated that the PowerPlex® 21 validation was being revisited to determine what work needed to be repeated now that the 3130xl B instrument was operating with a 5-second injection time. Also discussed was what information in the original validation document needed to be amended to clarify that some data were collected using a 3-second injection time and in some key areas 5-second and 3-second data needed to be separated out.

125. An email dated 23 July 2013 (document 13k) summarised which aspects of the validation had included 3-second data and which of these aspects needed to be revisited. This email also summarised the progress of rerunning casework plates at 5 seconds and noted that the OQI record had been added to almost all of the samples involved. A decision was made (document 13n) that all full volume casework batches would be rerun automatically. Half volume[48] casework samples would be assessed individually, and specific criteria (document 13n) were provided as to which would need rerunning or re-amplification at full volume, and which could be reported based on the original 3-second results. Another email dated 27 August 2013 (document 13v) revealed that this decision was made on the basis of half volume data generally producing taller peak heights than full volume data, thus offsetting

[48] While DNA profiling kit manufacturers generally recommend using full volume reactions, laboratories often include half volume testing during their validation studies. In addition to reducing costs, it is often observed that half volume reaction protocols can increase sensitivity and generate higher peak heights when compared to full volume reactions.

the effect of a shorter injection time, and that the half volume results obtained with a 3-second injection were often strong single source profiles.

126. In an email dated 1 August 2013 (document 13p) a concern was raised that rerunning negative controls at 5 seconds was yielding peaks that were above 15rfu[49]. It can be inferred from this email that these negative controls had been used in the baseline testing of the original PowerPlex® 21 validation but when run at 3 seconds were clear above 15rfu. This email acknowledged that drop-in data needed to be recalculated, however there is no evidence that the results of this reassessment were implemented. A spreadsheet called 'Drop-in & Contamin Alleles' was provided which includes results from 3 negative controls, 1 of which appears to be contaminated as 4 peaks are present, and the remaining 2 having 1 drop-in peak each. The heights of these peaks were not provided and it is unclear if these are the data referred to in the 1 August 2013 email.

127. On 2 August 2013 (document 13r) an email was sent with attachments[50] that included a review of the baseline on the 3130xl B instrument operating with a 5-second injection time and STRmix™ Model Maker data to assist with determining if the 3130xl B instrument could be used for casework again. Based on the information provided in the attachments the recipients unanimously agreed that the instrument could be used for casework again.

128. An email dated 6 August 2013 (document 13s) provided advice as to what wording should be included in any witness statements that had to be reissued as a result of this OQI.

5.2 WAS THE RESPONSE ADEQUATE IN RELATION TO CASEWORK?

129. The root cause was identified as being that the 3130xl B PowerPlex® 21 injection time setting was entered incorrectly at the beginning of the validation process and was not reviewed again. That the injection parameters on the 3130xl B instrument were not checked at all throughout the validation of the PowerPlex® 21 method, nor months later when the kit was introduced into casework is deeply concerning. This information is easily accessible in the Genemapper analysis software for every project run on the instrument.

130. Although it took some time to identify the injection time issue, the laboratory responded appropriately in quickly identifying affected batches and organising their rework. This appears to have been a massive undertaking.

131. That full volume batches were rerun automatically but profiles from half volume batches were reviewed individually is a curious decision. The reasoning supporting decisions as to whether or not half volume samples were to be reworked should be documented in the casefiles of any affected cases. The casefiles should have included a list of all the case samples affected and what action, if any, was taken in relation to each, and the outcome of the action. The details of the quality incident should have been disclosed in intelligence reports and/or witness statements of the cases affected. Much of this detail was missing from the Blackburn casefile and suggests this information may not have been included in other affected casefiles.

5.3 IMPACT ON BLACKBURN CASE

132. Approximately 41 samples from the Blackburn case were rerun. These samples can be found in spreadsheet 'ReGS_CM_workingspreadsheet' which includes notes as to whether the 5-second data changed the original interpretation. The spreadsheet 'Combined data' also contains the samples affected. All of the 41 Blackburn case samples originally run on

[49] rfu=relative fluorescent units, the unit of measure of peak heights in electropherograms.
[50] It is possible the attachments are included in the information provided, however the email does not contain their filenames.

the 3130xl B instrument were amplified in full volume[51] batches and therefore all were automatically rerun as a result of the OQI investigation. The laboratory made a business decision to replace all electropherograms from 3-second data with the 5-second electropherograms. This decision makes it difficult to independently review and compare the results obtained from the 3-second and 5-second injection time data.

133. Some of the Blackburn samples had already been reported in intelligence reports before the reruns were actioned. For most of the 41 samples there were no interpretation changes as a single source profile was obtained from each run. For at least one sample a partial profile became a full profile. For a small number of samples there were additional low-level peaks in addition to the full profile previously obtained, and these were deconvoluted and compared to the reference profiles submitted for this case. Any likelihood ratios providing support for inclusion or exclusion based on these trace results were calculated. Where an interpretation changed from that previously reported, it was reported in a subsequent intelligence report dated 13 Oct 2013 with a comment “Sample has undergone further processing”. There is no information about the OQI in the witness statements or intelligence reports. That the laboratory did not disclose the details of a quality incident that affected the outcome of the DNA profiling undertaken is troubling.

134. In the casefile only a few of the affected samples have a comment relating to the OQI. There is no summary of the samples affected which easily could have been taken from the spreadsheets. Not all of the electropherograms note that it is a replacement for a 3-second electropherogram. These documentation omissions make it difficult to identify which samples were affected and review the results of any rework action undertaken.

135. In summary, the laboratory took appropriate action in rerunning all the Blackburn case samples that had previously been run at the incorrect injection time. However, the casefile lacks adequate explanation of what samples were affected. This could easily have been remedied by inclusion of a list of the samples affected with actions taken and outcomes obtained. That a quality incident affecting the interpretation of profiling results in this case had occurred should have been disclosed on intelligence reports and the witness statement.

5.4 WHAT WERE THE IMPLICATIONS FOR THE VALIDATION OF POWERPLEX® 21 AND STRMIX™

136. Before discussing the implications for the validations undertaken by the QHFSS DNA laboratory staff it is useful to provide some explanatory information related to baseline noise, the setting of analysis thresholds and drop-in.

137. The CE instruments, from which DNA profiles are produced, generate a low level of baseline noise. The AT, also referred to as the Limit of Reporting (LOR), is a value at which peaks below it may be due to the instrument noise and above which peaks can, with reasonable confidence, be assumed to be genuine DNA profiling results or known artefacts that stem from DNA profiling (such as stutter or pull-up). The general convention for determining AT is to average the peak heights of observed baseline noise and then create a buffer region, which is commonly achieved by adding 10 standard deviations, thereby setting the AT well above the baseline noise.

138. There is some small variation in baseline noise within an instrument as some dye channels (i.e., the different fluorescent dye colours) may exhibit more noise than others and within the dye channels, more noise is generally observable at the low molecular weight loci (i.e., the

[51] According to page 3 of document ‘11. Project Report 131 – PP21 Post-implementation review…pdf’, half-volume amplifications were ceased 4 February 2013.

markers on the left side of an electropherogram). Variation can also occur between runs on a given CE instrument.

139. Baseline noise may also change when components of the instrument are replaced. Where two or more CE instruments are utilised by a laboratory, baseline variation is expected between the instruments. Baseline noise is also affected by the presence of DNA and high template DNA input can generate considerable increase of baseline noise and variability within and between dye channels. For this reason, AT is determined from data from both negative controls and known samples or positive controls containing a range of DNA inputs.

140. In what way does injection time influence baseline noise? Bregu et al (2013)[52] demonstrated that for blanks and negative controls baseline noise remains constant regardless of injection time. Put another way, injection time had an insignificant impact on instrument-related noise. When DNA is introduced to limit of detection testing the authors observed that there was an increase in noise for all dye channels and injection times for DNA target amounts up to 1ng, however these increases were relative and constant across all dye channels and at all injection times. This finding suggests that it is the presence of DNA that has the greater impact on baseline rather than other variables such as injection time, although longer injection times result in the introduction of more DNA into the CE instrument. Furthermore, with input DNA of 2 nanograms (ng) or more, there was a considerable increase of baseline noise and variability between dye channels, and this effect was exacerbated at longer injection times.

141. While setting the AT is a critical step in the validations of a CE instrument and the DNA profiling kits that are processed through the instrument, it is always a compromise. If AT is set high, then DNA profiling results may be excluded from the interpretation of results. If it is set low, then baseline noise may be included in the interpretation of results. Within reasonable ATs, there is the possibility that baseline noise may still occur above the AT. While the addition of 10 standard deviations to the average peak height of the baseline noise is intended to provide a buffer by shifting the AT well above the average heights of baseline peaks, it is not 100% fool proof, and this is especially true when an excess of DNA is present. Therefore, any peaks that are just above the AT must be evaluated with an understanding that they possibly could be noise instead of true DNA signal. Conversely, good morphology peaks below the AT but clear of other baseline noise may well be DNA, rather than noise.

142. Laboratories deal with this compromise in different ways. For example, if one dye channel exhibits a noisier baseline than others, a laboratory may choose to set its AT based on this noisier channel even though it likely overestimates the noise seen in the other dye channels, which could result in low level DNA results falling below the AT in those channels. Another laboratory may choose to average across all the dye channels even though this ‘average overall’ approach may result in baseline peaks from the noisiest channel exceeding the AT and perhaps being indistinguishable from genuine DNA profiling results. Another option is to set different ATs specific to each dye channel. Finally, as the presence of DNA also affects the baseline some laboratories adopt entirely different sets of ATs depending on whether the samples are low template or moderate to high template. If the laboratory utilises more than one CE instrument ATs may be set for each instrument individually or averaged data may be used to determine an AT that encompasses the variation across the instruments. Each one of these approaches has advantages and disadvantages, and they differ in complexity in terms of managing workflow. However, each is justifiable as long as the laboratory understands the limitations of the method adopted.

[52] Bregu J, Conklin D, Coronado E, Terrill M, Cotton RW, Grgicak CM, Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis. Journal of Forensic Sciences 2013;58;120-129.

143. Drop-in refers to fragments of DNA present in the laboratory environment that are inadvertently introduced into samples. Drop-in peaks may be detectable, especially if sensitive profiling techniques are used. Unlike contamination, drop-in[53] is not reproducible and is characterised by the presence of just 1 or 2 peaks. It should also be a rare event and laboratories that employ sensitive profiling techniques monitor drop-in rates as a general laboratory 'health check'[54].

144. The STRmix™ software is able to model the drop-in events that may be observed in a laboratory using a combination of parameters that are dependent on setting a drop-in cap and estimating the rate of drop-in. In version 1.05 of STRmix™ the cap is a peak height whereby a peak up to and including that height in an evidence input file[55] may be proposed as possible drop-in during a STRmix™ interpretation.

145. During the validation of a new DNA profiling kit, how the AT is set, its value and the injection times of the CE instrument are somewhat irrelevant to the drop-in cap. However, the same negative control data used to inform the AT can be used to also inform the drop-in cap, rate and STRmix™ parameters. The drop-in cap is based on the heights of peaks visible in negative controls that are above a certain threshold (15rfu or 20rfu, for example) and are believed to be drop-in rather than contamination or noise. The cap can be calculated in different ways. For example, if there are several occurrences of drop-in, a laboratory could use the average peak height of the drop-in peaks plus 3 (or more) standard deviations as the cap. If very few drop-in events are observed, then a laboratory may choose some value that exceeds the tallest peak height of the observed drop-in peaks.

146. During the initial PowerPlex® 21 validation the QHFSS DNA laboratory staff observed 3 drop-in peaks, the tallest of which was 21rfu and the remaining peaks were both 19rfu. It is unclear as to whether these peaks were derived from the 3130xl A instrument with a 5-second injection time or the 3130xl B instrument with the 3-second injection time or a combination of both, which was not clarified in the subsequent reissue of the PowerPlex® 21 validation report.

147. Given the very few observations of drop-in, all of which were below the 40rfu AT, the laboratory could have chosen to not model drop-in in STRmix™ at all. As the laboratory chose a drop-in cap of 40rfu in the conjunction with the parameters calculated by the STRmix™ developers, subsequent STRmix™ interpretations would consider only peaks in an evidence input file less than or equal to 40rfu to be potential drop-in. As the QHFSS AT was 40rfu, only peaks in an evidence input file that were exactly 40rfu in height could be proposed as drop-in by the STRmix™ software.

5.5 WAS THE RESPONSE ADEQUATE IN RELATION TO THE SETTING OF THE AT AND DROP-IN CAP?

148. During the original PowerPlex® 21 validation activities in 2012 the QHFSS DNA laboratory unknowingly included 3-second data from their 3130xl B capillary electrophoresis instrument with the 5-second data from their 3130xl A instrument. The laboratory analysed baseline data from both CE instruments using negative controls and known profile samples ranging from input values of 0.025ng to 0.500ng. In order to determine the AT, the peak heights and standard deviations across all dye colours and across both instruments were averaged and data were collated separately for full-volume and half-volume amplification conditions. The greatest amount of variation resulting in higher AT (or LOR) values was observed in the

[53] It can be difficult to determine what is contamination and what is drop-in. However, drop-in is usually determined from negative controls where there are no more than 2 or 3 peaks. More than this is normally classed as gross contamination.
[54] Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Science International 2000;112;17-40.
[55] This is the data file STRmix uses to perform its interpretations. It includes the DNA profiling results and their heights observed in a profile.

half-volume data from the 3130xl A instrument's green and yellow channels where the average AT approached 50rfu[57]. When the half-volume data from all the dye channels for this instrument were averaged the AT was calculated to be 42.68rfu. The full-volume data run on the 3130xl A instrument, and the full-volume and half-volume data run on the 3130xl B instrument produced much lower averaged ATs. The laboratory ultimately chose an AT of 40rfu to encompass the baseline variation from both instruments running full-volume or half-volume data. This approach was justified in the validation report as it avoided analysis errors that may occur if different ATs were set for each instrument and each amplification condition.

149. The 2012 validation report noted that the 3130xl B instrument produced lower baseline noise than the 3130xl A instrument and offered some possible explanations. If the laboratory had investigated this difference more closely the incorrect injection time may have been discovered much earlier.

150. That 3-second injection data were inadvertently included in the determination of the AT during the validation of the PowerPlex® 21 kit likely had little meaningful effect on the outcome. The taller average peak heights and greater variation obtained from the half-volume data run on the 3130xl A instrument with the correct 5-second injection time more so informed the decision to set the AT to 40rfu.

151. Once the laboratory became aware that during the original validation the 3130xl B had been operating with a 3-second injection time instead of a 5-second injection time, they amended the validation document. Table 4 of the 2013 version (V2.0) of the validation document[58] was amended to show that the two instruments were operating at different injection times, although it did not specifically state that the 3130xl A instrument had a 5-second injection time and the B instrument had a 3-second injection time. The same baseline analysis results as the previous validation document were presented in Tables 11 and 12 but there is no additional commentary about the difference in baseline noise between the instruments being possibly due to the different injection time conditions.

152. The 5-second data subsequently obtained from the 3130xl B instrument were presented differently in the 2013 (V2.0) validation document at Table 13 and in the report titled 'Summary report of baseline determination on 3130xl B after change in injection time', making it difficult to perform a direct comparison with the original validation data. The summary report also presented incorrect values in Table 2 for the last 2 columns in the Average +10SD row. However, assuming the average peak height and standard deviation values were correct in these tables then the 3130xl B 5-second data showed a slight increase in average baseline peak heights and variation, especially for the half-volume amplification condition. These values were still lower than that obtained from the 3130xl A CE instrument. Therefore, the original AT of 40rfu was deemed to be adequate for differentiating baseline noise from profiling results where input DNA does not exceed target amounts.

153. Drop-in was re-evaluated based on the 3130xl B negative controls rerun at 5 seconds, however the results of this re-evaluation do not appear to have been incorporated into the STRmix™ settings. An email (document 13p) recorded that there were more instances of drop-in in these data, but this was not addressed in the updated validation report.

154. Drop-in was being monitored by the laboratory after the implementation of PowerPlex® 21 and STRmix™ v1.05. Ongoing drop-in monitoring provides a useful data-source for

---

[57] Section 6.3 of Document 2.PowerPlex 21 – Amplification of Extracted DNA Validation – 14.12.2012.pdf

[58] Document 3.11.PowerPlex 21 – Amplification of Extracted DNA Validation V2.0-final.pdf

detecting changes in drop-in rates and contamination events. These data can be used by laboratories to re-evaluate drop-in parameters if significant differences are detected.

155. The document 'Drop-in re-assessment discussion 30-04-2013' indicated that drop-in had been assessed based on the negative controls on every full volume plate run between 7 February and 23 March 2013. It is clear from this document that there continued to be difficulties distinguishing drop-in from contamination or from carryover[59], which is another CE phenomena. According to the document 101 peaks had been identified in 562 negative controls, however 44 of these peaks, from 7 controls, appeared to be clear examples of contamination and/or carryover. Of concern, this document refers to a 'considerable carryover issue', which 'might not be politically expedient' if made public (for example if that data were made available to the STRmix™ developers to calculate the drop-in parameters).

156. The 'Drop-in re-assessment discussion 30-04-2013' document stated that if the contamination peaks and carryover peaks were included in the drop-in evaluation, then the new drop-in cap would increase from 40rfu to 136rfu. If the contamination peaks were excluded the drop-in cap would be 46rfu.

157. As drop-in is an entirely different phenomena to carryover, it is inappropriate to include carryover peaks in drop-in data used to inform the drop-in parameters in STRmix™. The STRmix™ drop-in parameters are not designed to mitigate the presence of carryover peaks in casework results. Ideally, carryover peaks should not be present in casework results at all.

158. The document recorded that a decision was made to incorporate the higher cap of 136rfu, and STRmix™ drop-in parameters were to be calculated with assistance from the STRmix™ developers. There is an email[60] to the STRmix™ developers asking for the parameters to be checked, to which there is a response, and seeking advice on what to do with the carryover peaks, to which no response is provided in the documentation made available. However, the drop-in section and settings remained unchanged between the December 2012 PowerPlex® 21 validation and the December 2013 reissue of the validation report, even though this would have been a suitable document in which to capture this work. More investigation is needed here beyond what could be done in preparation for this report.

159. The validation of STRmix™ v2.0.1 during the first half of 2014 would also have been a good time to re-assess the drop-in cap, rate and STRmix™ drop-in parameters. However, drop-in is not mentioned at all in either of the STRmix™ v2.0.1 validation reports[61] made available. An email dated 9 December 2014[62] suggests that re-evaluation of the drop-in cap and parameters still had not been undertaken.

160. In summary, that 3 second data was included in the determination of the 40rfu AT may not have had much of an affect given the choice to average across two CE instruments and data from half and full-volume reactions. Without the negative control data rerun at 5 seconds it is difficult to know to what degree the drop-in cap and parameters were underestimated.

161. More importantly, the failure to update the drop-in cap and parameters based on the data that emerged from ongoing monitoring is concerning.

---

[59] Carryover is defined as the physical transfer of DNA from one capillary injection into the next injection into the same capillary. In most cases carryover occurs below the background noise of the instrument and is therefore not detectable and if the CE instrument components are functioning correctly the amount of carryover is extremely small.
[60] Document 'Fwd RE Drop-in Calculations'
[61] Verification and Implementation of STRmix Version 2.0.1 – 06.2014, V0.1 & V0.2.
[62] Document '30. Corro – Drop in parameter – 09.12.2014.pdf'

162. While the effect of not reassessing the drop-in cap and parameters is likely minor for the majority of samples and cases, there are ramifications for the interpretation of low-level peaks. This is especially true given the laboratory's practice of calculating and reporting likelihood ratios for very trace results, as was done for multiple samples in the Blackburn case.

5.6 ADDITIONAL CONSIDERATIONS: CONTAMINATION AND CARRYOVER

163. It is clear that laboratory staff were concerned about the carryover, increased drop-in and contamination events that were being observed. Increased detection of drop-in and contamination is not unexpected when a laboratory transitions to a more sensitive DNA profiling kit such as PowerPlex® 21. For example, in the 2013 journal article "Environmental DNA Monitoring: Beware the transition to more sensitive typing methodologies"[63] scientists from the Victoria Police Forensic Services advised that more attention should be given to laboratory cleaning processes and more stringent cleaning procedures were required.

164. QHFSS laboratory staff were actively putting in place practices to reduce contamination events. An example was found in the 20 June 2013 meeting minutes[64] where there was a comment that the results of environmental testing were demonstrating that more frequent cleaning practices had been effective in reducing contamination.

165. However, minutes of an 'extraordinary' meeting dated 05 March 2014[65] and the associated PowerPoint presentation[66] indicate that laboratory cleaning and sample handling processes had not adequately reduced in-laboratory contamination. The large number of contaminated negative controls and the process to investigate these contamination events were discussed.

166. It was noted that the results from case samples in the affected batches were being reported <u>before</u> the negative control contamination investigations had been completed. This is not best practice as the investigation should include a consideration of how and whether any of the case samples were contaminated or potentially contaminated and how they were addressed. It is unclear if case results were being interpretated with or without consideration of the contaminant results present in the negative controls. Ideally, quality incidents of this nature should be disclosed as part of the reporting of the case results but it is unclear if this was being done. It is also unclear if the premature reporting of the case samples was done in order to meet turnaround times or simply due to a poorly defined process or a combination thereof. On a positive note, the goal of this meeting was to devise a process so that case results affected by a gross contamination event in the negative control were released only after investigation was completed (although it is unclear if the process was achieved.)

167. Another process change to come out of this meeting was to pass any batch containing a negative control with less than 3 peaks, because it was assumed that these peaks were due to drop-in rather than gross contamination and therefore would not be reproducible when reworked. This change in process would allow the associated case samples to be reported without delay. However, this process change is problematic when the results of the negative control are not included in the casefile; if the reporting scientist is unaware that there was a problem with the negative control then he/she cannot consider the negative control results when interpreting the associated case DNA profiles. It is possible that the negative control profiling information is available to reporting scientists through the laboratory information

---

[63] K.N. Ballantyne, A.L. Poy, R.A.H van Oorschot, Environmental DNA Monitoring: Beware the transition to more sensitive typing methodologies, Australian Journal of Forensic Scientists, 45, (2013), pp. 323-340
[64] Document '32. Meeting minutes – 20.06.2013.pdf'
[65] Document '05_5 March_Extraordinary Mtg_Minutes.doc'
[66] Document 'EXTN Rework Strategy Feb 2014.ppt'

management system, but information that a reporting scientist relies upon during the interpretation of the case results should be present in the casefile. Key batch information such as the performance of positive and negative controls is not present in QHFSS casefiles.

168. There is a troubling comment in the PowerPoint presentation on slide 7, “The Drop-in parameter provides an effective case management strategy for accounting for the possible presence of additional peaks.” As discussed previously, despite evidence of an increased rate of drop-in, and at peak heights greater than the 40rfu drop-in cap, the drop-in cap and parameters were not updated in STRmix™ v1.05 (nor in v2, apparently). Therefore, the drop-in cap and parameter as were originally implemented in late 2012 could not provide an effective case management strategy for dealing with the drop-in observed in the laboratory during 2013 and beyond. A better strategy is one that reduces drop-in events to begin with.

169. It is concerning that the laboratory had ongoing carryover issues in both their CE instruments. Carryover peaks are usually low in height, often indistinguishable from the background noise of the instrument. Meeting minutes dated 1 March 2012[67] included comments about increasing carryover issues, indicating that in early 2013 when the collation of drop-in data was being confounded by carryover peaks, carryover had been a detectable issue for at least a year. An OQI[68] from November 2011 addresses earlier occurrences of carryover however it is unclear if the carryover reported in the 1 March 2012 minutes were a continuation of the 2011 issues. It is also unclear if a new quality investigation was raised in 2012 to track the ongoing observances of carryover.

170. The 1 March 2012 meeting minutes stated that carryover was detectable in data from both CE instruments. The minutes noted that “…as carry-over is not occurring all the time, it is hard to detect” and advice from the manufacturer was to “not look below 50rfu.” These minutes detailed a number of troubleshooting ideas and suggested process changes, indicating that staff were committed to resolving this issue.

171. In an email chain dated 25 March 2014[69] QHFSS staff members discussed ways to mitigate the increased sensitivity, including optimising PowerPlex® 21 by reducing PCR cycle number, replacing their 3130xl CE instruments with the newer 3500 instrument and replacing the PowerPlex® 21 kit with the Globalfiler kit. While these emails did not refer to carryover directly, problems keeping the 2 CE instruments operational due to their age and lack of suitable replacement parts were discussed.

172. However, it is quite extraordinary that the carryover issue persisted well into 2013 and possibly beyond. It is an issue that should have been resolvable with assistance from the manufacturers and local support personnel. Although the heights of the carryover peaks are not recorded in the documentation provided their potential presence in conjunction with the increase in drop-in and gross contamination raise concerns about the interpretation of any low-level DNA profiling results in case profiles.

173. It is concerning that there was no process to check for cross-contamination between case samples processed in the same batch. According to SOP 31389V3[70] this capability was available but it was only used for reference sample batches.

174. The continued reporting of likelihood ratios for very trace and un-replicated results, as was done in the Blackburn case, despite the knowledge that there was an apparent increase in detectable drop-in, contamination and carryover is deeply problematic.

---

[67] Document ‘24. Meeting Minutes – 01.03.2012.pdf’

[68] OQI 30467 Carry over of fragments in Capillary Electrophoresis

[69] Document ‘39. Corro – PP21 optimisation – 25.03.2014.pdf’

[70] SOP 31389V3 – STR fragment analysis of PowerPlex 21 Profiles at Table 5.

6. OPPORTUNITIES FOR RETESTING

175. The Commission requested that during the Blackburn case review any samples suitable for retesting be identified. The recommendations provided by the author of this report and by Dr Wright have been collated into a table by Commission staff[71]. In addition to the recommendations collated in that table, the following suggestions should also be considered.

176. Given the interest in the very trace results obtained from L45, trace DNA sample from rear LHS upper leg area of Ms Blackburn's pants, it is strongly recommended that this sample undergo further testing using the male-specific Y STR test.

177. The discovery that the MultiProbe® II extraction method used to process several samples in the Blackburn case may not have recovered DNA optimally warrants DNA profiling of new samples from Blackburn exhibits. Rather than revisit DNA extracts of samples processed by QHFSS, re-examination by an independent laboratory of Ms Blackburn's shirt and the white 'Effekt' T-shirt is recommended. For Ms Blackburn's shirt DNA profiling of swabs or tapelifts of the bloodstaining may also assist in determining if a dye in the fabric was a factor in the poor-quality results obtained from the original QHFSS testing of the bloodstained fabric samples from this item.

178. For the white 'Effekt' T-shirt, this item could be bioscreened again, perhaps using Hematrace which is a test used for the identification of human blood. Depending on the results of this testing, further samples could be taken for DNA profiling.

179. According to the QPS Scientific Officer's formal witness statement72 there were several swabs, described as bloodstained, taken during his examination of Boddington Street that appear to have not undergone DNA profiling. These swabs, or a subset, could be tested.

[71] Document 'Comparison of Shandee Blackburn samples…recommended for re-testing…pdf'
[72] Page 7 of QPS Scientific Officer witness statement dated 03/12/2014.