STATEMENT

(CJ Act 1967, s.9: MC Act 1980, ss.5A(3) (a) and 5B; MC Rules 1981, r.70)

Statement of Dan E. Krane, PhD

Age: over 18.
Occupation: Associate Professor of Biological Sciences, Wright State University, Dayton, OH, 45385, USA and President/CEO of Forensic Bioinformatics, Inc., 2850 Presidential Drive, Fairborn, OH 45324, USA.

This statement consists of 94 pages, the first and last of which are signed and dated by me. It is true to the best of my knowledge and is based on information received through scientists at The Forensic Institute and those legally representing Mr. Sean Hoey. If any of this information changes then it may cause me to change my opinion. I understand that my first duty is to the Court to provide independent, unbiased opinion, and that I may be prosecuted if I have willfully stated anything that I know to be false or do not believe to be true.

Qualifications: I am an Associate Professor in the Department of Biological Sciences at Wright State University in Dayton, Ohio. I have a B.S. degree with a double major in Biology and Chemistry from John Carroll University (Cleveland, Ohio), and a Ph.D. from the Biochemistry program of the Cell and Molecular Biology Department of the Pennsylvania State University (State College, Pennsylvania). I have also done postdoctoral research using the tools of molecular biology to answer questions in the fields of population genetics and molecular evolution in the Genetics Department of the Washington University Medical School (St. Louis, Missouri) and in the Department of Organismic and Evolutionary Biology of Harvard University (Cambridge, Massachusetts). I have published more than 30 scholarly papers in a variety of topics including population genetic studies of the genetic diversity of human populations at DNA typing loci, of organisms exposed to environmental stressors, and the use of DNA typing in forensic science. I am also the lead author of a widely-used undergraduate textbook, *Fundamental Concepts of Bioinformatics*. I am a member of the Commonwealth of Virginia's Scientific Advisory Committee, a 12-member panel established by statute to provide oversight and guidance to the Virginia Department of Forensic Science (the crime laboratory for the Commonwealth of Virginia). I have testified in over 60 criminal proceedings that have involved forensic DNA typing (in 22 different states and in three different Federal courts within the United States as well as a Coronial Inquest in the State of Victoria in Australia).

Signed: Date: October 24, 2006

Low-copy number (LCN) DNA profiling in the case of Sean Hoey

Forensic DNA testing in general

Deoxyribonucleic acid, or DNA, is a long, double-stranded molecule configured like a twisted ladder or "double helix." The genetic information of all organisms is encoded in the sequence of four organic compounds (bases) that make up the rungs of the DNA ladder. Most DNA is tightly packed into structures called chromosomes in the nuclei of cells. In humans there are 23 pairs of chromosomes; half of each pair is inherited from the individual's mother, half from the father. The total complement of DNA is called the genome.

By some estimates, 99.5% of the genetic code is the same in all humans. The roughly 0.5% difference in the nucleotide sequence between two people are not evenly distributed across the human genome. To identify individuals, DNA tests focus on a few loci (plural of locus—a specific location on the human genome) where there is variation among individuals. Loci where these differences are found are said to be polymorphic because the genetic code can take different forms in different individuals. Each possible form is called an allele.

The late 1990s saw the advent of STR (short tandem repeat) DNA testing. STR tests combine the sensitivity of a polymerase chain reaction- (PCR-) based test with great discriminating power (profile frequencies are commonly as low as one in quadrillions or quintillions). PCR is a procedure that allows a small amount of DNA (which by itself would not be enough to type) to be amplified into an amount large enough for typing. It does this by making billions of copies of DNA fragments from a polymorphic area (or areas) of the genome. PCR is not a genetic test itself, but merely a tool to increase the amount of genetic material to be tested.

The "amplification" of DNA takes place in a test tube. The DNA that is extracted from each sample is placed in a separate tube, along with a mixture of primers, enzymes,

and other reagents. The tubes are then placed in a machine known as a thermal cycler, which can control their temperature precisely while going through a series of heating and cooling cycles. Each cycle has three steps. First, the tubes are heated to approximately 94 degrees Celsius. At this temperature the DNA denatures – that is, the double-stranded molecule "unzips" to form two complementary single strands. In the second step, the tubes are cooled to about 60 degrees Celsius. At this temperature the primers anneal (bind) to the single strands of DNA. The primers are single stranded-DNA molecules that are complementary to specific target areas on the single stands of human genomic DNA. The primers are designed to anneal at positions that flank the polymorphic areas to be amplified, thereby marking those areas. In the third step, the tubes are heated to about 72 degrees Celsius. At this temperature, an enzyme known as *Taq* DNA polymerase acts as a catalyst, causing single DNA strands in the areas flanked by the primers to bind to complementary bases that are floating in the solution. Each single strand of DNA from the marked areas thus becomes one side of a new double strand. When this process is completed, the number of identical double strands of DNA from the polymorphic areas is twice what it was at the beginning of the cycle.

This three-step cycle is repeated 28 times (in conventional Profiler Plus and SGM+ testing), doubling the number of copies of the target DNA each time, and producing literally billions of copies. The target DNA (from a polymorphic area, or areas), which was initially like a needle in a haystack of other DNA, is amplified to the point that there are far more needles than hay, at which point the needles can be typed using a variety of methods.

An STR is a DNA locus that contains a length polymorphism. At each STR locus, people have two alleles (one from each parent) that vary in length depending on the number of repetitions of a short core sequence of genetic code. A person with genotype 14, 15 at an STR locus has one allele with 14 repeating units, and another with 15 repeating units (Figure 1).
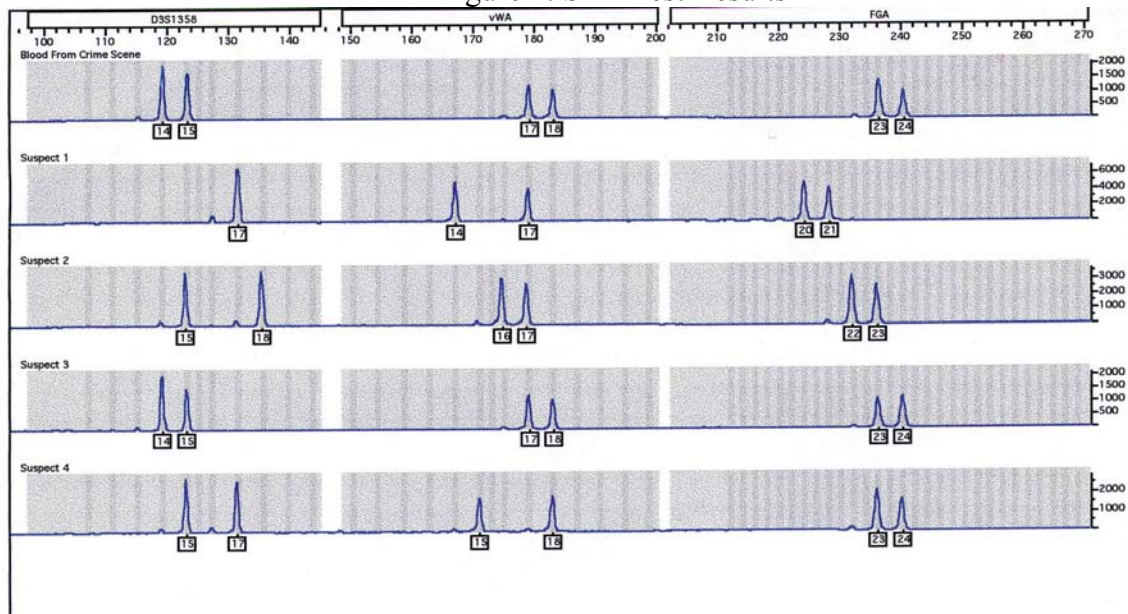
Figure 1: STR Test Results

Figure 1. Results of STR analysis of five samples: blood from a crime scene and reference samples of four suspects. This analysis includes three loci, labeled "D3S1358," "vWA," and "FGA." Each person has two alleles (peaks) at each locus, one inherited from their mother and the other from the father. The position of the "peaks" on each graph (known as an electropherogram) indicates the length (and hence the number of core sequence repeats) of each STR. As can be seen, the profile of suspect 3 corresponds to that of the crime scene sample, indicating he is a possible source. Suspects 1, 2 and 4 would be eliminated as possible sources in conventional STR-testing.

In 1997, the United State's Federal Bureau of Investigation (FBI) identified 13 STR loci that it deemed appropriate for forensic testing. Commercial firms quickly developed test kits and automated equipment for typing these STRs. Applied Biosystems Incorporated (ABI) developed what have become the most popular procedures including the PCR test kits known as Profiler Plus and SGM+ that simultaneously "amplifies" DNA from multiple STR loci and labels the loci with colored dyes. An automated test instrument such as the ABI 377 and ABI 310 Genetic Analyzers then separates the resulting amplicons by length (using electrophoresis) and uses a laser to cause fluorescence of the dye-labeled fragments. Software working with the output of a computer-controlled electronic camera detects the brightness and relative position of the fragments, identifies alleles, and displays the results as shown in Figure 1.

STR tests have greatly improved the capabilities of forensic laboratories, allowing highly specific DNA profiles to be derived from tiny quantities of cellular material. Test results often allow a clear-cut determination of whether a particular individual could or could not be the source of an evidentiary sample. However, experts have differed over interpretation of results in some cases even where conventional STR testing has been performed – particularly those involving mixed samples (DNA from more than one person) and low quantities of DNA.

PCR-based limitations of DNA profiling

Conventional STR profiling such as that performed with the Profiler Plus and SGM+ test kits is very widely used by crime laboratories around the world. Applied Biosystems, the manufacturer of the SGM+ test kits used in this case, recommends that a template amount (or starting quantity) of DNA should be between 1.0 to 2.5 nanograms (1 nanogram = 1×10^{-9} gram; ng) in order for their kits to produce accurate and reliable results. A diploid human cell contains roughly 6.6 picograms (1 picogram = 1×10^{-12} gram; pg) of genomic DNA such that 1 ng of human DNA comes from approximately 152 diploid cells. Use of more than 2.5 ng of DNA template is known to give rise to a variety of issues that can complicate the interpretation of DNA test results including: the presence of off-scale peaks; split peaks (such as those that arise due to a technical artifact known as +A; see Figure 2); and imbalance in the height of peaks between loci. Use of less than 1.0 ng of template DNA is also known to give rise to a set of issues that can complicate the interpretation of DNA test results including: imbalance of the height of peaks within loci; imbalance of the height of peaks between loci; and allelic drop-out (Figure 2) and drop-in.
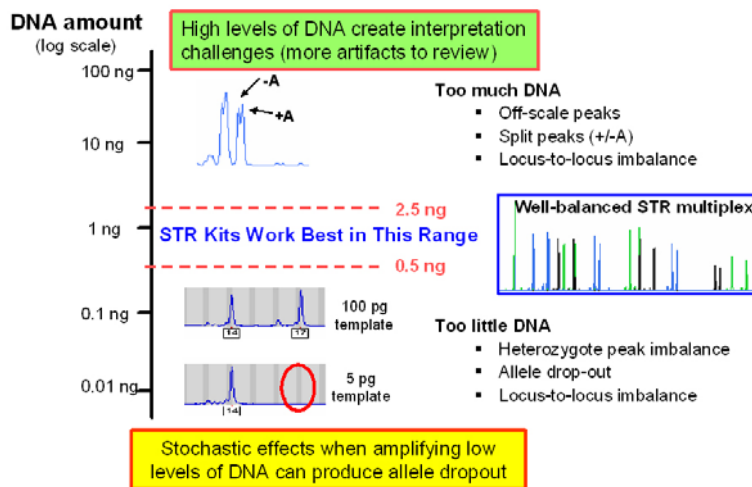
Figure 2. Impact of using non-optimal amounts of template DNA for PCR amplification. Derived from a presentation by Dr. John Butler (United States' National Institute of Standards and Technology, Office of Law Enforcement Standards) at the Midwestern Academy of Forensic Sciences annual meeting in Indianapolis, Indiana on October 11, 2006

As long ago as 1992, Walsh et al. ("Preferential PCR amplification of alleles: Mechanisms and solutions." PCR Meth. Appl. 1992;1:241-250) recognized that use of quantities of DNA below the recommended 1.0 ng starting amount could be problematic. At the heart of the problems associated with small quantities of template are what are called "stochastic" effects (sometimes also referred to as "preferential amplification"). These stochastic effects essentially arise from sampling errors that can occur when very few samples are made (much like those that might happen when blindly drawing black and white beans from a bag – a small number of draws, but not a large number, might suggest that all the beans in the bag are black even though they account for only 50% of a large number of beans in the bag) (Figure 3). With DNA templates arising from fewer than 150 human cells (and conceivably as few as five to ten cells) it is possible that one of two alleles at a locus will be amplified by the PCR process more than its counterpart (resulting in peak height imbalance or even allelic drop-out). It is also possible that stray alleles originating from just a few contaminating cells could be amplified preferentially, just by chance, relative to those that actually come from an evidence sample (resulting in allelic drop-in). Another commonly observed stochastic effect associated with small amounts of starting DNA template is an increase in the prevalence of a technical artifact

known as "stutter" (which occurs when the enzyme responsible for making copies of the STR regions either slips forward or back during the amplification process and makes a copy that is either one repeat unit shorter or longer, respectively, than what was actually present in the template DNA). Quite simply, if there is not enough starting quantity of DNA prior to the time the PCR is started, the PCR process can produce results that are inaccurate and unreliable.
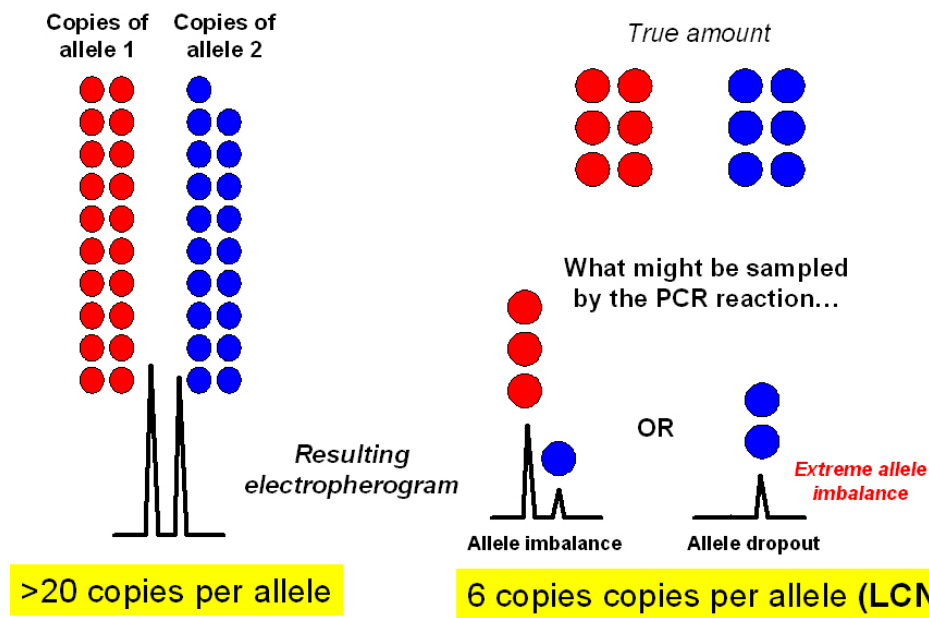

Figure 3. Stochastic statistical sampling during PCR amplification. When more than 20 copies of an allele are being amplified at the start of a PCR amplification it is unlikely that sampling errors will result in a large difference in the number of copies that are made of two different alleles. Sampling errors are much more likely to occur when fewer copies of the alleles are present at the start of the process. Derived from a presentation by Dr. John Butler (United States' National Institute of Standards and Technology, Office of Law Enforcement Standards) at the Midwestern Academy of Forensic Sciences annual meeting in Indianapolis, Indiana on October 11, 2006.

The propensity for stochastic effects in LCN analyses is well-documented in the scientific literature (i.e. Figure 4). Many have pointed to the problems associated with LCN testing but Dr. Bruce Budowle, a senior scientist of the United States' Federal Bureau of Investigation, summarizes the issues succinctly when he says:

"Because of the successes encountered with STR typing, it was inevitable

that some individuals would endeavor to type samples containing very

minute amounts of DNA. …. When few copies of DNA template are

present, stochastic amplification may occur, resulting in either a

substantial imbalance of two alleles at a given heterozygous locus or

allelic dropout." (Budowle et al., Low copy number – consideration and

caution. Proc. 12th International Symposium on Human Identification,

2001).

(Dr. Budowle also points out that "Mixture analyses and confirmation of a mixture are

not reliable with LCN typing, because of imbalance of heterozygote alleles, increased

production of stutter products, and allele drop-in can occur.")

Sensitivity Series - 32 cycles

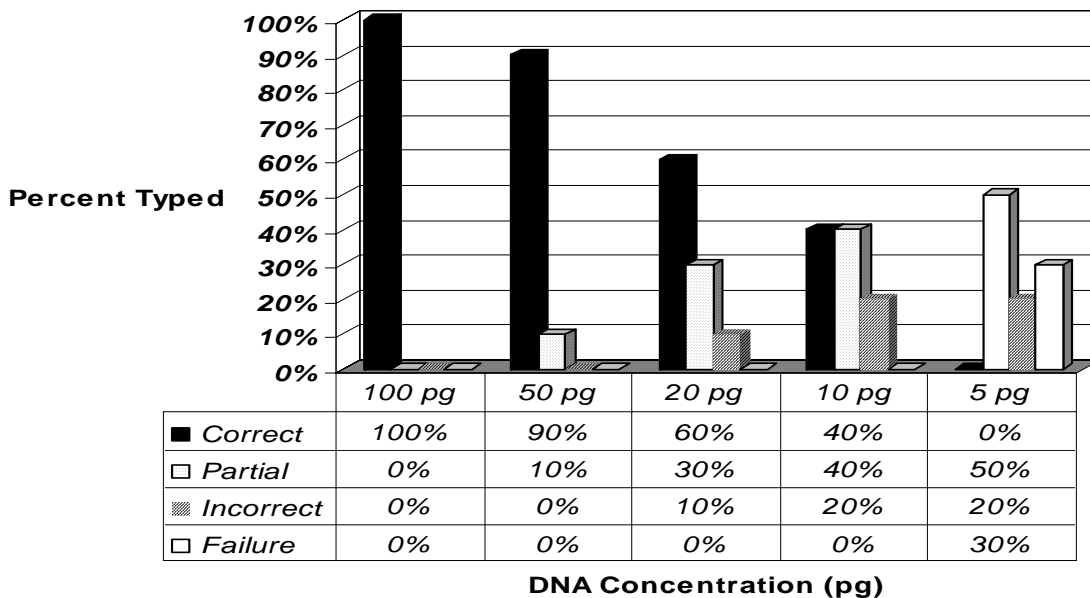| Percent Typed	100 pg	50 pg	20 pg	10 pg	5 pg
■ Correct	100%	90%	60%	40%	0%
□ Partial	0%	10%	30%	40%	50%
▦ Incorrect	0%	0%	10%	20%	20%
□ Failure	0%	0%	0%	0%	30%

DNA Concentration (pg)

Figure 4. Problems with obtaining correct allele calls at low DNA levels. Taken from M. Coble and J. Butler. "Characterization of new MiniSTR loci to aid analysis of degraded DNA." J. Forensic Sci. 2005; 50(1):43-53. Human genomic DNA with a very well-characterized concentration and a known DNA profile was used to deliver either 100, 50, 20, 10 or 5 pg of template DNA for LCN amplification with 32 cycles of PCR. Each sample was independently amplified ten times.

Because of the widely appreciated problems associated with stochastic effects in

LCN testing, the Applied Biosystems STR test kits have actually been deliberately

designed to fail when less than 125 pgs of DNA (that from roughly 20 diploid human cells) is used as template. This feature of the test kits specifically seeks to avoid these problems at the "stochastic limit of quantitation" (an amount of template that causes the chance of stochastic effects to be so likely as to make the test results incorrect) for the kits. The TWGDAM (Technical Working Group on DNA Analysis and Methods) validation of the AmpFISTR Blue test kit (J.M. Wallin et al. "TWGDAM validation of the AMpFISTR™ blue PCR amplification kit for forensic casework analysis." J. Forensic Sci. 1998;43(4):854-870) explored the use of 27 and 30 PCR amplification cycles. This validation study for the test kit that was the precursor of the Profiler Plus and SGM+ test kit ultimately settled upon using 28 cycles so that quantities of DNA below 35 pg gave very low or no peaks so as to avoid situations where peak imbalance results in only one detectable allele from a heterozygous pair and concerns regarding trace contaminants could be minimized. It was assumed that 35 pg constituted an "analytical threshold."

One of the easiest ways that a PCR-based test can be made more sensitive is by simply increasing the amount of DNA replicating enzyme present and/or the number of PCR amplification cycles that are employed (Figure 5). Even if concerns regarding stochastic effects can be set aside, increasing the sensitivity of any test comes always with the consequences of greater difficulty in distinguishing between signal and noise and other random factors. Any such a change in a testing protocol would require that rigorous developmental studies be performed. I have asked for but have not received validation studies performed by the Forensic Science Service that supports such a change.
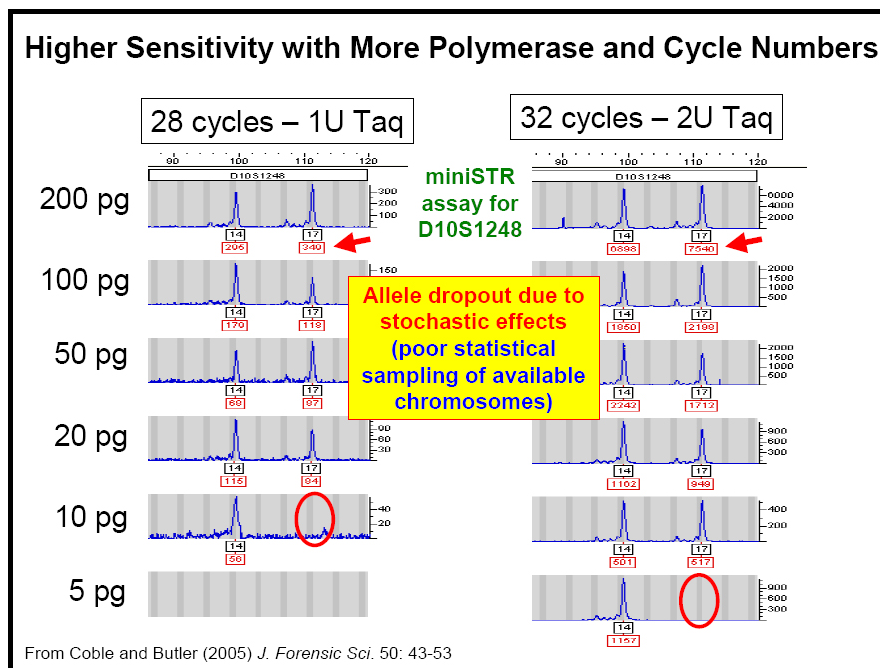
Figure 5. Higher sensitivity with more enzyme and rounds of PCR amplification. Taken from Coble and Butler (2005), *J. Forensic Sci.* 50:43-53. Derived from a presentation by Dr. John Butler (United States' National Institute of Standards and Technology, Office of Law Enforcement Standards) at the Midwestern Academy of Forensic Sciences annual meeting in Indianapolis, Indiana on October 11, 2006.

Integrity of material

Even under the very best of circumstances (virtually the opposite of those that require ultra sensitive LCN testing to be attempted), the presence of a DNA profile usually says nothing about the time frame or circumstances under which the DNA was transferred to an item. Further, the chance of "innocent" DNA transfer greatly increases as the amount of starting material for DNA profiling tests becomes smaller. Quite simply, even if one's DNA is found to be associated with an article of evidence, the great sensitivity of LCN testing causes very real questions to arise regarding both how and when that DNA was transferred.

DNA in quantities at LCN levels can easily be transferred from one article to another (e.g. from evidence sample, onto the analyst's lab coat or gloves, then to another evidence sample; or by having been stored together in a single package) – without the contributor having any knowledge that the transfer(s) has occurred. No DNA tests are currently able to distinguish between secondary transfer (such as the transfer of DNA through contamination events) or DNA present due to direct contact with an object. Similarly, DNA tests are not currently capable of distinguishing in any way between the presence of DNA due to contamination (such as could very easily occur through storing or opening the objects in the same location as items obtained from an individual) or direct contact between an individual and the object. Given that LCN analyses can conceivably generate results from as little material as a single cell of an individual, the only way to be confident that results have not been obtained solely through contamination is to demonstrate conclusively with continuity records that contamination is not even remotely possible.

The differing paradigms of conventional and LCN testing

The greater the possibility of persistence and/or transfer of DNA from and between people and items, the greater the reduction in the probative value of DNA test results. I am aware of at least one (Australia's State of Victoria's Coroner's Inquest into the death of Jaidyn Leskie – a complete copy of the Coroner's ruling can be found at:) www.bioforensics.com/articles/Leskie_decision.pdf) very well documented instance of DNA transfer between evidence samples in two separate cases that resulted in the generation of a spurious DNA profile (tested by conventional STR-approaches) without any associated failures of testing controls (such as reagent blanks and negative controls).

There should be no question but that the extreme sensitivity of detection of ultra sensitive LCN tests causes the possibility of persistence and/or transfer of DNA from and between people and items to be unquestionably much greater than that which occurs in more conventional DNA profiling.

Conventional DNA profiling's greatest strength comes ultimately from its great ability to <u>either</u> include or to exclude an individual as a possible contributor to an evidence sample. In conventional PCR-based DNA profiling, the presence of alleles in an evidence sample that are not observed in a suspect and/or the presence of alleles in a suspect that are not observed in an evidence sample that is unlikely to have experienced allelic drop-out would normally exclude a suspect as a possible contributor to an evidence sample. But, because LCN testing operates at or beneath a stochastic limit of quantitation for a PCR-based amplification system, any number of such differences between evidence samples and a suspect's reference can be easily (and reasonably) attributed to allelic drop-out, allelic drop-in, or other stochastic effects (Figure 6).

In conventional forensic DNA analysis, attempts are made to exclude a suspect, and only after failing to exclude, are inferences made regarding the rarity of the observed DNA profile. However, the statistical interpretation of LCN test results is significantly complicated by the expectation of stochastic effects. While inherent flexibility of interpretation of LCN testing results may be useful for the generation of investigative leads to which no statistical weights are attached, that flexibility of interpretation makes it unsuitable for establishing proof that an individual was directly associated with an evidentiary sample (let alone, what the time frame and circumstances of such an association may have been).
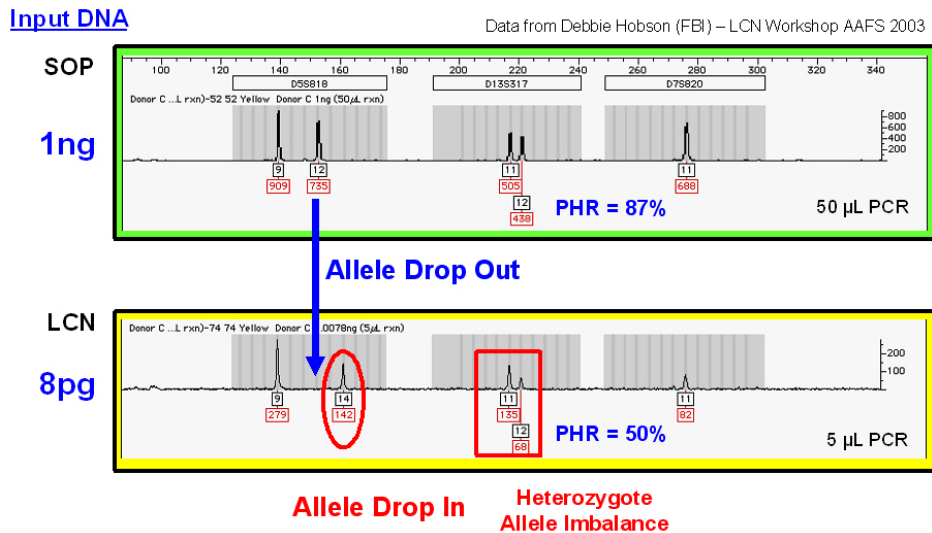
Figure 6. Comparison of conventional STR amplification and LCN amplification with the same test kit and source of human DNA. With conventional STR testing (SOP) the true DNA profile of an individual can be determined directly from an electropherogram. LCN testing of the same individual with the same test kit gives rise to a significantly different electropherogram. Derived from a presentation by Dr. John Butler (United States' National Institute of Standards and Technology, Office of Law Enforcement Standards) at the Midwestern Academy of Forensic Sciences annual meeting in Indianapolis, Indiana on October 11, 2006.

Quite simply, any of the differences between two DNA profiles that would have caused them to be deemed non-matching by conventional testing can be explained away by the expectation of stochastic effects in LCN testing (Figure 6). These factors, specifically addressed for this case in appendix A (Technical issues associated with DNA profiling), include but are not limited to: mixed samples; peak height imbalance; noise; pull-up; degradation/inhibition; allelic drop-out and drop-in; and stutter. At the very least, each of these factors can lead to alternative interpretations of the evidence and, consequently, to a compounded reduction of the probative value of the test results.

Detection instrument-based limitations of DNA profiling

Independent of problems associated with using quantities of template DNA at the analytical threshold of a test kit like SGM+ are problems that arise when an analytical system (such as a capillary electrophoresis machine) is used at the very limits of its sensitivity. Here again there can be difficulties in reliably distinguishing between signal from the sample and the spurious background signal of the system. This challenge is not unique to LCN analyses and is one that has been rigorously addressed in numerous other analytical disciplines. Good practice dictates that testing laboratories perform validation studies to establish criteria for distinguishing between reliable signal and noise/technical artifacts. Many DNA testing laboratories such as the Forensic Science Service have settled upon "minimum peak height thresholds" as a *de facto* approach to dealing with this problem even though these thresholds generally fail to consider variability in the sensitivity of instruments, reagents, and the skill of human analysts involved in the DNA profiling process over the course of time – all particularly important parameters given the extraordinary sensitivity to which LCN analyses aspire.

My colleagues and I have performed a validation study on an analytical chemistry-based approach to distinguishing between signal and noise in capillary electrophoresis-based DNA profiling. That study has been accepted for publication in the January issue of the *Journal of Forensic Sciences* and is attached to this statement as Appendix C. The alternative to minimum peak height thresholds that we describe involves establishing limits of detection and quantitation from information associated with the control experiments performed as part of each DNA profiling run. This approach has the distinct advantages of being run-specific (meaning that it is sensitive to the functioning of the testing system at the very time that evidence samples are being analyzed rather than to what it was at the time of a laboratory's initial validation studies) and totally objective (meaning that it is independent of any information associated with other electropherograms such as those from a suspect's reference sample).

Use of LCN testing

I am aware of only one crime laboratory in the United States that uses LCN testing approaches in its handling of evidentiary samples – the New York City crime laboratory. It is my understanding that its application of LCN testing differs in many ways from that of the Forensic Science Service but in three that I feel are particularly important: 1) the New York City laboratory has undertaken a rigorous developmental validation process; 2) it generates consensus DNA profiles on the basis of a minimum of three (not two) independent amplifications; and 3) it performs LCN testing exclusively as an investigative tool – not as a means of generating proof that an individual has been associated with an evidence sample.

Interpretation of LCN testing results

I have been asked to consider the LCN results generated in the case of Sean Hoey as if the concerns about LCN testing that I have described above did not exist, or had been addressed – which I do not consider to be the case. This statement is therefore based on the very strongly questioned premise that all is well with the LCN testing regime as applied by the Forensic Science Service in the case of Sean Hoey.

I understand that the FSS has utilized a minimum peak height threshold (occasionally as low as 25 RFUs) in their analysis. These minimum peak height thresholds are unusually low – especially troublesome in light of the extraordinary sensitivity of LCN testing and the extremely large amounts of background noise observed in many of their positive control samples. In place of a minimum peak height threshold, we have used a limit of detection (LOD) akin to those commonly used in other analytical disciplines. Positive control samples included with each run were analyzed in the fashion

described in our validation study (attached as Appendix C; accepted for publication in the January, 2007 issue of the *Journal of Forensic Sciences*). Where two (or more) positive control samples were available, the one with the highest LOD was relied upon in an effort to conservatively determine the testing system's minimum ability to reliably distinguish between signal and noise during the course of that experimental run.

I also understand from a review of Dr. Whitaker's previous testimony that the FSS requires that an allele be observed in both of the two runs of a sample before that allele can be reliably determined to be present on an evidentiary sample [e.g. – "it's been demonstrated again through research that this provides a more robust and reliable results (sic)"; "we only report or communicate DNA band (sic) which are observed in both of those tests, so the result has got to be seen twice before it's deemed reliably (sic) to provide as evidence"]. We have strictly applied this same criterion in our interpretations.

The strict application of a set of interpretation guidelines and our care to distinguish between signal and noise with a limit of detection derived from positive controls associated with each sample run has resulted in significantly different allele calls relative to those made by the FSS (Table 1). This is the case even though we have deliberately simplified our analysis to make it more consistent with that of the FSS by setting aside numerous other concerns (detailed and illustrated in Appendix B) about the information contained in the underlying electropherograms that would have been of concern for conventional SGM+ testing.

Sample	Reviewer	LOD 1	LOD 2	AMEL	D3	vWA	D16	D2	D8	D21	D18	D19	THO1	FGA
Sean Hoey				X Y	16 17	17 17	10 13	17 20	10 13	30 30	15 16	14 15	9.3 9.3	23 24
PM5-battery housing	FBS	331	1230, 58	X	16 17							14		
PM5-battery housing	FSS	25	50, 25	X Y	16 17	17	10 13	20	10 13	28 30		12 13 14 15	9.3	23 24
PM5-brown tape-1	FBS	82	56	X Y	16 17	17	13	17	10 [13]	30	15	14 15	9.3	[23], [24]
PM5-brown tape-1	FSS	50	25	X Y	16 17	17 F	13 F	17 F	10 13	30 F	15 F	14 15	9.3 F	24 F
PM5-brown tape-2	FBS	82	56	X Y	16 17	17	[10] 13	17 20	10 13	30	15 16	14 15	9.3	23 24
PM5-brown tape-2	FSS	50	25	X Y	16 17	17 F	10 13	17 20	10 13	30 F	15 16	14 15	9.3 F	23 24
PM5-blue-tape-swab-1	FBS	82	56	X Y	16 17	17	10		[10] 13	30		14 15	9.3	[23] [24]
PM5-blue-tape-swab-1	FSS	50	25	X Y	16 17	17 F	10 F		10 13	30 F		14 F (15)	9.3 F	(23) 24
PM5-blue-tape-swab-2	FBS	82	56	X Y	16 17	16 17	12 13		10 13	28 30	15	12 13 14 15 16	9.3	21 22 23 24
PM5-blue-tape-swab-2	FSS	25	25	X Y	16 17	17 F	(13)		10 13	(28) 30 F	15 F		9.3 F	(21)(22)23 24
PM5-other brown tape-1	FBS	255	1263	X Y	16 17	17	10	17	10 13			14 15	9.3	
PM5-other brown tape-1	FSS	50	25	X Y	16 17	17 F	10 13	17 20	10 13	30 F	15 16	14 15	9.3 F	23 24
PM5-other brown tape-2	FBS	255	1263	none										
PM5-other brown tape-2	FSS	50	25	missing										
MDH4-red-brown tape	FBS	255	137	X Y	16 17	16 17	10 13		10 13	30	15	14 15 17	9.3	
MDH4-red-brown tape	FSS	50	50	X Y	16 F (17)	17 F	(10)(13)	17 (20)	10 13	30 F	15 16	14 15	9.3 F	23 F (24)
MDH4-tube end	FBS	255	137	none										
MDH4-tube end	FSS	25	50	X Y	(16) 17 F	17 F								
MDH5-swab 1	FBS	56	137	X Y	16 17	17			10 13	30	15 16	13 14 15	9.3	21 24
MDH5-swab 1	FSS	25	50	X Y	16 17	17 F	13 F		10 13	30 F	15 16	14 15	9.3 F	24 F
MDH5-swab 2	FBS	56	137	X Y	16 17	17	10 [13]	17	10 13	30	15	14 15	9.3	23 24
MDH5-swab 2	FSS	25	50	X Y	16 17	17 F	10 13	17 20	10 13	30 F	15 16	14 15	9.3 F	23 24
JRJ3-tape 1	FBS	362	362, 364	none										
JRJ3-tape 1	FSS	50	50, 50	none										
JRJ8-tape 2	FBS	320	547	X Y	16 17	17	[10] [13]	17 20	10 13	30	15 16	14 15	9.3	[23] [24]
JRJ8-tape 2	FSS	25	25	X Y	16 17	17	10 13	17 20	10 13	30	15 16	14 15	9.3	23 24
FC/2-reachable bulb	FBS	2011	1502	X		17								
FC/2-reachable bulb	FSS	50	50	X Y	16 17	17	10 13	17 20	10 13	30	15 16	14 15	9.3	23 24

Table 1. Allele calls for the critical evidence samples in the case of Sean Hoey as made by FBS and the FSS. The first column of the table describes the evidence sample that was tested with the LCN approach. Columns labeled LOD1 and LOD2 list the limit of detection (in RFUs) determined from the sample's positive control(s) or the minimum peak height threshold that appears to have been used by the FSS. Allele designations in parentheses (#) for the FSS allele calls were determined to differ markedly in size from other peaks in that locus. Allele designations in brackets [#] for the FBS allele calls are shown only for seemingly unmixed samples that displayed dramatic differences (<60%) in peak height relative to the other allele observed at that locus in one or more of the runs of the sample.

The net result is that we frequently find that Sean Hoey is generally *excluded* as being a possible contributor to many of the 16 articles of evidence upon which LCN tests were performed. Two notable exceptions to this general trend are samples "PM5-brown tape-2" and "JRJ8-tape 2" where Sean Hoey cannot be excluded as being a possible contributor and the chances of a coincidental match are quite small. Issues associated with continuity and LCN testing in general need to be particularly carefully considered for these two samples. For instance, a control swab (QC00.0216.07) associated with a different item of evidence yet analyzed in the very same electrophoresis run as the PM5-brown tape-2 evidence sample exhibits a complete DNA profile (that also happens to have many alleles in common with Mr. Hoey). The presence of a DNA profile in a control sample suggests that contamination is a significant concern and raises serious questions about the reliability of any associated test results. It is not possible to determine when the contaminating DNA was introduced and it must be considered that tubes with other samples have been contaminated with other sources of DNA.

As mentioned above, many, many specific issues separate from those associated with LCN testing and limits of detection/quantitation in general have been identified with the electropherograms generated by the Forensic Science Services during the course of their testing of the materials associated with this case. An automated system (Genophiler®) used by a DNA profiling consulting company for which I am the President (Forensic Bioinformatics, Inc.) has drawn attention to each of those issues and I have had Jason Gilder, an employee of Forensic Bioinformatics generate an illustrated version of Genophiler's findings (attached as appendix B).

Appendix A (Technical issues associated with DNA profiling)

Mixed samples

In the case of a high-quality, single source sample and barring the possibility of error, STR analysis can provide compelling statistical evidence that an observed correspondence between an evidentiary sample and a particular individual is very unlikely to be the result of coincidence. However, many evidentiary samples are comprised of mixtures of two or more individuals' DNA and their interpretation can be significantly more challenging. Consider a locus where three alleles (such as the D3S1358 locus in QC01.0094.22, the PM5 battery housing, with a 16, 17 and 18 allele) are observed. Even if it is known that exactly two persons contributed genetic material to this sample, six different pair-wise combinations of genotypes are qualitatively consistent with the observation of these three alleles: (1) 16, 16 and 17, 18; (2) 16, 17 and 18, 18; (3) 16, 17, and 17, 18; (4) 16, 17 and 16, 18; (5) 16, 18 and 17, 17; and (6) 16, 18 and 17, 18. Interpretation becomes even more difficult when any assumption regarding the number of contributors to a mixed DNA sample is not made (e.g. the three alleles observed at the D16S539 locus could represent a mixture of three individuals with genotypes: 16, 16; 16, 17; and 17, 18). Unfortunately, the potential for alleles to be shared between individuals limits the ability of simple counting techniques (such as those that have been applied by Dr. Whitaker previously) to correctly infer the number of contributors to mixed samples (Paoletti, et al. "Empirical analysis of the STR profiles resulting from conceptual mixtures." J Forensic Sci. 2005; 50(6):1361-1366). These

inferences are particularly unreliable when seemingly discordant loci can be disregarded and/or allelic drop-out occurs at a high rate (discussed in greater detail below).

Peak height imbalance

Since a DNA sample usually consists of intact cells, the level of DNA present from a given contributor is generally constant. Numerous studies have shown that the signal associated with each allele in a given location should be roughly equivalent. General practice has found that "[t]he peak height ratio, as measured by dividing the height of the lower quantity peak in relative fluorescence units by the height of the higher quantity allele peak, should be greater than approximately 70% in a single source sample" (Butler, 2001, Forensic DNA Typing. San Diego: Academic Press). Deviation from peak balance is an indication of more than one contributor to a DNA sample. Consider the FGA locus from sample QC03.0158.11, exhibiting a 23 and 24 peak at 749 and 1,339 RFUs, respectively. The peak height ratio of 56% [very close to the maximum amount of imbalance reported in a report on LCN testing by Whitaker et al. ("A comparison of the characteristics of profiles produced with the AMPFISTR® SGM Plus™ multiplex system for both standard and low copy number (LCN) STR DNA analysis." Forensic Sci. Int. 2001; 123:215-223)] indicates a peak height imbalance and the possible presence of a secondary contributor.

Baseline noise

Any measurement made with a light-detecting instrument, such as a genetic analyzer is subject to at least some level of background noise – defined here as signal not associated with amplified DNA. Instrument-related factors that may contribute to

background noise in DNA testing experiments are typically run-specific and include (but are not necessarily limited to): the age and condition of the polymer and capillary being used; dirty capillary windows; and dirty pump blocks. Background noise may also differ between instruments due to differences in CCD (charged couple device) detectors, laser effectiveness and alignment, and cleanliness and alignment of the optical components. Many amplification-related factors that contribute to background noise (such as analyst skill and stocks of chemicals) are also run-specific and might be reasonably expected to have varying impacts over time.

These factors can create random fluctuations that are occasionally large enough to be confused with an actual peak or to mask actual peaks. For example, approximately the first half of the blue channel of sample QC01.0067.26 exhibits a raised baseline, with peak heights above 100 RFUs. The peaks of a minor contributor to this sample may be masked/obscured by this raised baseline.

It is worth noting that two negative control samples, QC01.0109.35 and QC01.0109.36, exhibit precisely the same peak information, including RFU values. Baseline noise generally arises from random processes and it is extremely unlikely that they could work to generate precisely the same values across an entire sample. The striking similarity of these two negative controls is much more likely to be the result of a data handling or processing error.

Pull-up

Pull-up (sometimes referred to as bleed-through) represents a failure of the analysis software to discriminate between the different dye colors used during the generation of the test results. A signal from a locus labeled with blue dye, for example,

might mistakenly be interpreted as a yellow or green signal, thereby creating false peaks at the yellow or green loci. Pull-up can usually be identified through careful analysis of the position of peaks across the color spectrum, but there is a danger that pull-up will go unrecognized, particularly when the result it produces is consistent with what the analyst expected or wanted to find. For example, the sample QC00.0216.06 exhibits potential pull-up originating from the X and Y peaks in the amelogenin locus and giving rise to the 9.2 and 11 peaks in D19S433 of 1,983 and 2,046 RFUs, respectively.

Degradation/Inhibition

As samples age, DNA begins to break down (or degrade). This process can occur rapidly when the samples are exposed for even a short time to unfavorable conditions, such as warmth, moisture or sunlight. Similarly, inhibition occurs when chemicals associated with an evidence sample interfere with the activity of the enzyme used to amplify DNA during the course of DNA profiling experiments. Degradation and inhibition create difficulties in interpreting differences between peak heights. Generally, degradation and inhibition both produce a downward slope across the electropherograms in the height of peaks because they are more likely to interfere with the detection of longer sequences of repeated DNA (the alleles on the right side of the electropherogram) than shorter sequences (alleles on the left side). Degraded/inhibited samples can be difficult to interpret in that the height of some peaks are reduced to the point that they are too low to be distinguished reliably from background "noise" in the data while other peaks from the same sample are still present. The samples associated with items FC/2, JRJ8, MDH4, MDH5, and PM5 exhibit appreciable indications of degradation and/or inhibition.

Allelic drop-out and drop-in

Each individual is typically expected to possess either one or two alleles (versions of DNA molecules that are recognizably different) at each locus (a location within the human genome) that is tested. If one allele is observed, the individual is assumed to be homozygous for that locus (meaning that both their mother and father contributed the same version of that locus to that individual). If two alleles are observed, the individual is heterozygous (meaning that their mother and father contributed different versions of that locus). DNA profile information can be compared locus by locus (in the case of the SGM+ test kit, ten variable loci are examined) between a suspect and an evidentiary sample. If a suspect has the very same alleles that are observed in an evidentiary sample, the suspect <u>may</u> be the source of the evidentiary material. If a suspect's alleles are all observed in an evidentiary sample but there are also alleles in the evidentiary sample that the suspect does not possess, the suspect <u>may</u> be a contributor to the evidentiary sample but only if the sample if a mixture of material from two or more individuals (in which case allelic balance is often used to resolve the mixture). If a suspect has one or two alleles at a locus that are not observed at the same locus in an evidentiary sample, the suspect can be excluded as a possible contributor to the evidentiary sample. The possibility of allelic drop-out however means that a suspect cannot be eliminated as a possible contributor to a possibly mixed sample even if they possess alleles that are not observed in the evidence sample. (For example, if Sean Hoey is a contributor to sample QC01.0109.13, associated with "FC/2's reachable bulb," allelic drop-out must have occurred in the D2S1338 locus in order to account for his 17 allele being absent.) And, the possibility of allelic drop-in means that a suspect

cannot be excluded as a possible contributor to a sample even if the sample is not a

mixture and contains alleles that he does not possess. In sum, *no one* can be excluded as

a possible contributor to an LCN sample because the allelic drop-out and drop-in cannot

be independently verified – the only evidence that these phenomena have occurred is the

"inconsistency" that they purport to explain (the alternative interpretation, that the

profiles actually do not match, is often more favorable to the suspect).

The weights attached to the ultra sensitive LCN DNA evidence by the FSS in this

case appear to take the possibility of allelic drop-out (but not drop-in) into consideration

only by using the formula $2p$ (rather than p^2) to calculate genotype frequencies at loci

where only one allele is observed. Such an approach assumes a very high rate of allelic

drop-out at loci where only one allele is observed at the same time that the presumption

is made that absolutely no allelic drop-out has occurred at loci where two alleles are

observed (even though it is possible that the samples considered are mixtures of DNA

from two or more individuals). Professor David Balding (Imperial College, London)

and I have discussed an alternative statistical approach for incorporating the possibility

of allelic drop-out (and drop-in) in the weighting of DNA evidence. At a conference I

hosted in Dayton, Ohio in August of 2006, Professor Balding proposed the following

approach: Let **A** represent the allele seen at a locus from a crime scene sample; let **AB**

represent the alleles seen at the same locus from a suspect (*s*). Since **B** is not observed

in the evidence sample, the suspect is normally excluded as a possible contributor.

However, with LCN DNA testing it is possible that *s* is the source of the evidence

sample and the **B** allele suffered allelic drop-out. Similarly, the true source of the

evidence sample could have any genotype that includes an **A** allele. The likelihood ratio

that the evidence sample came from a source other than the suspect relative to coming

from the suspect can then be described as:

different sources
vs.
same source

$$\frac{p_A{}^2 + 2\,p_A\,(1 - D_A)\,\Sigma_{x \neq A}\,p_x D_x}{(1 - D_A)\,D_B}$$

where D_x is the drop-out probability for allele X (under the conditions to which the

sample was exposed). If D is the same for all X, then:

$$LR = p_A{}^2\,\frac{1 - 2D(1\text{-}D)}{(1 - D)D} + 2p_A$$

Interestingly, the likelihood ratio becomes much greater than 1 (implying a source

other than the suspect is more likely) when the probability of allelic drop-out is either

small [as is the case with conventional SGM+ testing – Whitaker et al. ("A comparison of

the characteristics of profiles produced with the AMPFISTR® SGM Plus™ multiplex

system for both standard and low copy number (LCN) STR DNA analysis." Forensic Sci.

Int. 2001; 123:215-223) reports that not one instance of allelic dropout was observed with

conventional testing during an LCN validation study] or large [as is the case with ultra

sensitive LCN DNA testing – Whitaker et al. ("A comparison of the characteristics of

profiles produced with the AMPFISTR® SGM Plus™ multiplex system for both standard

and low copy number (LCN) STR DNA analysis." Forensic Sci. Int. 2001; 123:215-223)

report locus-specific allelic drop out rates as high as 14.8% when using 12 and 25

picogram amounts of template]. Allelic drop-in can be statistically incorporated to the

weighting of ultra sensitive LCN DNA testing results in a similar fashion (with similar

results).

Stutter

Stutter peaks are small peaks that occur immediately before (and, to a lesser extent, after) a real peak. Stutter occurs as a by-product of the process used to amplify DNA from evidence samples. In samples known to be from a single source, stutter is identifiable by the strength of its signal and position. However, it is sometimes difficult to distinguish stutter peaks from a secondary contributor in samples that contain (or might contain) DNA from more than one person. For example, the sample QC00.0209.17 exhibits a 16 and 17 peak in the D3S1358 locus at 348 and 2,337 RFUs, respectively. The 16 peak may be a relatively large stutter product from the 17 peak (14.9%) or be an indication of a secondary contributor to the sample.

Appendix B (Illustrated Genophiler® report)

This appendix contains an illustrated summary of issues that have been identified by Genophiler®, an automated and objective system that uses the GeneScan and Genotyper software. While many of these issues (e.g. peak height imbalance, low peak heights, locus drop-out) identified by Genophiler® can be caused by stochastic effects associated with low amounts of DNA template, it is also possible that they are caused by other significant problems and/or alternative interpretations of the evidence samples. The fact that so many issues have been identified for all of the critical evidence samples raises further questions regarding both the reliability and the interpretation of the LCN testing that has been performed in this case.

PM5 – Battery Housing C00.0055.20
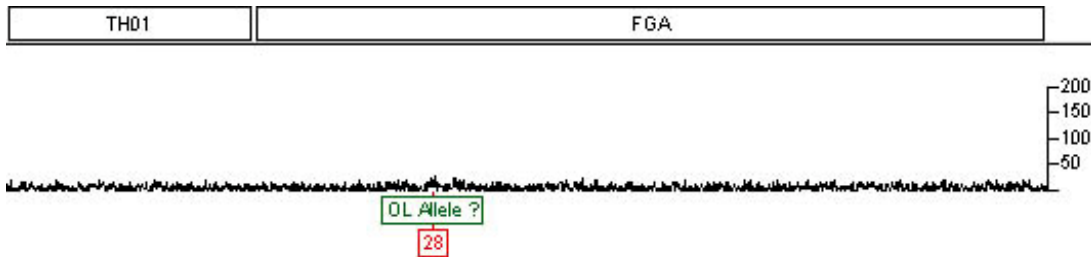
Has electropherograms that do not contain genotype information for 2 loci (THO1 and FGA). The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).
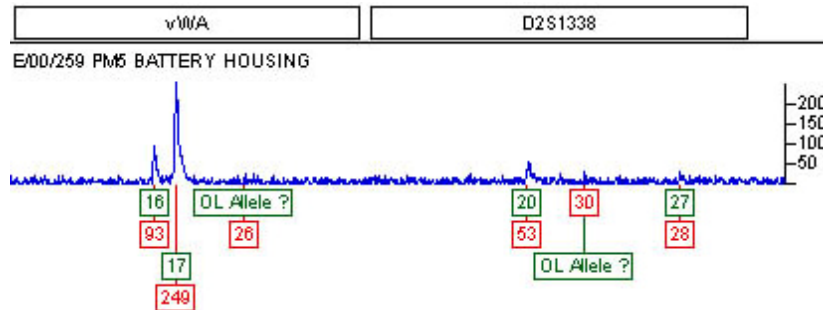
- Displays peak height imbalance at 2 loci (vWA, D2). The difference in the peak heights of the 16 and 17 alleles for the vWA locus (93 and 249, respectively) and the 20 and 27 alleles for the D2 locus (53 and 28, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
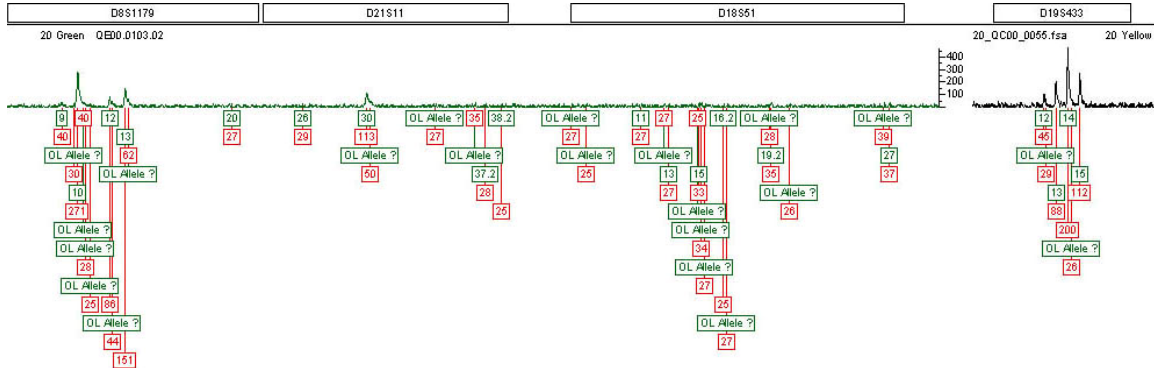
- May be the source of a mixture of two or more individuals. Several loci (D8, D21, D18, D19) appear to have more than two alleles. The additional peaks in this sample were found to be below the threshold of 150 RFUs, indicating that they are possibly caused by stochastic effects. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- Contains peaks with a low average peak height (93 RFUs). Peaks below 150 RFUs need to be interpreted with caution. The highest peak found in the sample is 465 RFU's.

PM5 – Battery Housing – QC00.0057.20

29
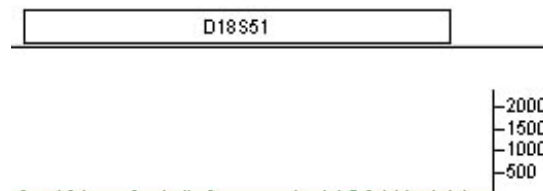
- Has electropherograms that do not contain genotype information for one locus (D18). The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).
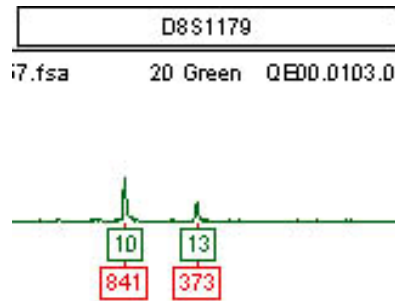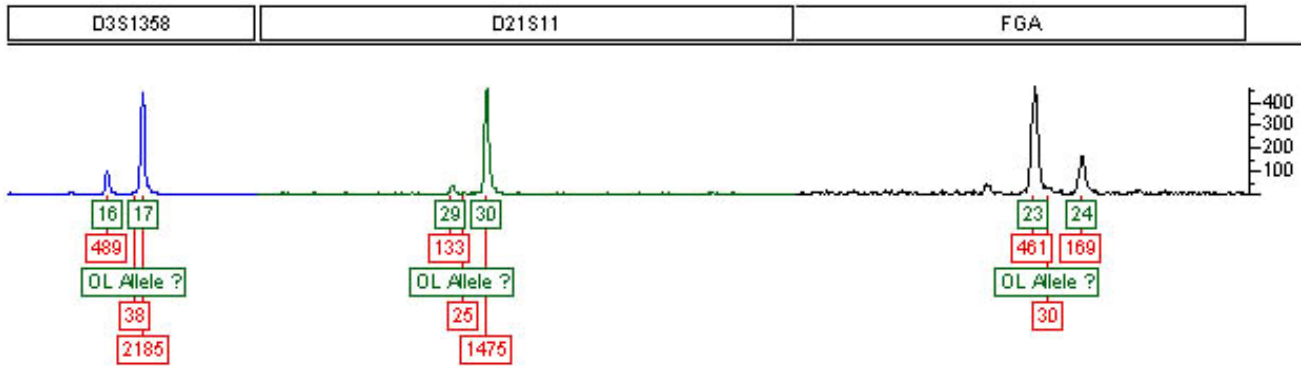

- Displays peak height imbalance at the locus D8. The difference in the peak heights of the 10 and 13 alleles for the D8 locus (841 and 373, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.
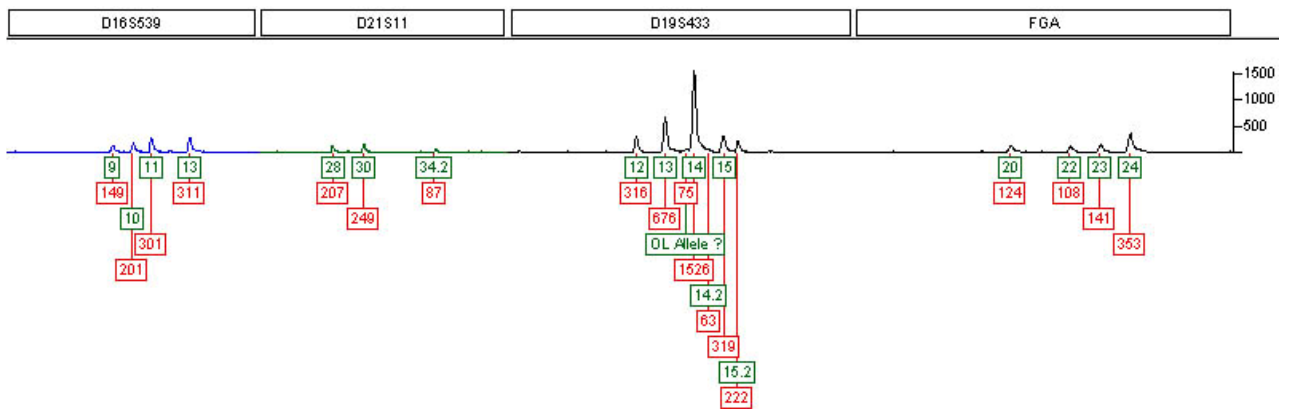
- May be a mixture of two or more individuals. Several loci (D16, D21, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

PM5 – Battery Housing – QC00.0094.22
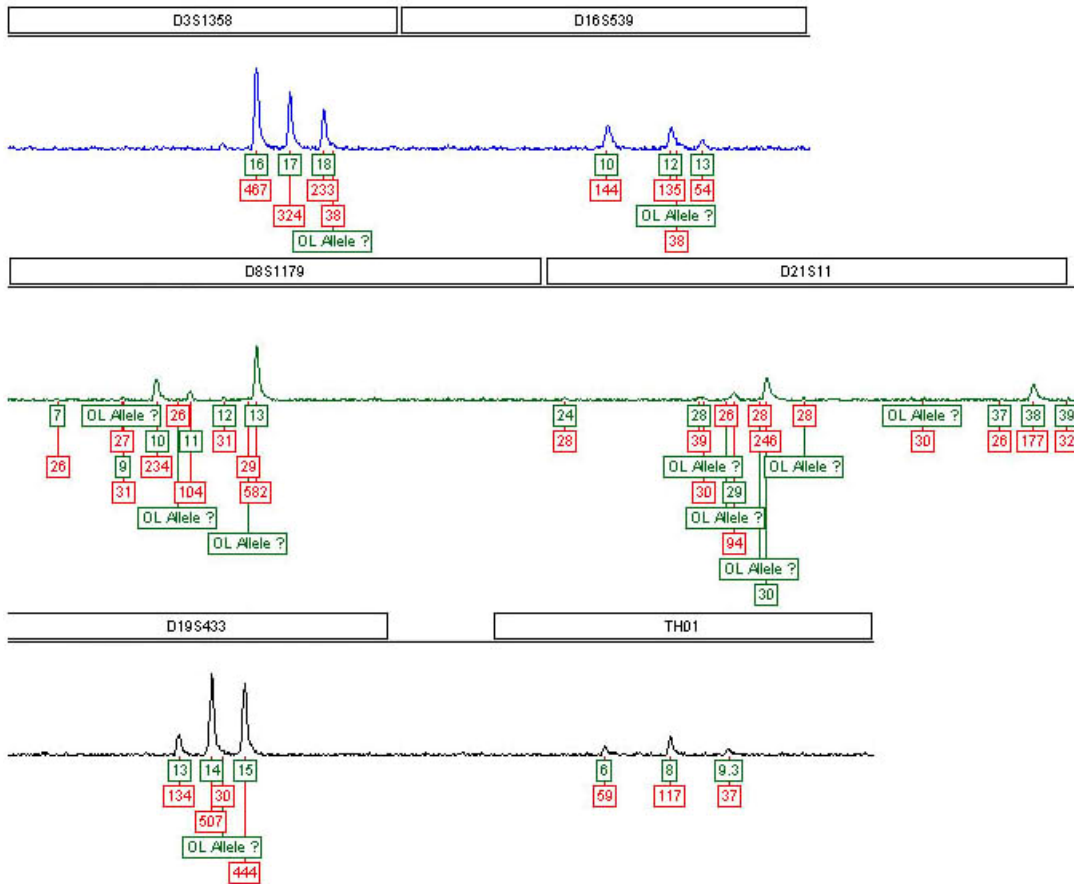
- <u>May be a mixture of two or more individuals</u>. Several loci (D3, D16, D2, D8, D21, D18, D19, THO1) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- <u>Contains peaks with a low average peak height (170 RFUs)</u>. Peaks below 150 RFUs need to be interpreted with caution. The highest peak found in the sample is 870 RFU's.

PM5 – Battery Housing – Consensus

The FSS consensus profile for D16 is 10, 13, but only sample QC00.0057.20 has both 10 and 13 reported.

PM5 – Brown tape swab 1 – QC00.0215.08
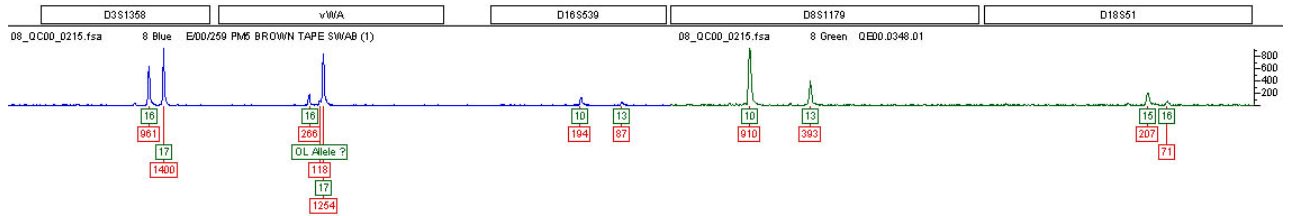
- Displays peak height imbalance at 5 loci (D3, vWA, D16, D8, D18). The difference in the peak heights of the 16 and 17 alleles for the D3 locus (961 and 1400, respectively), the 16 and 17 alleles for the vWA locus (266 and 1254, respectively), the 10 and 13 alleles for the D16 locus (194 and 87, respectively), the 10 and 13 alleles for the D8 locus (910 and 393, respectively), and the 15 and 16 alleles for the D18 locus (207 and 71, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. Two of the peaks in the imbalanced loci fall below the threshold 150 RFUs, indicating that they are possibly caused by stochastic effects.
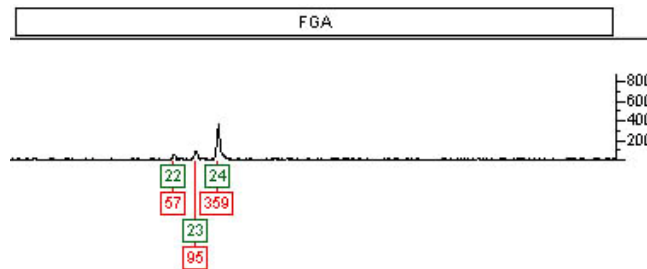
- May be the source of a mixture of two or more individuals. One locus (FGA) appears to have more than two alleles. The additional peaks in this sample were found to be below the threshold of 150 RFUs, indicating that they are possibly caused by stochastic effects. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

PM5 – Brown tape swab 1 – QC00.0216.17

- Displays peak height imbalance at 3 loci (D3, D21, FGA). The difference in the peak heights of the 16 and 17*^{O} alleles for the D3 locus (489 and 2185, respectively), the 29 and 30 alleles for the D21 locus (133 and 1475, respectively), and the 23 and 24 alleles for the FGA locus (461 and 169, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
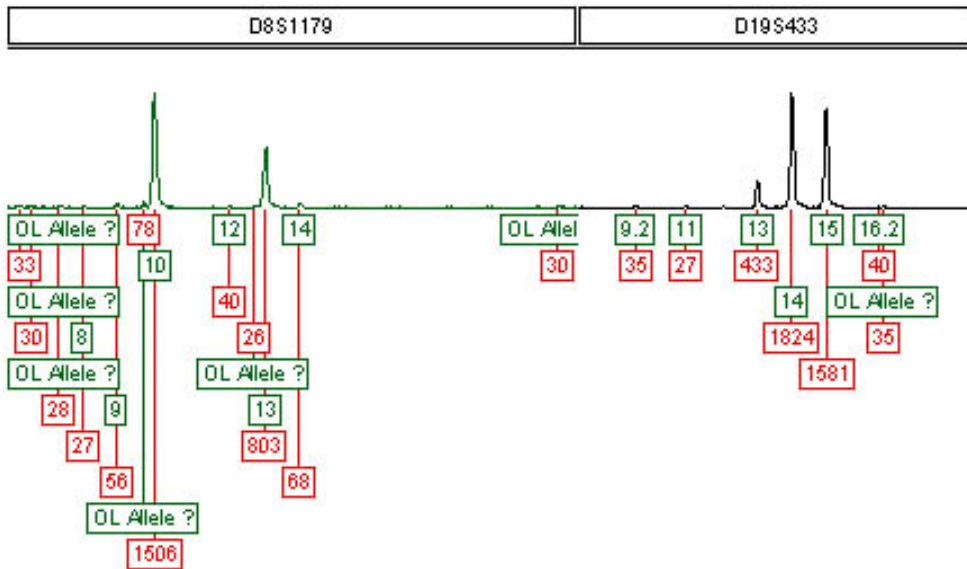
- May be a mixture of two or more individuals. Two loci, D8 (Allele 8 - 27 RFUs, Allele 9 - 56 RFUs, Allele 10 - 1506 RFUs, Allele 12 - 40 RFUs, Allele 13 - 803 RFUs, Allele 14 - 68 RFUs) and D19 (Allele 9.2*^{P} - 35 RFUs, Allele 11*^{Q} - 27 RFUs, Allele 13 - 433 RFUs, Allele 14 - 1824 RFUs, Allele 15 - 1581 RFUs, Allele 16.2*^{O} - 40 RFUs) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
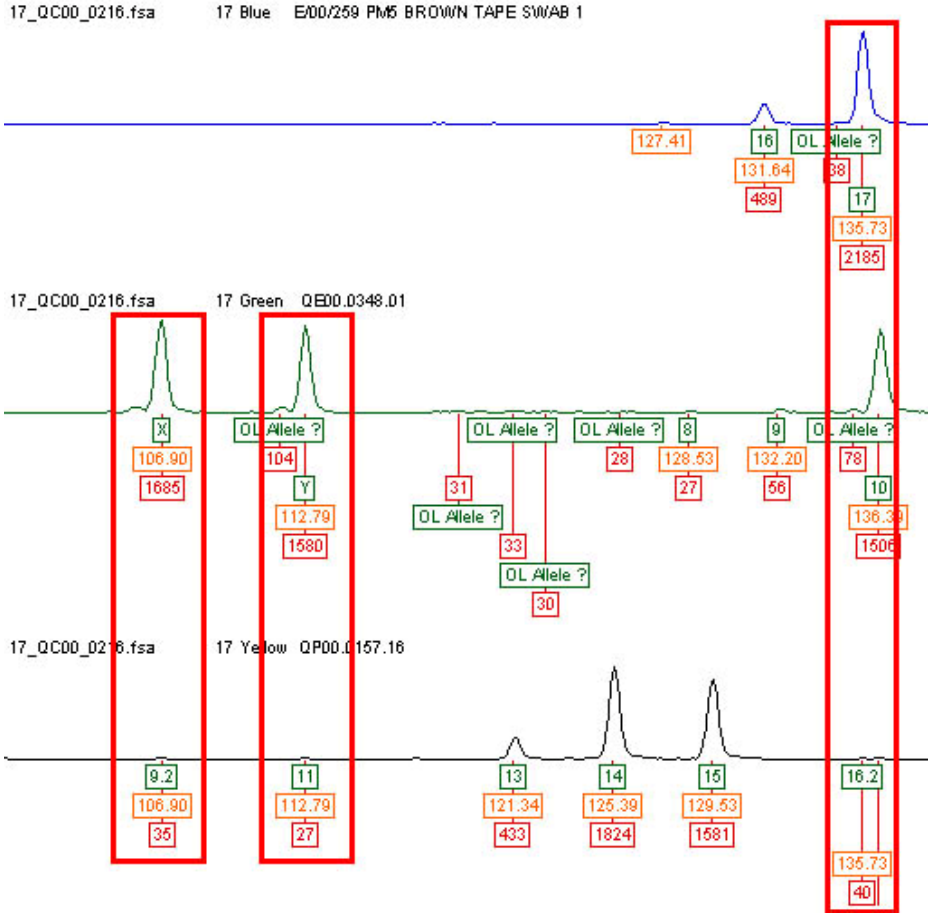
- Exhibits possible pull-up. The 16.2 peak at locus D19 with a size of 135.73 (40 RFU's) is possibly caused by a second peak (17) at the D3 locus with a size of 135.73 (2185 RFU's). The 9.2 peak at locus D19 with a size of 106.90 (35 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.90 (1685 RFU's). The 11 peak at locus D19 with a size of 112.79 (27 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of 112.79 (1580 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

PM5 – Brown tape swab 1 – Consensus

The vWA locus has a reported 16 peak in both samples, but they are both labeled as stutter, so it is removed from the consensus profile. The FGA locus has a reported 23 allele in sample QC00.0216.17, but the same peak is labeled stutter in sample QC00.0215.08. The 23 peak is not included in the consensus profile.

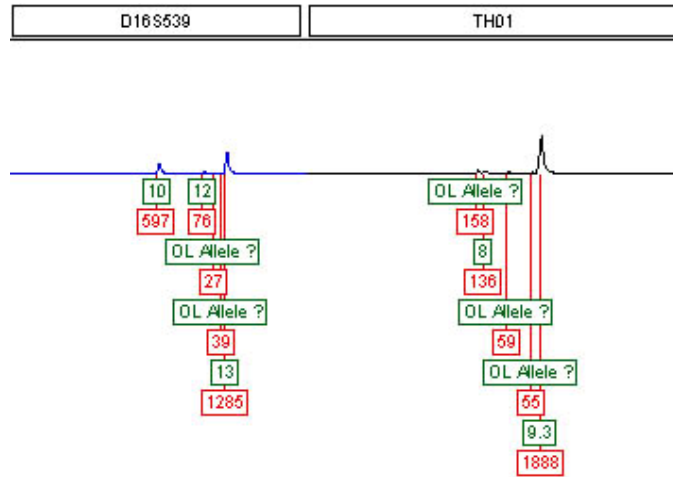**PM5 – Brown tape swab 2 – QC00.0215.09**
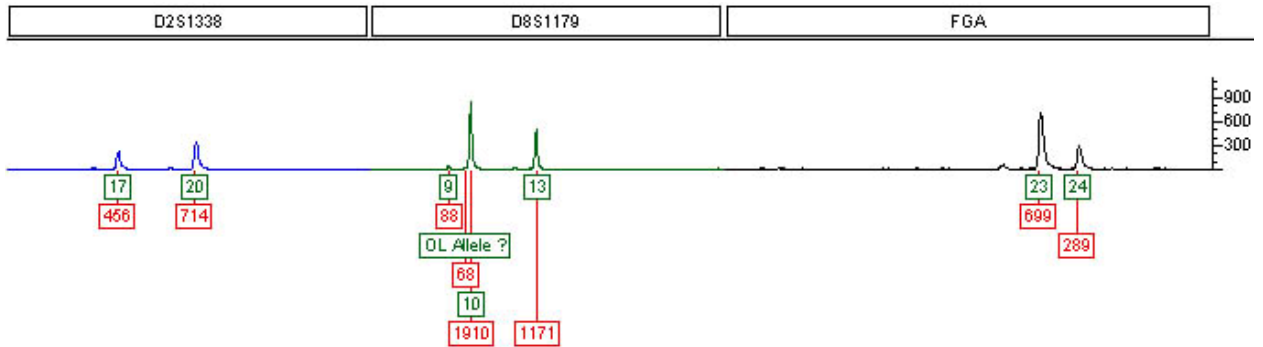
- Displays peak height imbalance at 3 loci (D2, D8, FGA). The difference in the peak heights of the 17 and 20 alleles for the D2 locus (456 and 714, respectively), the 10 and 13 alleles for the D8 locus (1910 and 1171, respectively), and the 23 and 24 alleles for the FGA locus (699 and 289, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.
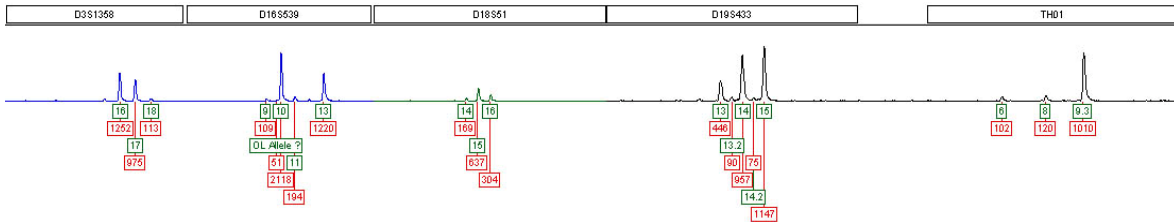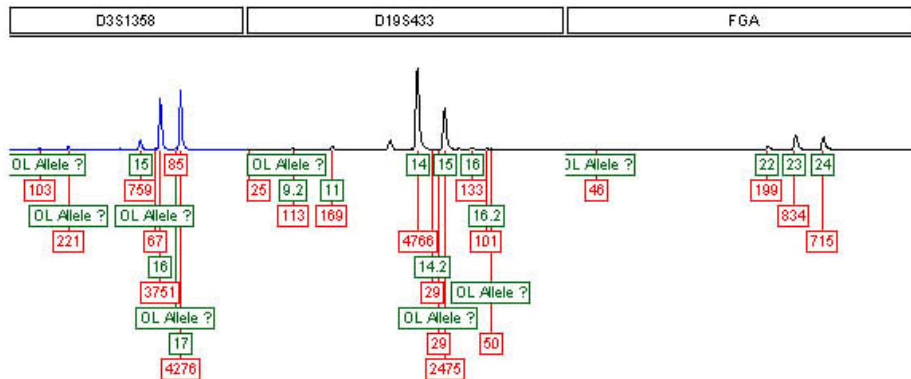

- May be a mixture of two or more individuals. Several loci (D3, D16, D18, D19, THO1) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

PM5 – Brown tape swab 2 – QC00.0216.18

D16S539	TH01

```
         10   12              OL Allele ?
        597   78                 158
             OL Allele ?          8
                27               136
             OL Allele ?      OL Allele ?
                39               59
                13           OL Allele ?
              1285              55
                                9.3
                               1888
```

- <u>Displays peak height imbalance at 2 loci (D16, THO1)</u>. The difference in the peak heights of the 10 and 13 alleles for the D16 locus (597 and 1285, respectively) and the 8 and 9.3*[S] alleles for the THO1 locus (136 and 1888, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.

D3S1358	D19S433	FGA

```
OL Allele ?    15   85    OL Allele ?      14  15  16     OL Allele ?      22  23  24
   103        759         25   9.2   11        133           46          199
OL Allele ? OL Allele ?   113   169            16.2                           834
   221         67                        4766   101                               715
               16                         14.2
             3751                          29  OL Allele ?
           OL Allele ?                  OL Allele ?
               17                          29   50
             4276                         2475
```

- <u>May be a mixture of two or more individuals</u>. Several loci (D3, vWA, D8, D21, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
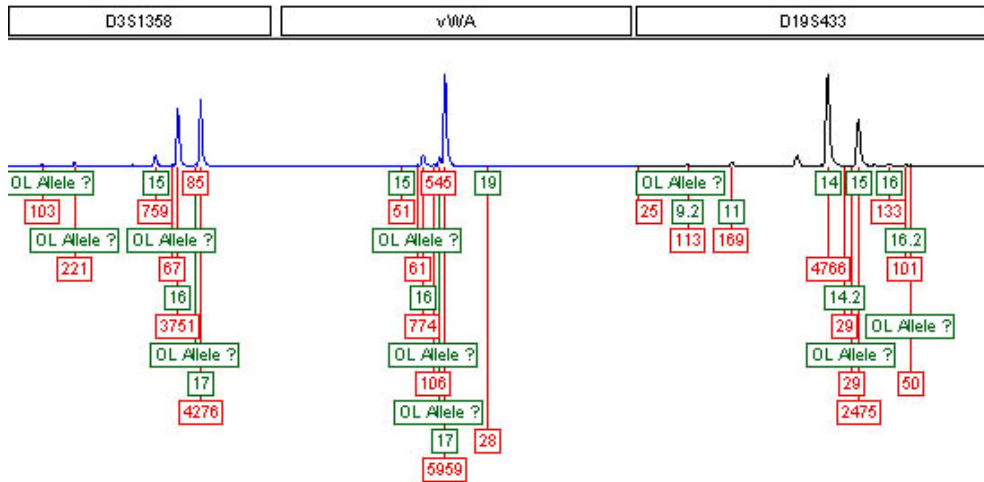
- <u>Has electropherograms that contain a few peaks (3) with high peak heights</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
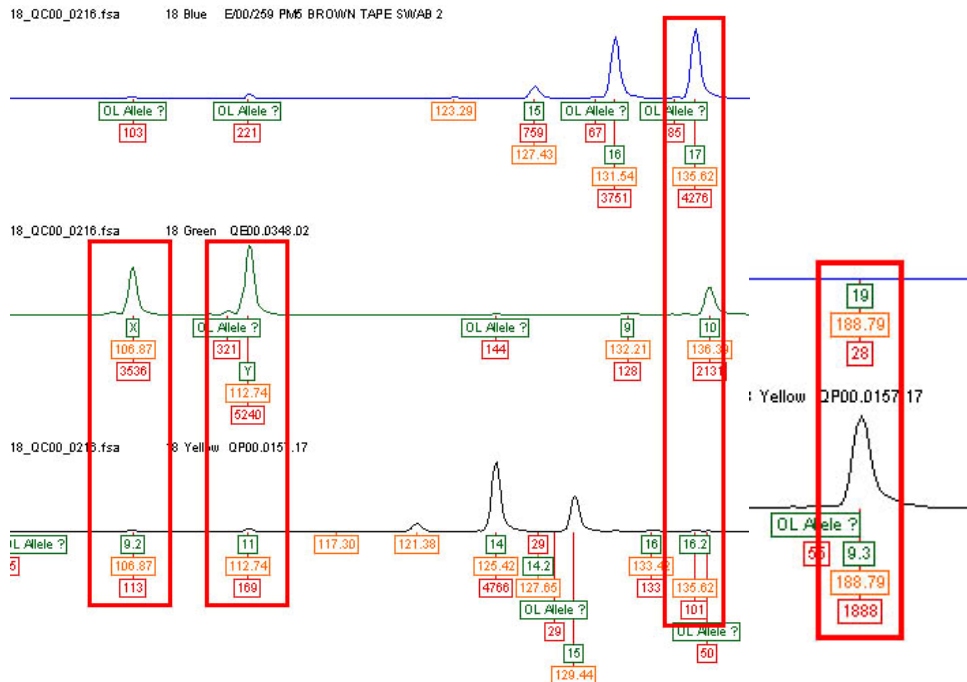

- <u>Exhibits possible pull-up</u>. The 16.2 peak at locus D19 with a size of 135.62 (101 RFU's) is possibly caused by a second peak (17) at the D3 locus with a size of 135.62 (4276 RFU's). The 19 peak at locus vWA with a size of 188.78 (28 RFU's) is possibly caused by a second peak (9.3) at the THO1 locus with a size of 188.78 (1888 RFU's). The 9.2 peak at locus D19 with a size of 106.87 (113 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.87 (3536 RFU's). The 11 peak at locus D19 with a size of 112.74 (169 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of

112.74 (5240 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

PM5 – Brown tape swab 2 – Consensus

The D3 locus has a reported 15 peak for both samples. Both peaks are labeled as stutter, so the peak is not included in the consensus profile.
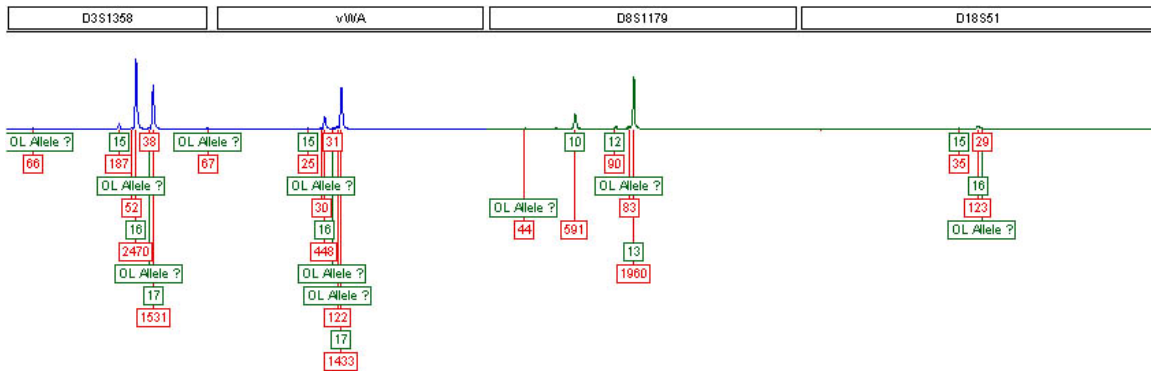
PM5 – Blue tape swab 1 – QC00.0215.12
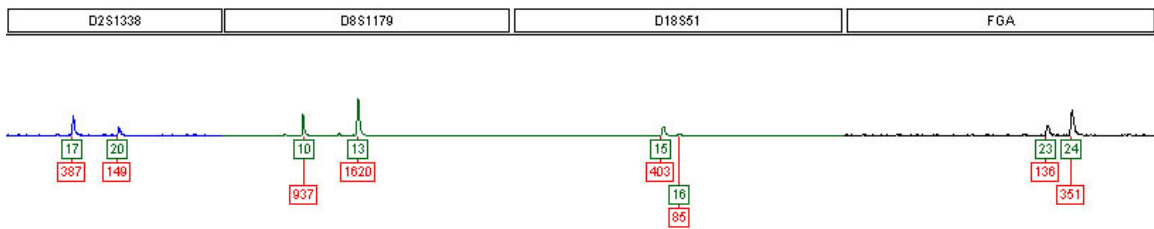
- Displays peak height imbalance at 4 loci (D2, D8, D18, FGA). The difference in the peak heights of the 17 and 20 alleles for the D2 locus (387 and 149, respectively), the 10 and 13 alleles for the D8 locus (937 and 1620, respectively), the 15 and 16 alleles for the D18 locus (403 and 85, respectively), and the 23 and 24 alleles for the FGA locus (136 and 351, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. Three of the peaks in the imbalanced loci fall below the threshold 150 RFUs, indicating that they are possibly caused by stochastic effects.
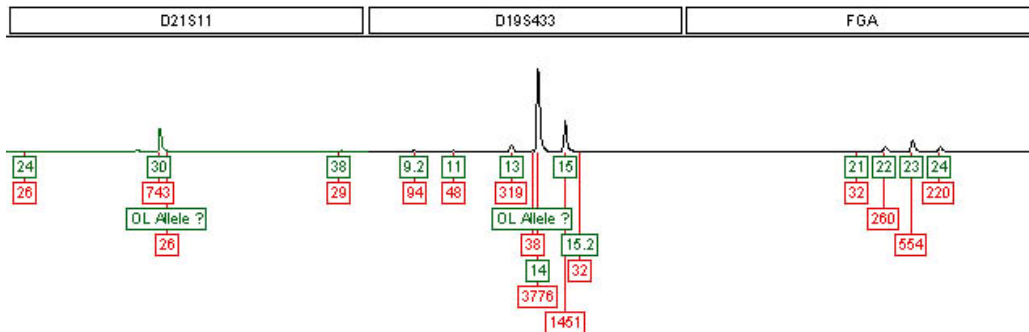

- May be a mixture of two or more individuals. Several loci (D3, D16, D19) appear to have more than two alleles. Interpretation of mixed DNA samples is

challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

PM5 – Blue tape swab 1 – QC00.0216.20

- <u>Displays peak height imbalance at 4 loci (D3, vWA, D8, D18)</u>. The difference in the peak heights of the 16 and 17 alleles for the D3 locus (2470 and 1531, respectively), the 16 and 17 alleles for the vWA locus (448 and 1433, respectively), the 10 and 13 alleles for the D8 locus (591 and 1960, respectively), and the 15 and 16 alleles for the D18 locus (35 and 123, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.

- <u>May be a mixture of two or more individuals</u>. Several loci (D21, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
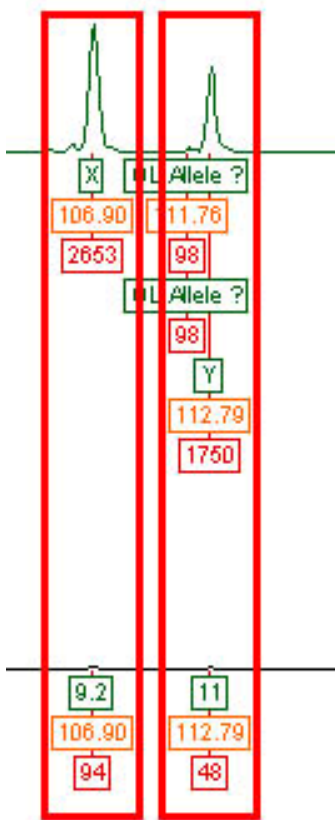
- Exhibits possible pull-up. The 9.2 peak at locus D19 with a size of 106.90 (94 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.90 (2653 RFU's). The 11 peak at locus D19 with a size of 112.79 (48 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of 112.79 (1750 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

PM5 – Blue tape swab 1 – Consensus

The FSS QC00.0215.12 labeled the 23 peak as stutter, yet it is included in their consensus profile.

PM5 – Blue tape swab 2 – QC00.0215.13

- <u>Has electropherograms that do not contain genotype information for one locus (D2).</u> The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).
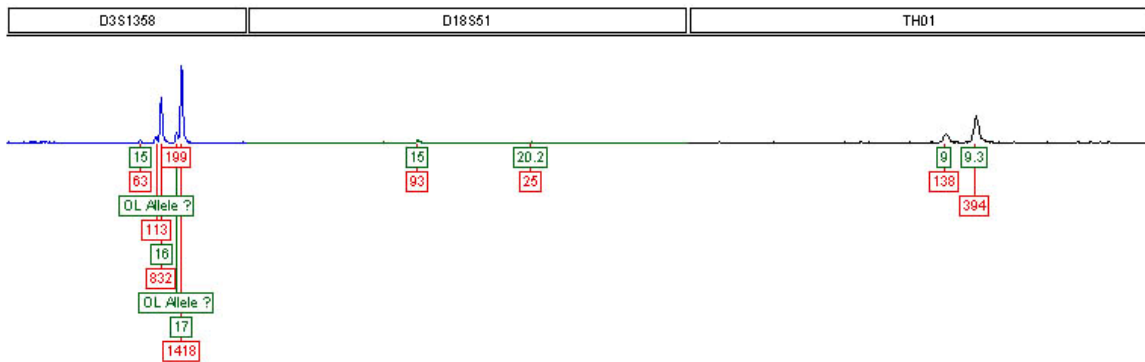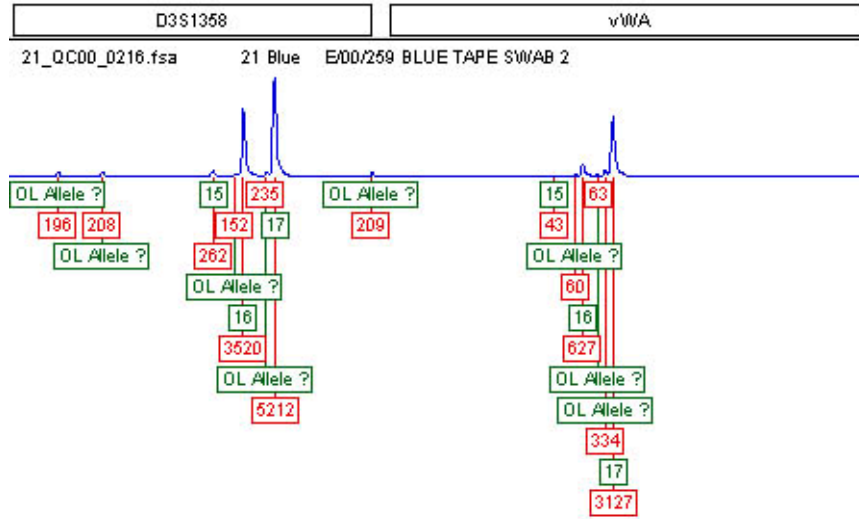

- <u>Displays peak height imbalance at 3 loci (D3, D18, THO1).</u> The difference in the peak heights of the 16 and 17 alleles for the D3 locus (832 and 1418, respectively), the 15 and 20.2 alleles for the D18 locus (93 and 25, respectively), and the 9 and 9.3 alleles for the THO1 locus (138 and 394, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
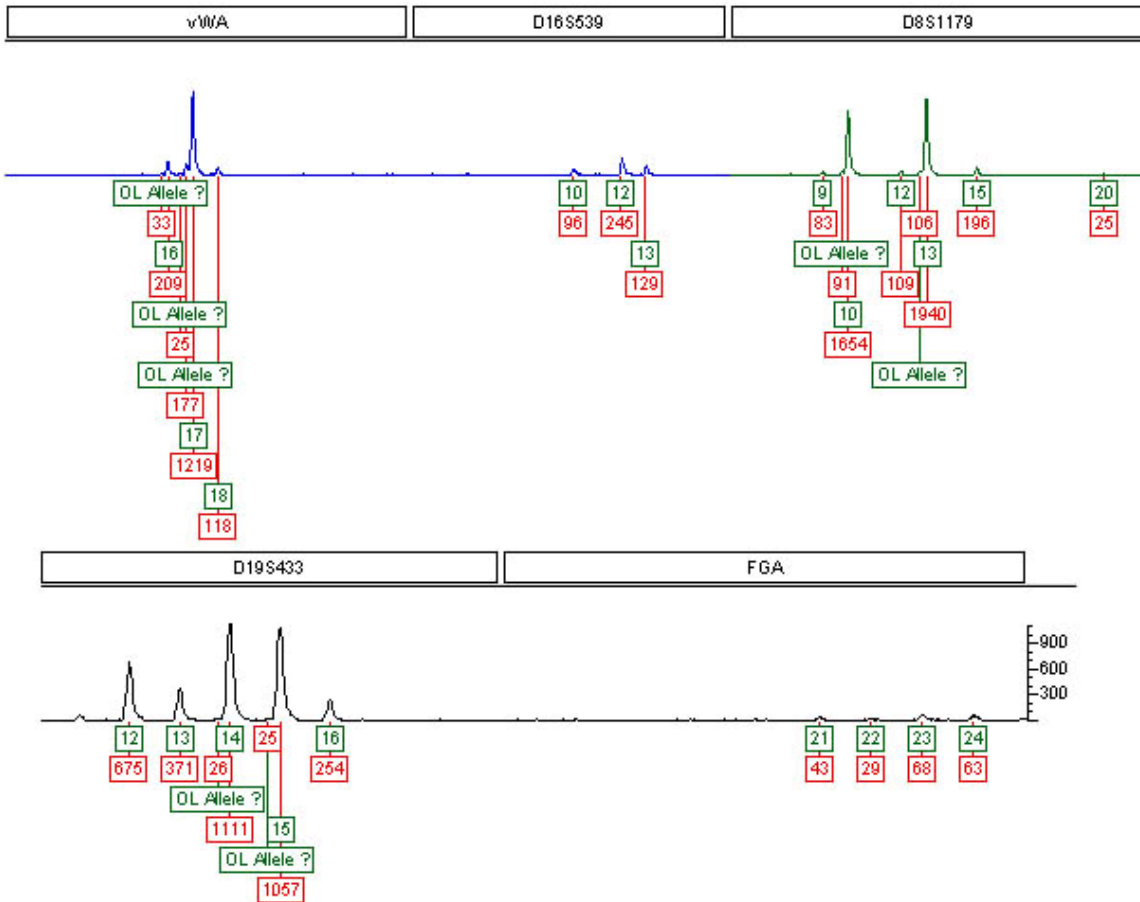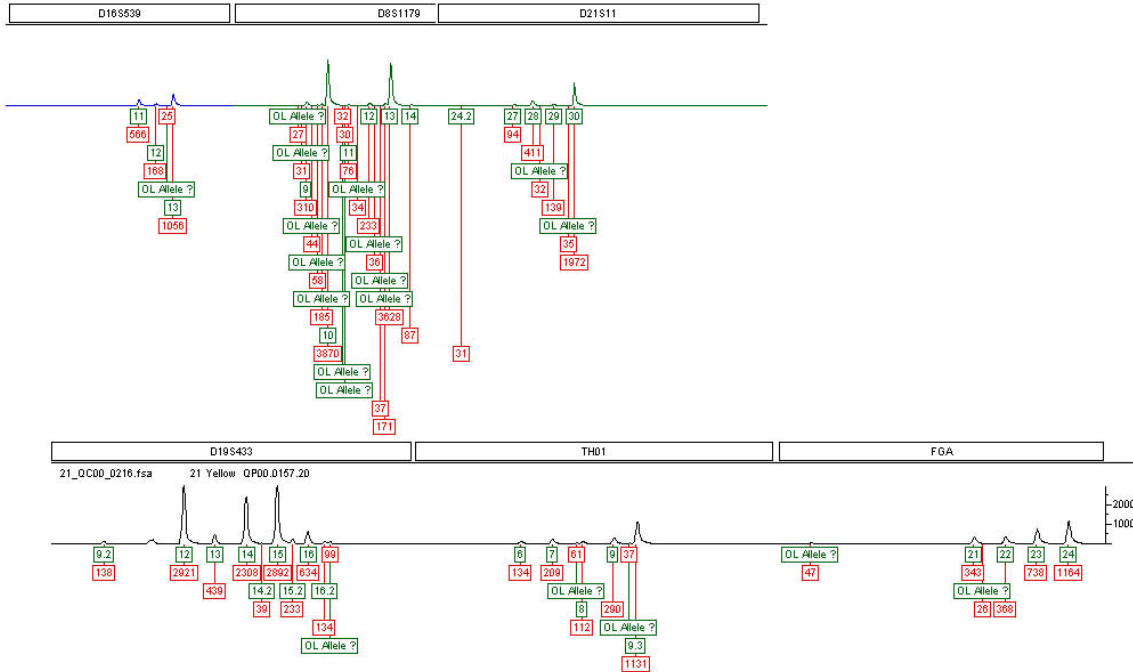
- <u>May be a mixture of two or more individuals</u>. Several loci (vWA, D16, D8, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

PM5 – Blue tape swab 2 – QC00.0216.21

- Displays peak height imbalance at 2 loci (D3, vWA). The difference in the peak heights of the 16*^{w} and 17*^{x} alleles for the D3 locus (3520 and 5212, respectively) and the 16 and 17 alleles for the vWA locus (627 and 3127, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.
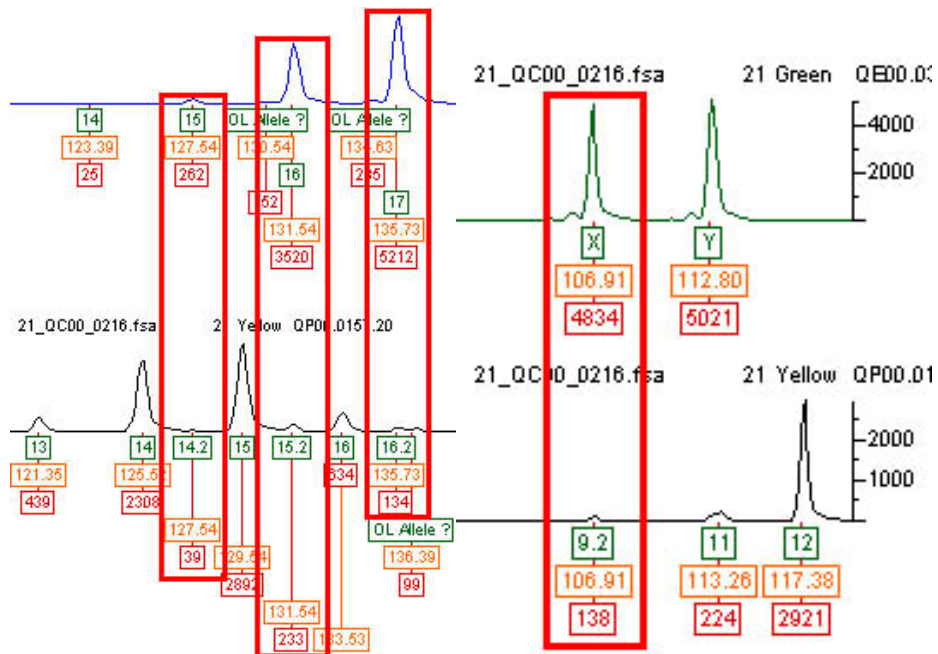

- May be a mixture of two or more individuals. Several loci (D16, D8, D21, D19, THO1, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with

DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
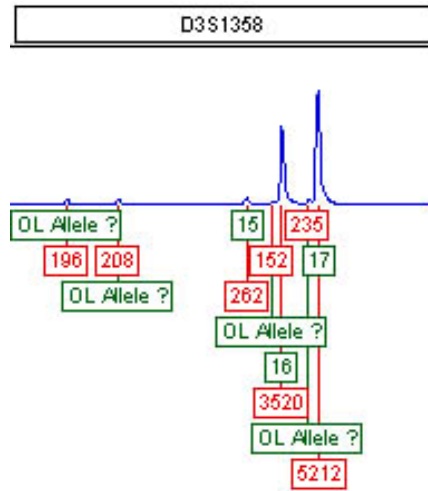
- Has electropherograms that contain one peak with a high peak height. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".


- Exhibits possible pull-up. The 14.2 peak at locus D19 with a size of 127.53 (39 RFU's) is possibly caused by a second peak (15) at the D3 locus with a size of 127.53 (262 RFU's). The 15.2 peak at locus D19 with a size of 131.54 (233 RFU's) is possibly caused by a second peak (16) at the D3 locus with a size of 131.54 (3520 RFU's). The 16.2 peak at locus D19 with a size of 135.73 (134

RFU's) is possibly caused by a second peak (17) at the D3 locus with a size of 135.73 (5212 RFU's). The 9.2 peak at locus D19 with a size of 106.91 (138 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.91 (4834 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.
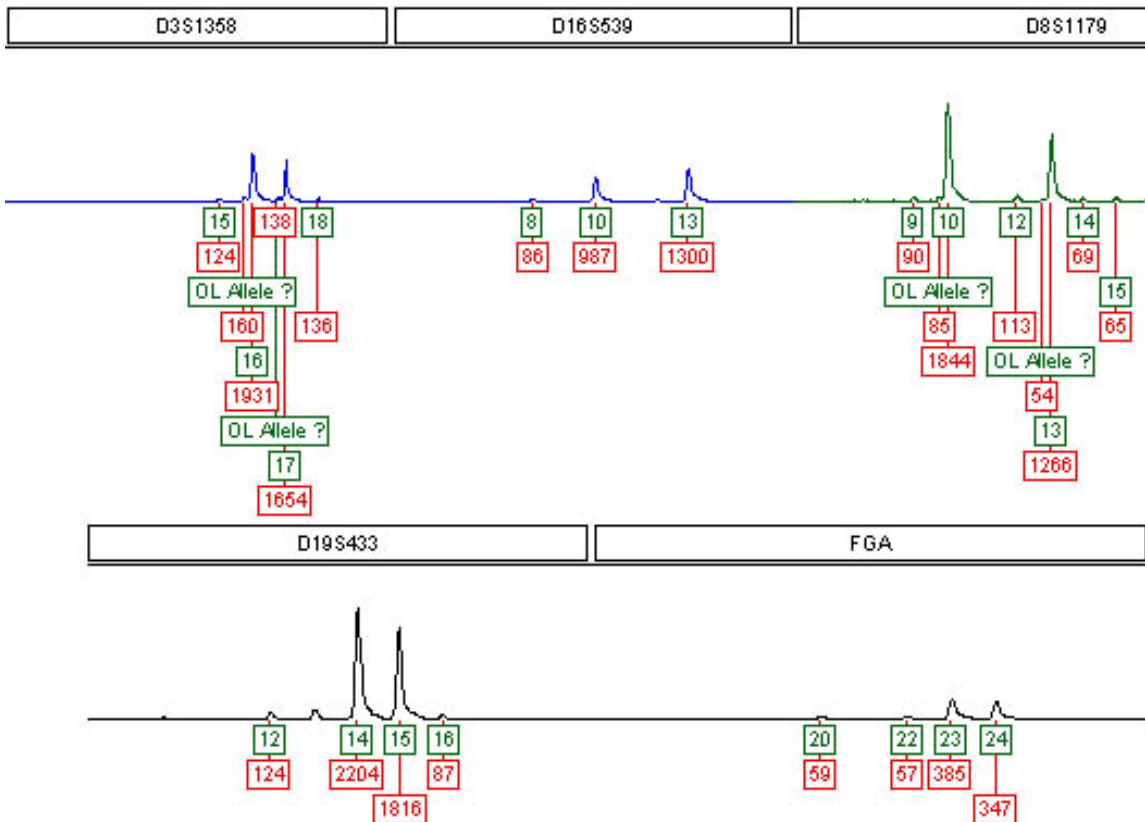
PM5 – Blue tape swab 2 – Consensus

Both samples have a 21 peak at vWA, but it is not included in the consensus profile because it is labeled as stutter in the sample QC00.0216.21. Both samples have 12, 13, 14, 15, and 16 peaks at D19, but no alleles are reported in the consensus profile. Both samples have a 9 and 9.3 profile at THO1, but the consensus profile is 9.3, F.
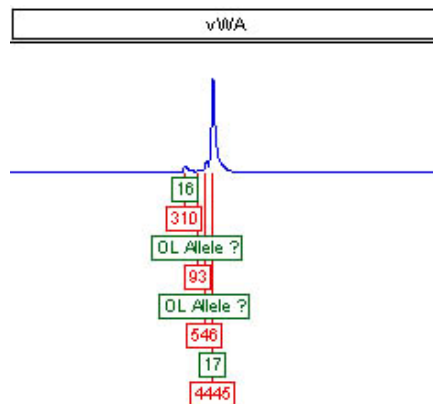
PM5 – Other brown tape swab 1 – QC00.0209.18
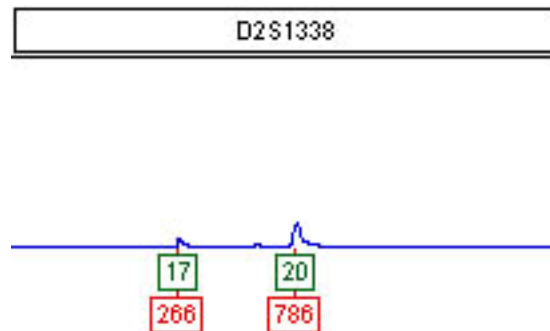
- Displays peak height imbalance at the locus D2. The difference in the peak heights of the 17 and 20 alleles for the D2 locus (266 and 786, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.

D3S1358   D16S539   D8S1179

15 138 18   8   10   13   9 10   12   14
124   86 987 1300   90   69
OL Allele ?   OL Allele ?   15
160 136   85   113   65
16   1844 OL Allele ?
1931   54
OL Allele ?   13
17   1266
1654

D19S433   FGA

12   14   15   16   20   22   23   24
124 2204   87   59   57 385
1816   347

- <u>May be a mixture of two or more individuals</u>. Several loci (D3, D16, D8, D19, FGA) appear to have more than two alleles. The additional peaks in this sample were found to be below the threshold of 150 RFUs, indicating that they are possibly caused by stochastic effects. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
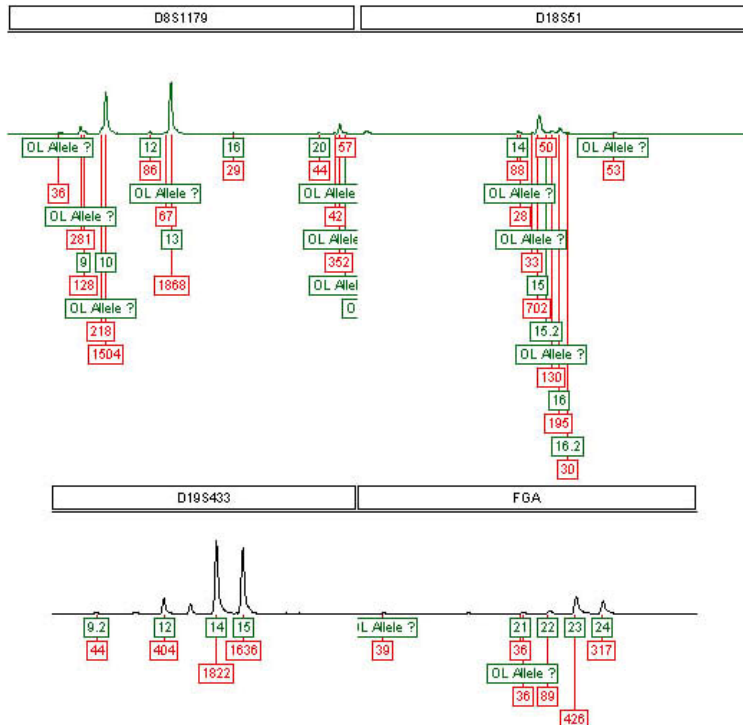
vWA

16
310
OL Allele ?
93
OL Allele ?
546
17
4445

- <u>Has electropherograms that contain one peak with a high peak height</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result,

it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".

PM5 – Other brown tape swab 1 – QC00.0217.09

D2S1338	D21S11

17 20   29   32.2
1511 662   82   52
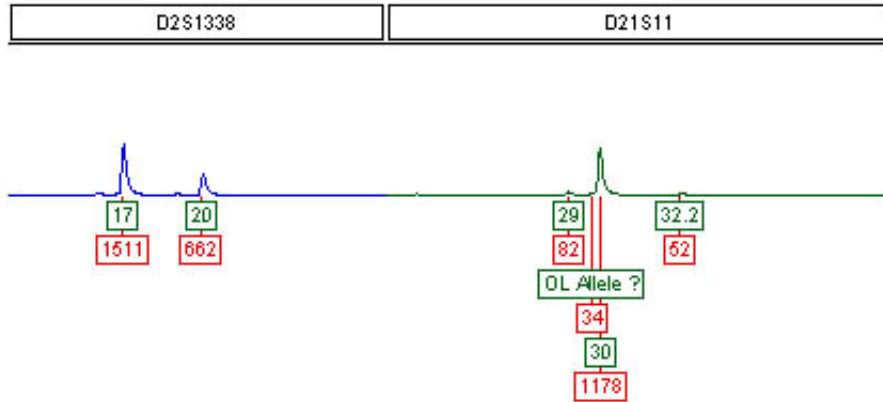
OL Allele ?
34
30
1178

- Displays peak height imbalance at 2 loci (D2, D21). The difference in the peak heights of the 17 and 20 alleles for the D2 locus (1511 and 662, respectively) and the 30 and 32.2 alleles for the D21 locus (1178 and 52, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.

D8S1179	D18S51

OL Allele ?   12   16   20 57   14 50   OL Allele ?
              86   29   44   88   53
36   OL Allele ?   OL Allele   OL Allele ?
OL Allele ?   67   42   28
281   13   OL Allele   OL Allele ?
9 10   352   33
128   1868   OL Allel   15
OL Allele ?   0   702
218   15.2
1504   OL Allele ?
                                       130
                                       16
                                       195
                                       16.2
                                       30

D19S433	FGA

9.2   12   14 15   lL Allele ?   21 22 23 24
44   404 1636   39   36   317
           1822   OL Allele ?
                                      36 89
                                      426

- May be a mixture of two or more individuals. Several loci (D3, vWA, D8, D18, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
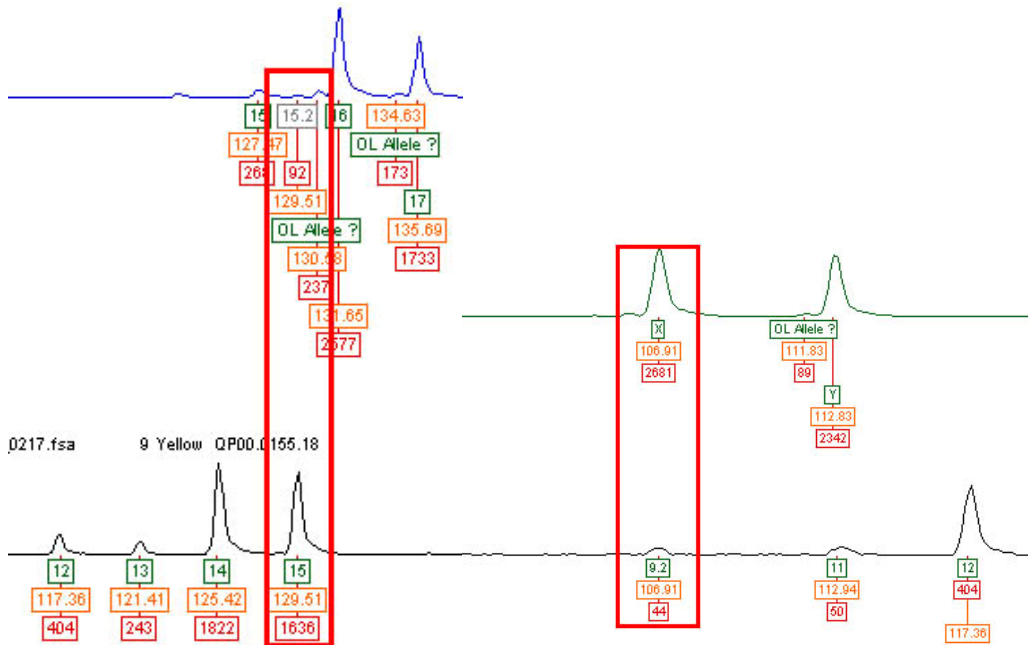

- Exhibits possible pull-up. The 15.2 peak at locus D3 with a size of 129.51 (92 RFU's) is possibly caused by a second peak (15) at the D19 locus with a size of 129.51 (1636 RFU's). The 9.2 peak at locus D19 with a size of 106.91 (44 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.91 (2681 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

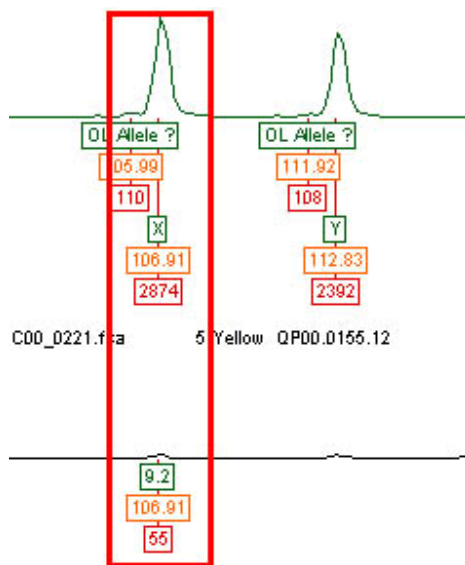**PM5 – Other brown tape swab 1 – Consensus**

Both samples have a 15 peak at D3, but the 15 in QC00.0209.18 is labeled as stutter, so it does not appear in the consensus profile. Both samples have 12, 13, 14, and 15 peaks at D19, but the consensus profile is 14, 15. Both samples have a 22 peak at FGA, but the 22 in QC00.0209.18 is labeled as stutter, so it does not appear in the consensus profile.

MDH4 – Red/brown tape – QC00.0209.12

- Displays peak height imbalance at 2 loci (D21, THO1). The difference in the peak heights of the 29 and 30 alleles for the D21 locus (64 and 465, respectively) and the 8 and 9.3 alleles for the THO1 locus (71 and 308, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. Two of the peaks in the imbalanced loci fall below the threshold 150 RFUs, indicating that they are possibly caused by stochastic effects.

- May be a mixture of two or more individuals. Several loci (D3, vWA, D16, D2, D8, D18, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to

consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

MDH4 – Red/brown tape – QC00.0221.05
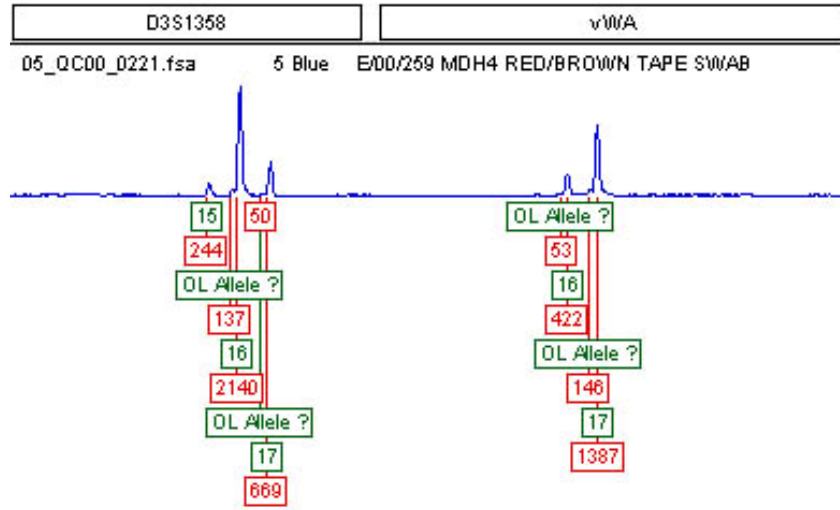
- Displays peak height imbalance at 2 loci (D3, vWA). The difference in the peak heights of the 16 and 17 alleles for the D3 locus (2140 and 669, respectively) and the 16 and 17 alleles for the vWA locus (422 and 1387, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.
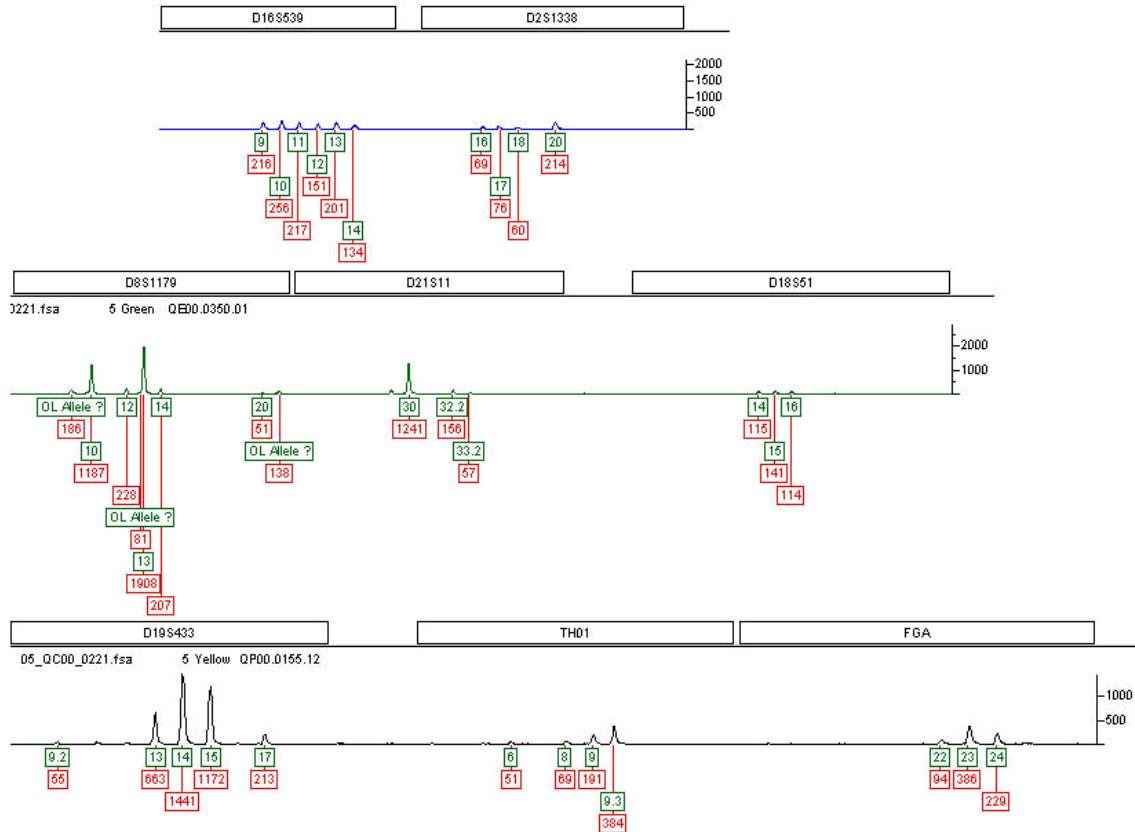
- May be a mixture of two or more individuals. Several loci (D16, D2, D8, D21, D18, D19, THO1, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- Exhibits possible pull-up. The 9.2 peak at locus D19 with a size of 106.91 (55 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.91 (2874 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

MDH4 – Red/brown tape – Consensus

Both samples have a 12 peak at D16, but the 12 in QC00.0209.12 is labeled as stutter, so it does not appear in the consensus profile. Both samples have a 29 peak at D21, but both peaks are labeled as stutter, so 29 does not appear in the consensus profile. Both samples have a 14 peak at D18, but it does not appear in the consensus profile. Both samples have a 17 peak at D19, but it does not appear in the consensus profile. Both samples have a 13 peak at D19, but both peaks are labeled as stutter, so 13 does not appear in the consensus profile. Both samples have a 8 peak at THO1, but it does not appear in the consensus profile. Both samples have a 22 peak at FGA, but both peaks are labeled as stutter, so 22 does not appear in the consensus profile.

MDH4 – Tube end – QC00.0209.14

D16S539	D2S1338
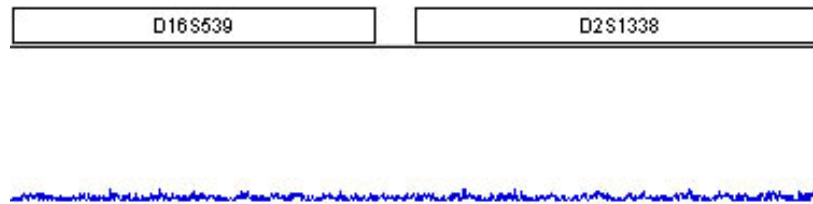
- Has electropherograms that do not contain genotype information for 2 loci (D16 and D2). The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).
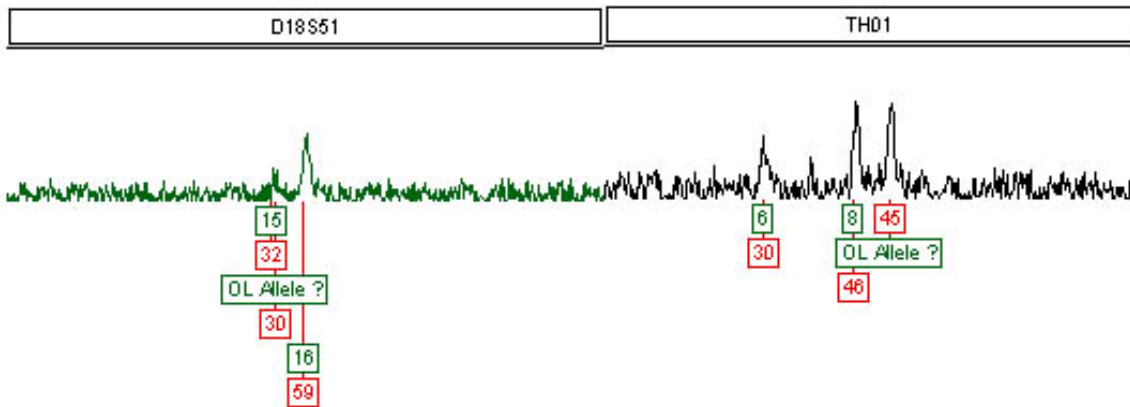
- Displays peak height imbalance at 2 loci (D18, THO1). The difference in the peak heights of the 15 and 16 alleles for the D18 locus (32 and 59, respectively) and the 6 and 8 alleles for the THO1 locus (30 and 46, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.
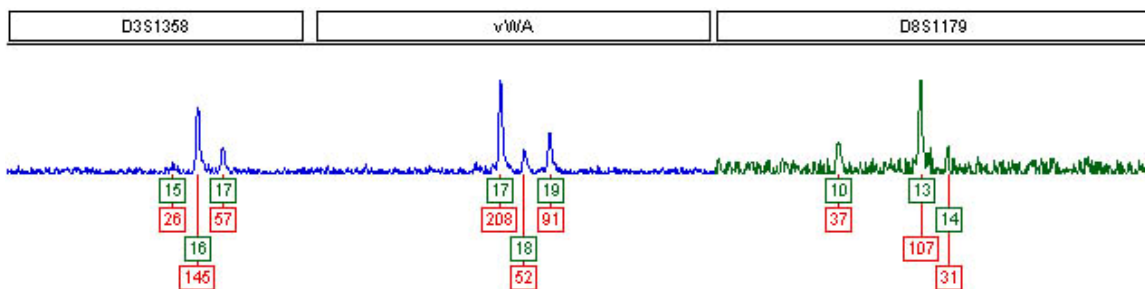
- May be a mixture of two or more individuals. Several loci (D3, vWA, D8) appear to have more than two alleles. The additional peaks in this sample were found to be below the threshold of 150 RFUs, indicating that they are possibly caused by stochastic effects. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- Contains peaks with a low average peak height (65 RFUs). Peaks below 150 RFUs need to be interpreted with caution. The highest peak found in the sample is 208 RFU's.

MDH4 – Tube end – QC00.0221.07

| D2S1338	D18S51	FGA

- <u>Has electropherograms that do not contain genotype information for 3 loci (D2, D18, and FGA).</u> The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).

D3S1358	D8S1179	TH01

16 | 8 | 76 | 7 | 8
428 | 361 | | 149
17 | 9 | 247
1071 | 249
OL Allele ?
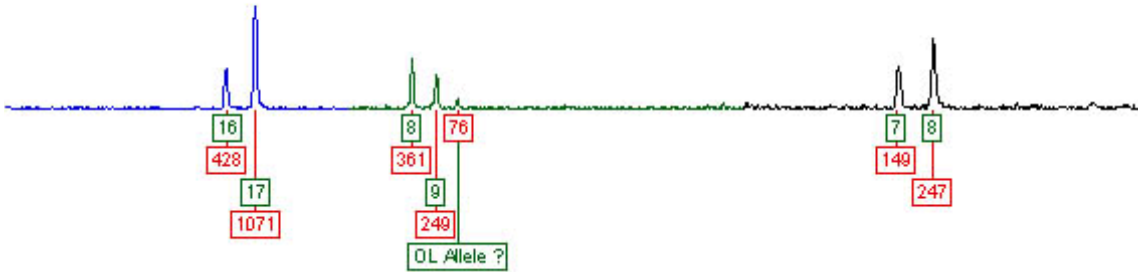
- <u>Displays peak height imbalance at 3 loci (D3, D8, THO1).</u> The difference in the peak heights of the 16 and 17 alleles for the D3 locus (428 and 1071, respectively), the 8 and 9 alleles for the D8 locus (361 and 249, respectively), and the 7 and 8 alleles for the THO1 locus (149 and 247, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.

vWA	D19S433

14 | 16 | 12 | 16
121 | 51 | 265 | 286
17 | 16.2
359 | 360
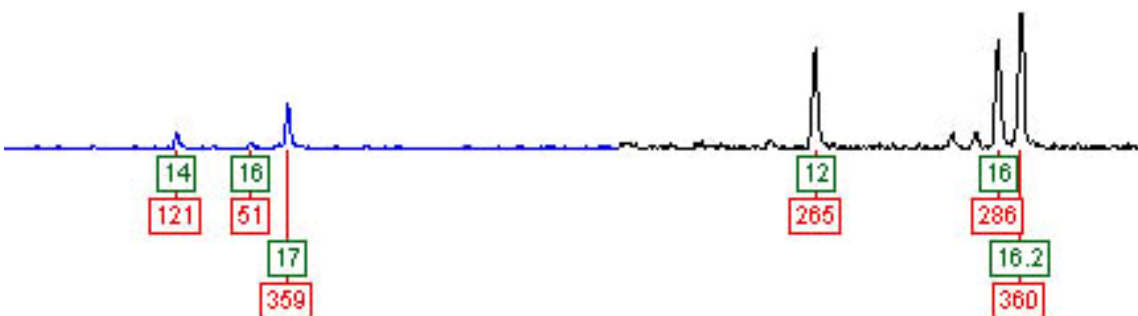
- <u>May be a mixture of two or more individuals</u>. Two loci, vWA (Allele 14 - 121 RFUs, Allele 16 - 51 RFUs, Allele 17 - 359 RFUs) and D19 (Allele 12 - 265 RFUs, Allele 16 - 286 RFUs, Allele 16.2 - 360 RFUs) appear to have more than two alleles. Some of the additional allele(s) may be technical artifacts or may be evidence that contamination has occurred. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- <u>Contains peaks with a low average peak height (349 RFUs)</u>. Peaks below 150 RFUs need to be interpreted with caution. The highest peak found in the sample is 1071 RFU's.

MDH4 – Tube end – Consensus

The FSS has labeled the 16 peak at D3 in sample QC00.0221.07 as stutter, but the allele appears in the consensus profile. Both samples have an 8 peak at THO1, but the consensus profile is blank.

MDH5 – Swab 1 – QC00.0216.05
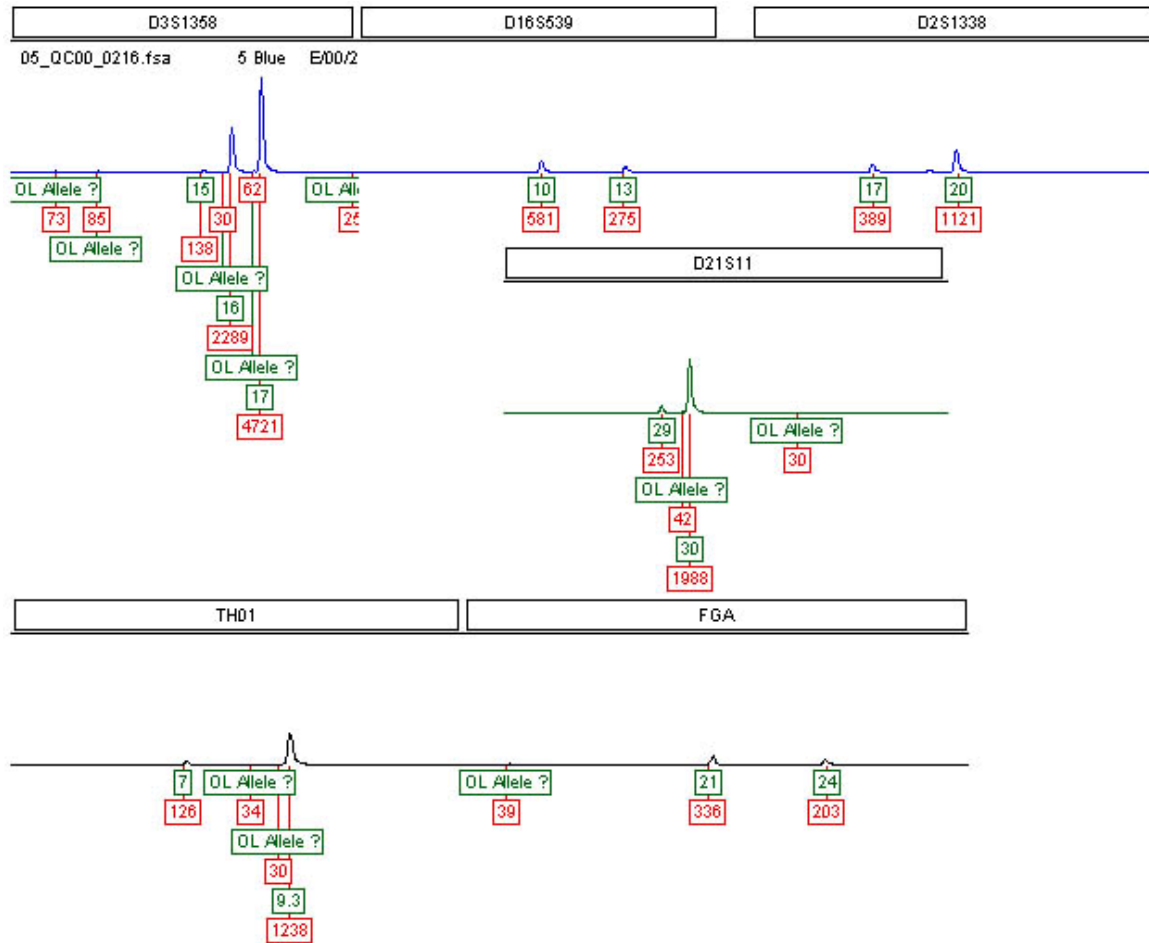
- Displays peak height imbalance at 6 loci (D3, D16, D2, D21, THO1, FGA). The difference in the peak heights of the 16 and 17 alleles for the D3 locus (2289 and 4721, respectively), the 10 and 13 alleles for the D16 locus (581 and 275, respectively), the 17 and 20 alleles for the D2 locus (389 and 1121, respectively), the 29 and 30 alleles for the D21 locus (253 and 1988, respectively), the 7 and 9.3 alleles for the THO1 locus (126 and 1238, respectively), and the 21 and 24 alleles for the FGA locus (336 and 203, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
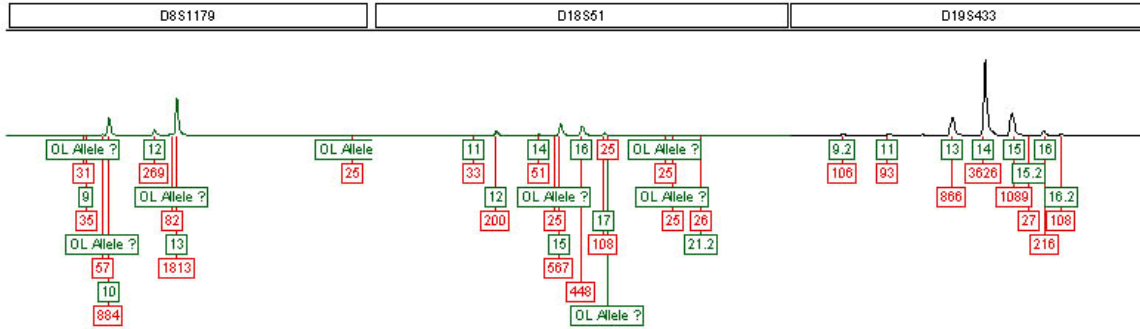
- <u>May be a mixture of two or more individuals</u>. Several loci (D8, D18, D19) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
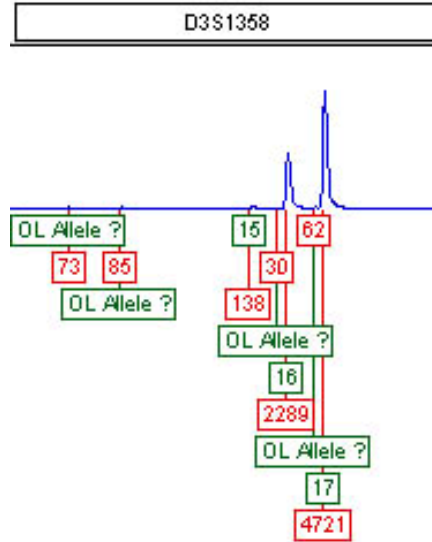
- <u>Has electropherograms that contain one peak with a high peak height</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
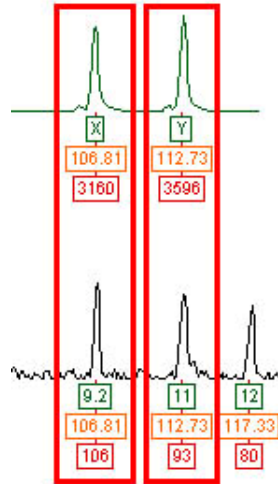
- Exhibits possible pull-up. The 9.2 peak at locus D19 with a size of 106.81 (106 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.81 (3160 RFU's). The 11 peak at locus D19 with a size of 112.73 (93 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of 112.73 (3596 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

MDH5 – Swab 1 – QC00.0221.19

- Has electropherograms that do not contain genotype information for one locus (D2). The absence of genotype information at one or more loci can be indicative of a number of different issues (such as: the use of inappropriate allelic ladders, degradation, or stochastic effects associated with small amounts of starting material).
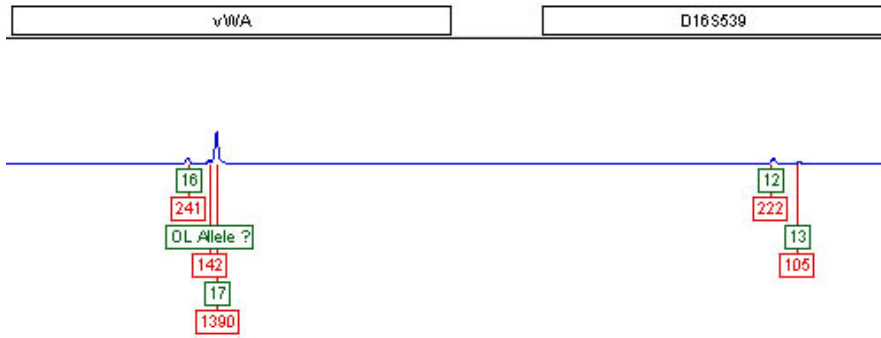
- Displays peak height imbalance at 2 loci (vWA, D16). The difference in the peak heights of the 16 and 17 alleles for the vWA locus (241 and 1390, respectively) and the 12 and 13 alleles for the D16 locus (222 and 105, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
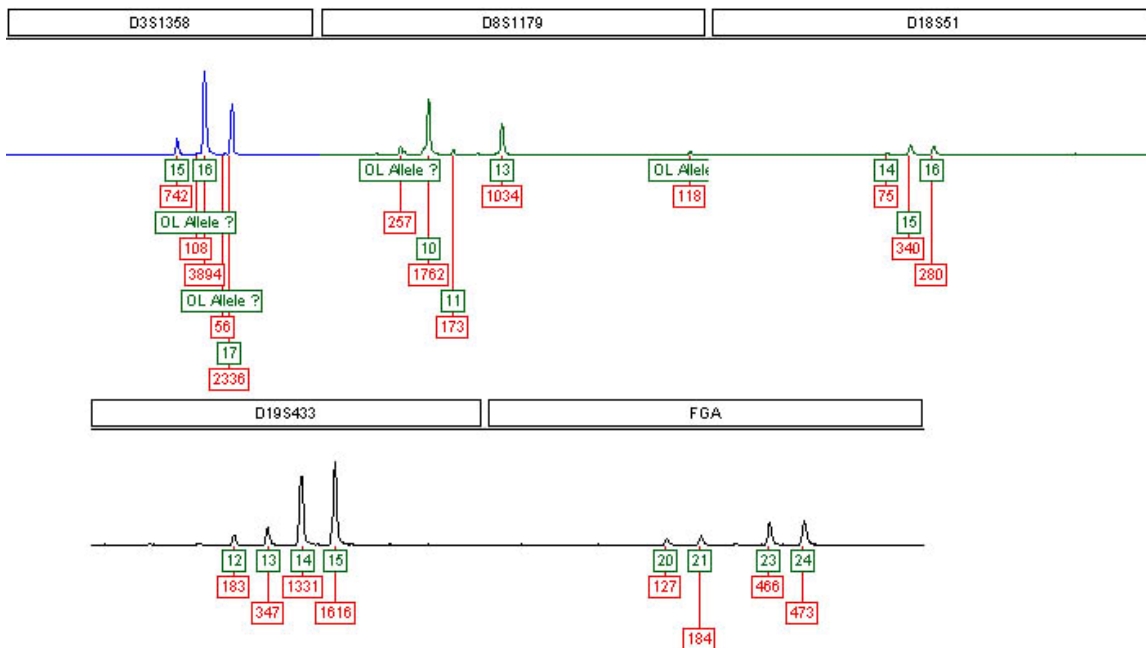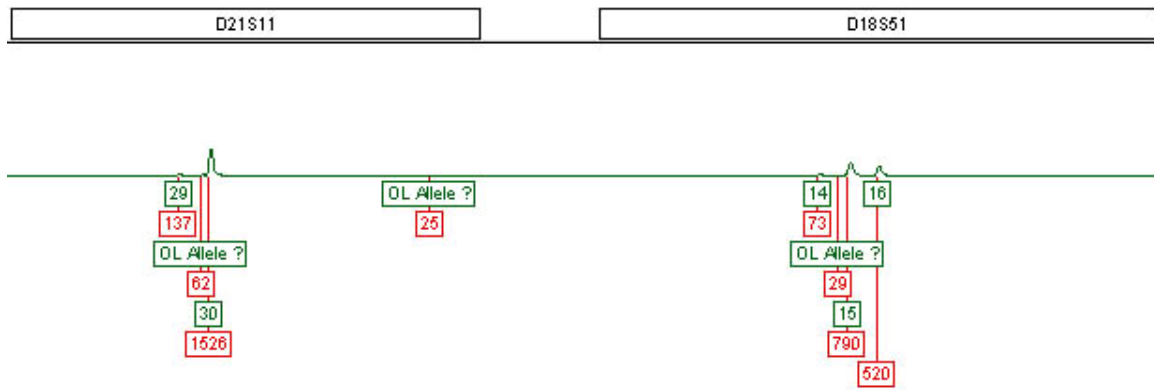

- May be the source of a mixture of two or more individuals. Several loci (D3, D8, D18, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

MDH5 – Swab 1 – Consensus

Both samples have a 15 peak reported at D3, but QC00.0216.05 has the 15 labeled as stutter. However, the FSS has still included 15 in the consensus profile. Both samples have a 12 peak labeled at D19, but QC00.0216.05 has the 12 labeled as stutter, so the allele does not appear in the consensus profile. However, there is no 13 peak reported in QC00.0216.05, so it is unclear as how the 12 peak could be considered stutter. Both samples have a 21 peak reported, but the allele does not appear in the consensus profile.

MDH5 – Swab 2 – QC00.0216.06

- Displays peak height imbalance at 2 loci (D21, D18). The difference in the peak heights of the 29 and 30 alleles for the D21 locus (137 and 1526, respectively) and the 15 and 16 alleles for the D18 locus (790 and 520, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.
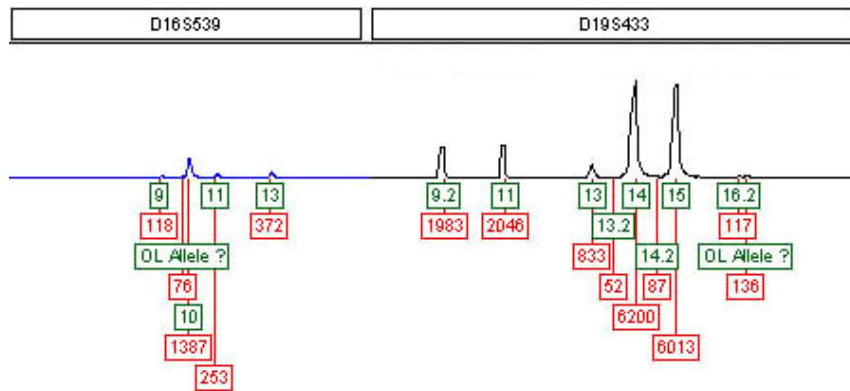

- May be a mixture of two or more individuals. Several loci (D3, vWA, D16, D19) appear to have more than two alleles. Interpretation of mixed DNA samples is

challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
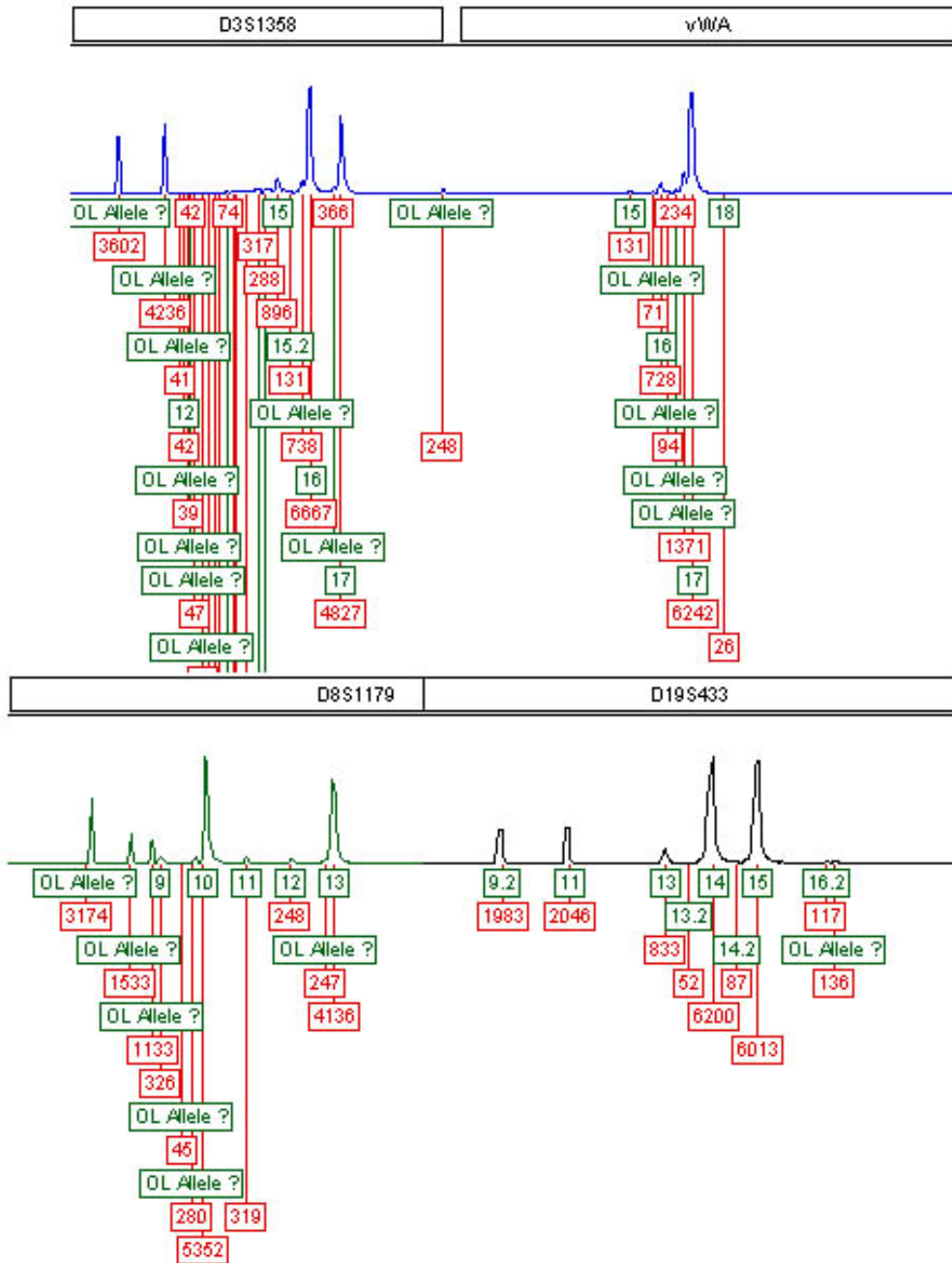

- <u>Has electropherograms that contain a few peaks (7) with high peak heights</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
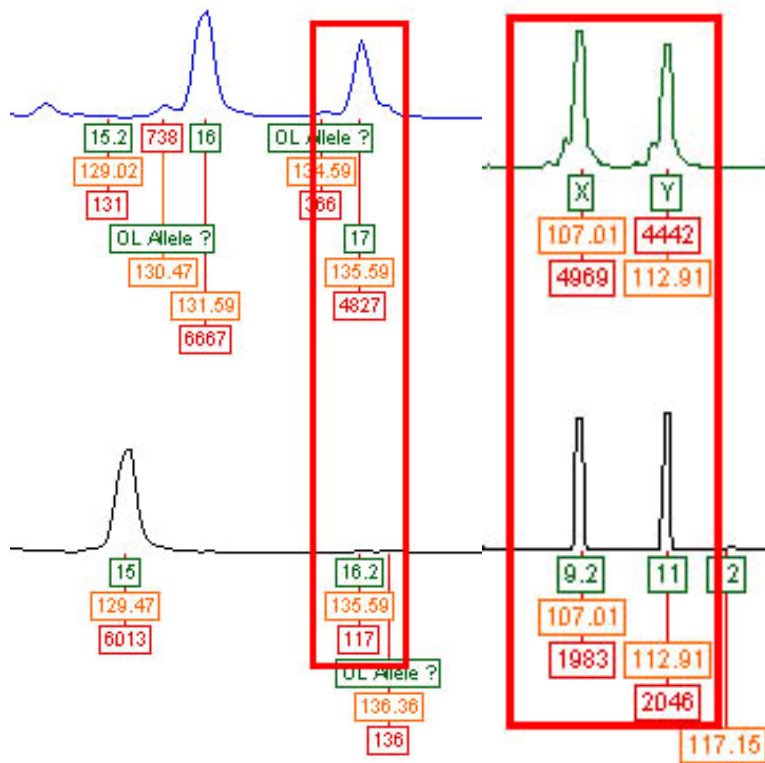
- <u>Exhibits possible pull-up</u>. The 16.2 peak at locus D19 with a size of 135.59 (117 RFU's) is possibly caused by a second peak (17) at the D3 locus with a size of 135.59 (4827 RFU's). The 9.2 peak at locus D19 with a size of 107.01 (1983 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 107.01 (4969 RFU's). The 11 peak at locus D19 with a size of 112.91 (2046 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of 112.91 (4442 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

MDH5 – Swab 2 – QC00.0221.20
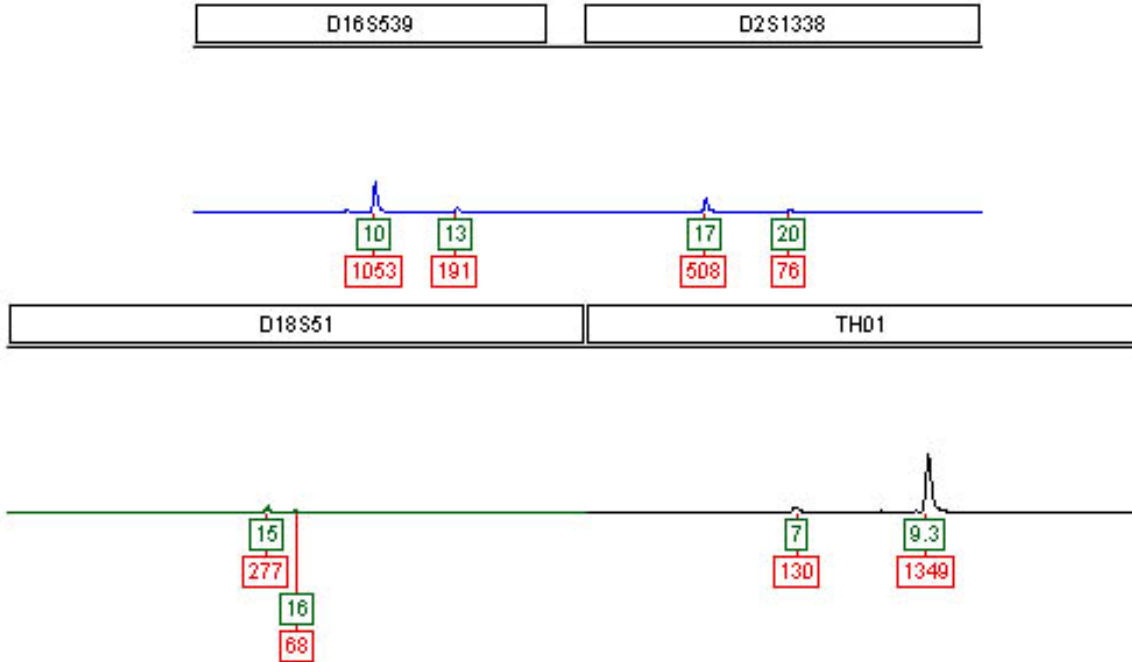
- Displays peak height imbalance at 4 loci (D16, D2, D18, THO1). The difference in the peak heights of the 10 and 13 alleles for the D16 locus (1053 and 191, respectively), the 17 and 20 alleles for the D2 locus (508 and 76, respectively), the 15 and 16 alleles for the D18 locus (277 and 68, respectively), and the 7 and 9.3 alleles for the THO1 locus (130 and 1349, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. Three of the peaks in the imbalanced loci fall below the threshold 150 RFUs, indicating that they are possibly caused by stochastic effects.
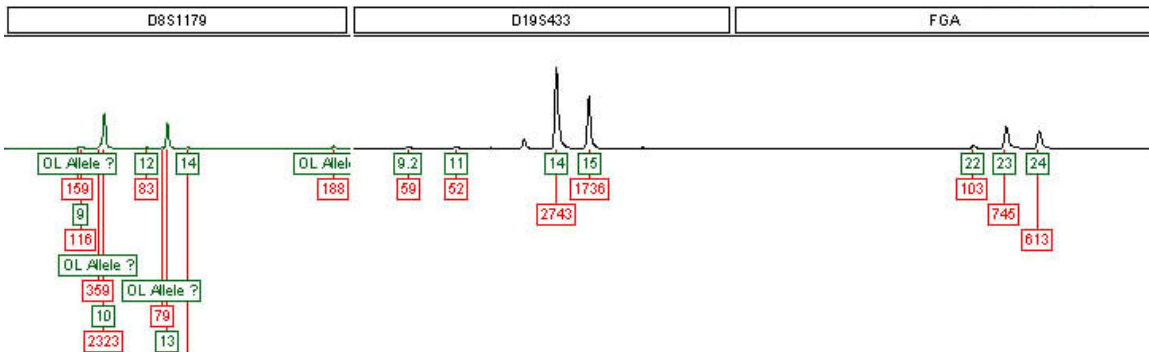

- May be a mixture of two or more individuals. Several loci (D3, vWA, D8, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with

DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
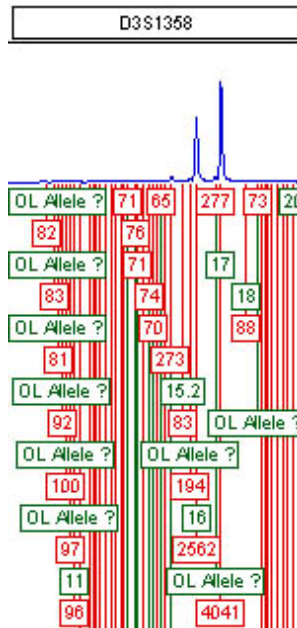
- • <u>Has electropherograms that contain one peak with a high peak height</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
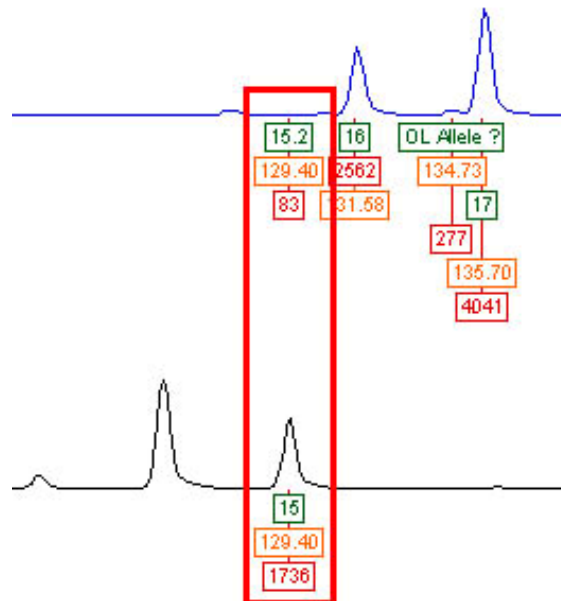
- Exhibits possible pull-up. The 15.2 peak at locus D3 with a size of 129.40 (83 RFU's) is possibly caused by a second peak (15) at the D19 locus with a size of 129.40 (1736 RFU's). The 9.2 peak at locus D19 with a size of 106.88 (59 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.88 (5367 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

JRJ8 – Tape 2 – QC03.0158.11
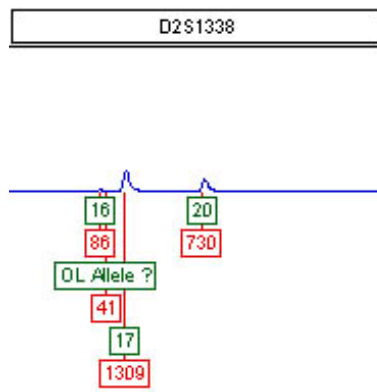
- Displays peak height imbalance at the locus D2. The difference in the peak heights of the 17 and 20 alleles for the D2 locus (1309 and 730, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.

- May be a mixture of two or more individuals. Several loci (D16, D8, D21, D18, D19, FGA) appear to have more than two alleles. Interpretation of mixed DNA

65

samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.
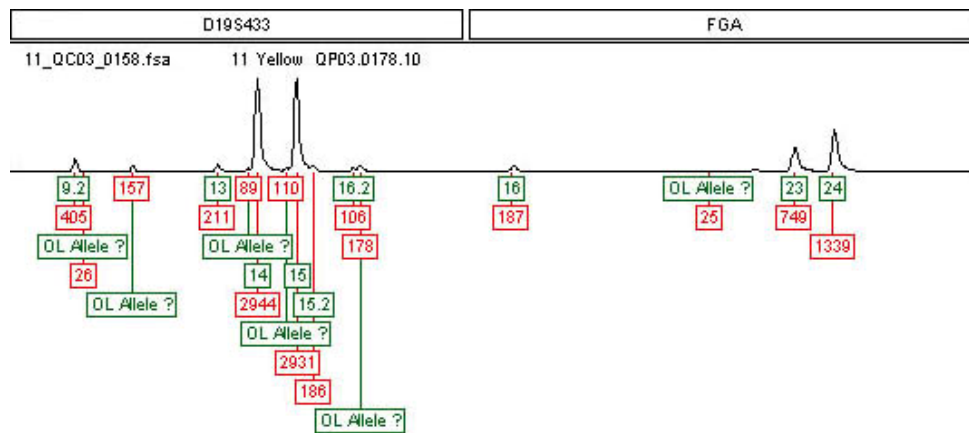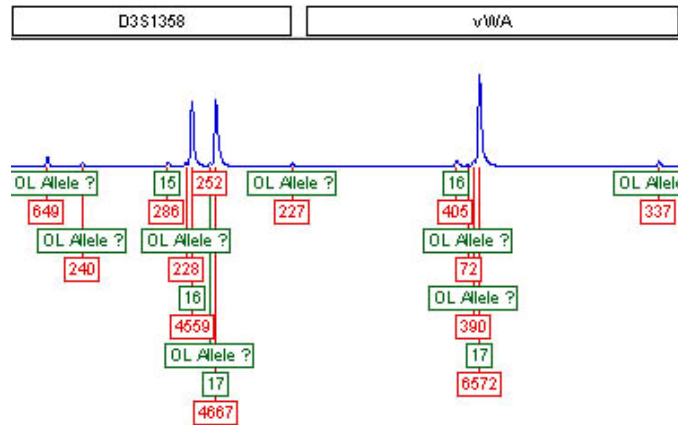

- <u>Has electropherograms that contain a few peaks (3) with high peak heights</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
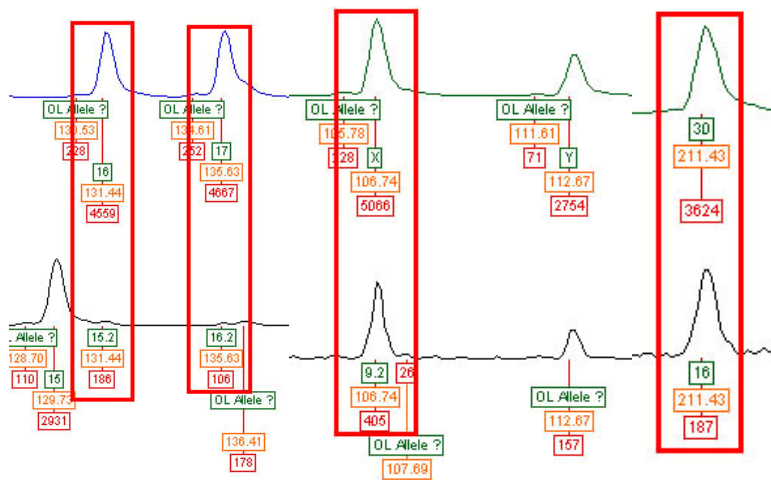
- <u>Exhibits possible pull-up</u>. The 15.2 peak at locus D19 with a size of 131.44 (186 RFU's) is possibly caused by a second peak (16) at the D3 locus with a size of 131.44 (4559 RFU's). The 16.2 peak at locus D19 with a size of 135.62 (106 RFU's) is possibly caused by a second peak (17) at the D3 locus with a size of 135.62 (4667 RFU's). The 9.2 peak at locus D19 with a size of 106.73 (405 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 106.73 (5066 RFU's). The 16 peak at locus FGA with a size of 211.43 (187

RFU's) is possibly caused by a second peak (30) at the D21 locus with a size of 211.43 (3624 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.

JRJ8 – Tape 2 – QC03.0159.11
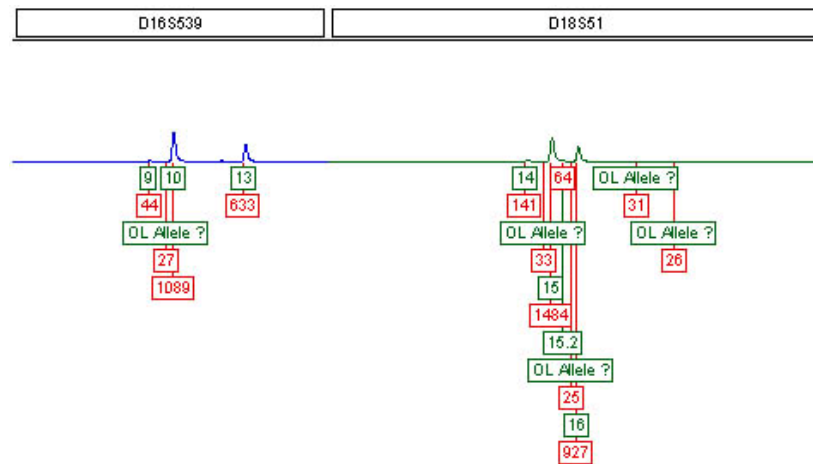
- Displays peak height imbalance at 2 loci (D16, D18). The difference in the peak heights of the 10 and 13 alleles for the D16 locus (1089 and 633, respectively) and the 15 and 16 alleles for the D18 locus (1484 and 927, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.

- May be a mixture of two or more individuals. Several loci (D8, D21, D19, THO1, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- <u>Has electropherograms that contain a few peaks (2) with high peak heights</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".

- <u>Exhibits possible pull-up</u>. The 14.2 peak at locus D19 with a size of 127.62 (97 RFU's) is possibly caused by a second peak (15) at the D3 locus with a size of 127.62 (228 RFU's). The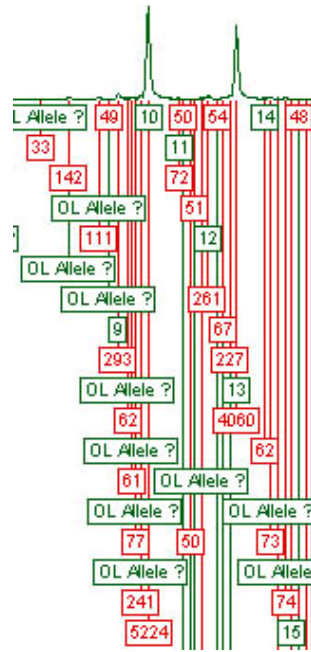 9.2 peak at locus D19 with a size of 107.00 (628 RFU's) is possibly caused by a second peak (X) at the Amel locus with a size of 107.00 (5352 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a

sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.
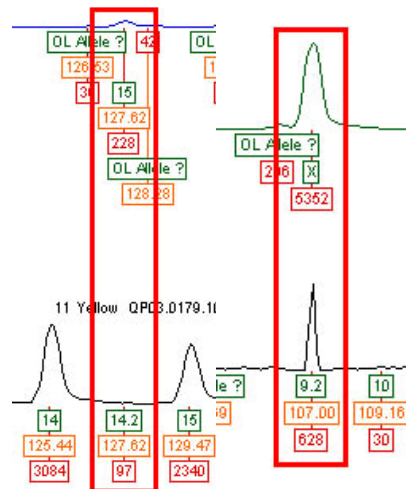
FC/2 –QC01.0067.26
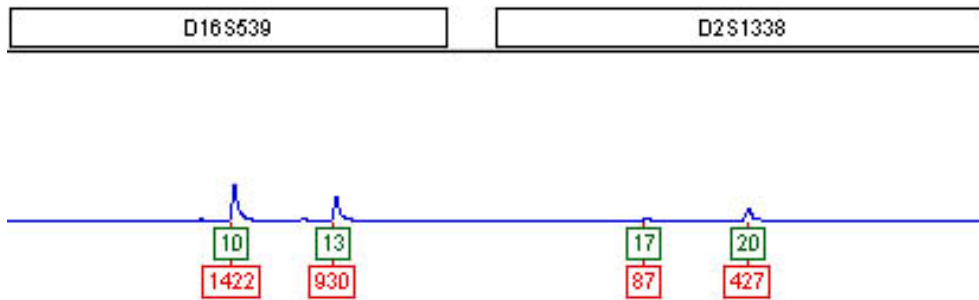
- <u>Displays peak height imbalance at 2 loci (D16, D2)</u>. The difference in the peak heights of the 10 and 13 alleles for the D16 locus (1422 and 930, respectively) and the 17 and 20 alleles for the D2 locus (87 and 427, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample. One of the peaks in the imbalanced loci falls below the threshold 150 RFUs, indicating that it is possibly caused by stochastic effects.

D8S1179   D18S51

OL Allele ? | 10 | 12   20 | Allele ?   15 | 17 | 18.2
54   79 | 60   146 | 62
OL Allele ?   OL Allele   16
78   87   60
8   OL Allele   64
61   321
OL Allele ? | 102
90 | OL Allele ?
OL Allele ? | 151
290 | 13
9 | 1880
203
OL Allele ?
109
OL Allele ?
384
2351

D19S433   TH01   FGA

9.2 | 11 | 14   7 | 8 | 121   21 | 23 | 24
93 | 91 | 15   593 | OL Allele ?   80 | 328
2866   267 | 9.3   256
1041   2632

- <u>May be a mixture of two or more individuals</u>. Several loci (D3, vWA, D8, D18, D19, THO1, FGA) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

- <u>Has electropherograms that contain one peak with a high peak height</u>. Peak heights above 4,000 RFU's may indicate that saturation has occurred. As a result, it may not be possible to detect otherwise significant peak height imbalances. High peak heights may also give rise to "pull-up".
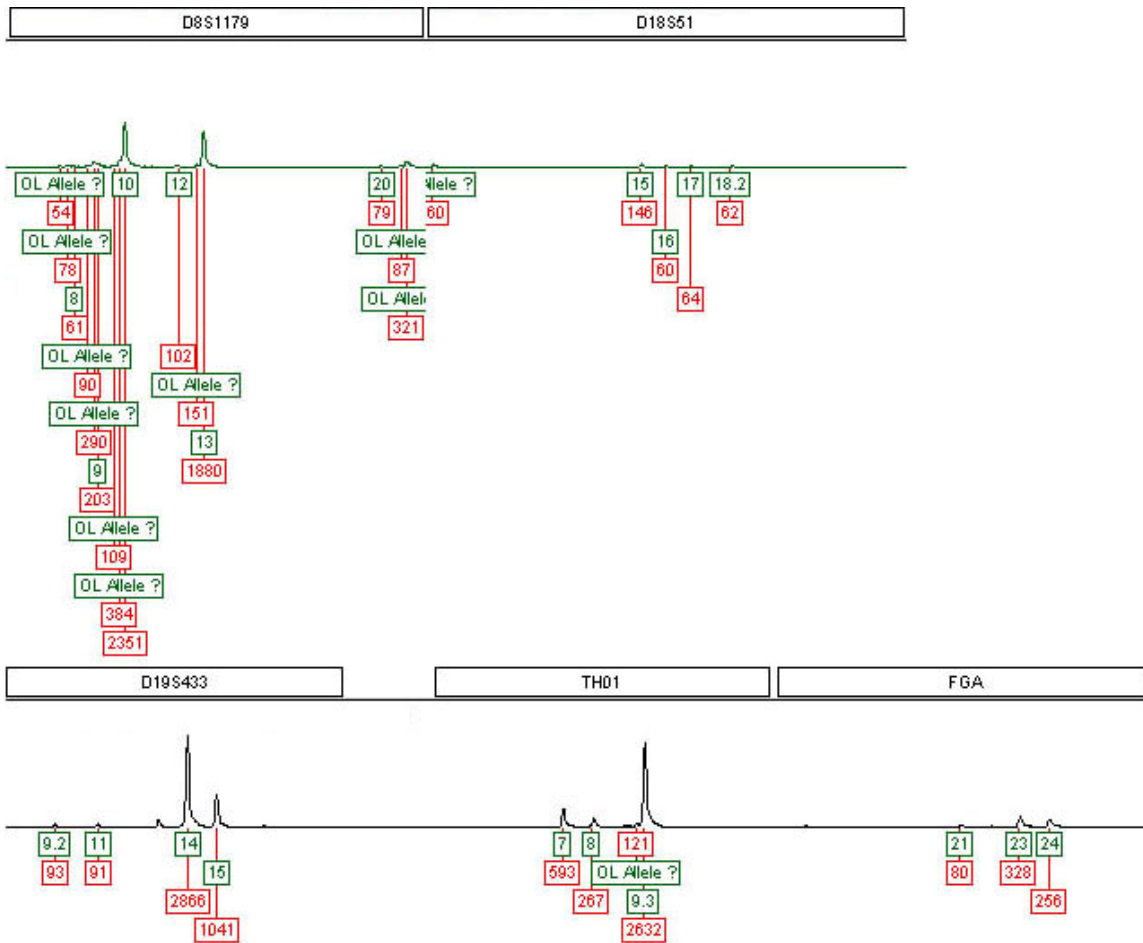
- Exhibits possible pull-up. The 15.2 peak at locus D3 with a size of 129.34 (164 RFU's) is possibly caused by a second peak (15) at the D19 locus with a size of 129.34 (1041 RFU's). The 20 peak at locus D8 with a size of 176.87 (79 RFU's) is possibly caused by a second peak (16) at the vWA locus with a size of 176.87 (665 RFU's). The 19 peak at locus vWA with a size of 188.81 (169 RFU's) is possibly caused by a second peak (9.3) at the THO1 locus with a size of 188.81 (2632 RFU's). The 11 peak at locus D19 with a size of 112.78 (91 RFU's) is possibly caused by a second peak (Y) at the Amel locus with a size of 112.78 (4639 RFU's). Pull-up is detected by seeing two peaks in two different dyes in a sample at approximately the same time. The large intensity of one peak "bleeds" into another dye sensor causing the appearance of peaks that are actually technical artifacts.
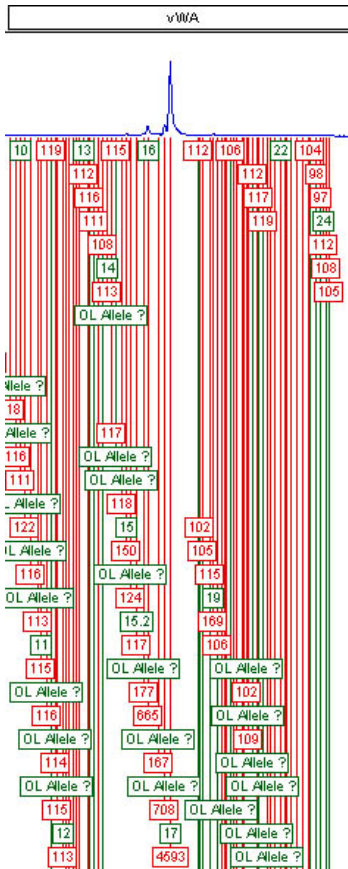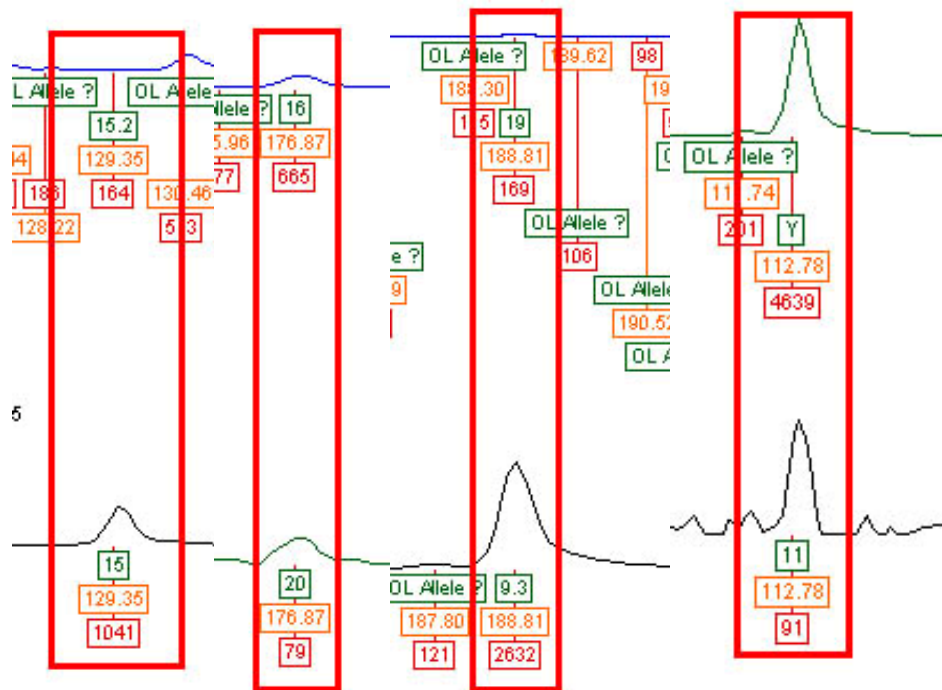
FC/2 –QC01.0068.26

D18S51

15
133
16
85

- <u>Displays peak height imbalance at the locus D18</u>. The difference in the peak heights of the 15 and 16 alleles for the D18 locus (133 and 85, respectively) could be the result of a technical artifact (such as primer binding site mutations), or be evidence of more than one contributor to that sample.

D3S1358   D16S539   D19S433

14 100 17   10 12   13 14 15
115   741 155   186 662
15   13   872
52   515
OL Allele ?
16
769
OL Allele ?
131
940
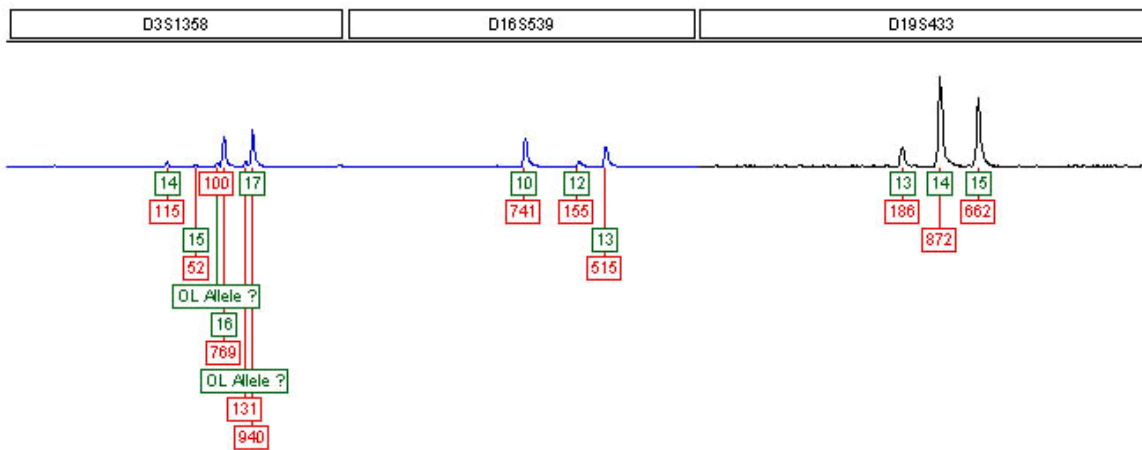
- <u>May be a mixture of two or more individuals</u>. Several loci (D3, D16, D19) appear to have more than two alleles. Interpretation of mixed DNA samples is challenging and attaching statistical significance to consistencies with DNA profiles of reference samples is difficult. A large number of alternative interpretations regarding potential contributors are possible.

Appendix C (Run-specific Limits of Detection and Quantitation for STR-based DNA Testing)

Run-specific Limits of Detection and Quantitation for STR-based DNA Testing

Jason R. Gilder,[1] M.S.; Travis E. Doom,[2] Ph.D.; Keith Inman,[3] M.Crim.; and Dan E. Krane,[4] Ph.D.

[1] Forensic Bioinformatics, Inc., 2850 Presidential Drive, Suite 150, Fairborn, OH 45324.

[2] Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435-0001.

[3] Forensic Analytical Specialties, Inc., 3777 Depot Rd., Suite 409, Hayward, CA 94545.

[4] Biological Sciences, Wright State University, Dayton, OH 45435-0001.

Corresponding author:
Dan E. Krane, Dan.Krane@wright.edu, ph (937) 775-2257, fax (937) 775-3320

1

Running header: Minimum peak height thresholds

ABSTRACT: STR-based DNA profiling is an exceptionally sensitive analytical technique that is often used to obtain results at the very limits of its sensitivity. The challenge of reliably distinguishing between signal and noise in such situations is one that has been rigorously addressed in numerous other analytical disciplines. However, an inability to accurately determine the height of electropherogram baselines has caused forensic DNA profiling laboratories to utilize alternative approaches. Minimum thresholds established during laboratory validation studies have become the *de facto* standard for distinguishing between reliable signal and noise/technical artifacts. These minimum peak height thresholds generally fail to consider variability in the sensitivity of instruments, reagents, and the skill of human analysts involved in the DNA profiling process over the course of time. Software (BatchExtract) made publicly available by the National Center for Biotechnology Information now provides an alternative means of establishing limits of detection and quantitation that is more consistent with those employed in other analytical disciplines. We have used that software to determine the height of each data collection point for each dye along a control sample's electropherogram trace. Those values were then used to determine a limit of detection (the average amount of background noise plus three standard deviations) and a limit of quantitation (the average amount of background noise plus ten standard deviations) for each control sample. Analyses of the electropherogram data associated with the positive, negative and reagent blank controls included in 50 different capillary electrophoresis runs validates that this approach could be employed to objectively determine run-specific thresholds for use in forensic DNA casework.

STR-based DNA profiling methodology is effectively at the theoretical limit of detection in that typable results can be generated from as little starting material as a single cell (1, 2). However, one of the most challenging aspects of forensic DNA analysis is the interpretation of low-level testing results where it is difficult to reliably distinguish between noise and signal from template DNA that is associated with an evidence sample (3, 4). This difficulty with minimal samples is often compounded by the consumptive nature of PCR-based DNA testing (5, 6) when material is unavailable for replicate testing. Forensic DNA testing laboratories typically endeavor to minimize the effect of baseline noise and stochastic artifacts by relying upon very conservative minimum peak height thresholds (commonly fixed in the range of 50 to 200 relative fluorescent units; RFUs) that are established during the course of their validation processes (7, 8, 9, 10). However, the conservative nature of these commonly employed thresholds can also arbitrarily remove from consideration legitimate signal from trace and secondary contributors to an evidentiary sample – matters of critical importance in many criminal investigations.

Any measurement made with a light-detecting instrument, such as a genetic analyzer is subject to at least some level of background noise (11) – defined here as signal not associated with amplified DNA. Instrument-related factors that may contribute to background noise in DNA testing experiments are typically run-specific and include (but are not necessarily limited to): the age and condition of the polymer and capillary being used; dirty capillary windows; and dirty pump blocks (12). Background noise may also differ between instruments due to differences in CCD (charged couple device) detectors, laser effectiveness and alignment, and cleanliness and alignment of the optical

components (10). Many amplification-related factors that contribute to background noise (such as analyst skill and stocks of chemicals) are also run-specific and might be reasonably expected to have varying impacts over time.

Many analytical disciplines aside from forensic DNA profiling have needed to rigorously account for background noise mixed with low levels of signal (13, 14). In the uncommon circumstances where background noise occurs at a constant level it can simply be subtracted from an analyzed signal to get true measurements of the tested material (11). It is much more common, however, for background noise, such as that associated with DNA testing results, to not be constant. In those instances, it is commonly assumed that noise magnitude is independent of analyte signal and that noise levels are distributed in a Gaussian fashion that can be effectively characterized with a mean and a standard deviation (11, 13, 14, 15). Two different signal-to-noise thresholds can be readily derived from the mean (μ) and standard deviation (σ) of the noise levels from a particular test and instrument: a limit of detection (LOD), and a limit of quantitation (LOQ) (11, 13, 14, 15). The LOD is the smallest quantity of analyte that the analytical process can reliably detect. LOD is expressed as a statistical confidence limit of noise error, usually 99.7% (i.e. three standard deviations) or:

$$LOD = \mu_b + 3\sigma_b \qquad (1)$$

where μ_b is the average amount of background noise and σ_b is the standard deviation associated with that value (11, 13, 14, 15). The LOQ represents the threshold beneath which measurements of signal strength cannot be reliably used to determine the relative quantity of detected analyte (e.g. because such measurements may include an appreciable amount of signal arising from background noise). LOQ is commonly expressed as the average background signal plus ten standard deviations (11, 13, 14, 15). or:

$$LOQ = \mu_b + 10\sigma_b \qquad (2).$$

Forensic DNA testing laboratories routinely test a positive control, negative control, and reagent blank with every DNA analysis run (7, 8, 9). While these controls are utilized primarily as sentinels for gross failures of the DNA testing processes, such as cross contamination of samples, as well as contamination or inappropriate activity of reagents, they also contain an abundance of subtle but important information about the running environment of the DNA testing system – particularly as it pertains to background noise. In this technical note, we describe a methodology that invokes generally accepted practices from other analytical disciplines and uses information associated with those ubiquitous controls to establish objective run-specific electropherogram peak height thresholds.

Materials and Methods

Baseline Noise Determination

Data for this study were obtained from 50 STR-based DNA testing runs generated by four analysts working at Forensic Analytical Specialties, Inc. (Hayward, CA) using the laboratory's validated standard protocols (e.g. no additional rounds of amplification were used as might be the case for low-copy-number analyses). All DNA profiles were generated with the Profiler Plus® commercial testing kit during the course of actual casework associated with approximately 150 cases conducted between 2004 and 2006. Each run was performed on the same Applied Biosystems 310 Genetic Analyzer and contained: a positive control; a negative control; and a reagent blank. A positive control consisted of template DNA from the 9947A immortal lymphoid cell line (16). This positive control DNA is provided by the manufacturer of the test kit and its STR

genotype is well characterized. Negative controls begin at the amplification step and contain all of the reagents used for amplification (but no template DNA). A reagent blank is a sample that contains all of the reagents used from the beginning of the extraction of a sample through amplification and typing, but again containing no template DNA. When a single run contained more than one injection of a given control, the last injection was used. No other information associated with a run (e.g. that associated with reference or evidentiary samples) was used. Electronic data files associated with these control samples (with any case-specific information removed) are available on the internet at: www.bioforensics.com/baseline/baseline.zip.

The National Center for Biotechnology Information's (NCBI) BatchExtract software (17) was used to obtain the trace and peak data from Applied Biosystem's GeneScan® sample files. BatchExtract provides the height (in RFUs) of each data collection point (DCP) for each dye along a sample's electropherogram trace. BatchExtract also provides additional information associated with labeled peaks, including the data collection points where GeneScan® considered peaks to begin and end. DCP regions containing a ROX size standard peak were excluded (masked) from consideration in all dye colors to avoid any complications from spectral overlap artifacts (i.e. pullup) (3, 4). A total of 296,592 DCPs associated with the 50 negative controls (μ = 5,932 DCP per run, σ = 131 DCP) and 297,315 DCPs associated with the 50 reagent blank controls (μ = 5946 DCP per run, σ = 87 DCP) remained for inclusion in subsequent analyses after masking was completed. Similarly, DCP regions (plus and minus 55 DCPs to conservatively account for potential stutter artifacts) associated with the expected alleles for the 9947A immortal lymphoid cell line (16) were also masked in all dye colors

for positive control samples. 120,762 DCPs associated with the 50 positive controls (μ =

2,415 DCP per run, σ = 198 DCP) remained for inclusion in subsequent analyses after

masking was completed. Shareware that performs these analyses (including masking) on

the output of BatchExtract is available at www.bioforensics.com (18).

Test Mixture

A two-person mixture was created by combining the genomic DNA of two

unrelated individuals with known genotypes in a ratio of approximately 10 to 1. The

major contributor was known to be a female with the following STR-DNA profile:

D3S1358 18, 18; vWA 16, 19; FGA 20, 21; D8S1179 13, 15; D21S11 32.2, 32.2;

D18S51 15, 17; D5S818 11, 12; D13S317 11, 11; and D7S820 8, 10. The secondary

contributor was known to be a male with an STR-DNA profile of: D3S1358 13, 17; vWA

17, 18; FGA 22, 24; D8S1179 11, 11; D21S11 28, 30; D18S51 12, 19; D5S818 11, 13;

D13S317 10, 11; and D7S820 11, 12. The electropherograms for the mixed sample were

generated with the same Applied Biosystems 310 Genetic Analyzer and protocols as

those used to generate the control samples described above.

Results

The distribution of baseline RFU level at each non-masked data collection point

(DCP) was generally Gaussian for each of the 50 analyzed negative, reagent blank and

positive controls (Fig. 1). Histograms displaying the distribution of all three controls for

all 50 runs included in this analysis can be found on-line at

www.bioforensics.com/baseline/baseline.zip. Differences in the average baseline levels

within each of the 50 analyzed runs were small between negative and positive control

samples (with an average difference of the averages of only 0.60 RFUs). Differences in the average baseline levels within each of the 50 analyzed runs were similarly small between negative and reagent blank controls (with an average difference of μ_b values of 0.41 RFUs) and between positive and reagent blank samples (with an average difference of μ_b values of 0.46 RFUs). While the inferred LOQ thresholds for all three controls were very similar within runs, average background noise values (μ_b) and standard deviations (σ_b) varied substantially between runs (Table 1) such that $\mu_b + 10\sigma_b$ (LOQ thresholds) derived from positive controls, negative controls and reagent blank controls ranged from: 27.7 to 75.7; 30.0 to 145.4; and 30.0 to 116.5 RFUs, respectively.

All of the combined average limits of detection and quantitation fall below 100 RFUs. Baseline values were found to be generally homogeneous in that the minimum and average limits of detection and quantitation were within three standard deviations of each other for each of the 150 analyzed controls. The maximum values for μ_b were generally similar in each of the three different control types, with a maximum observed difference within a run of only 8.8 RFUs (between a negative control and positive control).

Single averages and standard deviations for each of the 50 analyzed runs were also generated by considering all DCP values for a run together (i.e. independent of which of the three different controls they came from). Standard deviations for these larger data sets were generally smaller than those observed when each of the three controls were considered separately though the calculated LOD and LOQ values were very similar to those obtained by considering the three controls for runs separately (Table 1).

When considering the dye channels separately, the green channel, on average, exhibited the highest amount of baseline signal and the yellow channel exhibited the least (Table 2). The negative control containing the highest average baseline in the green channel exhibited a uniformly elevated baseline and was responsible for the single highest observed limit of quantitation (Table 1). The LOQ determined for this sample when information from all three color channels (145.4 RFUs) was found to be more conservative than the LOQ determined from the green channel alone (89.2 RFUs) such that no noise in the green channel would have been confused with signal.

A known mixed DNA profile from two unrelated individuals of an approximately 10:1 ratio was also examined using this methodology (Fig. 2). The negative control tested in the same analysis run as the mixture yielded a limit of detection (LOD) of 29 RFUs and a limit of quantitation (LOQ) of 77 RFUs. Eleven alleles (including the Y allele at the amelogenin locus) associated with the known DNA profile of the minor contributor were not labeled for this mixed sample when the GeneScan® default threshold of 150 RFUs was used. Eight alleles (including the Y allele at the amelogenin locus) associated with the male secondary contributor fall between the limit of quantitation and the commonly used 150 RFU threshold. Similarly, three additional alleles associated with the secondary contributor fall between the limit of detection and the limit of quantitation thresholds. The 17 allele (347 RFUs) at the D3 locus (which is in a stutter position relative to the major contributor's 3,509 RFU 18 allele at that locus) and the 10 allele (210 RFUs) at the D13 locus (which is in a stutter position relative to the major contributor's 2,670 RFU 11 allele at that locus) are the only alleles of the secondary contributor that are not labeled by Genotyper® when the threshold is set to the

limit of detection inferred from the negative control (29 RFUs) (Fig. 2). Two peaks with heights greater than the limit of detection that were observed in the blue channel were associated with pull-up from the green channel and were not considered.

Discussion

The similarity of the baseline levels of samples that were expected to have a high signal amplitude arising from analyte (template DNA in the positive controls) and those expected to contain little or no analyte (the negative and reagent blank controls) indicates that noise magnitude in STR-based DNA testing is independent of the analyte signal. Baseline levels for each of the three different standard controls included in each DNA profiling electrophoresis run were also very similar within runs, but differed widely between runs. These observations suggest that the baseline noise associated with capillary electrophoresis of DNA profiles is comparable to that encountered in other analytical endeavors and that generally accepted means of determining limits of detection and quantitation can be applied.

The samples analyzed in this study were primarily positive, negative or reagent blank controls. It should be possible to evaluate evidentiary or reference samples included in the same capillary electrophoresis run with the LOD and LOQ values inferred from these controls. Any peaks in evidentiary or reference samples that exceed these thresholds (such as those associated with the secondary contributor in the mixture containing DNA of two unrelated individuals with known STR-DNA profiles; Fig. 2) are unlikely to be due to baseline noise. All peaks above the threshold would then require

evaluation to ascertain whether they were signal from amplified genomic DNA, or if they may have originated from technical artifacts such as pull-up, voltage spikes or stutter.

It is worth noting that the maximum range of LOD thresholds (10.9 to 53.0 RFUs; Table 1) determined with this method in these 50 runs associated with casework performed by Forensic Analytical Specialties, Inc. is substantially below the minimum peak height threshold of 100 RFUs established by the laboratory during the course of their validation studies. Disregarding information associated with electropherogram peaks well above an analytical threshold of detection (and even above an analytical threshold of quantitation) might be considered abundantly conservative in some circumstances, given that DNA testing is a very sensitive process subject to a variety of technical artifacts such as pull-up, voltage spikes and stutter. However, in this abundance of caution, valid information about the presence of real DNA peaks is being discarded or ignored. In the instance of the mixture of two individuals with known STR-DNA profiles (Fig. 2) the lower levels of the LOQ and LOD allowed reliable recognition of alleles arising from the genomic DNA of a secondary contributor while the commonly used 150 RFU minimum peak height threshold did not. In some investigations (e.g. a mixture of a victim and perpetrator that was small enough to require consumption of the entire sample) the observation of alleles associated with a secondary contributor using the LOD threshold methodology described here could constitute critically important information that would have not been available if only conservative minimum peak height thresholds were used.

LOD and LOQ thresholds can be employed to reliably distinguish between noise and legitimate DNA signal. Two approaches can be taken with data gathered from intra-

laboratory collection of baseline data. The first is to use average LOD and LOQ

thresholds derived from both validation and current casework samples. These values

could be constantly updated. A second approach would involve the determination of

LOD and LOQ values for every run for use with the other samples within that run. In

either case, empirical statistically-derived values provide a more rigorous discrimination

between data contributed by noise and data derived from human DNA. Thorough

analyses of the data pertaining to baseline noise in control samples with software such as

NCBI's BatchExtract may help draw the attention of analysts to other important issues as

well. For instance, if one of the three control samples for a given run exhibits a larger

average and/or standard deviation of baseline levels than the others, it may be an

indication that that sample (and, perhaps the run with which it is associated) should be

evaluated with greater care. Similarly, controls with elevated average and standard

deviations of baseline activity might indicate the need for maintenance or replacement of

reagent stocks. BatchExtract is a freely available program (17) and its output can be used

with Forensic Bioinformatics' free baseline analysis program to determine the LOD and

LOQ for any control sample (18).

References

1. Findlay I, Taylor A, Quirke P, Frazier R, Urquhart A. DNA fingerprinting from single cells. Nature 1997;389:555-6.

2. Oorschot RAV, Jones MK. DNA fingerprints from fingerprints. Nature 1997;387:767.

3. Thompson WC, Ford S, Doom T, Raymer M, Krane DE. Evaluating forensic DNA evidence: essential elements of a competent defense review Part 1. The Champion 2003; 27(3):16-25.

4. Thompson WC, Ford S, Doom T, Raymer M, Krane DE. Evaluating forensic DNA evidence: Essential elements of a competent defense review Part 2. The Champion 2003; 27(4):24-8.

5. Leclair B, Sgueglia JB, Wojtowicz PC, Juston AC, Frégeau CJ, Fourney RM. STR DNA typing: increased sensitivity and efficient sample consumption using reduced PCR reaction volumes. J Forensic Sci 2003;48(5):1001-13.

6. Frégeau CJ, Bowen KL, Leclair B, Trudel I, Bishop L, Fourney RM. AmpFLSTR® Profiler Plus™ short tandem repeat DNA analysis of casework samples, mixture samples, and non-human DNA samples amplified under reduced PCR volume conditions (25 μL). J Forensic Sci 2003;48(5):1014-34.

7. DNA Advisory Board (DAB). Quality assurance standards for forensic DNA testing laboratories. Forensic Sci Comm 2000;2(3).

8. Federal Bureau of Investigation (FBI) Laboratory. National DNA Index System (NDIS) data acceptance standards. May 2005. See website http://forensics.marshall.edu/NEST/Nest%20PDFs/Documents/AppendB-NDIS-0505.pdf.

9. Scientific Working Group on DNA Analysis Methods (SWGDAM). Short Tandem Repeat (STR) interpretation guidelines. Forensic Sci Comm 2000;2(3).

10. Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B. Validation of STR typing by capillary electrophoresis. J Forensic Sci 2001;46(3):661-76.

11. Rubinson KA, Rubinson JF. Sample size and major, minor, trace, and ultratrace components. Contemporary instrumental analysis. Upper Saddle River: Prentice Hall 2000;150-8.

12. Applied Biosystems, Inc. (ABI). Chemistry reference for the ABI Prism® 310 genetic analyzer. Foster City:Applied Biosystems, Inc. 2000.

13. Anderson N. Determination of the lower limit of detection [Letter]. Clin Chem 1989;35:2152-3.

14. Thomsen V, Schatzlein D, Mercuro D. Limits of detection in spectroscopy. Spectroscopy 2003;18(12):112-4.

15. Arinbruster DA, Tillman MD, Hubbs LM. Limit of detection (LOD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused drugs. Clin Chem 1994;40:1233-8.

16. Frégeau CJ, Aubin RA, Elliott JC, Gill SS, Fourney RM. Characterization of human lymphoid cell lines GM9947 and GM9948 as intra- and interlaboratory reference standards for DNA typing. Genomics 1995;28:184-97.

17. National Center for Biotechnology Information (NCBI). BatchExtract. 2006. See website ftp://ftp.ncbi.nlm.nih.gov/pub/forensics/.

18. Forensic Bioinformatics, Inc. LOQ: baseline analysis software. 2006. See website www.bioforensics.com/baseline/loq.zip.

Additional information and reprint requests:
Dan E. Krane, Ph.D
Biological Sciences,
3640 Colonel Glenn Highway
Dayton, OH 45435-0001
Dan.Krane@wright.edu, (937) 775-2257, fax (937) 775-3320

TABLE 1—*The maximum, minimum, and average baseline levels observed in the set of reagent blanks, negative controls, and positive controls (determined from controls in 50 different runs). All values are in RFUs.*

Positive Control		μ_b	σ_b	$\mu_b + 3\sigma_b$	$\mu_b + 10\sigma_b$
	Maximum	6.7	6.9	27.4	75.7
	Average	5.0	3.7	16.1	42.0
	Minimum	3.7	2.4	10.9	27.7
Negative Control		μ_b	σ_b	$\mu_b + 3\sigma_b$	$\mu_b + 10\sigma_b$
	Maximum	13.4	13.2	53.0	145.4
	Average	5.4	3.9	17.1	44.4
	Minimum	4.0	2.6	11.8	30.0
Reagent Blank		μ_b	σ_b	$\mu_b + 3\sigma_b$	$\mu_b + 10\sigma_b$
	Maximum	6.5	11.0	39.5	116.5
	Average	5.3	4.0	17.3	45.3
	Minimum	4.0	2.6	11.8	30.0
All three controls averaged		μ_b	σ_b	$\mu_b + 3\sigma_b$	$\mu_b + 10\sigma_b$
	Maximum	7.1	7.3	29.0	80.1
	Average	5.2	3.9	16.9	44.2
	Minimum	3.9	2.5	11.4	28.9

TABLE 2—The maximum, minimum, and average baseline levels observed in each of three color channels for reagent blanks, negative controls, and positive controls. All values are in RFUs.

Positive Control	Dye	Minimum	Average	Maximum
	Blue	3.7	5.2	9.7
	Green	4.3	5.8	7.4
	Yellow	3.0	4.1	6.4

Negative Control	Dye	Minimum	Average	Maximum
	Blue	4.0	5.3	8.0
	Green	4.6	6.7	31.2
	Yellow	3.0	4.0	6.4

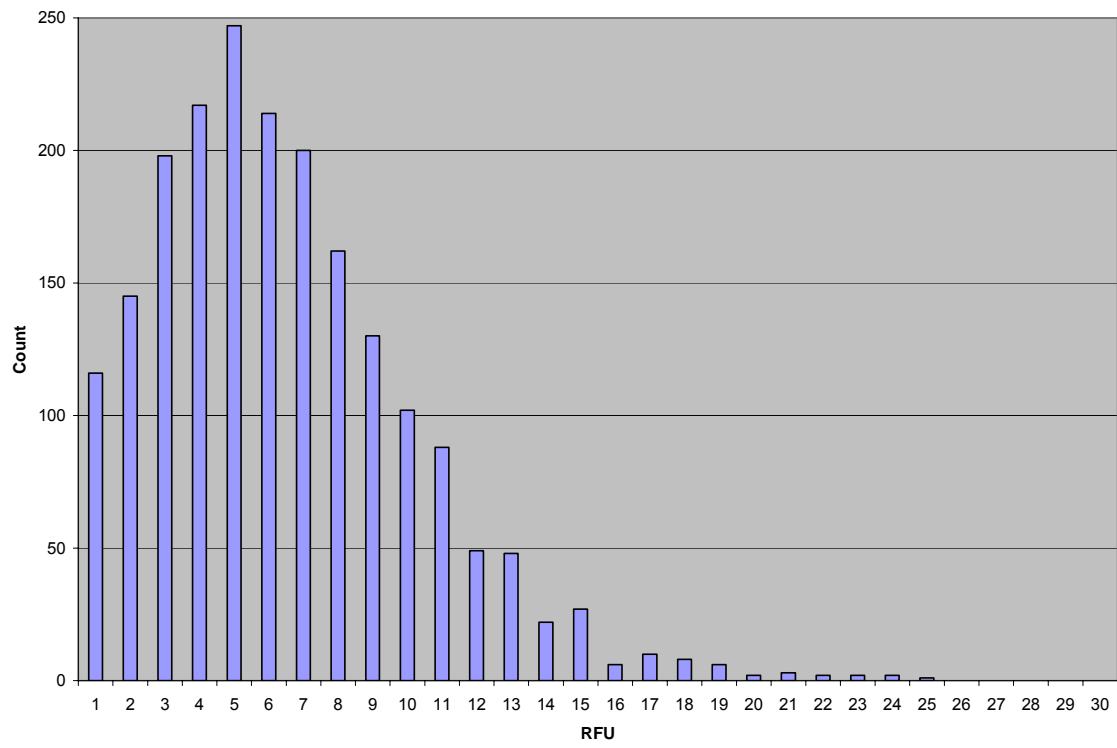| Reagent Blank	Dye	Minimum	Average	Maximum
	Blue	3.7	5.4	8.5
	Green	4.8	6.2	8.6
	Yellow	3.4	4.3	6.1

FIG. 1—*A representative histogram taken from the distribution of measured RFU levels at all non-masked data collection points in the first of 50 negative control samples after masking. This distribution is from a blue channel and exhibits an average baseline approximately equal to that of the population's average baseline signal (5.5 RFUs).*
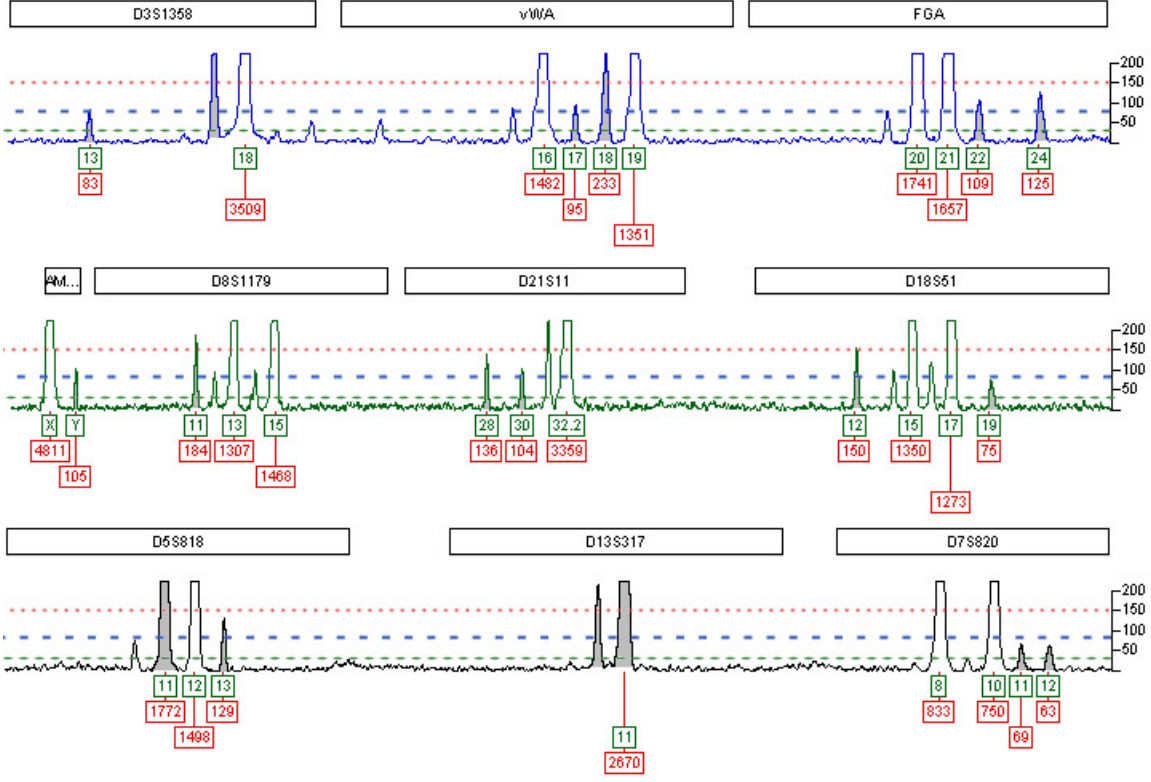
FIG. 2—*Electropherograms from an approximately 10:1 mixture of two reference samples. Three different thresholds are shown: a minimum peak height threshold at 150 RFU (dotted line); a limit of quantitation (LOQ) threshold determined to be at 77 RFUs from the negative control for this electrophoresis run (dashed line); and a limit of detection (LOD) threshold determined to be at 29 RFUs for this electrophoresis run (small-dashed line). Genotyper® assigned allele calls (with ABI stutter filters in place) are shown in boxes immediately below the electropherogram peaks while peak heights (in RFUs) are shown in boxes below those labels for all peaks with heights greater than the LOD. Peaks consistent with the known profile of the minor contributor are shaded.*

Signed: Date: October 24, 2006