IN THE UNITED STATES DISTRICT COURT

FOR THE WESTERN DISTRICT OF MICHIGAN

SOUTHERN DIVISION

UNITED STATES OF AMERICA,

Plaintiff, No. 1:17cr130

vs.

DANIEL GISSANTANER,

Defendant.

Before:

THE HONORABLE JANET NEFF,
U.S. District Judge
Grand Rapids, Michigan
Wednesday, May 23, 2018
Motion Proceedings

APPEARANCES:

MR. ANDREW BIRGE, U.S. ATTORNEY
By: MR. JUSTIN PRESANT
The Law Building
330 Ionia Avenue, NW
Grand Rapids, MI 49501-0208
616-456-2404

On behalf of the Plaintiff;

FEDERAL PUBLIC DEFENDERS
By: MS. JOANNA CHRISTINE KLOET
MR. PEDRO CELIS
MS. HELEN NIEUWENHUIS
Federal Public Defender's Office
50 Louis NW
Suite 300
Grand Rapids, MI 49503
616-742-7420

On behalf of the Defendant.

REPORTED BY: MS. KATHY J. ANDERSON, RPR, FCRR

May 23, 2018

PROCEEDINGS, 9:09 a.m.

THE LAW CLERK: All rise. Court is now in session. Please be seated.

THE COURT: Good morning, everybody.

MS. KLOET: Good morning, Your Honor.

THE COURT: This is the date and time set for hearing on two motions, the government's motion to exclude defense witness Nathan Adams, and the defense motion to exclude DNA evidence in case number 1:17cr130, the United States of America versus Daniel Gissantaner.

Can we please have appearances and introductions.

MR. PRESANT: Good morning, Your Honor. Justin Presant for the United States. With me at counsel table is Stephanie Miller who is a paralegal in our office and will be assisting with the presentation of evidence today.

THE COURT: Thank you.

MS. KLOET: Good morning, Your Honor. Joanna Kloet, Assistant Public Defender. To my left is Pedro Celis, our research and writing attorney who has filed an appearance on this case, Helen Nieuwenhuis who has also filed an appearance, Mr. Gissantaner, Emily Seale is a paralegal that will be assisting with the presentation of exhibits, and investigator Carlos Clay.

THE COURT: Okay. A couple of housekeeping matters

before we begin. First of all, Mr. Presant, your voice is soft, and Ms. Kloet your voice is soft, and secondly, we are going to be dealing with some relatively technical issues, I think, throughout this hearing. So my request is, first of all, that you speak directly into the microphones so that the court reporter and I can hear you, and that you don't hurry your speech. And, Mr. Presant, you are particularly guilty of that. It's hard enough to, for her, for the court reporter to take down your comments, but when it involves technical terms and so forth, it's even more difficult. And I would echo that for any experts who are going to be testifying today. We are going to be dealing with an important issue here. I want to have as much clarity on the record as possible. And I will say this. Be prepared for interruptions because I'm going to have I think a lot of questions.

So let's start -- I think logically we should start with the government's motion regarding Mr. Adams. I have read your memos. I have read all of the transcripts that the government has provided, and I think this: The testimony of Mr. Adams is properly offered. The scope of it is another matter altogether. And it is, it is clear to me from the response filed by Ms. Kloet last night that Mr. Adams's testimony is going to be narrowly tailored to his educational and experience background, at least if I understood her memo correctly. So I am going to allow him to testify today.

The scope of his testimony, the scope of his expertise I think I am capable of understanding and properly applying. The government may disagree, and of course, Mr. Presant, you're more than free to put your objections on the record. But I am comfortable, again, after having, particularly having read the I think there were five or maybe six transcript excerpts which were offered, which were very interesting, actually, I'm comfortable allowing Ms. Kloet to put him on the stand and hear what he has to say. Okay.

MR. PRESANT: Your Honor, if I may. Just to be clear, the government was not seeking to exclude him from this hearing today. Just to limit his testimony to what he was properly qualified on. The government is seeking exclusion at trial, but I think the Court has to hear what his testimony is before it can rule on that.

THE COURT: I understand. And I think, you know, it is a little bit unusual to have this kind of an issue come up at this stage, but I do understand where you're coming from, Mr. Presant, and it's actually I think good to have the issue out in front of us early so that we have some idea what we're dealing with.

So that being solved at this point, let's move on to the defendant's motion to exclude DNA evidence. Ms. Kloet, the ball is in your court.

MS. KLOET: Your Honor, as we discussed this with

Attorney Presant, and in light of the fact that the burden of proof is on the government, we have agreed that he would present his witnesses first. And I think in light of scheduling, that makes the most sense since people are coming from out of town.

THE COURT: Very well. What I would like you to start off with is a brief summary to put this issue in context. I've done some reading on the issues that are going to be presented, probabilistic genotyping, likelihood ratios and all of that. But what I know is that there is a potential, potentially broad range of topics within those designated areas, and so what I would like you to do, both of you, before we get to witnesses is to put this case in context in terms of what do you expect the issues to be surrounding the technical question of the DNA results that were reported in this case. I do think that we have, we can narrow this down so that at the point where a decision is necessary, it can be very focused in terms of what is or is not properly submitted to the jury, which is really our underlying goal here. What can we properly offer to the jury that is understandable, that is not prejudicial, and so forth. So let's start there. Give me a short -- as soon as I can find my legal pad to write on -- summary putting the case, putting the situation into context. Mr. Presant.

MR. PRESANT: Thank you, Your Honor. Of course it's the defendant's motion so it's, the government agrees with

Ms. Kloet that the government should present witnesses first. But of course it's her arguments, and I don't know if the arguments have narrowed somewhat from the time of the initial motion. Maybe they have broadened a little bit in some respects. But I'll do my best to address the Court's concerns and then of course we will be happy to take any questions from the Court before we call our first witness.

The main issue in the government's view is the admissibility of DNA analysis predicated on or derived from, rather, STRmix which is software that was developed by the New Zealand and Australian governments and is currently the market leader in the United States in terms of probabilistic genotyping software used by federal, state and local laboratories in order to analyze DNA mixtures.

The defendant in his initial motion raised a number of challenges to the presentation of such DNA analysis. The challenges I think begin logically in the pipeline earlier even before we get to the probabilistic genotyping software. There are some passing critiques of polymerase chain reactions, the process of replicating DNA in order for it to be analyzed. Also the use of capillary electrophoresis which is the process by which DNA fragments are separated by the genetic analyzer. In the government's view those are not substantial challenges because the Court will hear testimony that that technology has been in use for many decades. There's nothing new about that

technology in particular and perhaps Ms. Kloet doesn't challenge that anymore.

But the next step in the pipeline we get to the use of the software after the DNA has been analyzed chemically. And what the software does is it uses probabilistic genotyping, which is a subspecialty of statistics; it's really advanced statistics applied to biology, bioinformatics, if you will, and even though STRmix is relatively new, and software programs like STRmix are relatively new, probabilistic genotyping is built on really established principles, statistical principles that have been around some of them for more than a century, and the Court will hear testimony about that too.

And so what the software does is it analyzes that in order to produce a likelihood ratio that can be used in court as evidence.

The government's view is that at the time of trial all we will really need is the Michigan State Police forensic scientist who did the work in this case to testify to what the likelihood ratio was after she did the analysis. And she will testify during the proceeding too about what she did to come up with that likelihood ratio.

Now, after the Daubert motion was filed the defendant raised an additional challenge to the code, the underlying code, that's used by the software. It's different in kind from the previous challenges in that the previous challenges were

directed more towards the use of probabilistic genotyping, the use of the software, the decisions that the analysts made that led to that final likelihood ratio. The new criticism is, well, the code was not, was not coded in the preferred style of the defense witness. The defense witness, Mr. Adams, does have some expertise in computer science, and he conducted a code review subsequent to the last time we were in court on this matter and he looked at some of the code and he has some criticisms of the way it was coded. In the government's view those are stylistic preferences. He doesn't, he's never really created software like this before, and at the end of the day the key for this Court and that the other courts that have reviewed STRmix have largely relied upon is the validation of the software. That before the software is used it is tested, and much like Your Honor or, well, I'll speak about myself, much like I don't understand how an internal combustion engine in a car works, I can tell you when I get in my car and I drive it it gets me where I need to go, I can use it properly, it works safely. And you're going to hear testimony about how STRmix was coded and the coding decisions that were made that respond substantively to Mr. Adams's criticism.

But I think at the end of the day the key is was this software tested properly before it was put into use, and if it was tested properly, which the Court will hear testimony that it was, you can be confident that it does what it is supposed

to do and that it functions properly in this case.

THE COURT: Do you also understand that there may be, or, it is possible that there may be a challenge to the protocols used by the Michigan State Police or whoever applied this STRmix software to the DNA samples in this case?

MR. PRESANT: Yes. The Court will hear testimony during this proceeding, and the government anticipates there would be testimony at trial, that STRmix is a big step forward in that it imposes uniformity on how probabilistic genotyping is done, but there inevitably it is a human exercise and so there is, there are human judgment calls that go into operating the software. There are things that the forensic scientists, Amber Smith in this case, had to do in order to use the software, and I think at trial those are properly subject to cross-examination. She will testify to those during this hearing. She will testify to those calls she made at trial. There's of course been an offer to rerun the software using different parameters if the defendant had different parameters that he wanted to input into the software; that request has not been made yet. But there will be testimony -- I believe Ms. Kloet is challenging that and there will be testimony on that issue.

The one final issue that I haven't touched on yet is the challenge to the use of likelihood ratios themselves, statistical measures. There are different types of statistics

that are used in DNA cases. Likelihood ratios have been used for more than a decade by the Michigan State Police or at least about a decade in paternity cases. Likelihood ratios are appropriate for particular reasons that the experts will testify to. And I know Ms. Kloet raises a challenge to the use of likelihood ratios at all as a statistic, and so I think that will be the other issue the Court has to decide.

THE COURT: My understanding with regard to that challenge has more to do with the number of potential contributors to the DNA sample. And the limited reading I have had an opportunity to do suggests that there is a real difference in how one applies or uses likelihood ratios where there is a single contributor of the DNA and where there are multiple contributors, such as in this case where there are three. So, again, I think we are going to need to be very specific and focused on exactly what we are talking about in this case, both for my benefit and ultimately for the jury's as well.

So is there anything else that you need to put this case in context from the government's perspective?

MR. PRESANT: I don't believe so, Your Honor. I think everything the Court said is absolutely correct, and the government is prepared for the Court to interrupt as much as it needs to because the goal of course is for you to get your questions answered. And I appreciate the opportunity to

present evidence to that.

THE COURT: Thank you, Mr. Presant. Ms. Kloet.

MS. KLOET: Thank you, Your Honor. I brought this motion I think there are, it can be distilled into two main arguments:

First of all, the reliability of the software is at issue here. There are two components to that: The subjectivity of the data that's entered by the analyst, and that includes as the Court referenced the estimation as to the number of contributors, allele calls, or whether you determine is not an allele. And as well as the parameters that are entered into the program and the parameters used and set by STRmix and similar systems themselves.

The testing and validation of these programs is also a critical issue. There has not been a broad universe of testing, broad enough universe of testing done on complex mixtures involving ostensibly three or more contributors to that mixture, assuming there are in fact three or more in that mixture.

The testing is also not done in accordance with generally accepted software testing principles. And it's important when you're dealing with a program that concededly doesn't generate the same answer every single time because you'll never know the ground truth, as referred to in the industry. And that testing is not sufficiently independent as

required by those software standards.

The second part of this motion, Your Honor, is the immense and disproportionate influence that this type of statistic is going to have on a juror, on a non mathematician, non science based, someone without a science background.

The obscurity surrounding the effective, this type of information and the effective communication of -- the difficulty in effectively communicating this type of information is demonstrated and has been noted by experts even within the federal government itself.

And the witnesses here today, Dr. Howenstine if she testifies, will be testifying to the elements of these points, including the subjectivity of the calls that are made by the analyst and where issues can arise there, including in the analytical process itself; Mr. Adams will testify to software as it applied to Probabilistic Genotyping Systems in this case and software standards with which he is familiar; and Dr. Lund will talk about issues in the communication of this type of information.

So that's what I anticipate covering today, Your Honor.

THE COURT: Thank you. Okay. Mr. Presant, let's hear it.

MR. PRESANT: Your Honor, just one evidentiary issue that relates to the government's first witness. The

government's first witness is Dr. John Buckleton. He lives in New Zealand. The government anticipates he will be unavailable for trial, so if these issues do end up being relitigated at trial, the government will seek to admit his testimony from today's hearing under the Rules of Evidence. And I want to make sure that Ms. Kloet is on notice. We've previously discussed this, she's previously been advised of it, but I want to make sure that the record is clear so that the testimony can be developed appropriately from the defense perspective.

THE COURT: Very well.

MR. PRESANT: The government calls Dr. John Buckleton.

THE LAW CLERK: Please come forward over to the podium and turn toward me. You can set your materials down if you like.

JOHN BUCKLETON, GOVERNMENT WITNESS, WAS DULY SWORN

THE LAW CLERK: Please be seated. And state your full name for the record and any unusual spellings please spell those names.

THE WITNESS: My full name is John Simon Buckleton.

DIRECT EXAMINATION

BY MR. PRESANT:

Q Dr. Buckleton, where do you currently work?

A Largely in the United States, but the organization I work for is the New Zealand government. Mr. Presant, I have with me a typed glossary for the stenographer if that's any help, and I

have four copies if you and the defense would like them.

MR. PRESANT: I appreciate that very much. We can distribute those now or later.

THE COURT: Good idea to do it now, I think.

THE WITNESS: There is two pages. You have to separate them.

BY MR. PRESANT:

Q Dr. Buckleton, you work for a particular institute within the New Zealand government, is that right?

A I work for ESR which is a Crown Research Institute.

Q What does that mean, what's a Crown Research Institute in New Zealand?

A A Crown Research Institute is a New Zealand government owned institute. However, we have the expectation to operate in a fiscally prudent manner.

Q Ms. Miller, can we bring up Government's Exhibit 1, please?

MR. PRESANT: Your Honor, would you prefer electronic presentation of evidence?

THE COURT: Yes.

MR. PRESANT: All right.

BY MR. PRESANT:

Q And Dr. Buckleton, you can refer to the book or the screen if you like. Do you recognize Exhibit 1?

A Yes.

Q What is it?

A This is my curriculum vitae.

MR. PRESANT: Your Honor, the government moves to admit Exhibit 1.

THE COURT: Any objection?

MS. KLOET: No, Your Honor.

THE COURT: It's admitted.

BY MR. PRESANT:

Q Dr. Buckleton, I want to briefly go through your CV with you. Would you start for -- start by describing for the Court your academic background.

A I have a Ph.D. in chemistry from the University of Auckland, and DSc in forensic statistics. DSc is not a term familiar in the United States. It's a British Commonwealth term, and it's the highest academic qualification you can get in the British Commonwealth.

Q What about your employment record as it's listed here, what have you spent your career doing?

A From 1983 to the present, I have been employed in forensic science largely for the New Zealand government. However, I have had four periods of employment in the United Kingdom, and two periods in the United States.

Q What other positions have you held besides for the New Zealand government?

A Specifically referring to the United States, I was a researcher at North Carolina State University, then a visiting

scholar at NIST, the National Institute of Science and Technology, and then a researcher at University of Washington at Seattle.

Q Has your experience in forensic science focused exclusively on DNA or have you worked in other forensic disciplines as well?

A I have worked across a range of disciplines, but more recently my world has narrowed almost exclusively to probabilistic genotyping.

Q How long have you been working in forensic DNA analysis of any kind?

A I was involved in forensic DNA analysis from the very inception of it in 1998 when it was first used in case work in the United Kingdom.

Q At the bottom of this first page of your CV it looks like you testified before in a number of cases.

A I have testified in nine Frye Daubert or Kelly Frye hearings in the United States, once in the Netherlands, once in Scotland, once in Australia on STRmix, and multiple times in Australia and New Zealand on a range of topics.

Q And regarding your prior testimony in the United States, those were all in state court, correct?

A No. Some were in county court, I believe. I have never --

Q Let me ask the question this way. Is this your first time testifying in federal court?

A Yes.
Q Ms. Miller, can we go to page 2? What do we have here listed on page 2, Dr. Buckleton?
A I have listed some of the grants and major contracts that have been awarded to myself and others. So you can see three National Institute of Justice Research awards, and then I begin a list of paid plenaries and speaking engagements.
Q Page 4, please, Ms. Miller. I'm sorry, page 5. On page 5 you begin a list of publications, is that right?
A Yes.
Q And that goes on for several pages.
A Yes.
Q It actually goes almost to the end of the CV, correct?
A Yes.
Q So I want to ask you about your experience analyzing DNA mixtures. When did you first begin analyzing mixtures?
A In the early 1990s.
Q And how did you do it back then?
A It was done with a likelihood ratio but the process was manual.
Q And when did you first begin to develop STRmix?
A STRmix has very deep roots, but more specifically, the software was begun in May 2011.
Q What was the reason to begin to develop the software, what was the origination of the idea?

A There was a laboratory closure in Australia, and this led to a meeting of the senior managers of the Australasian laboratories. Australasia is a collective term for Australia and New Zealand. These managers were -- initiated a standardization project within Australasia and myself and Dr. Duncan Taylor were tasked with creating a software.

THE COURT: I'm going to interrupt here a minute, Mr. Presant. Dr. Buckleton, when you say that likelihood ratios were used in the early 1990s with a manual process, could you explain what you mean by that?

THE WITNESS: Yes, ma'am. Likelihood ratios have been used in forensic science since the 1940s, and in fact, even way back to the 1910s. Specifically answering your question, a likelihood ratio would be calculated for a mixture and the calculations would be done by hand, and a number of simplifying assumptions had to be made to do that, and we could realistically only do two-person mixtures at that stage.

THE COURT: Okay. Now, you're familiar to some extent with the American criminal justice system, right?

THE WITNESS: Yes, ma'am.

THE COURT: Okay. And it is not uncommon for a jury to be presented with a simple statement that says, the DNA analysis shows that it is a match for the defendant. Okay. What does that mean to you, and how is that different from this likelihood ratio?

THE WITNESS: We have been encouraged away from that type of statement by such groups as the National Forensic Science Commission, now disbanded. That is a yes/no statement, and takes to the witness the task of determining certainty from probability.

I'm taught that such judgments should really be left to the factfinders, and I am trained to try and provide the factual and opinion evidence to enable a factfinder to make such a decision.

THE COURT: Well, if we tell the jury the defendant's DNA was found on the gun, just as an example, what does that mean to you as a forensic scientist in terms of how that communication was arrived at? Does it mean anything at all?

THE WITNESS: I think the United States Federal Bureau of Investigation have done source attribution statements which are what you're speaking to for a great many years and that is where they say that with certainty this DNA came from that person. Such a statement can still be made from a likelihood ratio if desired. In this particular case we wouldn't meet the federal limit of certainty. However, I understand the evidence is still very probative.

THE COURT: What is the federal limit of certainty?

THE WITNESS: I think it was a multiple of the population of the United States. I think it was something like, you know, I think it was ten times the population of the

United States. But I can have that checked. There is a person who may know that in the court. I can check that at a break.

THE COURT: Okay. Thank you.

BY MR. PRESANT:

Q Thank you, Your Honor. Dr. Buckleton, what have your responsibilities been in rolling out STRmix to various laboratories that might be interested in using it?

A I'm one of the three developers of STRmix. And I have been responsible for training and support and development. During that time I have run of the order of 70, four-day training sessions in the United States for different groups around the United States, and I have worked with people on the setting of procedures and guidelines.

Q What's your and ESR's financial interest in STRmix?

A I have no personal financial interest in STRmix. I do not benefit from STRmix financially at all.

Q You're paid a salary?

A I'm on my New Zealand government salary. ESR takes some of but not all of the proceeds of STRmix, but it is not required to pay a dividend to anyone. Their money is used for its support work, development work, and, for instance, it's funding my court appearance here.

Q And you're appearing here today voluntarily, correct?

A Yes.

Q The government has covered, the United States government

has covered some travel expenses but we are not paying you to be here, right?

A That's correct.

Q You have held positions with governing bodies, bodies that govern DNA analysis, correct?

A Yes.

Q What are those organizations?

A The primary organization in the United States and Canada is SWGDAM. That is an FBI sponsored group of North American scientists and invited guests, and they are the premier guideline setting organization for forensic DNA. I have also been involved in some others, but that is the primary one in the United States.

Q Well, let's start with SWGDAM which is the scientific working group for DNA analysis methods, correct?

A Yes.

Q What did you do for SWGDAM?

A I was on SWGDAM from 2013 to 2017 where I retired from it in the hope of relocating back to New Zealand. During that time I operated as a member of SWGDAM. With specific regard to probabilistic genotyping, I sat out on all those discussions to avoid the appearance of undue influence.

Q But otherwise you reviewed various issues with respect to DNA and you voted on them or you consulted on them?

A I'm not a voting member. There are votes given, for

instance, the FBI has a large block of votes. I was not a voting member but I was involved in all the discussions and recommendations and drafting processes that were put to me.

Q You testified that you worked at NIST for some time period?

A Yes.

Q What time period was that?

A That was from October 2014 to December 2016.

Q What did you do at NIST?

A I was a visiting scholar, and my work there involved work on forensic evidence interpretation and in particular, footwear.

Q Have you worked --

THE COURT: What does that mean?

THE WITNESS: Footwear, shoe print work, Ma'am. I was deliberately staying away from DNA to again avoid the appearance of undue influence.

THE COURT: Thank you.

BY MR. PRESANT:

Q What about other organizations besides SWGDAM and NIST?

A The other organizations that are the primary guideline setting bodies that the United States tends to take cognizance of are the International Society of Forensic Geneticists, and I was on the DNA commission that wrote the probabilistic genotyping guidelines. OSAC, O-S-A-C, which is the Organization of Science Area Committees, which I was on until

2017, and that has not made a work product in this field. PCAST, the President's Council of Advisors On Science and Technology, I'm not on that but I was invited for discussions with them on that.

Q What did you do for OSAC?

A I was on various committees, including the proposition setting document and the DNA subcommittee.

MR. PRESANT: Your Honor, at this time the government offers Dr. Buckleton as an expert in forensic DNA analysis, probabilistic genotyping, and software development.

THE COURT: Ms. Kloet, did you wish to voir dire Dr. Buckleton on his qualifications as an expert in those three fields?

MS. KLOET: The three fields, can you repeat the three fields?

MR. PRESANT: Forensic DNA analysis, probabilistic genotyping, software development.

MS. KLOET: I would like to ask a few questions.

THE COURT: Thank you.

MS. KLOET: Dr. Buckleton, your background or your education is not in computer science, correct?

THE WITNESS: Yes, it is.

MS. KLOET: Your formal education is not computer science, is it?

THE WITNESS: Some of it is.

MS. KLOET: Do you have a degree in computer science?

THE WITNESS: I have undergraduate papers in it but not a major in it.

MS. KLOET: Okay.

THE COURT: What is your undergraduate degree in, Dr. Buckleton?

THE WITNESS: Chemistry, Ma'am.

THE COURT: That's your Ph.D. as well?

THE WITNESS: Yes.

THE COURT: Do you have a master's degree in something different?

THE WITNESS: Chemistry. Three degrees in chemistry.

THE COURT: Okay.

MS. KLOET: You didn't write the code in STRmix, did you?

THE WITNESS: No.

MS. KLOET: You don't write computer code.

THE WITNESS: I have.

MS. KLOET: Do you do that on a regular basis?

THE WITNESS: I haven't done that regularly since the mid-'90s.

THE COURT: Who did write the code in STRmix?

THE WITNESS: The version in question in this case was written almost totally by Dr. Duncan Taylor of Adelaide, Australia.

THE COURT: T-A-Y-L-O-R?

THE WITNESS: Yes, ma'am. He's top of the glossary there as well.

MS. KLOET: You don't currently engage in forensic analysis.

THE WITNESS: Excuse me, my mind was elsewhere for a minute. What did you say?

MS. KLOET: That's okay. You don't currently engage in forensic analysis or wet work as it's sometimes called, do you?

THE WITNESS: No, I haven't been in a laboratory since '02.

MS. KLOET: Do you have a degree in statistics?

THE WITNESS: My DSc could be argued to be in statistics, and I take a position of professor of statistics on first of June this year.

MS. KLOET: Have you ever held a position as a professor of statistics before?

THE WITNESS: I have held a position as professor in statistics departments before.

MS. KLOET: Have you ever held -- or your degree is not in mathematics, is it?

THE WITNESS: I have a minor in mathematics and I'm reasonably strong in mathematics.

MS. KLOET: As I understand it, the government asked

to qualify him in forensic DNA, Probabilistic Genotyping Systems, and computer software. I think to the extent to which he needs to clarify his statements or testify as to Probabilistic Genotyping Systems, I'm fine with him testifying as to the underlying forensic DNA principles and software principles. But I'm not comfortable qualifying him as an expert in computer science software.

THE COURT: Mr. Presant.

MR. PRESANT: Redirect or argument, Your Honor?

THE COURT: Redirect, please.

MR. PRESANT: Thank you, Your Honor.

BY MR. PRESANT:

Q Dr. Buckleton, would you tell the Court a little bit more about your role working with Dr. Taylor and the process of developing STRmix?

A Certainly. I think some of the comment is fair. My formal training in computer science was in the early '80s.

THE COURT: Were there computers back then?

THE WITNESS: Yes, ma'am, there were. But they were much bigger than they are now.

I have actually in the late '80s published a chapter in a book on artificial intelligence, and during my research career I have been involved in quite a number of computing projects. The genesis of probabilistic genotyping dates back to a mathematical solution I developed in 1999, so I'm actually

the father of probabilistic genotyping. And the mathematics and the majority of modern probabilistic genotyping software either comes directly from me or dates back to work I've done.

My work with Dr. Taylor varied from side-by-side work on algorithms and testing to work vicariously across the Tasman Sea which separates Australia from New Zealand on the same subject. And initially we were the first two developers. It's much of the mathematics in the early version are mine, or mine and Dr. Taylor's combined.

Subsequently, we added a third developer, Dr. Jo-Anne Bright, who is very strong in quality systems, specifically in ISO 17025, which is the governing standard for most of forensic science.

BY MR. PRESANT:

Q And I believe your testimony was STRmix was first began to be developed in 2011, correct?

A 2011.

Q So that was seven years ago?

A Yes.

Q And over the past seven years, have you been continuously involved with Dr. Taylor and other individuals who have been doing the line-by-line coding?

A Yes.

Q And you've advised them on how to code for it?

A Yes. Algorithmic work, yes, coding.

Q Is that still part of your responsibilities today to work with the people who are developing the software?
A Yes.
Q Now, Ms. Kloet also asked you questions about your experience doing work in a wet lab. You don't do that anymore, right?
A No.
Q What was your experience doing work in a wet lab generally, getting your Ph.D. in chemistry, and doing your other academic work in chemistry, and specifically, in forensic DNA analysis?
A I think Ms. Kloet is fair to suggest that my experience is almost ancient history now in the wet lab. However, my experience in the actual foibles of DNA analysis is completely current, and for instance, I wrote with two others the textbook on DNA interpretation.
Q But I guess with respect to the ancient work, you actually did used to work in a wet lab, right?
A Yes, I did.
Q So you've done --
A Thousands of cases.
Q -- DNA extractions?
A No. No. That tends to be split in science and technical work. So the actual extractions and running the robots tends to be done by technicians, and at that point in my career I wasn't that. So we tend to get the process, to the process

post production of the e-gram, which is the electropherogram.

Q That's currently or even back when you were being trained academically in your early career?

A In my early career it was chemistry and it wasn't DNA. DNA, the advent of DNA was 1987, and my university days were in the '70s.

MR. PRESANT: Your Honor, at this time I would reoffer Dr. Buckleton in those three categories: Forensic DNA analysis, probabilistic genotyping, and also software development. Based on the experience he's testified to, I would also note that he's a fact witness here with respect to the development of the software because he was actually doing it.

THE COURT: I do accept Dr. Buckleton, and as I said earlier, I think that I can sort out areas where he's been quite candid. His expertise dates to the 1980s, but it does sound to me like it's very current, and he will be recognized as an expert in those areas.

MR. PRESANT: Thank you, Your Honor.

BY MR. PRESANT:

Q Dr. Buckleton, would you begin just by telling the Court -- well, you've testified already to the initial reasons for developing STRmix. But would you tell the Court a little bit about the history of the development of the software from the very first version to where we are today?

A The first versions of STRmix were never intended for release beyond Australasia. It was to be an in-house tool for the Australasia laboratories. And that applies to the whole 1.0 series which goes from version 1.0 through to version 1.08.

At sometime in about 2012 various U.S. agencies, particularly the U.S. Army and Californian Department of Justice heard about this, and made representations to my management that eventually led to us releasing a version for United States use, in fact for international use. And that was in January 2014.

I was involved in the development of that and in training in the United States and internationally at that time.

The first release was the 2.0 series, and we actually skipped 2.1 and 2.2, and the next series is the 2.3 series, which is the version we are speaking of today.

MR. PRESANT: I'm sorry to interrupt you. But specifically the one at issue in this case is 2.3.07, correct?

THE WITNESS: Yes.

BY MR. PRESANT:

Q I'm sorry, continue.

A I have often described 2.3 as the version I always wished we had made if I had envisaged what it might be back in, for instance, the year 2000. This is what I wished we could have made.

And I was very proud of it and it contains a number of

really solid features such as treatment of relatives, which I believe to be a very important matter.

The release of 2.3.07 actually predates many of the guidance documents, including SWGDAM. So we came out in October 2015, the SWGDAM guidelines came out in July 2015, so it post dates SWGDAM, but it predates the others.

The validation of 2.3.07 is done to the SWGDAM guidelines, which are the primary guidelines of which the United States takes cognizance. And I just remind you it's an FBI sponsored agency.

Subsequently we have either met or back, back engineered I guess so that we meet and exceed all the other guidelines, including now the PCAST guidelines. With the possible exception of some aspects of the new Forensic Regulator Guidelines.

BY MR. PRESANT:

Q So that was 2.3, the 2.3 series which is at issue in this case.

A Yes.

Q Where are we now, what version are you working on now?

A We are currently debugging 2.6.

Q What does that mean, debugging?

A Debugging is an iterative process of testing and coding where you run various trials, test the components for compliance, and if they fail compliance, diagnose the problem

and rectify.

Q Well, I want to talk to you about some of that process in a little bit. But first I want to ask you is 2.6 going to be the last version of STRmix?

A No.

Q What -- when do you stop? When will you stop developing STRmix?

A The future is often opaque to me. However, sequencing is one of the next matters of great interest, so at the moment we don't use DNA sequencing, we use a different method called STR typing. Sequencing is likely to be the next phase, and people are also interested in the Y. chromosome. So I imagine we will be developing versions for those.

Q Is it unusual for software to be operational while development of new versions is still going on?

A Yes.

Q It is unusual?

A No, not unusual. Sorry. It's completely usual. Completely misspoke, I'm sorry.

Q Well, I just want to make sure.

A Absolutely normal. I mean that's why there are all different Microsoft versions.

Q And what about apps on your phone?

A Absolutely. Updates come all the time.

Q And that's kind of what the process is that you're in

currently with STRmix, correct?

A Yes.

Q The fact that you're developing new versions mean that older versions didn't work?

A No. The engine of STRmix was in place very early, the core Metropolis-Hastings algorithm and the Monte Carlo Markov Chain was all in place very early, and has changed very little during that time.

Q Will you describe that for the Court, the way STRmix actually works that engine you're talking about; what is the Markov Chain Monte Carlo engine, how does probabilistic genotyping work, how does it get to that final likelihood ratio?

A Yes. Thank you. Markov was a Russian mathematician working on Russian poetry, and his particular interest was in whether the next letter would be a vowel or a consonant, and he published his work in 1906. I mention this to demonstrate this is not novel technology.

The Monte Carlo aspect is named for the Monte Carlo casino in Monaco and was a name used by the Los Alamos scientists during the development of the atomic bomb in the Second World War to describe the process. They had to conceal the true nature of what they were doing so they used the code Monte Carlo.

The Monte Carlo and Markov Chain processes were

synergized in the '50s to '70s to become Monte Carlo Markov Chain and you're at liberty to say that backwards. You can say Markov Chain Monte Carlo with equal accuracy.

It has become a dominant and mainstream methodology for solving this type of complex problem from the '70s onwards, and is through many fields such as physics, engineering, geoscience, medicine, and a great many others.

It works roughly like a game of hot and cold where as children you may have taken a step and your parents said whether you were hotter or colder, and using that feedback you would try and find your way to whatever your goal was. And a Markov Chain is quite similar to this, and, as I said, it's very well established mathematics. This process is used to tease apart the mixed data and essentially make a list of plausible single source genotypes that may have contributed to that mixture.

STRmix is in two completely separate parts; the first part does what I've said, which is termed deconvoluting a mixture, the second part does the relatively trivial action of assembling the likelihood ratio.

BY MR. PRESANT:

Q So what are the inputs into the engine that does this work in the first step?

A The inputs into STRmix are a set of data that comes straight out of analysis of the electropherogram, often with

GeneMaker or GeneMarker, which are two commercial softwares for processing e-grams.

Q Can I just interrupt you there? Ms. Miller, will you bring up Government's Exhibit 6, please? Now, Dr. Buckleton, you didn't review the data that was used in this particular case, correct?

A No.

Q I'll represent to you that this is one of the electropherograms that was generated by the Michigan State Police lab. My question for you is do you recognize it as an electropherogram?

A Yes.

Q Would it assist you in describing what electropherograms are to the Court and how they are input into STRmix?

A Yes, I think it would, thank you.

Q If there's a portion that you would like blown up, would you just, it's a touch screen, you can just highlight what you would like zoomed in on, and Ms. Miller will blow it up for the Court.

A Gosh, all this new tangled technology. Can I have the D3 locus blown up, please? I have put a small black dot there. The data for STRmix consists of the set of allele names and peak heights. So if we look at D3, you'll see a small text box there with the numbers 14 then 118.84, then 2213.

Q That's right there that you're looking, correct?

A Yes, it is. So the 14 is the allele name, so that is the 14 allele at D3.

Q What's an allele?

A Technically a variant of a gene but we use it to mean a variant at a in this case noncoding region.

Q So D3 right here is the locus or the gene, right?

A Well, it's not a gene. That's why I'm making that somewhat pedantic distinction. Gene is a coding DNA and these aren't.

Q So it doesn't produce protein, right?

A It doesn't produce a protein. In fact D3 isn't even transcribed. So D3, the 3 means on chromosome 3, so the human chromosomes are numbered in decreasing size order, so chromosome 1 is the largest, 2, then 3. So this is D3. Then the number 1358 is simply a sequence number in which those loci were found. They are noncoding regions of DNA, sometimes incorrectly termed junk DNA. And because they don't code, humans can be very variable at them. We obviously cannot vary much in the important elements such as arms. It's selectively disadvantageous to not have two arms. However, here you can have almost anything you want and nothing bad happens.

So we see at the D3 locus a number of peaks, and the genotype software has detected peaks at 14, 15, 16, 17, and 18. The next number, 118.4 is the molecular weight of that allele. And that is used later in the mathematics. And then the 2213 is the peak height of the 14 allele, and if you come down to

the 15 allele, you can see it's 3848, therefore a little bit higher, and if you go up and look at the picture you can see that indeed the 15 is a little bit higher than the 14.

These are the data that are input into STRmix. I'll highlight the subjective elements of data processing at this point due to the indications earlier that that was of interest.

BY MR. PRESANT:

Q If I can just interrupt you to ask you a couple of clarifying questions. These alleles are sometimes referred to as short tandem repeats, right?

A Yes.

Q That's the STR in STRmix?

A Yes.

Q Are these numbers 14, 15 or 16 significant with respect to the number of repeats?

A Yes, those are the number of repeats.

THE COURT: Let me interrupt then again, Mr. Presant. Where do these numbers originate that are put into the STRmix software?

THE WITNESS: These numbers originate from the standard DNA typing process.

THE COURT: Which is the sample?

THE WITNESS: The sample is taken and extracted, it is then amplified using an amplification process that has often been likened to a molecular photocopier. So you take one copy

of the DNA and copy it into two, then you copy those again into four, and we do this plausibly with Fusion 29 rounds of copying until we now have millions of copies of the DNA from which we started. That DNA has incorporated a fluorescent tag into it during this copying process, and it is synced down a capillary, and it's often likened to putting a platoon of soldiers through a swamp. The little light guys will get out the other side first and the big heavy guys will take a bit longer to get out. That's exactly what happens. The light fragments come off the column first, and then later the middle sized, and subsequently the big ones.

THE COURT: And this is what the technicians do that you described earlier, is that right?

THE WITNESS: Yes. My part of the process would usually start from the e-gram at which we are looking.

THE COURT: That somebody else produced.

THE WITNESS: That's correct.

THE COURT: And they would produce that -- you know, I really am trying to get, Dr. Buckleton, at the whole process from start to finish. The law enforcement collects something that they suspect may have DNA of a defendant or a proposed defendant on it. That object, whatever it is, in this case it was a gun, in other cases it may be fabric, whatever it is, then goes to a technician who does what with it to get to these numbers? This is the wet lab process?

THE WITNESS: Yes.

THE COURT: Okay.

THE WITNESS: However it's very highly robotic now. So the, you know, the human element is quite small now. But the processes are extraction, and that takes the DNA off whatever it's on, so it probably was on a swab from the gun, I imagine.

THE COURT: Okay.

THE WITNESS: So a swab is almost like those things you might use to clean your ear. You swab the gun, the DNA is now on the swab, extraction is a process of taking it off that into a liquid. That liquid is then quantified to see how much DNA is in there. And from that quant, an aliquat or a small portion is taken to this amplification process, which again is in a machine and the machine is called a thermocycler and it is the one that does these 29 cycles of amplification. The amplified product is then usually taken and placed on a robot, which will take an injection from that liquid and inject it into this capillary I've been speaking of which will then separate the fragments by size. As they come off, as each one comes off, it will make a peak in here, and you're looking at the blue line so you can see those pictures of blues so that's fluorescing blue. If we go down in the electropherogram, I'll hit the green line. So this is coming off at the same time but these are fluorescing green because they have incorporated a

green tag, and I believe Michigan State Police are using 5C which means five colors. So there are five colors or five lanes that will be in this e-gram.

THE COURT: And to go back just a little bit. Quantification, that tells you how much DNA is present in what measurement?

THE WITNESS: It tells you how much DNA is present in your extract as a concentration, how much DNA per unit volume. And is disappointingly inaccurate. It is an approximate quantification and at the best it only gets human DNA and doesn't account for other such things as degradation.

THE COURT: And what, what is the measurement of the DNA?

THE WITNESS: It's usually in nanograms.

THE COURT: Nanograms. Okay.

THE WITNESS: But it could be in picograms which is a thousandth of a nanogram.

THE COURT: And is there a minimum quantity of measurable DNA that can go through this amplification process and on to the further analysis that you've described?

THE WITNESS: No. There is no minimum. If you have no DNA or very little DNA, you simply get no peaks. Nothing bad happens. You just simply get no peaks at the other end.

THE COURT: Okay.

THE WITNESS: Most people amp negative quants -- which

is amplify, short for amplify, A-M-P -- negative quants, and quant is short for quantification.

THE COURT: Thank you.

MR. PRESANT: Ms. Miller, if we can zoom in here where we were before again. So, Dr. Buckleton, you've explained what these different numbers mean in terms of the STRs and the peak heights. When STRmix is reading this electropherogram what is it doing with that information?

THE WITNESS: These are the inputs for the start of the Markov Chain Monte Carlo process.

BY MR. PRESANT:

Q Now, before I interrupted earlier you were going to testify about the decisions that the specific forensic scientist using the software has to make in addition to inputting the electropherogram information. Would you please tell the Court what those decision points are?

A The decision points currently in the process for this version are certain peak analysis decisions which I'll outline in a minute, and the number of contributors. The peak analysis decisions are to remove spikes, pull up and forward stutter, and we are on Fusion so they also have to remove one exotic stutter.

These are standard decisions that are being made in forensic DNA typing since its inception in the mid-'90s. And there is nothing novel in that at all.

The other decision point is the number of contributors that is an input into STRmix. Generally speaking, changing the number of contributors does not materially affect the result, and if it does, it affects it in the conservative direction. So I have been using the phrase you're either correct or conservative with respect to the number of contributors.

THE COURT: Could we go back just a second? There are, in this case there are going to be, if I understand your testimony correctly, and the motion of the defense, there are two discreet areas of inquiry in terms of the usefulness of this information. The first area of inquiry is in getting these peaks, this electropherogram. Okay. What are the variables that can affect the validity of the electropherogram? Is deterioration of the sample, for instance, is that an issue? Is there the potential for human error? Is there potential for mechanical error? Focus if you would for me on that part of the process.

THE WITNESS: Thank you, Ma'am. There is definitely a, there is always an element of risk from human error. We are all human. Sample swapping is the one that's been of greatest concern over the years. I can't speak to that particular aspect for Michigan State Police and hope someone else can speak to that, but in my experience enormous care is taken in laboratories to minimize any risk from human error.

You mention degradation of the profile. DNA is an

actually very fragile molecule, and degrades and bacteria, UV light, and thermal shock are all things that can damage the DNA.

This is no longer a problem. The evolution of probabilistic genotyping tells us that if a sample is degraded the software will reliably report that that is uninformative, there is no information left at that particular locus. So that is no longer a risk factor, although certainly once upon a time it was.

I don't know what else.

THE COURT: With regard to the equipment, you mentioned robotic, the process becomes robotic at a certain point. What equipment is used and can failures, what are the typical failures in the equipment, if any?

THE WITNESS: I understand Michigan State Police are on a 3500 machine. This is the most modern of the capillary machines. It's a fine tool. No endorsement implied. I think other witnesses can probably speak to the risk, but this is a very fine piece of equipment.

THE COURT: And it would require a certain amount of maintenance and so forth, I assume.

THE WITNESS: Yes. Yes, they do. And the robots throw little fits every now and then and throw the samples everywhere.

THE COURT: Okay. And so that, that is the first

discreet part of this process. The second discreet part of the process is when STRmix comes in to the picture.

THE WITNESS: Yes.

THE COURT: And tell me what happens then. You get this electropherogram. What happens then?

THE WITNESS: The software that is drawing these pictures can also output this as a file, and that file is output and taken as the input for STRmix.

THE COURT: And what does STRmix do with that input?

THE WITNESS: STRmix takes that input and some settings that have been set during the laboratory validation, and one human input, which is the number of contributors, and then begins the deconvolution process to try and see what various combinations in the situation we are looking at of three people, which combinations of three people could explain this profile. There will be many different combinations of three people that could explain this profile, and it will attempt to find all of those.

THE COURT: Okay. Mr. Presant.

MR. PRESANT: Thank you, Your Honor. So there are the different inputs, and you mentioned just before we were discussing those inputs that there are several layers of conservatism built into the software.

THE WITNESS: Yes.

BY MR. PRESANT:

Q Would you tell the Court what those layers are and what you mean by conservative?

A In reverse order, conservative is a term used in forensic judicial work to mean conceding in the interest of the defendant. So conservatism is generally viewed as a good thing inasmuch as it concedes rational doubt in the interest of the defendant.

The three layers in this version of STRmix are an inherent conservatism of the population genetic model. The model itself is conservative on about the ratio of 99 to 1. So about 99 percent of the time it understates the value.

Q So let's just go into a little bit more detail there. When you're talking about the population genetic model, there are empirically studied known frequencies of these specific alleles in different populations, correct?

A Yes.

Q And the software incorporates data of those known frequencies in order to do its mathematical calculations?

A Yes.

Q All right. And what did you mean by the 99 to 1 with respect to that data?

A We have to predict the occurrence of a genotype at more than 20 loci. It's almost certain this genotype is either very rare or doesn't exist at all, and there is no way we can measure it directly. It is -- the estimate is generated by a

population genetic model. That model assumes a certain level of population subdivision, that is, that the human population is not completely homogenous and mixed; that we are genetically subdivided; and it makes a concession in the interest of the defendant; and in empirical measurements its concession is such that about 99 percent of the time it understates the true value. That is the first level of three.

Q Just to be clear, when you're talking about subdivisions, those are racial subdivisions based on empirical studies and frequencies?

A They are usually sub racial. So I believe Michigan State Police are using the NIST data so we could talk about the African American population, the Caucasian population and, for instance, the Hispanic population. If we talked about the Caucasian population, even that is not homogenous inasmuch as there are people of stronger Spanish origin or Italian origin or anything else. It is that level of subdivision about which we are speaking.

Q So that's the first manner in which STRmix is conservative, right?

A Yes.

Q You said there were other ways it's conservative as well.

A That's right. There are two further ways in use in this version. The --

Q I'm sorry to interrupt you one more time. But when you say

this version, for all your testimony thus far and going forward, we are assuming it's 2.3.07.

A Yes.

Q Unless otherwise stated.

A Should I just stop mentioning that?

Q No. I think you can just say this version. I just want to make sure the record is clear you're talking about 2.3.07, the one used in this case.

A Yes, I am.

Q I'm sorry, continue.

A I notice that MSP are using a coancestry coefficient of one percent for the African American group. That means the chance that any two genes are identical by it because of relatedness is taken as one percent. That is a conservative value. The global average is about .8 of a percent. And one percent is higher than .8.

Then the third level is that they have applied a 99 percent probability interval. A probability interval is not strictly the same as a confidence interval but no harm will happen if we think of that in that way. So we could think of it as the lower bound out of a 99 percent confidence level.

To summarize, there are two layers of conservatism that can be quantified as being at 99 to 1 in favor of the defendant, and one further layer that is qualitative and that is they have used a conservative coancestry coefficient.

Q So what you mean by that conservatism as it relates to this case, if I told you that the Michigan State Police came up with a likelihood ratio of 49 million to 1 in this particular case, is it accurate to say that you think the true likelihood ratio is actually probably higher than that, but the software has made those conservative decisions to make the number at the small end of the range?

A That's correct.

Q I want to shift to a different topic, Dr. Buckleton, which is validation. In the development of the software, what sorts of --

THE COURT: Before we go there, Mr. Presant, I want to, I would like to clarify just a little bit.

MR. PRESANT: Of course, Your Honor.

THE COURT: What exactly does that ratio represent, 1 out of 49 million?

THE WITNESS: The likelihood ratio is a very hard number to get your head around, Ma'am. And that's the communication issue that has been a problem.

However, again nothing bad will happen if you think the chance of someone else having this profile is 1 in 49 million. That will give the same value to the evidence.

If you want it strictly, this profile is 49 million times more likely if the named person is a donor than if he is not.

THE COURT: Okay.

MR. PRESANT: So the numerator, the top number, 49 million, is sometimes called the prosecution hypothesis, right?

THE WITNESS: Yes.

BY MR. PRESANT:

Q That you would expect to see this particular profile that was found on the piece of evidence in this case if the defendant was a contributor, right?

A Yes.

Q And the denominator, the bottom number, the 1 is sometimes called the defense hypothesis, right?

A Yes.

Q And that's you would expect to see this profile if the defendant were not a contributor, right?

A That's correct.

Q Three randomly selected people.

A That's correct.

Q And so the likelihood ratio is just the ratio of those two hypotheses, how much more likely is the prosecution hypothesis as opposed to the defense hypothesis.

A Pedantically correcting the language. How much more likely the evidence is if the prosecution hypothesis is correct.

Q I very much appreciate the clarification. It's very easy to get --

A It is very easy.

Q -- confused linguistically.

A This is an area that is sometimes referred to as the prosecutor's fallacy and the most minor misstatements of the probability phrasing can lead to what's called the prosecutor's fallacy.

Q And that's why forensic scientists are trained very carefully in terms of how they testify?

A Yes.

Q All right. Well, I want to come back to likelihood ratios a little bit later on, but first would you describe for the Court the validation processes that STRmix went through during development and as it's been implemented in various laboratories.

A Validation happens in two stages: Developmental validation and internal validation.

Developmental validation we run a number of tests. The primary one of which I wish to speak at the moment is we repeat the calculations by a separate method, usually by hand. So we repeat a great many of the calculations by hand and check that we get the same answer as the software.

Other requirements are to run a number of samples of true donors and false donors and make sure the performance of the software is correct.

There are quite a number of tests performed, and I'll be most happy to go to them, but I'll try and summarize them.

We essentially follow the SWGDAM guidelines to the letter, and we have published our developmental validation process in the peer reviewed literature.

BY MR. PRESANT:

Q Let's look at a few of those publications. Can we bring up Exhibit 4, please? What is Exhibit 4, Dr. Buckleton?

A Exhibit 4 is the largest ever validation work. It's a compilation of the internal validations of 31 laboratories across the United States and internationally. It was in response to a PCAST request for a much larger number of samples to be processed and published.

THE COURT: This is a validation of the STRmix software?

THE WITNESS: Yes.

MR. PRESANT: And some of the people you've testified about here today are authors on the paper; there is you at the end and Dr. Taylor, Dr. Bright.

THE WITNESS: So those are the three developers, and then you can see representatives of a large number of labs across the United States and internationally.

BY MR. PRESANT:

Q And this was an example of a peer reviewed published compilation of internal validations, right?

A That's correct.

Q I'm sorry if I missed it but would you just distinguish

between internal validation and developmental validation?
A Developmental validation is done before we release the software and it's often called a debugging process and we iteratively try to repeat calculations and get the same answer until eventually we get a locked down version of the software in which we have repeated the number of calculations we see for yourselves. And then we begin the SWGDAM process, what we begin and finish the SWGDAM process.
Q Is the SWGDAM process the internal validation?
A No. It is that as well. It's very clearly both. And once we release a version or someone such as Michigan State Police uptake a version, they will then perform the internal validation. They won't tend to do the by hand calculations; they take a little bit of training to be able to do. But they will certainly do a great many samples and other tests of the software to see that it's fit for purpose.
THE COURT: And the STRmix software remains proprietary, correct?
THE WITNESS: I don't know what proprietary means.
THE COURT: Not available freely.
THE WITNESS: No. It's sold. However, we do respect the defendant's Fifth Amendment rights and will make the software and anything else available to the defense.
BY MR. PRESANT:
Q All right. And how about Exhibit 5, please?

A Exhibit 5 is a publication of the internal validation of STRmix done at the Federal Bureau of Investigation laboratory in Quantico, Virginia.

Q What about Exhibit 23?

A This is not specifically relating to STRmix. This is the DNA Commission of the International Society for Forensic Geneticists recommendations for validation. I'm one of the authors, and STRmix meets the requirements of this document.

Q Have you published more peer review journal articles on STRmix besides the three we just reviewed?

A I wouldn't count this one as being on STRmix. There are 33 publications in the peer reviewed literature that I count as pertaining to STRmix.

Q Can we bring up Exhibit 12, please? And go to the second page 2. Is this the list that you just testified to?

A Yes.

MR. PRESANT: Your Honor, the government moves for the admission of 4, 5, 12 and 23.

THE COURT: Ms. Kloet.

MS. KLOET: I have no objection, Your Honor.

THE COURT: They are admitted.

BY MR. PRESANT:

Q So I want to ask you some specific questions, Dr. Buckleton, about the operation of the software. You've already described the hot and cold process by which it reaches

its final number, right?

A Yes.

Q If you go through that hot and cold process again, the Monte Carlo engine process, with everything else the same, all the same inputs, the same electropherogram, the same number of contributors identified by the analysts, the same population genetics data, will you get the same answer the second time you got the first time?

A No.

Q Why not?

A The Monte Carlo effect. So Monte Carlo is a gambling venue, and the process is like that. It is using random numbers and the randomness of the process creates a slightly different answer each time. That answer is usually quite well clustered and almost always within one order of magnitude.

Q What does that mean, one order of magnitude?

A A factor of ten. So ten times higher or ten times lower.

Q So if you ran this process again and again, I'm representing to you that Michigan State Police's result in this case was 49 million to 1, the worst you would expect the software to do in terms of the lowest number it could come up with the second time would be approximately 5 million to one?

A Yes.

Q As a result of that, do you think it's fair to say that while you can have a low degree of confidence in the precise

answer of the likelihood ratio, you can have a high degree of confidence in the approximate likelihood ratio?

A Yes.

Q Does STRmix have a way of measuring the precision with which the likelihood ratio was attained in a particular case?

A Not in this version.

Q Not in 2.3.07?

A But in subsequent versions it does.

Q What sort of diagnostics or analysis are available for 2.3.07?

A STRmix outputs a range of diagnostics; the one you might be referring to the Gelman-Rubin diagnostics. But there are several others.

Q And what's the Gelman-Rubin?

A The Gelman-Rubin informs us whether the various chains of the Markov Chain have converged. So we tend to run either four or eight chains. And we check whether they have obtained the same sample space.

Q Is convergence of the various chains a good sign or a bad sign?

A It's a good sign.

Q Let's talk about the number of contributors. Is there a limitation on the number of contributors that the software can handle?

A There is no mathematical limitation. There are hardware

limitations. And we become unable to run due to memory constraints at about five. Most laboratories in the United States have stopped themselves at either four or five.

Q In a mixture as displayed on an electropherogram, can you determine at least approximately which contributor contributed more than the others?

A Yes.

Q And the one who contributed the most is sometimes called the major contributor?

A Yes.

Q And the least is the minor contributor?

A In America, yes.

Q In America. It might have other names elsewhere. Does STRmix take into account the difference in the peak heights in doing its mathematical analysis?

A Yes.

Q Is there a limitation or a lower bound in the software itself on what, how low the minor contributor can be, what percentage of the total DNA mixture the minor contributor can be in order for the software to still do what you would expect it to do?

A No. There is no lower limit. As the minor contributor becomes smaller, the answer will tend towards inconclusive, and that is correct, you end up having no information about that person because he or she is so small, and STRmix will reliably

report the answer inconclusive.

Q So the likelihood ratio itself then reflects in a way the uncertainty created by the smaller proportion of DNA; as the DNA gets lesser and lesser in the mixture of the minor contributor, the likelihood ratio is also going to tend lower as well, is that what you just testified to as well?

A That's correct. That's very nicely put.

Q Thank you. So you've testified a little bit about this already but I want to talk about the process for implementing STRmix in a new laboratory. How many laboratories in the world have purchased a license to STRmix?

A Purchased. I don't know in the world, but in North America 65 percent of laboratories have purchased STRmix.

Q 65 percent of all forensic laboratories.

A Yes.

Q How many of those are actually using it?

A There are 31 laboratories, all laboratories systems live with STRmix in the United States.

Q Can we bring up 13, please, Ms. Miller? Do you recognize 13, Dr. Buckleton?

A Yes.

Q What is it?

A It is a download from my web page showing the laboratories using STRmix and the date they began case work.

Q So the columns are here are the areas of the world USA,

Australasia, and the rest of the world, is that right?
A Yes.
Q Then the rows are the years and the specific laboratories with their dates.
A Yes.
Q Michigan State Police are right here on February 22nd of 2016, is that right?
A Yes.
Q Can we go to page 2, please? And it looks like the most recent lab in the U.S. to roll it out was in Houston, Texas just this year.
A Yes.
Q So this is work that is still going on, you're still training additional laboratories and staff to use STRmix?
A Yes.
Q You said that there's training process for specific forensic scientists?
A Yes.
Q You've personally delivered some of those trainings?
A Yes.
Q What is involved in that week-long training of the forensic scientists? What topics do you cover about how to use the software?
A So it's not a week, it's four days. I'm sure it feels like a week. We cover theory and practice of use of STRmix. So we

certainly go through the theoretical aspects of the Markov Chain, and the calculations, and students are expected to do one calculation by hand, and certainly operation of the software, use of the diagnostics, and we also speak quite a bit about how to phrase the LR and report it to court.

Q Do you train them on the inputs that they will have to make based on their interpretation of the electropherogram?

A Yes.

Q Are you involved in the internal validation in new laboratories as they come online?

A No, not typically. We actually have 12 people working for STRmix now, and many U.S. labs outsource their work to the people back in New Zealand, but I'm usually only involved if some issue arises.

Q Have you been involved in the trainings of the Michigan State Police?

A Yes.

Q How many have you done with MSP?

A Two.

MR. PRESANT: Your Honor, the government moves 13. I should have done it earlier.

THE COURT: Ms. Kloet.

MS. KLOET: Your Honor, I would like to know who created that document.

MR. PRESANT: I believe the witness testified that it

was a download from his website.

THE COURT: Let her voir dire on that exhibit, please.

MS. KLOET: I may have missed that, Dr. Buckleton. But did you create the exhibit that's displayed on your screen?

THE WITNESS: Yes.

MS. KLOET: And when did you create that?

THE WITNESS: It's a live document. I update it regularly. I don't know when I first created it.

MS. KLOET: When did you last update it?

THE WITNESS: April this year.

MS. KLOET: How do you collect information for inclusion on this document?

THE WITNESS: Some of the laboratories write to me telling me they have gone live, others I find out retrospectively they have gone live and have to ask them.

MS. KLOET: Do you find out firsthand or do you hear through the community that they have gone live?

THE WITNESS: I always check firsthand.

MS. KLOET: Thank you.

THE COURT: Any objection to the exhibit?

MS. KLOET: No, Your Honor.

THE COURT: It's admitted.

BY MR. PRESANT:

Q On the topic of market share, Dr. Buckleton, so you said 65 percent of North American labs approximately have purchased

licenses.

THE COURT: How much do the licenses cost, Dr. Buckleton, do you know?

THE WITNESS: No. I don't. I try to ignore the commercial aspects. But I think it's about 27,000 US dollars for your first license.

THE COURT: And is that for unlimited use of the software?

THE WITNESS: Yes.

MR. PRESANT: You have 30 or so labs actively using the software in North America. What about competitors? Do you know how many competitors products are used by forensic laboratories in North America?

THE WITNESS: I estimate or from information I have had is approximately seven are using TrueAllele, and two are using Lab Retriever. That's spelled just like the dog.

MR. PRESANT: Are you familiar with a recent paper written by Steven Lund and Hari Iyer?

THE WITNESS: Yes.

BY MR. PRESANT:

Q Can we bring up 15, please? Do you recognize Exhibit 15 as the paper we are talking about, "Likelihood Ratio As Weight of Forensic Evidence: A Closer Look"?

A Yes.

Q Have you read the paper?

A Yes.

Q What are the general arguments made by Drs. Lund and Iyer in the paper?

MS. KLOET: Your Honor, we have Dr. Lund here. I would object that he's the best person to explain his arguments as expressed in that paper.

MR. PRESANT: Your Honor, the witness has recently published, actually just yesterday became available, a response in a peer review journal of this paper. I suppose there are two ways to do it. We could recall him on rebuttal after Dr. Lund has testified, or we could let him testify now as to what his understanding of what the paper is and then talk about his published response to it.

THE COURT: Let's get it over with now. But, you know I think you need to keep this within some fairly narrow discussion. We are going to hear apparently from Dr. Lund. He does -- I haven't read the whole article but I've also read the, that little monograph that talks about it. So let's not go too far afield with the examination of the paper.

MR. PRESANT: Thank you, Your Honor. Would you describe, Dr. Buckleton, concisely if you can the main arguments made in this particular paper?

THE WITNESS: In my view, Drs. Lund and Iyer make two points: First of all, we should not impose our view of the evidence on a court, and that is certainly a point I completely

accept. And second, we should take care to understand the uncertainties in our process, and that's again a point I accept.

BY MR. PRESANT:

Q And you've written a couple things in response to it, correct?

A Yes.

Q Can we look at Exhibit 18, please? And what's Exhibit 18?

A This is a statement I made for the specific case.

Q And it treats both the Lund paper as well as arguments raised by Mr. Adams that we will get to in a second?

A Yes.

Q And can we bring up Exhibit 28, please? What's 28?

A This is the accepted publication of the response to Drs. Lund and Iyer.

Q And you published this paper along with others just recently, right?

A Yes.

MR. PRESANT: Your Honor, the government moves 18 and 28.

THE WITNESS: Ms. Kloet.

MS. KLOET: No objection, Your Honor.

THE COURT: They are admitted.

MR. PRESANT: Would you describe or summarize, rather, Dr. Buckleton, the responses that you've published in the two

exhibits we just reviewed? Well, I'm sorry, that you published in Exhibit 28 and you've also written in Exhibit 18 that you prepared for this case. Specifically, with respect to Dr. Lund. We will talk about Mr. Adams in a second.

THE WITNESS: Dr. Lund and Iyer's argument has in my view correctly been likened to a straw man argument. A straw man argument is where you argue against something that people aren't even doing, and in mine and my coauthors' opinion they are arguing against a process that quite simply no one advocates at all. And specifically, no one that I know of would impose their view of the evidence on a court.

I fully understand my duty is to explain both the estimations I make and the uncertainties in that so that the factfinders can draw a correct opinion as to the weight of evidence.

BY MR. PRESANT:

Q So that's the first argument with respect to the role of the courts in understanding scientific evidence, right?

A Yes.

Q And what about the second critique they have about incorporating uncertainty into models as they are presented in court?

A Well, they make the not too startling suggestion that if you change the inputs you will change the outputs. This is barely news. I wish to add that both these doctors are my

personal friends, so I don't wish to be too harsh.

Q You worked at NIST with them, correct?

A I did. They were several doors up the corridor from me and we are friends.

But, yes, I'm perfectly happy to explain the uncertainties in the process. In fact, I argue not only would I explain them here but I've published them.

Q And you've also testified here and explained elsewhere in the literature that STRmix in coming up with its final number attempts as well as it can to reflect uncertainty and to make decisions, you've made decisions in developing it in the conservative direction, so any uncertainty is resolved in favor of the defendant.

A That's correct.

Q Now, in reading the Lund/Iyer paper, can I ask you was it specific to forensic DNA analysis or to STRmix in particular?

A It's, it's neither of those. It's a very general paper speaking completely about modeling for likelihood ratios.

Q General to forensic science, right?

A Yes. It could apply to firearms or footwear or any other thing.

Q And they advocate for testing the same evidence with a number of different models and then explaining all of the decisions that were made to the jury, correct?

A Possibly even infinite models.

Q They say it's possible to, right, create an infinite number of models, right?
A Yes.
Q But in a finite criminal trial, a criminal trial that has a limited amount of time, it's not possible to introduce infinite number of models to the jury, right?
A That's true. And one would have to really doubt or I would doubt my own ability to convey all that information meaningfully.
Q What do you mean by that?
A Well, if I sat here and gave an infinite number of possible answers and possible models, we would be here for a very long time, and I suspect the clarity of my presentation would fail.
Q And the jurors certainly wouldn't be experts in probabilistic genotyping or in the biological and chemical processes that go into DNA analysis that were used by the developers of STRmix or the lab in making specific decisions that led to the selection of the particular model, correct?
A That's correct.
Q Let's turn to another topic. Can we bring up Exhibit 17. Do you recognize Exhibit 17?
A Yes.
Q What is it?
A It's the report of Mr. Nathaniel Adams.
Q The report he wrote for this case in particular, right?

A Yes.
Q You understand it's subject to a protective order?
A Yes.
Q Did you review this report?
A Yes.
Q Did you write a response to it?
A Yes.
Q And that response is also contained in Exhibit 18?
A Yes.
Q What's your familiarity with Mr. Adams?
A Mr. Adams and I are cordial acquaintances. I would like to say friends. We have exchanged e-mails a number of times and I spent four days with him in training in Ventura, California.
Q Has he been involved in reviewing STRmix prior to this case?
A Yes.
Q Has he reviewed the source code for STRmix prior to this case?
A Yes.
Q Did you read the report he wrote in that case?
A Yes.
Q So would you summarize for the Court then your view of Mr. Adams's report here in what's been marked Exhibit 17? And if you would like to refer to your written statement in Exhibit 18, or any of the other supporting materials just ask

me, we will bring it up.

A Thank you. Mr. Adams makes a number of sensible and valid points many of which we have listened to. They tend to go to the cosmetic or theoretical appearance of code style and documentation, and at no point has he actually questioned the or found fault in the functionality.

His argument is that testing or I believe his argument to be that testing is inadequate for validation, and that we really should pay a lot more attention to code quality. Again, I would say he has some veracity to that. But his conclusions just massively go beyond the data he's used for those.

Without diminishing the importance of code quality or documentation, STRmix is the single most tested and trialed software in use in this field. It's been tested by multiple groups in multiple situations. And I would suggest that there is a very good basis for accepting its reliability.

BY MR. PRESANT:

Q All right. Mr. Adams makes much of these IEEE standards, correct, in this report?

A Yes.

Q And what is your view of the applicability of the IEEE standards to probabilistic genotyping software like STRmix?

A So IEEE is the most commonly used standard setting body in computer science. It's not without its faults. For instance, adherence to IEEE has actually caused a bug in Excel. And

Microsoft itself complies with some but not all of IEEE. However, moving to Prob Gene, none of the guidance bodies have suggested adherence to IEEE. The ISFG, International Society of Forensic Geneticists, do make some suggestions for including core computer science principles, and in many cases I was the proponent of those. The Forensic Science Regulator also makes suggestions that it should be developed within a quality system, and I would suggest that it is developed within that quality system.

So in summary, I would say that Mr. Adams makes some valid points. But that whilst we are listening to these, they do not diminish the value of the extensive testing that's been done on STRmix.

BY MR. PRESANT:

Q Do his points he makes, which are valid as you've testified to, give you concern about the functionality of the operation or rather the functionality of version 2.3.07?

A No, they don't. His, some of his points are completely valid and I recognize them immediately. Some he's just got the wrong end of the stick or not been given the correct data.

Q So let's talk about some of those. One thing that Mr. Adams says in his report is that he didn't review certain documents, he didn't review something called Github. What's your reaction to those points he's making?

A Some of those things we could have made available to him.

They do exist. And some we did make available and he didn't recognize them as the document he was seeking, and some he is asking for do not exist.

Q Can we go to page 6 of Exhibit 18?

A Yes.

Q We will bring it up on the screen. But it's in front of you as well. This is the appendix to the statement you wrote for this case, is that right?

A Yes.

Q And what are you trying to do in this appendix?

A This is a point by point response to Mr. Adams's commentary.

Q So let's go through them, if you can summarize each of them one by one so that the Court has the opportunity to ask you any clarifying questions it might have about them. You can start with 4.1.3.

A Okay. Well, Mr. Adams I believe is just simply incorrect in his assertion that this will affect the deconvolution. If we come to test script, I just need to check. So the test script is clearly not in the format he would have liked but we believe contains the information that he sought. And I think there was some confusion over what expected result is. So the expected result is a defined thing. For instance, reproducible means we've been able to get the same number by a different, separate and independent mechanism.

Issue trackers relate to the use of Github, I think, and we do indeed use Github. And we don't know as yet whether we could technically give him access to just the single branch for 2.3.07. It seems likely he actually wants the entire back history of STRmix which we'll have to see if we can technically achieve.

MR. PRESANT: Ms. Miller, we are on the bottom of page 7.

THE WITNESS: The change log does identify the changes, and we think the deficiency here is that Mr. Adams would have actually liked the change logs all the way back to inception. Which we may be able to give him some but not all of those because this process began at some point.

Dates of changes, Mr. Adams has assumed that's the end date. It's actually the beginning date of this document, and I have put down the timeline for that segment.

The run-time checks, Mr. Adams is mathematically incorrect, and it's disappointing to see him speaking on the subject.

BY MR. PRESANT:

Q Well, will you explain why he's mathematically incorrect, or how do you know that?

A A Markov change should be a memory less walk and a Gaussian walk that proposes an illegal value should still be counted.

Focus on reproducibility. Reproducibility means we

have been able to reproduce the result.

Q Before we get too far ahead from the random walk, can we go briefly to Exhibit 27? What's 27, Dr. Buckleton?

A It's a chapter from a basic text on Monte Carlo Markov Chain.

Q And can we go to I believe it's page 19 of the PDF? It's page 185 of the textbook chapter. Zoom in on this gray box. What's the significance of the part we have just blown up there in the gray box?

A That is a part that explains that you should count Gaussian walk proposals that are outside the constrained limits.

MR. PRESANT: G-A-U-S-S-I-A-N, did I spell that correctly?

A You did.

Q And is this what Mr. Adams misapprehended in the criticism in 4.2.2.2?

A Yes.

Q Let's go back to -- well before we go back.

MR. PRESANT: The government moves to admit 27, Your Honor.

THE COURT: Ms. Kloet.

MS. KLOET: If I could ask a couple questions as to its genesis.

THE COURT: Sure.

MS. KLOET: Dr. Buckleton, was this document provided

to you by the government or did you provide it to the government?

THE WITNESS: I provided it to the government.

MS. KLOET: Okay. Is this an excerpt from a textbook that you reviewed in a course that you were enrolled in?

THE WITNESS: No.

MS. KLOET: Where did you find this textbook?

THE WITNESS: This was sent to me by Dr. Duncan Taylor when I asked him for formal proof of this particular rule.

MS. KLOET: And Duncan Taylor is the co-creator of the STRmix program, correct?

THE WITNESS: Yes.

MS. KLOET: He's a geneticist?

THE WITNESS: I'm not sure. His degree, his doctorate degree was studying STRs in Australian snakes.

MS. KLOET: Would it be fair to say he's a forensic scientist?

THE WITNESS: Yes.

MS. KLOET: Thank you. I don't have any objection.

THE COURT: It's admitted.

MR. PRESANT: Can we go back to page 8 of Exhibit 18, please? I interrupted you, Dr. Buckleton, when you were on 4.2.6, Focus on reproducibility.

THE WITNESS: So we use the word reproducible to mean we are able to reproduce the result by a separate method which

is usually by hand in Excel.

Recreation testing means testing against previous versions and we do indeed do that, and that's outlined in the R&T report.

Testing and verification. So we actually set the changes for each version in a thing called a change request, and I think Mr. Adams would prefer this to be a vastly more technical document. And that's certainly something we could do.

STRmix defects. Mr. Adams is unaware of the process for notifying stakeholders. He didn't ask for it, and we would have been happy to tell him that all stakeholders are notified of all defects as soon as we diagnose them and measure the extent of the troubles.

Published. So indeed both the STRmix web page and my own web page outline defects, and we openly disclose them to the community. I'm not sure what he wants further. He calls them plain language descriptions, which they are. I'm sure we could provide him if requested with some more technical data.

I'm not sure what he means by Unpublished but known to ESR. I don't think we have any. We don't do that. That's not part of our philosophy.

The bug known to MSP. Mr. Adams has misunderstood something.

BY MR. PRESANT:

Q What has he misunderstood?
A Well, it's not a bug. The creation of a locus with a zero when all other loci are performing well is a diagnostic usually indicating an input error.
Q What about 4.3.4.3?
A So this is entitled Garbage collection. Mr. Adams suggests that it is improperly coded. It's actually a completely unnecessary function and has been removed. But it had no effect on 2.3.07.

Code style. He wishes to know the author of the code. Well, the author of nearly all the code or possibly even all the code is Dr. Duncan Taylor. But the three words, Admin, Owner and Dude, refer to computers on which Dr. Taylor was working.

Object oriented principles is the next one. And Dr. Adams argues that our use of them is limited and I think he says something slightly rude. Where is it? Anyhow, he was critical of our use of it and possibly with some justification. We have certainly upped the use of this, but this goes to style rather than functionality.

Q Is that a way to view most or if not all of his criticisms of STRmix is that they are stylistic preferences that don't actually affect the way the software is operating?
A That is what I am going to say. I believe his comments are stylistic and do not affect the operation. Mr. Adams's actual

argument would be they are stylistic and although he can't detect that they could affect the operation.

Q Are you aware of any probabilistic genotyping software that meets Mr. Adams's stylistic preferences?

A I don't believe any of them would get close. But we would get the closest.

Q Now, one of his criticisms was the identification and disclosure of miscodes, right?

A Yes. Is it a criticism. Yes, carry on, please.

Q Can we bring up Exhibit 14, please? What is 14, Dr. Buckleton?

A 14 is a download from my web page of the miscodes detected to date in STRmix.

MR. PRESANT: Government moves to admit 14, Your Honor.

THE COURT: Ms. Kloet.

MS. KLOET: I just have a couple brief questions. Dr. Buckleton, when was this document created?

THE WITNESS: This document is a live document and I've been updating it regularly. Updated it most recently I think on the weekend just gone.

MS. KLOET: This past weekend.

THE WITNESS: Yes, I believe so.

MS. KLOET: When was it first created?

THE WITNESS: I don't know.

MS. KLOET: When was it first uploaded to the website?

THE WITNESS: I don't have that information, I'm sorry.

MS. KLOET: Do you have an approximate idea?

THE WITNESS: Well, this is my web page. The individual things are released individually. This was a compilation that I think I made about six months ago.

MS. KLOET: Are you the only individual who has added information to this document?

THE WITNESS: The information is compiled by me from information given to me by others.

MS. KLOET: Okay.

THE WITNESS: So you could argue that others have had input into it.

MS. KLOET: Just so I'm clear, this document has been on the website available to the public for how long?

THE WITNESS: In various forms for six months.

MS. KLOET: Okay. Thank you.

THE COURT: Any objection to the Exhibit?

MS. KLOET: No, Your Honor.

THE COURT: It's admitted.

BY MR. PRESANT:

Q Dr. Buckleton, on Exhibit 14, the two, I'm sorry, there are more than two pages. On the multiple pages of Exhibit 14, I think you identify seven miscodes, is that right?

A Yes.
Q How many of those miscodes impacted the version at issue in this case, 2.3.07?
A One.
Q Which one is that?
A Number 3.
Q Can we just zoom in on number 3? Would you explain to the Court how this miscode was identified and what significance if any it had on the calculation of the likelihood ratio in version 2.3.07?
A This miscode was detected by testing, and it has an error in the conservative direction, and that error is very minor and occurs in very rare circumstances.
Q So if that error had an impact in this case, where 2.3.07 was used, if anything, it would have made the likelihood ratio smaller than it should have been?
A Yes.
Q Now you're on version 2.6 currently, right?
A Debugging, yes.
Q Have you identified bugs or miscodes in 2.6?
A Yes.
Q And is 2.6 being worked on by Dr. Taylor or have other people come in and taken over the coding?
A The coding is now done by professional coders, and I think would meet many of Mr. Adams's requirements now.

Q But you still found bugs, right?
A 416 to date.
Q Can we bring up again Exhibit 18, page 4 this time? This table at the top. So you found 416 bugs to date in 2.6 which is being worked on by professional coders you said, right?
A Yes.
Q What does that tell you about the significance of bugs in code to the operation of the code?
A Professional coders are not, there will be bugs and they have to be found by testing.
Q Like the testing that's done in validation?
A Yes.
Q Now, Mr. Adams also makes much of the need for code to be open source. Are you familiar with that criticism?
A Yes.
Q What's your view of it?
A When I first heard of open source, I thought it was a fantastic idea, and my liberal civil service background makes me think that's a really good concept, which is you make the code available publicly. Over time I have decided it's a terrible idea. And I'm going to give multiple answers but this also appears on my web page.

First of all, it actually means anyone can download and use the code without training. And this truly horrifies me. You cannot buy STRmix without training, and in fact many

people have asked us that we take more care for the labs who are using STRmix rather than less. So just throwing it out into the public domain and saying have a go is really akin to tossing the keys to your new Ferrari to your teenage son. And I just think it's just a disaster waiting to happen.

The other things I've noticed is that the open source are very light in documentation. Mr. Adams comments about our lack of documentation. Well ours would be orders of magnitude ahead of their's.

They also would not meet IEEE coding standards, and seriously I start to wonder if there is even a safe activity for forensic science.

Q Turning to a different topic. You're familiar with and you've testified here today about the PCAST report, right?

A PCAST report was released in 2016, October 2016.

Q And the PCAST report touched on a number of areas of forensic science, one of which was probabilistic genotyping, right?

A Yes.

Q And would you summarize for the Court what PCAST said about probabilistic genotyping?

A The key finding in the PCAST report is finding 3. Is that going to appear here or shall I go from memory?

Q I have not marked it as an exhibit.

A Okay. I can do it from memory, I hope. Finding 3 has two

parts. The first part says that the community should move rapidly to more objective methods such as probabilistic genotyping. And part B is that probabilistic genotyping has demonstrated its validity up to three-person mixtures in which the minor is at least 20 percent.

Q What's your view of those findings by PCAST?

A First of all, those findings were amended. Do you want me to go to the amendment or just say my view of it?

Q I would appreciate it if you would summarize the amendment first then go to your view.

A So PCAST continued to work on this. Subsequent to their report coming out, they engaged in conversation with me and others, but I think predominantly me, by e-mail and phone and then eventually by meeting. And they outlined to me at that meeting, which was highly disrupted, they outlined to me what they would like to be done to meet their standard of proof, and that is now being done.

The amendment in particular came out in January 2017 and amended certain small aspects of that wording I have given to you.

Q And what's your view of the amended version of the PCAST finding?

A At the time of its appearance it was a massive understatement of the power of the system. And this came about because of their requirement that proof be published in the

peer reviewed literature. It is very hard to publish internal validation work in the peer reviewed literature. It is boring and journals often have a policy against it. However, there are now two very large internal validation publications on STRmix which makes in my view STRmix the most tested and most published in peer reviewed software of this type available.

Q We reviewed one of those publications earlier.

A Looked at them both. We looked at the FBI internal validation and the 31 laboratory compilation.

Q That's Exhibits 4 and 5.

A Yes.

Q What about these limitations, the three-person mixture and the 20 percent for minor contributor, was PCAST right about those?

A PCAST amended the word minor to person of interest. And that was correct. And that was at my request that they do that. And the fact that they made that fairly obvious mistake of wording is problematic. PCAST do suggest they expect the range to expand as new studies are published, and I would suggest that those studies have been published and therefore we have met PCAST requirements.

Q In terms of the number of contributors and the sensitivity down to the minor contributor?

A That's right. We have gone up to five-person mixtures. The study is massive, it's 28, 25 mixtures. And massive

number, 28 million non donor tests. It's a massive study.

Q When STRmix is internally validated by a particular laboratory, are those things they study how the number of contributors that it can analyze in its system and the percentage or the limitation on the minor contributor that can be used?

A Yes.

THE COURT: What exactly does that mean, Dr. Buckleton, in your view, that "Probabilistic genotyping has demonstrated its validity up to three-person mixtures in which the minor is at least 20 percent." What does that mean?

THE WITNESS: So the three-person mixture is probably obvious. That suggests it hasn't demonstrated it for four or five-person mixtures, as at their writing. And the minor being 20 percent would mean that at least 20 percent of the DNA must come from the smallest donor.

THE COURT: What is the significance if that number is less than 20 percent?

THE WITNESS: As at writing, PCAST would say that that hasn't been validated. They would have been incorrect. However, they are the President's council and I must show them due respect. And I have asked them what they would need to extend that statement, and they outlined to me what they would need and we have done it.

THE COURT: What have you done?

THE WITNESS: We have done a compilation that goes well below 20 percent, down to effectively zero percent. And four and five-person mixtures. And we have published that.

THE COURT: So if I understand you correctly, using STRmix you can validly demonstrate the presence of DNA of up to three people or up to five people where the minor or the person of interest, as you put it, has contributed what?

THE WITNESS: Well, zero, Ma'am. But it wouldn't demonstrate their presence. If there was zero of the person, it would answer inconclusive. So STRmix will reliably report when it knows and when it doesn't know.

THE COURT: If it reports one percent, what does that mean?

THE WITNESS: It will give a very low LR and suggest that it has a high uncertainty about whether that person is there or not.

THE COURT: Okay. Thank you. Mr. Presant, how long how much longer do you anticipate your direct examination?

MR. PRESANT: Maybe just ten minutes.

THE COURT: Okay. We will take our, well, I am going to have one question after you're done and then we will take our break.

MR. PRESANT: Thank you, Your Honor. Dr. Buckleton, STRmix was excluded by a court in one case, is that right?

THE WITNESS: Not STRmix but the evidence, yes.

Sorry. But pedantic.

BY MR. PRESANT:

Q I think it's an important clarification so I appreciate you making it. What case was that in which it was excluded, evidence was excluded?

A That was People versus Hillary in New York State.

Q Were you involved with that case?

A Yes, I was.

Q Would you describe for the Court the analysis in that case and what led the Court to exclude the evidence?

A The case involved the fingernail scrapings of a deceased 12-year-old boy. They presented as a very imbalanced two-person mixture where the minor contributor was very small, and the work was done by the New York State police. New York State police did not have STRmix and do not have parameters for STRmix, so I borrowed parameters for the analysis from a Toronto lab who used the same kit and chemistry. This was candidly disclosed to the Court. And His Honor deemed that to be inadequate. And I do not relitigate that. That is his decision, not mine. He did say that STRmix was generally accepted and had been validated, but that I had used it improperly in that case by the act of borrowing samples, borrowing parameters.

Q Have you tried to approximate the number of cases in which STRmix has been admitted or rather evidence derived from STRmix

has been admitted?

A We have a lot of trouble getting this information. But we think conservatively it's, it's been admitted without an admissibility hearing, so just used --

Q Used in court.

A In the United States, several thousands.

Q And it has faced several challenges where there was an admissibility hearing?

A Nine by me, and some by other people.

Q And have there been any other cases besides the New York case that you just described in the United States where evidence derived from STRmix has been excluded?

A No.

MR. PRESANT: Nothing further, Your Honor.

THE COURT: Thank you. I do have one, and you may want to think about this over the break if you're not prepared to answer it at this point. But early in your testimony you talked about communicating the results of STRmix to a factfinder or to a Court and the necessity to do that accurately and properly. And in your discussion of the training that you provide to a purchaser of your software, that you emphasize how to phrase the likelihood ratio and report it to the Court.

THE WITNESS: Yes.

THE COURT: If this evidence is going to be submitted

to a jury in this case, how would you articulate an instruction to the jury of how to use it or how to understand it or how to apply it? Because that's really --

THE WITNESS: Yes.

THE COURT: -- my task.

THE WITNESS: No, I understand. I'm honored by you asking me my opinion. And you're very correct to highlight this as an important point.

Most of the discussions of statistical evidence in court relate to understanding the result, and that's not just likelihood ratio, that's any statistical evidence. And there's a lot of evidence that it's done poorly.

In this particular case, the statistic is substantial, and if the statistic is deemed reliable and put to court, one would have to say it's something like very strong evidence to suggest the presence of the accused in this mixture. We tend to stop short of certainty, but this is very strong evidence.

THE COURT: Well, that's my concern. Because I think the jury needs to be armed with the ability to understand both sides of that. By saying to the jury the ratio, the likelihood ratio is 49 million to 1, without more that's really game over for most lay people.

THE WITNESS: 49 million to 1 is very strong evidence. It is. I don't know what -- I mean game over may be an American term. I think I can obviously glimpse what it means.

I mean we hold open a very small chance that someone else may produce evidence of this strength. In fact, that chance is roughly 1 in 49 million. Obviously that's not quite certain. It has to be zero to be certain. But it's amazingly strong evidence.

THE COURT: Thank you, Dr. Buckleton. Okay. It's 11:30. We are going to take 45 minutes and come back here at 12:15 and start with cross-examination of Dr. Buckleton.

THE LAW CLERK: All rise. The court is in recess.

(Recess taken, 11:32 a.m.; Resume Proceedings, 12:23 p.m.)

THE LAW CLERK: All rise. Court is back in session. Please be seated.

THE COURT: Dr. Buckleton, first of all, obviously you're still under oath. But I just want to follow up a little bit on the last discussion that we had before we adjourned.

You aren't proposing, are you, that I tell the jury that the results of probabilistic genotyping are powerful evidence, are you?

THE WITNESS: I believe the evidence in this case is powerful, is that what you mean?

THE COURT: No, that's not what I mean.

THE WITNESS: Probabilistic genotyping is just a tool, and it can only do, it can only do things if the evidence is there.

THE COURT: Okay. Ms. Kloet, are you ready for your cross-examination?

MS. KLOET: Yes, Your Honor.

CROSS-EXAMINATION

BY MS. KLOET:

Q Dr. Buckleton, you didn't write the software at issue in this case, did you?

A No.

Q I'll repeat the question. If could you re answer it. Dr. Buckleton, you didn't write the software that was at use in this case, did you?

A No.

Q I believe it was your testimony that Duncan Taylor did, correct?

A I think either all or most of it.

Q You also testified today that Duncan Taylor is a forensic biologist, correct?

A Yes.

Q He's not a computer scientist.

A No. He's a scientist who programs.

Q You don't need a degree to program software, do you?

A No.

Q Okay. You testified in a Michigan State case a couple years ago called People versus Mohammed?

A Yes.

Q In fact, that was the first time I believe you ever testified in a Daubert hearing.
A No.
Q It wasn't?
A No.
Q When was the first testimony?
A 1995.
Q And that was in a United States Daubert hearing?
A In South Carolina.
Q Did that have to do with probabilistic genotyping?
A No.
Q Okay. The one in Muhammad is the first time you ever testified in a Daubert hearing involving probabilistic genotyping.
A Yes.
Q Okay. Thank you. And you were asked during that hearing by defense counsel, "You're not a computer guy, are you?" And you responded, "No." Is that correct?
A Possibly, yes.
Q Okay. I understand based on your earlier testimony that STRmix version 2.6 is now being coded by a professional coding company, correct?
A Yes.
Q And it's still in the testing process, fair to say?
A Yes.

Q And during that testing process 416 miscodes have been identified. Right?
A I'm going to say yes.
Q I think that was in the chart.
A Yeah, it's at least 416. I'm not sure miscodes is the right word but certainly the tenor of your question is correct.
Q Okay. Thank you. So there were 416 miscodes, or for lack of a better description, similar problems or issues identified and this was after bringing in professional software developers, correct?
A Yes.
Q How many miscodes were identified in 2.3.07 during the testing process?
A I don't have that information.
Q How can you assure the Court that there are no other errors to version 2.3.07 that was used here?
A No one can ever make such an assurance.
Q You also testified regarding the assistance of peer reviewed articles that relate to STRmix.
A Yes.
Q Almost all of those articles are in one journal, isn't that right?
A That's the journal I tend to submit to, yes.
Q And that journal would be "Forensic Science International"?
A No.

Q What journal is that?
A "Forensic Science International Genetics".
Q Thank you, I misspoke. Now that's not a software engineering journal, is it?
A If one were writing for the forensicology community why would one publish in a software engineering journal?
Q From that is your answer no?
A I think no.
Q It's not a computing journal, is it?
A No.
Q Okay. You testified today that you're familiar with the IEEE standards.
A I know of them.
Q Okay. So you know that one of the recommendations in the guidance from IEEE is that software should go through independent validation and verification processes.
A Yes.
Q You also testified that you are a member I believe of the ISFG?
A Yes.
Q Could you remind me again what that acronym stands for?
A International Society of Forensic Geneticists, but I particularly meant I was on the panel of the DNA commission.
Q Okay. One of the recommendations of the ISFG in the recent article that was presented to you by the government was that to

incorporate the IEEE standard of independent testing, correct?
A We probably meet that standard.
Q Do you agree that that is what the article itself stated?
A Yes, I think it did, yes.
Q Okay. And, in fact, the United Kingdom Forensic Science Regulator that you previously testified about, they too recommend independent testing of software.
A Yes. But I think we meet that requirement.
Q Okay. The version in 2.3.07 or version 2.3.07, pardon me, that was used here, that did not follow the independent review according to the IEEE standards.
A Yes, I think it did.
Q In what way?
A Do you have IEEE and particularly the addendum on the independent review?
Q I can --
A The full requirements for independent review.
Q I can bring up the article that was previously, that references the IEEE standard.
A I need the actual IEEE I think.
Q I don't have it.
A I do on my laptop. I'm sure Mr. Adams has it.
Q I think -- let's table that for a minute and move forward. I'm referencing the, I'm talking about the reference that was made in the article that was previously admitted as an exhibit.

A Which article would that be?
Q I'm sorry?
A Which article?
Q The ISFG article.
A Yes, I think we meet the IEEE validation or verification they call it requirements.
Q Okay.
THE COURT: What do you understand independent verification to mean?
THE WITNESS: I understand it differently from IEEE. So IEEE says such things such as the facilities and resources should be separate, the personnel should be separate, and the funding should be separate. And I think we achieve that.
I would actually mean they should be separate from the developers altogether. And to some extent we have met that as well.
BY MS. KLOET:
Q Your testimony just now is to some extent. To what extent?
A So in particular, the group who do the validation or verification do not do the programming. They are a separate group. So in that regard, we maintain independence from the coders. That independence has softened slightly by the fact that we find bugs and you get into an iterative process of talking with the coders. Now speaking of complete separation, at least two groups have done some quite extensive by hand

verification of the STRmix outputs.

Q So your statement just now is that they have done verification of the STRmix outputs. Does that mean they have duplicated the mathematics in the program?

A Yes, yes.

Q How many calculations are there in a STRmix analysis?

A A typical analysis would be hundreds of millions.

Q So is it your testimony today that hundreds -- the hundreds of millions of mathematical equations that are required for STRmix to run once have been repeated in toto?

A There are not hundreds of millions of equations. There are hundreds of millions of calculations, and a sample of them would have been tested.

Q So it was a sample that was tested, not all of the calculations.

A You can't even begin to do all of them.

Q Okay. Thank you. You testified just a moment ago about the independence of the review of the software for validation of the software. How many of those reviews involved a review of the source code?

A None.

Q Has anyone reviewed the source code?

A Only Mr. Adams.

Q Okay.

MS. KLOET: If I could confirm what he is viewing. I

can confirm that's what I intended to bring up and I could use my paper copy as a reference. If that pleases the Court.

THE COURT: What are we after here?

MS. KLOET: Your Honor, I have a copy of the STRmix comparison reports, the seven-page document that was generated in this case and I would like to be able to refer that to him for the next set of questions.

THE COURT: Was that introduced by the government?

MS. KLOET: I don't believe so.

MR. PRESANT: It was not during his testimony, but we do have it in our exhibit book so we could bring it up if that would be helpful.

THE COURT: Would you do it, please?

MS. KLOET: Thank you.

MR. PRESANT: Government's Exhibit 26, I believe.

BY MS. KLOET:

Q That looks the same to me, Your Honor. So I can continue.

THE COURT: And this is the report generated by Dr. Buckleton in this case.

MS. KLOET: Not precisely, Your Honor, but I can have him identify it for you if you wish.

THE COURT: Okay.

BY MS. KLOET:

Q Dr. Buckleton, can you identify this document?

A It's STRmix output.

Q Did you create this personally?
A No.
Q Okay. Thank you. Do you have any reason to doubt it was the STRmix report that was generated in this particular case?
A No.
Q Okay.
MS. KLOET: I don't know if you want to stipulate to that with me since it's your exhibit. That was the report in fact generated in this case.
Your Honor, I have asked the government if they will stipulate if this was the report generated in this case, seeing as it was a Government's Exhibit.
THE COURT: Mr. Presant.
MR. PRESANT: I mean we didn't mark it. I think the Court will hear testimony from another witness that it was produced in this case. I would rather that witness lay the foundation for it.
MS. KLOET: Dr. Buckleton, I would like to have you look at page 3, please. What's the title of that page or the chart on that page, more specifically?
THE WITNESS: Per Locus Likelihood Ratios.
BY MS. KLOET:
Q What does that mean?
A That is the likelihood ratio for each locus.
Q And when you say locus, you mean each specific spot on the

DNA strand that is viewed for purposes of DNA analysis in this case, correct?

A Yes.

Q Okay. Were there any loci that were left out of analysis here?

A Yes.

Q Which were those?

A Can you just expand this and show me loci I think it's 19 and 20 or 18 and 19? So it's DYS391 and D8S1179.

Q Okay. Does STRmix or can you tell from this document the reasons that those two loci were left out of the analysis?

A I can't tell from this document, but I otherwise know.

Q Okay. Assessing a number of contributors is a necessary first step to running a STRmix analysis, right?

A Yes.

Q All right. Now, the number of contributors is always an estimate, is it not?

A Yes.

Q Okay. I want to talk a little bit about -- I don't think we have covered the concept called the analytical threshold. Can you explain to the Court what an analytical threshold is?

A Other than the fact that an analytical threshold has absolutely nothing to do with STRmix. It's a threshold set by the laboratory below which we ignore the data.

Q So it's set by the laboratory. So the laboratory then

would be ignoring data under the --
A Well, ignoring is a big word. We will not use that data. We will look at it sometimes for number of contributors information.
Q What is the implication if a number is below the analytical threshold?
A A peak is below?
Q Correct.
A We won't use it in the analysis.
Q Okay. So if the number is below that, the RFU in a particular allele or peak, as you said, is below the analytical threshold it could be thrown out?
A It will be thrown out.
Q It will be thrown out of the analysis. Is it possible that that low peak could actually be evidence of the presence of true DNA?
A Many of them are exactly that.
Q And yet it's thrown out. Correct?
A That's correct. And that's not a new policy. That is a policy that's been in place for well over 20 years and is a very good policy. It stops us using data that is confusing or likely to cause issues. Whilst we throw out some good data, we throw out a lot of bad data.
Q I think I heard you testify just now you could possibly throw out what you characterized as good data.

A Most of it will be. But some of it is also confusing, bad, and will derail the process. And this is a very good policy. It's sanctioned by SWGDAM and every major policy group.

Q So could some of that data that is thrown out, so to speak, could it lead potentially to an underestimation of the number of contributors in a mixture?

A Yes.

Q And that in turn could potentially affect the likelihood ratio generated.

A Generally if you get the number of contributors wrong, you get a conservative likelihood ratio.

Q Generally speaking.

A Yes.

Q Okay. There are competing programs to STRmix, correct?

A Yes.

Q And isn't it true that some of those programs choose to incorporate that data below that threshold?

A Yes.

Q One of those, that program, an example of it would be TrueAllele, correct?

A There are two I know of that do that.

Q What's --

A TrueAllele is one of them.

Q Thank you.

THE COURT: What's the other one?

THE WITNESS: Lab Retriever, which is spelled like the dog.

MS. KLOET: Thank you. You testified on direct to a case you were involved in, New York versus Hillary.

THE WITNESS: Yes.

BY MS. KLOET:

Q And that was I believe you testified that STRmix analysis was not admitted or the likelihood ratio generated by the STRmix analysis in that case was not admitted, correct?

A That's correct.

Q Now, in that case they also ran another analysis using TrueAllele, correct?

A Unbeknownst to me, TrueAllele had been run prior to my involvement.

Q You're aware of that now, however?

A Yes, very much so.

Q So you're aware that that program concluded that in fact the defendant in that case was excluded from the mixture.

A No, that's not so.

Q What was the conclusion in that case?

A Essentially inconclusive.

Q So the conclusion for TrueAllele was inconclusive and what was your, the STRmix conclusion as to a likelihood ratio?

A Am I allowed to answer this?

THE COURT: Why would you not be allowed to answer?

THE WITNESS: I don't know. It was precluded then. I'm about to give it now. Shall I do so?

THE COURT: Yes.

THE WITNESS: The likelihood ratio I think was 300,000 from memory.

BY MS. KLOET:

Q Okay. And that was excluded. Correct?

A Yes.

Q Thank you.

A We have subsequently diagnosed where TrueAllele went wrong if you're interested.

Q Not at this stage, thank you. I would like to move on to the next chapter. I think we are getting a little bit outside of the scope of this case.

I want to talk about drop-in. Are you familiar with the concept of drop-in?

THE WITNESS: I invented the word.

BY MS. KLOET:

Q Okay. Can you explain it for us then, please?

A Yes. Drop-in relates to the appearance of alleles in a non reproducible fashion in the electropherogram.

Q So in lay person terms, what does that mean, is it a single allele in the --

A We have a model of alleles snowing from the ceiling. So you imagine that alleles are falling from the ceiling and

occasionally fall into your Eppendorf tube. And I can not spell Eppendorf.

Q Is it an Eppendorf --

A It's the little plastic thing that we have our extractor.

Q So that's something you would find in a lab?

A Yes.

Q Okay. So is drop-in a form of contamination?

A Yes.

Q So drop-in could take place in a lab, correct, as you just described?

A It would only take place in a lab. If it happened prior to the taking of the sample, we wouldn't consider it drop-in.

Q And a lab will typically have filters and protocols to address the risk of contamination, wouldn't it?

A Yes.

Q Okay. So I just had this document, it's the same one as before. And you identified it as a report generated by STRmix in this case.

Can you please turn your attention to page 4 of this document? What's the title of this chart present on this page?

THE WITNESS: Parameters.

BY MS. KLOET:

Q What are parameters?

A Is there a hard copy? Because mine is sideways and bleary. Can I get a hard copy?

THE COURT: Should be in the book.

THE WITNESS: Is this number 5 or 6?

MS. KLOET: It's K in defense exhibit.

THE WITNESS: I think we have the equivalent. But can someone help me?

MR. PRESANT: 26 in the government book.

THE WITNESS: I'm good, I think.

MS. KLOET: You may need the other one going forward.

THE WITNESS: I'm good now. Thank you.

BY MS. KLOET:

Q So we are looking at the same document as before, right?

A Yes.

Q All right.

A I think.

Q And you just testified that the title of this chart is parameters. Can you explain what parameters are, what that means for purposes of this document, please?

A These are the input parameters for STRmix.

Q What do you mean by input parameters?

A They're the parameters that are input into STRmix.

Q Who inputs them?

A They usually sit during the laboratory validation and they are usually locked as a default. If any of them change they become bold and italic.

Q Okay. So are these figures that are input by the local

lab?
A Yes.
Q Okay. Not by STRmix itself.
A No.
Q Do some of these parameters have an impact on the STRmix analysis potentially, not in every case but potentially?
A I think nearly all of them would.
Q Okay. Fair enough. In the third column, the second row, there's a parameter called drop-in cap.
A Yes.
Q What is the value for that?
A 400.
Q What does that mean?
A Peaks above 400 will not be considered drop-in.
Q So the converse of that would be peaks below 400 would or could be considered drop-in, correct?
A Would be proposed potentially as drop-in. One of the proposals.
Q Okay. If you can turn to the next page titled Evidence Input Files. The very first locus listed there is D3S1358. Are there any of the alleles present at this particular locus that are under the drop-in cap?
A Yes.
Q Which one is that?
A 17.

Q All right. How about at the next locus, any under the drop-in cap?
A Yes.
Q How many?
A Two.
Q All right. How about at locus D10S1248?
A Two.
Q Two. Okay. If I told you there were 11 peaks on this, in this document that were at or lower than 400, would you have any reason to question that?
A No.
Q Okay. So there are 11 peaks that are noted as present here but in fact potentially could have been drop-in, right?
A Yes.
Q And your testimony earlier was that drop-in could, is essentially contamination or could be attributed to contamination, true?
A Yes.
Q Okay. Thank you. Back to the page we were just on. Right under that row we were looking at drop-in cap. I'm looking at the row entitled drop-in frequency.
A Yes.
Q What's the value there?
A .3453.
Q Okay. You said this was a parameter that was inputted by

the lab following local validation, right?
A Yes.
Q So the frequency of drop-in as a result of their own testing was .3453 or 34.5 percent, true?
A Not really.
Q What do you mean by not really?
A Looks like they fitted the gamma distribution which is a good idea or an exponential distribution which is a good idea. So the 3453 may have come from a spreadsheet I made that gets the tail area correct. But am I right witness for this? I didn't do any of this work.
Q I'm just asking about where you believe those figures came from and what they signify.
A They have come from the internal validation which a subsequent witness will speak to.
Q These are parameters that are used in the software that you co-developed, correct?
A Yes.
Q Okay. Thank you. Speaking of contamination, I think you used the analogy snowing from the ceiling --
A Yes.
Q -- for how the alleles might arrive in a mixture in a lab or in a sample in a lab. The DNA can arrive on an object and therefore in a sample without any sort of direct physical contact, isn't that true?

A Yes, I think that's true.
Q Okay.
A This has nothing to do with STRmix. Or prob gene or --
Q Would you agree it has to do with potential contamination?
A Yes. Is that an area you would like to qualify me as an expert on?
Q I believe you're qualified as an expert in forensic science for purposes of this hearing. So --
A I have studied contamination.
Q Okay. What I'm getting at, so the DNA that is present in the sample that was analyzed here could potentially be explained by, could potentially be there even without direct contact, isn't that true?
A Really there are better witnesses. I don't know the quant. And I mean you're getting me to say things like possible where things are, you know, some things are possible but enormously unlikely.
Q Well, drawing upon your experience and expertise in forensic science.
A I haven't studied this case at all. I don't know where the sample is from. I don't know the quant value. There are better witnesses for this evidence.
Q Okay. Do your best as a qualified expert in forensic science. My question is, could DNA arrive in a sample without direct contact between an individual and that sample?

A Yes, you could drop blood on something or you could ejaculate on something or you could spit on something.

Q Okay. Thank you. That's as far as I'll take it on that point.

I believe this document was previously displayed to you but for purposes of the record could you identify it again please?

A It's an e-gram.

Q If I told you it was the e-gram or the electropherogram for this particular case, for the sample taken in this case, would you have any reason to doubt that?

A No.

Q Okay. Did you have a chance to review other materials in this particular case including Mr. Gissantaner's own EPG?

A I have seen it. I haven't reviewed it. It was present at a preparation meeting yesterday. So I've seen it.

Q Okay. Did you have -- do you agree that his -- at every loci that was analyzed here there are some loci where both of his alleles are not present?

A I haven't done that work. But I could it for you, but I haven't done it.

Q Okay. I'm just going to take you from one locus and we will go from there. So have a look at CSF which is in the bottom half of this first page, just to the left of the -- I'm looking at this one right here. So what alleles are present at

this locus?
A 7, 9, 10 and 11.
Q Okay. I'm going to pull up another document quickly. Was this one of the documents that you had a chance to look at yesterday?
A It was present and I have looked at the blue lane only.
Q So if I were to -- would you have any reason to doubt that this is the EPG for Mr. Gissantaner's reference sample?
A No.
Q Okay. So let's look at, have a look at this locus right there. I'm trying to look at CSF. Second circle. Okay. So this one right here. So what two alleles are present there?
A 9 and 12.
Q Okay. So that would represent ostensibly the alleles that Mr. Gissantaner carries at that particular locus, true?
A Yes.
Q So 9 and 12. Can we go back to the previous exhibit, please? So if you could look back again at that same locus, I believe you just agreed that the alleles that are present on Mr. Gissantaner's EPG were 9 and 12. And you testified earlier too that the alleles that are present at the same locus are 7, 9, 10 and 11. So there's no 12 present in this sample, is there, at least as displayed by the EPG?
A There is no 12 peak above the analytical threshold.
MS. KLOET: Okay. Thank you.

THE COURT: What does that mean?

THE WITNESS: There is no peak on which we would choose to rely for the 12 allele that is concordant with Mr. Gissantaner's 12 allele. So there are two possible explanations. The 12 of Mr. Gissantaner has dropped out, or Mr. Gissantaner is not a donor.

MS. KLOET: Are you ready for me to change the display, Your Honor?

THE COURT: Yes, yes.

BY MS. KLOET:

Q So this is Defense Exhibit K, the comparison report that you've already identified. Page 3. I would like you to have a look at page 3, please, which she is about to display. Just a minute while she orients the page. So I have some highlights here, but what I would like you to look at is the highlights under CSF, that locus we were just looking at.

A Yes.

Q Okay. So at loci CSF, the third column we have there, what does that represent?

A The likelihood ratio for that locus.

Q For that particular locus, okay. What's that number?

A 4.2.

Q Thank you. Now, a likelihood ratio greater than one, what does that signify?

A Support for the prosecution proposition.

Q Okay. So to put it another way, it's an inclusive piece of information or inclusive likelihood ratio?
A That's correct.
Q So here you have a likelihood ratio greater than one, and therefore inclusive at a location where you don't see, as we just established, all of Mr. Gissantaner's alleles, correct?
A Yes.
Q If this number, 4.2 was the likelihood ratio at each locus on the 21 reporting loci, do you have, can you give us an estimate of what the likelihood ratio would be generated from?
A I can give you the exact answer.
Q What is that?
A It's 4.2 to the pair of 21.
Q Okay. That's a 4.2 with 21 zeros after it?
A No, not quite. It's about --
Q Can you tell us what that is?
A Do you want me to do this in my head? I think I can.
Q Would it be fair to say that you take the number 4.2 and you move the decimal 21 places?
A About 150 billion.
Q 150 billion. Okay. So 4.2, although it seems small has a pretty large impact on the overall likelihood ratio if it were carried across all the reporting loci?
A Yes.
Q 150 billion, is that what your testimony was?

A I'm doing math in my head which isn't a good plan.
Q I have a different number but I'll go, they are both quite large.
A Somewhat might have Excel open. They can put it in quickly.
Q But you would agree that's a fairly large likelihood ratio.
A Yes.
Q And that's generated from a number that's only a couple of integers above 1, right, at each locus? I'll rephrase. Might be a confusing question. So what does a likelihood ratio of 1 indicate, is it informative?
A No.
Q Okay. And here we have a likelihood ratio of 4.2. That's only 3.2 above an uninformative likelihood ratio number, right?
A Well, the answer is yes.
Q Okay.
A It's a poor, it's a poor thought process.
Q But if you were to have that particular, although it's only a difference of 3.2, at least on a likelihood ratio scale, if you were to use that 4.2 and input the per locus likelihood ratio for each one of these rows in this particular column, you would generate a likelihood ratio statistic at least in the billions.
A Yes.
Q Okay. Thank you. You talked about stutter in your direct

exam. I would like to discuss that just a little bit. What is stutter exactly to refresh, to the extent we didn't already cover it to refresh the Court's recollection?

A Stutter is an artifact of the amplification process.

Q What kind of artifact do you mean? How does it manifest?

A In Fusion 5C there are three predominant stutters. The most common is the minus 1 repeat stutter which creates a peak, one repeat smaller than the allele.

Q So I'm sorry to interrupt you. I just want to make sure we don't get too much at once. So when you say one repeat shorter, would that mean then if a peak was 12 in order for it to be, if backward stutter took place as you just described, there may be evidence of an 11 present?

A Yes.

Q Thank you. Okay. Go ahead. And you were talking about three different types of stutter.

A The other is forward stutter which is one repeat larger.

Q Okay.

A And we have a minus 2 base repeat stutter happening at D1 in Fusion.

Q Okay. So to summarize, forward stutter is one repeat longer, backward stutter is one repeat shorter, true?

A Yes.

Q Okay. And that manifests as an increase in the allele number that's present in the EPG, if -- potentially if that

took place?
A No, it shouldn't. We have got mechanisms in place to avoid that.
Q That's how it could appear, potentially, without those mechanisms?
A If someone were to do forensic science without those mechanisms, yes, they would get the answer wrong.
Q Okay. Thank you. What is forward stutter modeling?
A Modeling the forward stutter I spoke of.
Q Okay. When you say modeling, what does that mean?
A We essentially predict a peak height for forward stutter and compare it with the observed peak height and do our standard calculation for the probability of that occurrence.
Q Okay. Thank you. Did the version of STRmix that was used in this case, 2.3.07, take into account the possibility of forward stutter?
A No.
Q Do the later versions of STRmix including the one you're currently developing take into account that possibility?
A Yes.
Q Thank you. So I'm going back to the electropherogram for the sample in this case. If you could turn to page 2. I would like you to look at this second column of boxes here. So the second column of boxes, I'm referencing to the left of that line there, those are the alleles present at locus D12S391,

correct, as they appear on the electropherogram?
A Yes.
Q Okay. The 19 peak here, the 19 allele is, the RFUs, or the peak height as you described it earlier, are 4649, true?
A Yes.
Q The 20 peak just below it or just next to it, would be, I'm not saying this specifically, but a 20 peak could be in forward stutter position to the 19 peak.
A It is in the forward stutter position.
Q Okay. Thank you. But by comparison to 19 the amount of RFUs at 19 which we just established was 4649, the RFU at 20 are quite low at 648 in height.
A Yes.
Q So potentially the 20 that's present there could be forward stutter to the 19 that's present in larger amounts.
A No.
Q Your testimony --
A You would never get a forward stutter that big.
Q Why do you say that?
A It's almost always less than two percent of the parent. But it is also in the back stutter position of the 21.
Q The 20 you mean?
A Yeah, it's between the 19 and the 21.
Q So is it your testimony today it's absolutely impossible?
A It's absolutely impossible for that 20 peak to be all

forward stutter.

Q Okay. How about we look at peak 21. Now, that particular, the RFU at that particular peak are 1970, correct?

A Yes.

Q And the 21 peak is or the 20 peak is in backward stutter position to that 21 peak, right?

A Yes.

Q Is it possible that that 20 peak is a sum of the forward stutter from 19 and the back stutter of 20?

A I have to go and do some math. I would say -- see I don't have the stutter ratios for D12 in my head but I'll just try and do it. Really why are you making me do mental arithmetic in court? Why haven't we done prepared work on this?

MR. PRESANT: Your Honor, I'm going to lodge an objection now on that basis which is he has already testified that he has not reviewed the specific work done in this case. The government is going to call a witness later on who has done the work in this case and can testify about these particular electropherograms. I think Ms. Kloet has had a lot of leeway to question Dr. Buckleton about these issues, but I think her examination would be better focused on how the software functions as opposed to what was done in this particular case because he hasn't reviewed it. It's not fair to make him do that on the stand.

THE COURT: Ms. Kloet.

MS. KLOET: Well, Your Honor, this is information that's incorporated into the software program that Dr. Buckleton created himself. This is -- he's familiar with the parameters. I'm asking him to simply read the numbers that everyone in the courtroom can review. And he's familiar as a forensic science expert about what, of concepts of forward and backward stutter. The question I just asked didn't require him to do mathematics. I just asked him if it was possible under the theories and principles of forensic, of forward and backward stutter that this particular allele in the amount that's there could potentially be a combination of forward and backward stutter. Not conclusively yes or no, just potentially. I don't think that requires math.

MR. PRESANT: Well, I think the testimony is he would have to refer back to other materials in order to say conclusively. So I don't think there is a lot of probative value in the question. Is it something possible hypothetically or conceivably. I think she should ask the witness what he knows about and what he is prepared to testify about here today.

THE COURT: I think you need to move on, Ms. Kloet. I really do. I think you are getting a little bit deep in the woods here with regard to Dr. Buckleton's expertise and the meaning of these documents which he did not produce.

MS. KLOET: I'll move on, Your Honor. And I can ask

another witness to hopefully answer some of those questions that he's not in a position to do today.

Based on your -- so is it your testimony for purposes of clarification that you have not reviewed the EPGs in this case or you just did briefly yesterday?

THE WITNESS: I saw them. I don't think I could say I reviewed them.

BY MS. KLOET:

Q Okay. Can you pull up J, please? You are familiar with what a STRmix report looks like, correct?

A Yes.

Q General format. Okay. I hope so. And you're aware to the extent you have some familiarity and from the cover page of the document I'm displaying here that the determination of the number of contributors here was three, right?

A Yes.

Q Okay. No reason to dispute that.

A No.

Q Okay. Or at least that determination.

So what is this document that I'm looking at here?

A That's the opening page of a STRmix output.

Q Okay. And if you see towards the top under the STRmix brand stamp there is a comments. It says deconvolute. What does that mean?

A Mine is different from yours.

THE COURT: Are you on Exhibit J in the defense book?

THE WITNESS: I must not be. I think I'm sorted now, thank you. In any case, deconvolute is, tends to mean exactly what it says, it's asking STRmix to deconvolute the mixture.

BY MS. KLOET:

Q And by deconvolute, could you just break it down for us?

A It's the process I spoke of earlier which is essentially to run the Markov Chain and ascertain those sets of genotypes that could plausibly explain the mixture.

Q Okay. And on this front page there is a chart entitled Summary of Contributors, correct?

A Yes.

Q And in that chart there is a row entitled mixture proportions.

A Yes.

Q What does it say as to the proportion of the mixture contributed by 1, what is the figure there?

A 68.

MR. PRESANT: Same objection, Your Honor. The witness didn't prepare this document. If Ms. Kloet has questions about how the software works or why it generates a document like this, I think those are fair for Dr. Buckleton. The document speaks for itself and there will be other witnesses who did prepare the document who will testify about what it means.

THE COURT: I think she is heading in that direction,

Mr. Presant. And he can obviously relate what is on the page.

MS. KLOET: I am, Your Honor. I don't expect very many more questions on these particular documents. With respect to contributor 2, what's the proportion of the mix?

THE WITNESS: 25.

BY MS. KLOET:

Q And contributor 3?

A 7.

Q Okay. Dr. Buckleton, do you recognize this document?

A Yes.

Q What is it?

A This is the raw output from STRmix.

Q The raw output you said?

A Yes.

Q If I could call your attention to the first page, the section entitled contributor order giving highest likelihood ratio.

A Yes.

Q Do you agree that this report generated by STRmix tends to indicate the strongest support for the sample labeled LS15-3777-2AX as contributor 3?

A Yes.

Q So if I were to tell you that 2AX represented Mr. Gissantaner, would you agree that the strongest likelihood ratio for him would be at contributor 3 according to this

report?
A Yes.
Q Okay. Thank you. Earlier in your testimony you referenced diagnostics, is that true?
A Yes.
Q And STRmix generates several diagnostics. What was their purpose?
A To help diagnose the performance of the software in this particular run.
Q So it can help tell the analyst whether or not the run was good, I guess, for lack --
A Yes.
Q And the one you testified to earlier was the Gelman-Rubin diagnostic?
A Yes.
Q Can you just roughly describe what that means?
A Have all the different chains gone to the same places.
Q Okay. I guess I lied. I had one more reference to this. So back to J. So this is the document I was just showing you that the STRmix report that we went over with the contributors listed. Under the section entitled Run Information, is there any information about the Gelman-Rubin diagnostic you just explained?
A Yes.
Q What's the figure for that?

A 1.41.
Q 1.41. Is that a potential warning indicator that this may have not been a good run?
A It's above 1.2 which is a limit many people would use.
Q So if it's above 1.2, then potentially this could indicate there was an issue or concern with the run?
A With running it for longer.
Q A lab can run STRmix more than once, right?
A Yes.
Q Okay. Thank you. I want to talk about the validation studies that you covered pretty extensively. I'm not going to cover them as extensively. You mentioned SWGDAM. SWGDAM has published guidelines, right?
A Yes.
Q But guidelines aren't the same as standards, are they?
A There are no standards.
Q What is the difference?
A To some extent standards are mandatory. Whereas guidelines are not.
Q Okay.
A But for all intents and purposes, SWGDAM guidelines are mandatory in the United States.
THE COURT: Would it be fair to say that a guideline is more like a range of behavior or --
THE WITNESS: Yeah.

THE COURT: -- values as opposed to a standard which is a single --

THE WITNESS: Yes.

THE COURT: -- statement.

THE WITNESS: Yes, that's well put.

BY MS. KLOET:

Q When was the STRmix software first commercially available in the United States?

A January 2014.

Q And when were the SWGDAM guidelines published?

A Can you check for me? I think it's July 2015.

Q That sounds about right. I have June.

A It could be June.

Q So you completed the developmental validation or engaged in a significant amount of it and sold the product in the United States but in the absence of formal standards under SWGDAM.

A Yes.

Q Okay. Thank you. And when SWGDAM eventually did issue those guidelines, didn't they specifically instruct not to apply those guidelines retroactively?

A They may have. I think they usually do that. But it has no, it has no traction. A Court can deem that it should be taken cognizance of whether SWGDAM thinks so or not.

Q Okay. Talking about the peer reviewed publications in this case. The peer reviews of these validation studies of yours

did not involve a duplication of the exact tests that you completed on the STRmix program, correct?

A No.

Q So they were an evaluation of the studies or tests that you completed, true?

A They usually about paper exercise. So they actually get the paper electronically and read it. Very few people actually repeat the calculations, but I do tend to do that, but most referees do not.

Q Okay. Thank you. The internal validations that you referenced, you indicated that Nathan Adams is the only individual who has reviewed the source code.

A Yes.

Q So impliedly then the Michigan State Police didn't review the source code during their validation.

A No, they did not.

Q Okay. Someone who has to review the source code would have to get permission from ESR, is that true?

A Permission is automatic on signing of the NDA, so they don't need to get specific permission.

Q Okay. I mean you don't make it publicly available.

A No, we don't.

Q Do you recognize this document, Dr. Buckleton?

A Just give me a minute, please. Yes. This is our access to STRmix that appears on the STRmix web page.

THE COURT: We are at C?

MS. KLOET: CC, two Cs. All right. So this is a 3-page document, right?

THE WITNESS: Yes.

BY MS. KLOET:

Q On the first page I would like to call your attention to the fourth full paragraph. It begins, "Where STRmix has been used."

A Yes.

Q Okay. So this paragraph delineates what's available for inspection, true, for defense review?

A This delineates what we would automatically hand over.

Q Okay.

A We would also hand over extraneous material if requested and if our lawyers allowed us to.

Q Okay. Do you agree that there are four items listed in this bulleted list in that paragraph?

A Yes.

Q And that's the source code, a limited trial version, developmental validation records, and the user's manual, correct?

A Yes.

Q Github repository is not on that list, is it?

A No.

Q You charge defense to do this review, don't you?

A No.
Q ESR charges defense to do this review.
A No.
Q So is it your testimony today that if defense counsel requested this information that that would be free?
A I think you have to pay the costs of the supervision and it doesn't go to ESR at all. And an installation fee by the installer of the software and that doesn't go to ESR at all.
Q What if I told you it was $100 an hour to access the source code.

MR. PRESANT: I'm going to object on relevance grounds, Your Honor. I think this is a Daubert hearing about the reliability of the software. Dr. Buckleton developed the software. I think he's testified already today that he doesn't handle the business end of it. He's certainly not the lawyer for ESR. I'm not sure there has even been a foundation that he developed this document. And I don't understand the relevance of it considering the code review has already taken place and the Court is going to hear testimony about that in order to determine the issue here which is whether or not the software is reliable.

THE COURT: I don't understand the relevance either, Ms. Kloet.

MS. KLOET: Your Honor, whether the software has been reviewed and validated and the extent of those reviews and

validation goes directly to how accessible this program is. And I believe Dr. Buckleton has already testified that there are some open source and previous software versions for probabilistic genotyping software, and his position previously was that this be allowed open source. So I think it goes directly to access, which goes directly to whether or not it's been sufficiently validated and tested. And as he -- I'll stop there.

THE COURT: Well, it strikes me that whether it is free for purposes of validation is irrelevant. The question is has it been validated independently of Dr. Buckleton and the company for which he works, or the agency for which he works. Right?

MS. KLOET: The question is the breadth of the validation, the depth and the breadth, as it applies to different mixtures from independent review, and his testimony tended to suggest today that the review was in fact done internally. So access to the source code would enable independent reviewers to identify issues or problems in the coding that might manifest themselves during the operation of the program itself.

THE COURT: Well, is it your position then that charging a hundred dollars an hour is somehow limiting independent review? Why don't you just ask him, has, have there been independent reviews of this software including the

source code. Isn't that really the question?

MS. KLOET: I think he's testified to that today, Your Honor, that Mr. Adams has done that and he's the only one. So I can move on to the next set of questions.

THE COURT: Okay.

MS. KLOET: My point was to emphasize the lack of access.

G. Dr. Buckleton, in 2008 you coauthored an article about the likelihood ratio and the random man not excluded as means of presenting evidence. Isn't that true?

THE WITNESS: Yes.

BY MS. KLOET:

Q Is this the article that you authored?

A Yes.

Q Now, in the article fair to say that you weighed these two approaches of presenting evidence against each other, right?

A Yes.

Q And in fact you came to two conclusions: One of those was that likelihood ratios are different -- difficult to present in court.

A Yes.

Q Specifically, if you would like to follow along, on page 344, in the second column, second paragraph, can you read the last paragraph or last sentence, pardon me, for me starting with there is.

A I stand no chance of reading this at the moment. Can this be made bigger?

THE COURT: Can you look at it in the book?

THE WITNESS: Maybe I could. What is it?

THE COURT: It's G.

MS. KLOET: Here we go. Exhibit G. The final paragraph, final sentence in that paragraph, could you read that for me, please?

THE WITNESS: Yes. "There is considerable evidence that likelihood ratios are harder to understand and that they may be slightly more prone to the prosecutor's fallacy."

BY MS. KLOET:

Q Okay. Thank you. The paragraph right below that, if you could pull that up, please. Can you in the middle of the paragraph you see the word furthermore. Could you read that sentence?

A "Furthermore, if likelihood ratios lead to better scientific decision making then they should be used."

Q What's the next sentence?

A "Whether they are then presented to court is a secondary decision."

Q Thank you. This was an article that you coauthored, correct?

A Yes.

Q I'm just going to ask this final question and I think it's

done. It's not related to this document necessarily so you can pull it down.

What is a conditioning profile?

THE WITNESS: A conditioning profile is a profile that can reasonably be assumed to be present under both the prosecution and defense scenario.

BY MS. KLOET:

Q Okay. Could there be more than one conditioning profile?

A Yes.

Q So does that mean if there's an additional reference sample added in the analysis that it would impact the outcome, the outcome of an analysis potentially?

A Yes.

Q Would it tend to increase its reliability if you knew for certain who else was in a mixture?

A They are all reliable. It would increase its ability to discriminate true from false contributors.

Q Okay.

MS. KLOET: Thank you. That's all I have, Your Honor.

THE COURT: Any redirect, Mr. Presant?

MR. PRESANT: Thank you, Your Honor.

REDIRECT EXAMINATION

BY MR. PRESANT:

Q Dr. Buckleton, Ms. Kloet asked you some questions about the IEEE and you said it would be helpful if you could refer to the

manual.
A Yes.
Q Do you recall that?
A Yes.
Q I wanted to offer you the opportunity if you want to see anything in the manual.
A Only if you can get me directly to the independence clause which I think we may meet. I may have brought it to your attention yesterday. I can't remember. I don't think it's page 17, and I think you only printed to there.
Q May I approach, Your Honor?
THE COURT: Yes.
MR. PRESANT: I don't think I can get you there but I'll give you the opportunity to do it. If it's not important we will move on.
THE WITNESS: All right. Perhaps I can try and answer your next question while I look.
BY MR. PRESANT:
Q Ms. Kloet also asked you some questions about the analytical threshold and then some subsequent questions about drop-in and drop-out. You recall those?
A Yes.
Q She asked you a number of questions about drop-in/drop-out stutter. Is it fair to say those are all biological phenomena?
A They are all biological phenomena. They are not peculiar

to STRmix and they have been in existence since the inception of DNA analysis.

Q Do you think when you're building a model that reflects the existence of those phenomena it's necessary to have training in forensic analysis or forensic biology in order to be able to understand how they should properly be modeled?

A I think it's advantageous, yes.

Q Now, you also testified I believe that the analytical threshold is not an issue for STRmix.

A It's not an issue peculiar to STRmix. We have had analytical thresholds in the United States again since the inception. And they are highly accepted by courts all over America, and it's just standard usage.

Q And STRmix uses analytical thresholds.

A Yes.

Q Ms. Kloet asked you a series of questions about contamination being source of drop-in. Are there other causes of drop-in besides contamination?

A Drop-in we think are fragments of cells present in the laboratory often after the cleanup process. It's a very rare phenomenon, and it is a form of contamination.

Q But it's not necessarily mishandling of evidence contamination.

A No. It's good sensitivity. It's almost impossible to completely get rid of. And it's reasonably innocuous. It's

single alleles, and it would be almost inconceivable they could ever build to form any form of false inclusion.

Q Ms. Kloet asked you about the drop-in threshold of 400 I believe it was.

A Yes.

Q How did you come up with that number, 400, or did you come up with that number?

A I think MSP would have come up with it. And it tends to be a number higher than the highest drop-in they have seen.

Q That's because it's specific to the laboratory where the work is being conducted.

A Yes, yes.

Q Just to be clear, STRmix in fact models for the different biological phenomenon we just covered, drop-in, drop-out and stutter, those are all things that are incorporated into the mathematical model?

A Yes. This version models only back stutter, not forward stutter.

Q Ms. Kloet also asked you about the particular data in this case of an example of drop-out or that 12 allele I believe that dropped out, right?

A Well, it dropped out if it's from Mr. Gissantaner, and we shouldn't make the assumption it is from Mr. Gissantaner. That's the question, not the conclusion.

Q But the likelihood ratio models both of those

possibilities, that it is from him and that it is not, correct?
A Absolutely. That's exactly what it does. It's weighed those against each other, and in their particular locus, the colony specific factor locus which we were speaking of, it's the rarity of the 9 allele that's driving that locus specific allele of 4.20.
Q She also asked you a series of questions about 4.2 being close to 1, and the product rule getting you to a number in the billions. And I believe you started an answer saying that wasn't a logical way to think about it. I just wanted to give you an opportunity to explain why that is.
A Sure. So 4.2 could be viewed as 3.2 bigger than 1, or over 4 times 1. And it's the multiplier I ask you to think of. So if I could lower your chance of developing cancer by a factor of 4, would you call that immaterial?
Q I wouldn't, no.
A Well, I would suggest you shouldn't. It's a massive factor.
Q And if you could do that 21 times that compounded on one another.
A If you could get a factor of 4 in 21 major diseases, I would take it.
Q Now, she also asked you about this diagnostic being 1.41.
A Yes.
Q You said a commonly used threshold is 1.2.

A Yes.
Q Is that a mandatory threshold?
A No, it is not. It's a soft threshold. Recall this is a convergence diagnostic, not a has it run properly diagnostic. So you can get the right answer from something with 1.41, and often more complex mixtures do have higher GRs.
Q And I think you started to say something along the lines of well it could have run for longer.
A Yeah.
Q What did you mean by that?
A You can set STRmix to run for longer.
Q So if someone were unhappy with that particular number, they could choose to run it again, run for longer and that diagnostic might get lower?
A It would get lower.
Q But that's a judgment call that whoever is running it has to make.
A We have done a fairly massive trial so we have picked on GRs that were of the sort of 1.4 and then run them for longer, and the GR does come down and the likelihood ratio seldom moves at all. So I would predict if you ran this for longer the GR would come down but the likelihood ratio would be similar.
Q Finally, Ms. Kloet showed you sentences out of context from that 2008 paper regarding likelihood ratios. I wanted to give you a chance to elaborate on that. Has your thinking on

likelihood ratios changed since 2008?

A No. They are not so out of context, to be fair to Ms. Kloet. I think I just accept that presenting likelihood ratios is difficult. The wording is very awkward. Her Honor has very correctly picked up on this. It is awkward. And it is an area which were we to go to trial we could give some proper attention to. I spend a lot of my time thinking about how to present this evidence.

Q But the bottom line opinion you have is that it is useful evidence that can be presented to a jury in court?

A Likelihood ratios are so much more powerful than other methods. We spoke earlier about this particular case and what represented certainty. If we had done this in the olden days with CPI, it would have been completely impossible to interpret this profile. This profile, the CPI is such a blunt tool it would have not been able to interpret this profile. We now have a powerful tool that can produce a reliable statistic for this type of profile. So we are obligated both in the interests of victims and falsely accused people to try and present this powerful evidence properly.

Q In the best way you can.

A In the best way we can in conjunction with the legal participants to do the best job we can.

MR. PRESANT: Thank you, Your Honor.

THE COURT: Ms. Kloet. Any recross?

MS. KLOET: Your Honor, I don't believe I have any recross. But I did want to address the issue of exhibits. The witness identified the six exhibits that I offered in the course of my cross-examination of him. Three of them have already been part of the record through the motions and responses filed. That's Exhibit J, H, K. The other three which is CC, G and F are not part of the record and I would move to admit those now seeing as they have been identified and discussed on the record.

THE COURT: Mr. Presant.

MR. PRESANT: What were those again?

MS. KLOET: CC, G and F. Those are the three that are not in the record and the other three are the EPGs and the STRmix reports.

MR. PRESANT: CC and G the government is fine with. I think F the foundation hasn't been laid yet. But I think the foundation, I think it will come in through another witness. But he didn't create it so I don't think he can testify that it was the one used in this case. If she wants to admit it simply --

THE COURT: Did you say F?

MR. PRESANT: F I believe, yes. I mean he can I think if it's just being admitted right now for the purpose of saying this was, this looks like something that would be output by STRmix, that's fine. But another witness is going to have to

testify to what it is. And it's marked in the Government's Exhibit book too. So for all intents and purposes it can come in. I just want to be clear that he hasn't testified this is what was done in this case.

THE COURT: CC and G then are admitted without objection. And F I don't think it is properly offered at this time. So let's wait until we get the proper witness in front of us. Okay?

MS. KLOET: Okay. Thank you.

THE COURT: Before I excuse you, Dr. Buckleton, I want to, first of all, take a quick look at my notes.

In terms of, well, first of all, under Daubert it's the duty of the Court to determine the reliability of the evidence. And if I understand, again what limited reading I've been able to do in this area, there are for purposes of addressing the jury, we have to talk about how was the sample collected, how was it handled after it was collected, and then how was it analyzed in the lab and using the software that's in question here, STRmix. Do you agree with all of that?

THE WITNESS: I'm not certainly not going to debate with Your Honor the legal interpretation of Daubert.

THE COURT: Okay. Well but in interpreting the results that are going -- assuming the evidence is deemed to be admissible and relevant, and not unduly prejudicial, in interpreting the results that are going to be presented, those

areas have to be independently addressed by the testimony and the evidence.

THE WITNESS: Yes.

THE COURT: Okay. And they should lead to a question or an answer to the question of are these results reliable based on existing scientific principles.

THE WITNESS: Yes.

THE COURT: And your testimony is that they are. That these, that when we look at, at least from your perspective, when we look at what was done in this case, you don't have any opinion with regard to the collection, you have no opinion with regard to the handling, but you do have an opinion as to the analysis that was done leading to the probabilistic figure of 49 million to 1.

THE WITNESS: Yes.

THE COURT: That is exclusively where your areas of expertise come into play in this case.

THE WITNESS: Yes.

THE COURT: And you are satisfied to the best of your scientific knowledge that these results are reliable.

THE WITNESS: Yes.

THE COURT: Okay. Now, the one thing that I am still not entirely clear on based on what I've read is this idea of the interpretation of the results. Could you please speak to that?

THE WITNESS: Yes. So I have in the break been able to clarify that the FBI have stopped doing source attribution, so they do not say this DNA came from him anymore. When they did do it, their number was 700 billion. That was the number I said I would try and find for you. And even then they produced the number first and then, so they have always tried to produce a statistic. I won't go into any depths but this case would have been very hard to interpret without the modern software.

But if we come to the concept of what does 49 million mean. 49 million should be combined with the other evidence, and that is not my job to do. I do not know whether there is any other evidence or whether there's evidence that points either towards or away from Mr. Gissantaner. I have no idea.

If the other evidence were neutral, then you multiply that by 49 million and the odds are now 49 million to 1 that this DNA came from Mr. Gissantaner.

But if the odds are not neutral, if there is massive evidence in his favor, or if he's plucked randomly from the street as being one in a million people, then you should start from a position of 1 in a million and multiply that by 49 million, and that comes to simply 49 to 1.

I'm certainly not going to tell you how to interpret the other evidence. I don't even know what that is. But you must form that view of what the other evidence is, combine it with the 49 million, and then you get the overall value.

THE COURT: But my question is this. Again, based on the reading that I've done suggests that there is a certain amount of interpretation of the data that is done when these results are produced; not, not interpretation for the jury to make or for me to make, but the people who produce these results do some individual interpretation that may be subjective.

THE WITNESS: Yes. So that it is subjective. So it's utterly routine. It's been happening for years. And it's not to do with STRmix. The subjective elements are some artifact management, so that is removal of spikes, and pull up, and forward stutter. These are artifacts of the PCR in imaging process, and this is standard stuff. It's been happening for years. And we all follow the SWGDAM guidelines on how to do it.

Then the other one is assigning a number of contributors. When you assign a number of contributors you are either correct or incorrect. If you're correct, then you are correct. If you are incorrect, you produce a conservative number. So there is nothing bad that can happen by getting the number of contributors wrong. You simply get additional conservativeness.

THE COURT: Okay. I think I understand. And that's all the questions I have. Thank you, Doctor. You may step down.

THE WITNESS: Thank you, Ma'am.

THE COURT: Mr. Presant.

MR. PRESANT: The government calls Jeffrey Nye.

JEFFREY NYE, GOVERNMENT WITNESS, WAS DULY SWORN

THE LAW CLERK: Please be seated. And state your full name for the record, spell your last name.

THE WITNESS: Full name is Jeffrey Burn Nye. Last name is spelled N-Y-E.

DIRECT EXAMINATION

BY MR. PRESANT:

Q Mr. Nye, where do you work?

A I work for the Michigan State Police Forensic Science Division.

Q What's your current position with the Michigan State Police?

A My current position is the assistant director of the Forensic Science Division.

Q Let me bring up Exhibit 2, please, Ms. Miller. Do you recognize 2?

A Yes, I do.

Q What is it?

A My curriculum vitae.

MR. PRESANT: Government moves to admit 2, Your Honor.

THE COURT: Ms. Kloet, any objection to Mr. Nye's CV?

MS. KLOET: No, Your Honor. I have seen it.

THE COURT: It's admitted.

BY MR. PRESANT:

Q As we follow along here, Mr. Nye, would you please describe for the Court your educational background?

A Certainly. I have a Bachelor of Science Degree in Biochemistry from Michigan State University. I have a Master's Degree in Crop and Soil Sciences from Michigan State University. And then I also have a specialization, another Master's in Environmental Toxicology, and then I have an additional course in molecular biology from Lansing Community College.

Q What did you do after you finished your graduate work in 1995?

A After I finished my graduate work I worked for a company called Michigan Biotechnology Institute. Essentially that organization would take laboratory research from universities and scale them up and if we could prove that they were viable corporations or companies around that, then we would spin them out as private companies. So research.

Q And then here at the bottom we pick up in 1996.

A Yes.

Q And that's when you moved to the laboratory at the state police?

A That's correct.

Q And have you been there ever since in various capacities?

A I have.
Q What are the various positions you've held with the Michigan State Police?
A I started out as a forensic scientist conducting DNA testing at the Lansing laboratory. I did that for a number of years, then I promoted to a unit supervisor where I oversaw other forensic DNA examiners. Did that for a number of years, then I promoted to a DNA technical leader position which is sort of a quality assurance manager for that discipline with respect to training and validations and things of that nature. I did that for approximately nine to nine and a half years. And then about two years ago I promoted to the assistant director for the Forensic Science Division.
Q What are your responsibilities as the assistant director for forensic science?
A So as I described my position before that, it was as a technical leader for the DNA discipline. Forensics within the Michigan State Police there is a number of different disciplines: Firearms, controlled substances, all the different forensic disciplines. Each discipline has a technical leader. And all those technical leaders report to me as the assistant division director and quality assurance manager.
Q What's involved in being the quality assurance manager?
A Quite a bit. So we as a laboratory system are accredited

to certain standards and guidelines. And it's to ensure that our operation, which is eight laboratories and approximately 260 scientists and technicians, comply with standards and we meet those accreditation standards, and then also to establish policies and procedures as far as how our operation operates and complies with those standards.

Q The Court has heard testimony today about a forensic governing body called SWGDAM. Are you familiar with SWGDAM?

A I am.

Q How are you familiar with it?

A I am a member of SWGDAM.

Q Would you describe for the Court what the structure is of SWGDAM and what its role in the forensic science community is.

A So by congressional act, SWGDAM is sponsored by the Federal Bureau of Investigation. And they convene twice per year for about a week at a time, January and July of each year, a variety of experts from around North America and then some international guests as well. The spectrum basically runs the gamut of practitioners like myself, researchers, a whole host of individuals basically to get together to address current concerns within forensic DNA and biology testing to help develop guidelines and standards specific to that discipline.

Q How many people sit on SWGDAM?

A Roughly I would say between 40 and 50 individuals.

Q And that's primarily in the United States but also from the

whole world?

A Correct. There are also other invited guests. There is a distinction between being an invited guest and being a member. Invited guests are able to contribute to the work product, but the members actually vote on accepting the final work product.

Q So if there are only 50 total then every state laboratory doesn't have someone who is a member of SWGDAM, correct?

A That's correct. They work very hard to make sure they represent the full range of laboratories within the United States. So there are federal laboratories such as the FBI or the Army crime lab or ATF, there is state laboratories that are represented, then there is also local laboratories geographically distributed as well as size distribution as far as how big of an operation that they have.

Q How does SWGDAM decide who should be invited as a member?

A Probably by recommendation. And I believe that they have a separate committee that votes on those recommendations as to extend an offer of an invitation for an invited guest or to offer somebody as being a member.

Q How long have you been a member?

A Goodness. Eight or nine years, I'm going to guess.

Q And you're a voting member so you vote on standards?

A I am.

Q Bring up Exhibit 19, please. Do you recognize 19?

A I do.

Q What is it?
A It's the "Guidelines for Validating Probabilistic Genotyping Systems" that was developed by SWGDAM.
Q Were you on SWGDAM when these were developed?
A Yes.
Q Were you involved in the subcommittee that developed them?
A I was not.
Q Did you vote on whether to adopt these guidelines?
A I did.
Q Did you vote in favor of them?
A I did.
MR. PRESANT: Your Honor, the government moves to admit 19.
THE COURT: Ms. Kloet.
MS. KLOET: No objection, Your Honor.
THE COURT: It's admitted.
BY MR. PRESANT:
Q Now, there's been some testimony, though we haven't pinned down an exact date of when these were adopted. Can you zoom in on the portion I just highlighted? Reviewing the part that I've just blown up, do you recognize a date that these were adopted?
A I do.
Q And when does it say they were approved for posting on the SWGDAM website?

A June 15th of 2015.

Q And so would you just summarize for the Court the significance of these guidelines as they relate to the use of probabilistic genotyping in the United States?

A Yes. So there was some discussion on the previous witness about what guidelines actually represent, and I would indicate the guidelines are basically recommendations. They are recommendations that if you for this particular guideline, if you want to validate a probabilistic genotyping software application, that you should follow these guidelines to properly validate the system. Dr. Buckleton had mentioned that there's a strong incentive to follow these guidelines as more of like standards, but there's not really a standard associated with these, it's just a strong recommendation that you follow these.

THE COURT: What is the incentive?

THE WITNESS: Just consensus in the community that these are -- there's not an incentive like you're going to get an audit finding when we undergo audits. There's not a monetary incentive. There's not an incentive to a particular government resources or anything like that. It's just a, it's driven in everyone that we do quality work. If you want to subscribe to doing that quality work, then it's highly recommended that you follow these guidelines.

THE COURT: And if you say that you follow them, does

anybody determine whether you do follow them or have followed them?

THE WITNESS: So throughout accreditation, the accreditation process, there is a number of different standards that we follow. So when I mean standards, they are shalls. You shall do these things. So one organization that we are accredited to is an acronym called ANAB, it's the National Accrediting Board, and there's a host of standards associated with that particular organization and one is related to validation. There is another set of standards that SWGDAM also produces called the Quality Assurance Standards for Forensic DNA Testing Laboratories. So it's exactly how it's described, these are quality assurance standards. If you're a forensic DNA testing laboratory in the United States, you need to follow these. And that, there's a, there's an expectation that you follow those because it gives you access to other government resources such as data bases and other resources that the government maintains. So if you don't follow those standards, there's actually ramifications for it. And those standards also have directions specifically on how you conduct a validation.

In a very general sense, the validation needs to cover certain criteria. A little bit was discussed this morning about precision, and reliability; they have to cover the range of samples that you would normally cover in the laboratory

during testing, you know, there's a whole host of standards.

And so your question is how does somebody check to make sure you follow those standards. We go through extensive audits every single year. For the ANAB standards, we are audited once every four years for what they call a full assessment. And that assessment for our system, for the Michigan State Police, can be as many as 75 independent assessors coming in from around the country and around the world and evaluating our practices and policies and procedures against these standards.

That occurs every four years. And then every year, so year 2, year 3, in between those there is also what they call surveillance visits where smaller contingency of assessors will come in to evaluate to make sure you're continuing your accreditation activities to make sure it's a part of your culture of your organization.

Aside from that accreditation process is also the quality assurance standards for forensic DNA testing and that's an every other year assessment.

So every other year we bring in, again, external assessors to evaluate specifically the DNA program to ensure that we are complying with the standards that are in place, and then on the opposite years, we are required to do our own internal review of our guidelines with those standards.

So continually there's some sort of an assessment

going on with the laboratory.

THE COURT: Thank you.

BY MR. PRESANT:

Q With all those qualifications is it fair to say that to the extent that there is a document governing forensic DNA analysis, specifically probabilistic genotyping, and the validation of those systems in the United States, that this Exhibit 19 is that governing document?

A Yes.

Q Let's go back to Exhibit 2, please. Let's go to page 2. I have highlighted, Mr. Nye, just a portion of the trainings portion of your CV. Right?

A Correct.

Q And there are a couple on here that say STRmix training, right?

A Yes.

Q And so I want to ask you before you talk specifically about those, what training in general and these in particular you have had on STRmix.

A So before I get into STRmix, just probabilistic genotyping in general, there's what I would probably call less formal training that occurs. So when I go to a general conference related to forensic DNA testing, there is often presentations by experts in the field about probabilistic genotyping, you know, could have been somebody related to STRmix, or it could

be related to a competitor to just kind of get a general understanding of the system. And then more formally, after we purchased STRmix it included formal training. Dr. Buckleton had mentioned it was four days. And, yes, it felt like more than a week. And that occurred in the spring of 2015 was our initial training event that included I think it was roughly, I forget if it was 10 or 20 people, it was in that class, and then we conducted a second class for an additional number of examiners in the summer of 2016.

Q Let's go to page 4. You have been published in several peer review journals on topics including DNA analysis, is that right?

A Yes.

Q Now, Mr. Nye, was there ever a time where you were cited for approving a report without proper technical review?

A Yes, there was.

Q Would you describe for the Court what happened in that instance and approximately when that occurred?

A I'm sorry, the last part.

Q When it occurred approximately.

A So we talked a little bit about accreditation, and one of the requirements as well as our own internal policies and procedures is that before a laboratory result is published that it undergo a review by a peer. They call them technical and administrative reviews. And approximately six years ago I

inadvertently published three to four reports without the proper reviews being done. And this was the result -- I used the word inadvertently because it was a result of a software application that we use and under my specific profile with that software application, there were some additional privileges in that software application to issue reports that I didn't fully understand.

Q And you received some sort of citation for it?

A I did, yes.

Q Has the Michigan State Police promoted you since then?

A Yes. I have, so at that particular time I was the DNA technical leader and I have since been promoted to the assistant division director responsible for quality assurance. And then I've also received a number of different awards from the department as well.

Q Does that incident have any bearing on the adoption of STRmix by the Michigan State Police that you're here to testify about today?

A Not whatsoever.

MR. PRESANT: Your Honor, the government offers Mr. Nye as an expert in forensic science and DNA analysis.

THE COURT: Ms. Kloet.

MS. KLOET: Just a moment, Your Honor. No objection, Your Honor.

THE COURT: Thank you.

BY MR. PRESANT:

Q Mr. Nye, the Court had some questions this morning just about the overall process kind of soup to nuts from the time the sample comes into the laboratory until the time that the laboratory issues its reports about how forensic DNA analysis is conducted, specifically in a DNA mixture case like this one. I thought it would be helpful --

THE COURT: You're talking awfully fast.

MR. PRESANT: I appreciate you pointing that out. I often don't even notice, Your Honor.

It would be helpful I think if you started out by summarizing for the Court the process that the Michigan State Police use for that pipeline.

THE WITNESS: Yes.

THE COURT: Before you do that, and sort of in conjunction with that, could you pretend that I'm a jury of one, could you tell me, go through the process of DNA genotyping, if that's what it's called, before this probabilistic method was developed, and after it was developed. Can you lay those two things side by side so that I can understand the difference?

THE WITNESS: It's going to be a long explanation.

THE COURT: That's okay.

THE WITNESS: Please refrain me if I go into too much detail.

THE COURT: Okay.

THE WITNESS: So previous to probabilistic genotyping and now in probabilistic genotyping, everything in the laboratory is the same. Everything is. So when I say everything, I'm talking about the DNA extraction method, the purification process, the DNA quantitation method to determine how much DNA we have, the DNA amplification, the actual separation and visualization of the different DNA fragments are all exactly the same.

The difference from let's say three years ago to today is just the tools that are used to assist the examiners in interpreting mixed DNA results. And I'm, I like Mr. Presant tend to speak relatively quickly and I'm choosing my words very carefully because I want to stress that the tools that are used may be different, but the approach and the modeling and the interpretation itself is very, very much the same.

I want to point out that STRmix is really just a tool applying methods and modeling that was previously used manually.

The significant difference with STRmix or any probabilistic genotyping system is that because it's computer based it allows the user to incorporate more data because the amount of data that was available exceeded what the average human could actually efficiently use and analyze.

THE COURT: So is it true that before the development

of this probabilistic genotyping with software like STRmix, the determination of DNA matches and so forth was essentially manual, is that correct?

THE WITNESS: Yes.

THE COURT: There was no machine intervention that either produced results or interpreted results, is that fair?

THE WITNESS: There may be some tools just simply like Excel spreadsheets and some different things of that nature to just help with the math. The math is very simple to do. It's just more than what I would care to do and sit down and do it with a pen and a piece of paper and a calculator.

But the approach is very, very much the same. Dr. Buckleton had talked about certain artifacts. Stutter, we heard this morning about forward stutter and back stutter, and pull up, and all these different things. Those are still things that have always been there, and they have always been used in the interpretation, it's just now that there's actually a method that can, there is a tool that can be used to assist the DNA examiner to do it.

I might try and offer an example that might illustrate it a little bit better. So three years ago and prior for many years when you had a DNA mixture, the statistical approach that was very common in the United States was something called a CPI. Combined probability of inclusion. And the approach in that method is to only look at the DNA types that are present.

So Dr. Buckleton had testified that there are DNA types. There is a DNA size, and then the height of the peak which corresponds to the relative amount of DNA that's in that mixture.

And so the CPI only looks at the presence or absence of a particular DNA type. It does not use the size of the fragment or the height or relative amount of each fragment.

THE COURT: So that the conclusion that would be reached pre three years ago would be the defendant's DNA was found on this gun, period. Or it was not found on this gun, period. Right?

THE WITNESS: It entirely depends on the data that's available and present to interpret. And I think that that was one thing I wanted to clear up from this morning.

I think there was some comment that we used 21 genetic markers. That number has varied over the years. The more data that you have, the more significant the statistic is. Regardless of which type of statistic you're generating, whether it's a likelihood ratio or whether it's a CPI or any other type of statistic, the more data that you can use, the more significant that interpretation, the statistical interpretation becomes.

And so your question is entirely based upon how much DNA data that you have to work with. So if you're targeting 21 genetic markers, let's say, and you have a very limited amount

of DNA that you're working with and you get data, useable interpretable data at let's say 3 or 4 genetic markers of the 21 that you attempted, then your conclusion is going to be quite different in that and the significance of it. You may not be able to say his DNA is on the item or not on the item. It's more of the significance is going to be quite a bit different. Versus if you have a full range of genetic data at the 21 markers that you look at, then the significance of that statistic and the strength that you can apply to your statements related to that are quite a bit different.

So there's quite a bit of discussion around the 40 or 50 million to 1 number that we have on this particular case. I have testified as low as 10 or 1 as far as a likelihood ratio. And that does not have the same significance. In your words, I would not be able to go in and say with any certainty that an individual's DNA was on a particular item with a likelihood ratio of that level.

Versus I have also testified to a likelihood ratio of a duodecillion to one ratio. My comments to the jury would be significantly different between those two situations. So here we have a likelihood ratio in the 40 to 50 million to 1, and I agree with Dr. Buckleton that's very strong evidence. And I also want to add because I think many of your questions this morning were how this is going to relate to the jury. And I know Mr. Presant's questioning, a lot of questioning will get

into this later.

MR. PRESANT: We can do it now. That's fine with me.

THE WITNESS: So a likelihood ratio is just a different way to relate to the jury the strength of the evidence. And although Dr. Buckleton is an excellent instructor and an excellent researcher, myself as a practitioner I have to hone my craft, so to speak, about how I communicate to the jury. Because that's the end result of my work. I'm as much a communicator as I am a scientist. And we hire people and we work with them to develop how they interact and how they communicate to the jury. So a likelihood ratio is just a different way to relay that information to the jury and that is something that we develop over a period of time. So we are essentially two years into this process of moving from a frequency based statistic to a likelihood ratio based statistic, and we have been very cautious in how many of our scientists that we are training in this method so that we can sort of craft and mold how we present this information to the jury.

And your question about likelihood ratios, one way that you could present what a likelihood ratio is to a jury is to try and present it on a topic that they can actually grasp. DNA is hard for a lot of people, and so I like to bring in the idea of weather. Everybody pays attention to the weather. When I get up in the morning I look outside and it dictates

what kind of day I'm going to have. And so one of the things that's actually in the training program that Dr. Buckleton runs through with new STRmix users is an example where if you look at the weather report and there was an 80 percent chance that it was going to rain today. The opposite of that is going to be there is 20 percent chance that it's not going to rain today. So that's basically your two competing hypotheses. Your prosecution hypothesis is going to rain, your defense hypothesis, it's not going to rain. So the ratio of 80 to 20 is four. And the way that it would be worded, and hopefully in front of Dr. Buckleton I can somewhat closely get it right, is that given the evidence, it is four times more likely to rain than not rain. And that's a very easy understandable way for a jury to understand a likelihood ratio.

And when you apply that to forensic DNA, it becomes a bit more information to digest, but basically you give what your hypothesis is for the prosecution, which might be worded something like I have a mixture of three people, the prosecution hypothesis is Mr. Gissantaner, plus 2 unknown individuals, and then the defense hypothesis would be it's three unknown individuals, not Mr. Gissantaner. And it's the ratio between those two which turns out to be approximately 40 or 50 million to 1 in favor of the prosecution.

So that's -- you kind of have to build up to it without just putting out a number and saying it's a likelihood

ratio. You need to build into it a little bit.

THE COURT: Thank you.

BY MR. PRESANT:

Q We need to follow up on what you were just talking about.

How long has Michigan State Police been using likelihood ratios in one form or another in its testimony?

A So there are a number of different statistical methods to use and it really depends on which situation you're dealing with. And what I mean by situation, that means the type of evidence or the type of crime that you're working with. So if you have in its simplest terms a single source DNA profile, the very common method is the random match probability. That's a frequency. How often would you expect to see a DNA profile in a population. Then you get into a mixture where you have more than one person in a mixture. We currently use probabilistic genotyping, STRmix as that. And then another situation would be is if you're dealing with a criminal paternity case. So if there is a sexual assault that results in a child or a fetus, there's a criminal paternity case and that child or that fetus becomes the evidence which is essentially the joining of the victim and the potential assailant. And it's very common for probably I would say the last ten plus years that we use a likelihood ratio for criminal paternity cases, and then for -- we don't do them -- but for noncriminal paternity cases, just parenthood, likelihood ratios are very common in the United

States, in Michigan specifically.

Q Are there any other statistics that are frequently used in DNA cases besides likelihood ratio or the random match probability?

A CPI would be one. I referenced that previously.

Q That's the combined probability of inclusion?

A Yes. And then the antithesis of that it would be the combined probability of exclusion. Pretty much the same statistic. And then there are some laboratories in the United States that used likelihood ratios for even non paternity type of cases, and it's certainly much more prevalent in Europe and other areas of the world. It's just been delayed in getting into the United States.

Q Would you agree that one advantage of a likelihood ratio is that it incorporates uncertainty by using two competing hypotheses?

A It does.

Q All right. So you testified that everything really up to and including the generation of the electropherogram by the genetic analyzer is pretty much the same whether it's a probabilistic genotyping case or not, right?

A Correct, yes.

Q Does that process change at all whether the source of the genetic sample is from blood or from skin cells or from another type of cell?

A The process?
Q Yeah, the extraction process, the whole laboratory process, does it change based on the source of the DNA?
A It does in one particular instance. If it's an item of evidence that results from a sexual assault, whether it's semen present, the extraction process is slightly different. But otherwise, everything is done the same.
Q Now two steps of the process are PCR, or, and then second, capillary electrophoresis, right?
A Yes.
Q PCR is the multiplication of the DNA that's present so it can be viewed, right?
A Yes.
Q And capillary electrophoresis is the sorting of the different fragments of DNA, right?
A It is.
Q Then those are imaged by the fluorescent tag we heard testimony about, right?
A That's correct. So during the amplification process each DNA fragment is tagged with a fluorescent label, and then as they are sorted in the capillary, they pass a detector, a camera, from which they could detect the fluorescence of each fragment as it goes by, then that detection is represented visually with an electropherogram or an e-gram as was described earlier today.

Q Now, in the defendant's Daubert motion, he makes some passing arguments that there should be some question about the reliability of those two steps, PCR, and capillary electrophoresis because before those things existed DNA analysis could be done, right?

A It could be done, yes.

Q What's your view on how established PCR and capillary electrophoresis are both in forensic DNA analysis specifically, and all DNA lab work generally?

A So I have worked for the state police for 22 years and when I first came into the department PCR was not used. It was a different technology at that particular time called RFLP, restriction fragment length polymorphism. P-O-L-Y-M-O-R-P-H-I-S-M. And it was approximately 1998 when we switched from the RFLP technology to a PCR based analysis method. So we are at this point in time 20 years into using PCR as a method. Really the only thing that's changed in 20 years with respect to PCR is the number of genetic markers that we look at, and then there's been some increases in sensitivity of the systems that we use.

Approximately in 1998 when we first started using PCR we used a gel based system to separate out the DNA fragments, and then approximately in 2000 I believe it was we switched from a gel based system to a capillary electrophoresis. Which for all intents and purposes is the same as a gel based system

only it's carried out in a small capillary. The system is really the same. And then it's been ever since 1998 to 2000 it's been all capillary electrophoresis since then.

That's our system in forensic DNA analysis. The use of PCR and the use of capillary electrophoresis is throughout vast portions of research, medical applications, parentage, ancestry, a whole host of other applications all use the same technology that we are using today.

BY MR. PRESANT:

Q And how does PCR compare to its predecessor, RFLP?

A So RFLP technology I think at the time we were looking at five genetic markers. And those genetic markers had more statistical power per marker. They were more discriminating per genetic marker than what the STRs are. But it also required a very large sample size. There was some discussion about DNA quantitation this morning. And we can get DNA results from a very, very, very small sample in a nanogram or picogram range, but in RFLP technology we would require as much as 500 nanograms of DNA to get a result at that particular time.

So the sample size was orders of magnitude higher. We needed more in order to get a result. But then the other significant portion is that the RFLP technology required DNA that was in very good condition, had not been degraded, inhibitors, different things like that. It really required a

very good sample. And the PCR technology with capillary electrophoresis allows us smaller sample sizes; it allows us to deal with samples that have been degraded or insulted in some way. And there is also sort of an efficiency standpoint, it's less hands on, it's more automated. Because of the increase in sensitivity there's more samples coming to the laboratory and we can get more work done when automation is going on.

BY MR. PRESANT:

Q Do I take from your answer then that capillary electrophoresis is also an improvement upon gel based separation of fragments?

A Significantly, yes.

THE COURT: Mr. Presant, I'm thinking about taking a break at about 2:30. Because I would like to go as long this afternoon as we can. So if you are getting ready to switch topics, or if you want to wrap up a particular topic, now is the time to think about it.

MR. PRESANT: I appreciate that, Your Honor. I'm happy to have the Court interrupt me whenever the Court is ready for a break.

THE COURT: Okay.

MR. PRESANT: Mr. Nye, so now I would like to ask you about the decision to begin to use probabilistic genotyping in general and to choose STRmix in particular because you were the person who made that decision, correct, for the Michigan State

Police.

THE WITNESS: I was, yes.

BY MR. PRESANT:

Q And when you looked at where you were kind of in the progression of forensic DNA science over time and you look at the market of options of probabilistic genotyping software, what led you to make the consumer decision that you ended up making?

A Certainly. So my position as the DNA technical leader was to not only train individuals but seek out new technologies and make sure that we are current on everything we can be with what's available in the community, and certainly probabilistic genotyping offered an opportunity for us to improve our service. And specifically, the combined probability of inclusion statistic that was available was not, we did not use the entire set of data that was available to us, and it also resulted in a large number of analyses that would end in an inconclusive result because of that inability to use the full gamut of data that was available to us.

So there was an interest in furthering it because we knew, as did the community in general, that we weren't using the full amount of data that was available to us. So probabilistic genotyping would have been an advancement, another tool available to us to use our expertise in an area that we couldn't currently use it within. So then I started

looking at the different opportunities, the different applications that were available, and there were two commercial ones. So basically the first decision that I looked at was do I use a commercial off the shelf product, or do I use what they would call a free ware system which is basically just a free system, and the question is how much support are you going to get. Are they going to continue to develop the product, make further advancements with the product, were they going to be able to help train, do all those different things. And when you're looking at a free software system, those resources are not available to you.

So if you make the first decision that you're going to look at a commercial off the shelf system that has those resources available, there is essentially two. One would be TrueAllele, which is an American company, and the second being the STRmix which is from New Zealand. So then you start and evaluate the application itself, what type of resources, where the rest of the community is going, the educational opportunities with understanding and learning because it's not enough to just have a piece of software where you can click a button and get an answer out the other side. You actually as a practitioner need to understand enough of the method to be able to convey that to the jury so that they can make the decisions that they need to make. And at the end of the day we felt strongly that STRmix was the system that we wanted and that's

the process that we went.

Now, I want to back up for one second because I skipped over an important part.

So the Michigan State Police Forensic Science Division is made up of approximately 90 DNA examiners. It's a pretty good number. And with CPI or other more manual interpretations for mixtures, I believe Dr. Buckleton mentioned it, and I'm going to mention it again, is that when you put human intervention into something you're going to get some variability on how mixtures are interpreted. And going with an additional tool such as STRmix allowed us to standardize that approach a little bit. There is still some subjectivity, I think it's significantly less than prior to probabilistic genotyping, but from a management perspective it allowed us to standardize the approach and the level of service that we are providing to our customers, if that makes sense.

THE COURT: I think we are going to take our break here, Mr. Presant. We will come back in about 20 minutes.

MR. PRESANT: Thank you, Your Honor.

THE LAW CLERK: All rise. Court is in recess.

(Recess taken, 2:29 p.m.; Resume Proceedings, 2:55 p.m.)

THE LAW CLERK: All rise. Court is back in session. Please be seated.

THE COURT: Mr. Presant.

MR. PRESANT: Thank you, Your Honor. Mr. Nye, before the break you had testified about your lab's decision to use STRmix.

THE WITNESS: Yes.

BY MR. PRESANT:

Q And my question for you is when did the lab actually start using STRmix?

A In February of 2016.

Q And you testified earlier that only certain analysts were trained to use it?

A Correct.

Q What sort of training program did you develop for the people selected, the forensic scientists selected to use STRmix?

A So we required that they would have received the four-day training from Dr. Buckleton's group, as well as additional literature readings in our laboratory, and then a practice set of samples before we deemed them to have sufficient training. And then before they actually started using STRmix in case work there's a required competency test that I developed that they had to pass, and then once they started using STRmix in case work then they went through our normal proficiency testing program.

THE COURT: Can I ask one question here, Mr. Nye? When you say that a practice set of samples, does that mean a

set of samples where the DNA contributor is known and they run it, run the sample to determine if they can agree with what you already know, is that it?

THE WITNESS: Yes, correct.

THE COURT: Okay.

BY MR. PRESANT:

Q So what's involved in the ongoing proficiency testing?

A So one of the accreditation standards is that everybody specifically in DNA do external proficiency tests twice per year. An external proficiency test is a mock case that's administered by a private company which I don't even know the answer to and each person has to complete that mock case proficiency test as though they would a regular case and then report the results to that external organization and they get a response back at some point in the future as to whether they successfully completed the proficiency test or not.

Q The Court has heard testimony today about the importance of internal validation before a forensic laboratory begins to use STRmix. In addition to training these analysts, did the Michigan State Police undertake an internal validation of STRmix?

A Yes, we did.

Q Prior to the internal validation of STRmix, in addition to the -- or strike that. Prior to the internal validation of STRmix, had you been involved with the internal validation of

other forensic tools?

A Yes.

Q So you were familiar with the process for conducting internal validation?

A Yes, depending on what you're validating the process may be slightly different. But certainly, yes, that was a major portion of my responsibility as a DNA technical leader was to either complete it myself or to certainly oversee it.

Q And what's the theoretical purpose behind internal validation?

A So in my view there is really two different types of validation that occurs. There is a developmental validation which is typically done outside of the laboratory sometimes by the researcher, in this case ESR or other entities, commercial entities sometimes. That's a developmental validation and there is a requirement for an internal validation to make sure that the product or method works as expected in your hands, in your specific situation within your laboratory. And so we conducted an internal validation of STRmix.

Q Different laboratories have different conditions and instrumentation and procedures and so it's important to make sure the tool works in that particular environment.

A That's correct, yes.

Q Go to Exhibit 10, please. Do you recognize Exhibit 10?

A I do.

Q What is it?
A It is a summary of our validation, our internal validation of STRmix in cooperation with our amplification chemistry which is PowerPlex Fusion that Dr. Buckleton had mentioned this morning.
Q Page 2, please. That's your signature at the bottom?
A It is.
Q It's dated February 22, 2016?
A It is.
MR. PRESANT: Government moves 10, Your Honor.
MS. KLOET: No objection.
THE COURT: It's entered.
BY MR. PRESANT:
Q So on this page 2 there's a paragraph here right underneath the table, is that right?
A Yes.
Q And what's that paragraph basically saying?
A So that paragraph is my declaration that the internal validation was completed and enclosed within the document that follows, meets as my determination as the DNA technical leader that meets the additional guidelines that were published by SWGDAM for validating Probabilistic Genotyping Systems.
Q Page 3, please. Now, you have some background information in this document, it's what, like a 40 or 50-page document, right?

A It is.
Q So we don't need to go page by page, but I would like to direct your attention to page -- start with page 5. You discuss the concept of threshold here that's been talked about earlier, is that right?
A Yes.
Q And then on the next page stutter.
A Correct.
Q Page 11, please. There are drop-in parameters here, right?
A Yes.
Q And then you set the drop-in cap right here is 400, the Court has heard testimony about that earlier today too, right?
A That's correct, yes.
Q The drop-in frequency, there is also testimony about that figure, that was determined by this document during the internal validation?
A It was.
Q You discuss a concept called saturation here, is that right?
A Yes.
Q Let's go to page 20, please. What's being described here in this table 4, adjudicated cases?
A So the period of time that I was summarizing or completing the validation, the internal validation on STRmix within our system was around about the same time that the Daubert hearing

was ongoing in Muskegon County. That was discussed earlier this morning. And although not required, there was some testimony from a defense expert in that particular hearing that it's very difficult to create samples in the laboratory much like you just explained a moment ago where we know the result because we created the samples, but that's not always the best reflection of a sample type that we would see from a crime scene. So in other words --

THE COURT: Is that correct?

THE WITNESS: I think it has some merit. The difficulty is that if you don't create the sample yourself, you really don't know what the ground truth of the sample is. So I think there is benefit from doing both. The samples that are created in the laboratory are typically coming from very nice samples, DNA that's extracted from blood, typically from some of your own employees, and we create mixtures of different proportions, different contributor numbers, we have a lot more control over that situation. But the advantage of looking at actual evidence from crime scenes is that they have gone through a bit more I'll use the word insults. Maybe they are dirty, they have been degraded, they are on different substrates. Maybe a cotton swab versus some blue jeans, versus a cigarette butt. I mean they represent a better range of things.

So we did both. Although not required. And this

table in the exhibit shows a number of cases, and we used adjudicated cases with the understanding that to the best of our ability if the adjudication has been complete that that's the closest we can get to ground truth of whether somebody contributed to a sample or not. And we ran those adjudicated cases through STRmix and then compared the STRmix results to the adjudication and/or the statistical value that was provided previously to understand whether it's giving us -- it being STRmix -- is giving us information that is supportive of what was developed previously.

BY MR. PRESANT:

Q What was the conclusion from that analysis?

A So the overall conclusion is that if we were able to provide a statistical estimate with our standard, whether it be CPI or random match probability, STRmix was able to support that conclusion. Although it's a different type of statistical approach, it's a likelihood ratio as opposed to a frequency, it was still supportive of the end conclusion. The other thing that we could conclude out of it was that there was a significant number of samples or cases where we were unable to provide a result, a conclusion, because the statistics and the modeling, it was that manual approach, whereas if we ran it through STRmix we were able to provide some supportive information either through exclusions or inclusions. It would go both directions.

Q You tested firearms as part of this historical analysis?
A I believe so, yes.
Q So if we look at this one right here, the third entry of swabs from rifle, right?
A That's correct.
Q The original statistical value reported was 1.2 septillion?
A Correct.
Q And what was STRmix's conclusion upon reanalysis of that sample?
A So on that particular sample I just want to point out that was a two-person mixture and the 1.2 septillion represents AA frequency. So the proper way to word that particular type of statistic is that if you were to go out into the population at random, select people at random, you would expect to find a person that would match a particular donor within that sample 1 out of every 1.2 septillion individuals. So very, very significant.

The STRmix results, which again is expressed as a likelihood ratio, gave a value of approximately 4.5 times 10 to the 24th times more likely that an individual is included in that mixture than not included. So both very strong indications of inclusion, just two different ways to represent the results.

Q Page 22, please. Zoom in right here. So this is another example of a firearm right here, a shotgun trigger, right?

A Yes.
Q Would you describe the results of the analysis of this case?
A So this case is a partial mixture, so what that means is that we have DNA from more than one person but we don't have a full complement of all the data for all the genetic markers. We estimated that to be a two-person mixture, and our initial conclusion was that there is nothing applicable, it was inconclusive. And our result with STRmix was a likelihood ratio that was less than 1, which would support the defense hypothesis that somebody was not included within that.
Q Page 4, please. These were a number of other firearms you tested it on right here, right?
A Yes, that's correct.
Q Page 25. More firearms, all those revolvers right there, correct?
A Correct.
Q Let's jump ahead to page 29 now. Let's zoom in on the chart on just the first few lines of the text below. Mr. Nye, would you discuss for the Court what was being analyzed in this portion of the validation study?
A Certainly. So this is a lab created mixture, two individuals in a 1 to 1 ratio. So equal proportions of each contributor to the mixture. And then depending on the hypotheses that would be set up within the STRmix software, we

would result in different likelihood ratios. So the significance of the comparison would be predicated upon the hypotheses that you would use. So as an example, in that first block under number 15, sample 15, if we had a set of hypotheses where the prosecution hypothesis was that it was a two-person mixture, that included item number 15 or person number 15, and some unknown contributor, and then the defense hypothesis would be that it would be two unknown contributors, our likelihood ratios would be essentially ten to the 20th on an order of magnitude. So significant.

If we took that exact same mixture and looked at the other contributor, so our hypotheses would be it's item or person number 17 and an unknown contributor, versus two unknown contributors, our likelihood ratio would again be in the ten to the 20th area.

If we change our hypothesis on the prosecution hypothesis, we are going to say that that two-person mixture is made up of individual number 15 and individual number 17, but the defense hypothesis is that it's two unknown individuals, you can see that the significance of the statistic jumps to ten to the 50th power.

So the point in this is is that depending on the hypothesis that you use or the hypotheses that you use, it can change the significance of the statistic significantly.

And this might be a good point to explain how we come

up with our hypotheses that we use in our report, our analytical reports that we issue.

So based on information that we are provided when the case is submitted to us from law enforcement, we will use our best judgment as to the hypothesis that we will use in our analytical report. But we also have a statement in every single report that upon request we will rerun the analysis with another set of hypotheses.

So let's say there's a couple scenarios. Let's say an individual is sexually assaulted and we find out that on a particular orifice swab there's three, it appears that there's three donors to the swab. We could run that as the victim, the female victim being a contributor in that sample and two unknowns, I'm sorry, with the suspect and an unknown, versus the victim and two unknowns, and then later on if we come to find out that there was a consenting partner that she had sex with around the time of the assault, we can change our hypothesis to say that it was the victim, the suspect, and this elimination individual versus the victim, the elimination, and some unknown person. So we can change the hypotheses as we go, and it has a bearing on the statistics that's provided. And of course we open up that opportunity to defense counsel as well; if they want to propose other alternatives, we will certainly run the software with their alternative and issue a report then we can talk about the merits of their hypothesis when trial

comes.
Q But you can at least do whatever calculation they wanted to do then the parties could debate which set of hypotheses were more reasonable.
A That's correct, yes.
Q All right. So after this two-person mixture analysis was done, in this text below you start talking about another analysis, a three-person mixture analysis, right?
A Correct.
Q It says here you did it using what ratio of contributors?
A So a 10 to 5 to 1. So you have a major donor, an intermediate donor, then a minor donor. So basically trying to run the range of what we would see. Rather than like the one above which was equal contributions of the DNA to the mixture, we look at a three-person mixture with different contributions from each person.
Q So the minor contributor here, the 1 would represent roughly just under 7 percent of the sample right, 116th, did I do that math right?
A Correct, yes.
Q So can we go to the next page, please? And let's start with just the chart. This is the result of that three-person analysis, is that right?
A Correct. So the layout is similar. Different hypotheses along the bottom. And then the likelihood ratio or the log

likelihood ratio represented there, so they range anywhere from 10 to the 10th power, to as much as high as 10 to the 60th power, depending on the hypothesis that you set up.

Q And because those numbers are also large your conclusion was that?

A It was significant support for inclusion.

Q That STRmix could work on a three-person mixture with these ratios, minor contributor down to just below 7 percent.

A Yes.

Q Let's zoom out and go to the text right here. You next did validation on a four-person mixture, correct?

A Correct.

Q What were the ratios there?

A A 10 to 5 to 1 to 1, so that would be a major donor, intermediate donor, and two minor donors.

Q The minor donor in that case would be roughly 1 --

A 17th.

Q 17th, so a little bit smaller than before the minor donor?

A Correct.

Q And if we go to the next page. You see the chart. What were your conclusions about this chart?

A That it still models the four-person mixtures in those ratios very well. You can see that as you become a minor donor in a four-person mixture, that because the software incorporates more combinations in different possibilities that

it has an effect on the likelihood ratio. So as there is more combinations and more possibilities that it considers, the likelihood ratio becomes smaller, which is understandable.

Q Mr. Nye, are there any other portions of this validation study you would like to touch on?

A There was some discussion previously about repeatability, or the variation between one run and another run. And I did a study related to that. I forget right now the exact results of that study but I think Dr. Buckleton had mentioned that it could be circle around a ten full plus or minus. I wouldn't disagree with that. There is some variation. That's not unexpected. The similar variation that we would see with other statistical methods as well. The random match probability which has been used since, gosh, the early to mid 1990s is not uncommon to explain to a jury that the true value is ten fold higher or ten fold lower from the number that you are presenting. It's an estimate for a reason.

I also looked at partial data, so if you took a single source DNA profile and just continually removed data, so basically mimicking drop-out, that it has an impact on the likelihood ratio. So the less data that you put into the system, the lower your likelihood ratio can be.

So that would mimic something similar to the sample on this case where we have some drop-out or loss of DNA types that could be contributed to the defendant. That that actually

impacts the likelihood ratio to a less significant statistic.

I'm sure there was more. But those are the couple that come to mind.

Q Just one more I would like to touch on. Can we go to page 36, please? Zoom in right here. In addition to the charts we reviewed earlier about the various person mixtures and other ratios, you also tested other combinations of two, three and four-person mixtures?

A Yes.

Q And the ratios are contained in this paragraph here?

A Yes.

Q So one of them is 10 to 10 to 5 to 1?

A Correct.

Q So that would mean the minor contributor would be less than four percent?

A Correct.

Q One in 26. What's your, what was your conclusion about the efficacy of STRmix on those types of mixtures?

A So it still works very well. As Dr. Buckleton and I agree, four-person and five-person mixtures it works very well. It's just you get to the point of the computing power and how long each run is. We set a policy that we will not go over four-person contributors. I do know other laboratories in the U.S. that will go to five, but I would agree that most are in that four to five-person range. We did not test five. But it

works very, very well at four. And the more minor, the minor component is, so in that four percent range, because it has more possible combinations that it's looking at, the likelihood ratio is lower. And that's okay. That's a fair representation of that particular sample.

Q Turning to another topic. Is your laboratory accredited and audited?

A It is.

Q And does the use of these Probabilistic Genotyping Systems, STRmix in particular, comply with your accreditation standards?

A It does. We have not had any findings in that particular arena.

Q Are you aware of state commissions on forensic science that have reviewed probabilistic genotyping in general?

A Yes.

Q And have some of them approved of the use of probabilistic genotyping?

A They have, yes.

Q Let's go to Government's Exhibit 11. 11 isn't found in the book, Your Honor, because it's some 4, 500 pages long. We provided a copy to the Court but we just printed it once. Do you recognize 11, Mr. Nye?

A I do.

Q What is 11?

A Our procedure and training manuals for the DNA program.

Q And what sorts of things does this manual cover?
A So it's divided up into two major categories, a training manual, which has some relevancy to this proceeding but it also, the other section is our procedure manual, basically our instructions to our examiners on how they analyze evidence, how they interpret it, how they report it, makeup of chemicals and QC and different things of that nature.
Q Does it cover topics like handling of evidence in a laboratory?
A It does.
Q Are there strict procedures in place to try to minimize the likelihood of cross contamination?
A There are.
Q Does it cover decisions that a particular forensic scientist has to make when analyzing an electropherogram?
A It does.
Q And does it provide guidelines for using STRmix in the Michigan State Police setting?
A Yes, it does.
MR. PRESANT: Your Honor, the government moves to admit 11.
THE COURT: Ms. Kloet.
MS. KLOET: I have no objection. This is the full policy manual? I just don't have a copy in my binder. Is this the whole manual?

MR. PRESANT: It's, can we go to the last page? I don't know how many pages there are. That would probably be the best way to get an idea of the scope.

THE COURT: 446.

MS. KLOET: If that reflects the whole manual, I'm fine with it.

MR. PRESANT: I believe it does.

MS. KLOET: Okay.

THE COURT: Do you have any objection?

MS. KLOET: No objection, sorry, Your Honor.

THE COURT: It's admitted.

BY MR. PRESANT:

Q Now, while we are on the topic of possibility of contamination, is that something you worry about as a supervisor of DNA analysis at the Michigan State Police lab?

A Yes.

Q Why do you worry about it?

A So when I described previously the changes in technology over the 22 years that I've been in the system, the sensitivity of what we do has increased dramatically. And when I'm referencing sensitivity, what I'm referencing is the ability to detect DNA at smaller and smaller and smaller quantities. And so one thing that you have to be concerned about is that when your sensitivity increases in your laboratory analysis, that you're actually detecting DNA that's from the crime scene and

not cross contaminating with other items or contaminating with DNA from yourself or other exogenous items. So it's something that we work very closely with. There are accreditation standards that speak to some of these items such as training and policies and procedures, and facilities, evidence handling, there's a number of accreditation standards that speak to this, and then also we have our own requirements to try and minimize it.

Q Are you confident in the procedures that have been put in place in order to handle DNA evidence at the MSP lab?

A I am.

Q Now, a closely related topic is this idea of touch or transfer DNA. Are you familiar with that concept?

A I am. Depending on the terminology that you use, it may mean very different things. But we will see.

Q What do those mean to you if I say touch DNA or transfer DNA?

A So in my world when I talk about touch DNA, I'm talking about DNA that can be recovered from an item that was handled for a short period of time. So we could be talking about a doorknob, a pen, I mean a whole host of different things. That continuum of what people want to consider touch DNA can range quite a bit further than that. Some people may consider touch DNA a garment that you're wearing. Or more intimately held items than something that's just touched for a brief period of

time. So there is not a lot of consensus on what you consider to be touch DNA, so we have to be somewhat cautious of what we lump into that type of explanation.

Q What about the idea of DNA transfer and the related ideas of primary transfer, secondary transfer, are you familiar with that terminology?

A I am.

Q What does it mean to you?

A Well, when you get into primary and secondary transfer, again, there is some differences of opinion as to what that actually means. But in a very general sense when we talk about transfer of DNA, just me picking up this cup could be I'm transferring my DNA to the cup. Now, when you start to get into secondary transfer and tertiary transfer, if you pick this cup up after I have handled it and then go to the door, are you going to transfer my DNA from this cup to another item. And that's a more extended version of transfer that can be discussed.

Q Are you aware of literature where people have attempted to study those types of DNA transfer?

A You chose your words very well. Attempted because it's very, very difficult to model that type of situation. And there have been studies, and they have come up with a variety of different conclusions as to how prevalent or how rare that type of situation is.

Q Can we bring up Exhibit 9, please? And go to page, well, first on page 1, you recognize this as a laboratory report or the format of a laboratory report produced by Michigan State Police?

A Format, yes.

Q You didn't produce this report?

A I did not.

Q Can we go to the second page, please? Can we zoom in on that table? There's been testimony today, Mr. Nye, about how we communicate likelihood ratios to a jury. What is the significance of the table that I'm showing you here?

A So when we produce a likelihood ratio statistic, it's a numerical form. In this particular case, 40 or 50 million to 1. And I had previously explained that we have approximately 90 DNA scientists just within our one organization, let alone any other experts that might review that material or testify to those findings, and for me 40 or 50 million to 1 may mean something quite different to somebody else.

And so we created a policy of a verbal equivalent of a numerical value. So in other words, to try and standardize or bring some normalization to how we testify with the Michigan State Police. So if you get a value that's let's say a likelihood ratio of greater than 0 to 99, I think I've mentioned I have testified before to a value that was 10. What does that 10 actually mean? A 10 to 1, 10 times more likely

than not, to me represents something that's relatively uninformative. You're very close to 1 which I think Dr. Buckleton had mentioned is uninformative, to me that has less value than a likelihood ratio that's a duodecillion to 1. And so this is our effort to try and bring a verbal equivalent to a numerical value to again assist the jury and assist our scientists to normalize how they respond to how strong or weak the evidence is.

THE COURT: Isn't that to some extent invading the province of the jury?

THE WITNESS: I think what it's doing is, DNA is very good at, DNA analysis is very good at determining whether, whether you have a result that can support an inclusion or exclusion. I think the significant portion of that is is that DNA testing can't fully explain how the DNA was deposited on that item or how it came to be found there. I think that was what Dr. Buckleton was trying to explain is that when you have a likelihood ratio, it helps inform a part of it but there's other items and investigative information that fully informs the jury as to the significance of that.

So our effort here is not to provide any more information to the jury other than to normalize how we present the data and the significance that we see in it so that we don't have wide variations of what a million to one means versus me versus somebody else.

BY MR. PRESANT:

Q So you helped develop this verbal equivalency table?

A I did.

Q And the top level is 10,000 or greater, correct?

A It is.

Q So that would be three orders of magnitude or a thousand times lower than the tens of millions of range that we have had testimony about here today like 50 million, right?

A That's correct, yes.

Q The purpose of this equivalency table is to impose uniformity, correct?

A It is.

Q That's part of the purpose of the policy and procedures manual that we just looked at, Exhibit 11?

A It is.

Q That's part of the reason that you testified you chose to adopt PG system such as STRmix in the first place to impose uniformity on the way that DNA mixture analysis is done, correct?

A Certainly one portion of it, yes.

Q And are you aware of the labs in the United States, the forensic laboratories in the United States that are doing probabilistic genotyping analysis, what percentage of them are using STRmix as opposed to some other form of software?

A I don't have access to the same level of information that

Dr. Buckleton does. But when I go to different conferences, go to SWGDAM, also a member of the OSAC through NIST, and have just individual conversations with different laboratories or representatives from those laboratories, there is a significant portion of people that either have purchased it and implemented it, or have purchased it and are going through the validation process right now. The validation process is extensive. We talked about the end date of February of 2016. But it was actually started, the validation formally was started in about July or August of '15. So it takes quite a period of time. And certainly a year from now there will be significantly more laboratories that will have implemented it because they are undergoing their internal validation right now.

MR. PRESANT: Thank you very much. Nothing further, Your Honor.

THE COURT: Cross-examination.

MS. KLOET: Thank you, Your Honor.

CROSS-EXAMINATION

BY MS. KLOET:

Q Good afternoon.

A Good afternoon.

Q Just to refresh my recollection, when did the MSP or when did you decide to purchase STRmix on behalf of MSP?

A I don't know if I can clarify exactly when I decided to purchase it. But I can say that in I think it was May of 2015

I think it was was when we had our first training event. So the purchase would have been shortly before that time period.

Q Okay. Thank you. During the decision making process I believe you testified that you read developmental studies with respect to STRmix, developmental validation studies?

A I read a number of different literature, citations, spoke with a number of different individuals, received demonstrations of the software application as it existed at that particular time. I did quite a bit of I guess I'll use the word research before I ultimately made my decision.

Q Of the studies that you read specifically, can you recall how many of them involved complex mixtures of three or more contributors?

A I think many of them do. That's where STRmix or other probabilistic genotyping applications excel. And so I'm sure a large portion of them had mixtures that were more than two individuals.

Q You said many. Do you have any idea how many?

A I don't recall. It's not something I tracked.

Q Okay. I believe you testified the date that STRmix went live was approximately February of 2016.

A Correct.

Q Okay. The time that you purchased STRmix in or before May of 2015, were you aware of any errors in the program?

A Not that I specifically recall. I mean at that particular

point in time -- so let me back up just a second. So when we implemented in February of 2016, I believe we were laboratory number 9 in the country to actually implement. I may be wrong on that. We were a very early adopter. So if you back up a year from the time that we implemented to the approximate time of when we purchased, there wasn't as much information out there about like coding errors and things of that nature. I think Dr. Buckleton had mentioned about his personal website and different websites that they are posting a lot of that information, and a lot of those weren't available. I don't recall if I ever had a discussion with the distributors of the application about whether there was coding errors. I don't recall whether I did that or not.

Q Okay. And you just referenced Dr. Buckleton's publishing of a few errors. And I believe that one of those was admitted as an exhibit or document representing those through the government's direct exam of John Buckleton.

A I believe so, yes.

Q Were you aware of those specific errors?

A I was aware of the specific error as it relates to the one of our version.

Q So is it your testimony today there was only one potential error in your version?

A I think that was Dr. Buckleton's testimony.

Q Okay. I'm going to help refresh your recollection if we

pulled up that document.

A Sure.

MS. KLOET: Your Honor, can I ask the Court to pull up Government's Exhibit 14?

THE COURT: I think we are going to ask the government's tech to do it.

MS. KLOET: Thank you.

BY MS. KLOET:

Q So this is the document that was previously admitted with respect to the errors in STRmix. I would like you to look at these paragraph 3. This paragraph indicates that there was an issue with the version that MSP employed, correct?

A Yes, if it's version series 2.3 and we use 2.3.07.

Q And were you aware of that particular issue at the time that you purchased STRmix?

A I'm not sure that they were aware of it at that particular time. I was not aware or if I was aware, because it's such a rare instance on a specific scenario, I'm not sure it would have necessarily impacted my decision to purchase or not purchase it. I don't recall specifically.

Q So you don't remember if you were aware of this particular error.

A I do not.

Q Okay. Thank you. Did you become aware of those errors after the purchase of the program?

A I don't remember if I was aware of it before it so I don't know if it was after or not, I'm sorry.

Q Did you first become aware of it today?

A No. No. If that's your question, no, I knew of it before then.

Q That is my question. Thank you. I would like to talk a little bit about the training at MSP with respect to STRmix. You came up with the competency test for the analysts with respect to STRmix, correct?

A Yes.

Q And you administered that test to the 90 or so, I guess the portion of the 90 or so analysts that are --

A Correct.

Q Okay. Thank you. Now, at the time you did this, had you passed any competency tests yourself before administering it?

A No. It's sort of a chicken and egg kind of thing. You can't competency test yourself if you're the one that's administering the competency test. So as a DNA technical leader there is no requirement that I become competency tested before I administer the competency test to others. The real test, so to speak, is I make a mixture up that represents the range of testing that we do and I know the answer to it because I made it up myself. That's the real essence of the test.

Q Okay. Had you taken and passed any competency test on STRmix?

A No.
Q Thank you. I would like to talk a little bit about the internal validation MSP completed in this case. You referenced adjudicated cases earlier.
A Correct.
Q And I believe you suggested that you utilized those because they tended to be more representative of real life samples?
A They can be.
Q And why is that?
A For the reasons I explained previously which was that they are more representative or they can be more representative of evidence types that we get in the laboratory; so they have been subjected to different conditions, there's varying amounts of DNA, varying numbers of contributors, on different substrates. And what I mean by substrates is the surface that they are on, whether it be a fabric of a particular kind or whether it's on a hard surface like a gun, or just a whole host of different conditions that are sometimes a little bit hard to mimic in the laboratory.
Q Thank you. And you are comparing your results in STRmix to the adjudicated conclusion, correct?
A Correct. It's the only other source of information that we have because we, we didn't deposit the DNA on the item so we don't actually know whose on the item or in what proportions that they are on the item. So that's the closest we can get to

a comparison. So my position was it's better to do that comparison knowing that there is some shortcomings of that comparison than to not do the comparison at all. That there was some value in doing that.

Q So you're comparing it to your conclusion not to ground truth, so to speak, in those cases?

A Right.

Q Thank you. And I believe you testified that STRmix is validated for up to four potential contributors?

A Yes.

Q So four contributors for purposes of this version of STRmix at MSP is the boundary of acceptable use.

A That is the boundary that I placed on it, yes.

Q Okay. But isn't it true that the true number of contributors is always unknown?

A It is always unknown, and I would agree with Dr. Buckleton that you can estimate it. It's better to be conservative than, or correct than to be incorrect. But I would also clarify that that estimation, I don't know if I'm going to use this term correctly, but it's an informed estimation. The estimation of the number of contributors is not a trivial matter. We spend a fair amount of time in our policies and procedures explaining what things to consider when trying to make an estimation as to the number of contributors. And we train quite a bit around that. And it comes through training and experience and

knowledge and procedures.
Q Did you publish the MSP validation summary, internal validation or study in any peer reviewed journals?
A No. Our positions as forensic scientists, as practical practitioners isn't really to go out and publish material.
Q Did you -- did any third party duplicate any of the tests that you performed during the validation?
A No.
Q Would you mind pulling up your Exhibit 9?
MR. PRESANT: Exhibit 9 wasn't admitted.
MS. KLOET: It wasn't admitted?
BY MS. KLOET:
Q What is this document? I know you already looked at it but just for clarification.
A It would be a laboratory report reference a specific laboratory number and date and time for an agency.
Q And you previously looked at the second page of this document, right?
A I believe so, yes.
Q And we were talking or you were testifying about the chart that's on this second page. You are the individual who created these guidelines, right?
A Yes.
Q Okay. You determine these cutoff values?
A So the cutoff values for the verbal equivalent were derived

from a publication which I don't have with me today but I could certainly provide it to the Court if necessary. It was a journal publication, peer reviewed journal publication not related to DNA but it was related to other forensic evidence. I believe it might have been maybe latent prints, and where they do likelihood ratios and have a verbal equivalent for a likelihood ratio. I used that as a guide. This is not verbatim out of that publication. There were too many different statements, there is too many different categories. I believe that the publication might have had maybe eight or nine different categories. It's like the equivalent of saying very, very, very strong evidence. It just how many times can you say the word very. So I condensed a number of them and then I don't recall if I specifically adjusted the range or the numbers, but that was where it was derived from.

And then I will also offer that in the forensic community for those that have adopted probabilistic genotyping, the use of a qualitative equivalent, some laboratories use them, and some not. And those that do use them they generally are consistent amongst them. But it's something that is being looked at in the community to see if we can standardize those verbal equivalents. So my effort to standardize within our laboratory, there is an interest to standardize that across the forensic DNA community as well.

Q So you made your own personal judgment call with respect to

the creation of this qualitative equivalent chart, fair to say?
A Yes.
Q You did that based on a study that had nothing to do with DNA.
A Forensic latent print examination.
Q Not DNA, fair to say?
A Correct.
Q And that was published in a peer review publication.
A I believe so, yes.
Q But you changed what was published in the peer review publication for MSP's purposes, true?
A I certainly condensed the number of categories because I thought it was overly confusing to have that many different categories. And I don't recall whether I changed the likelihood ratio cutoffs for the verbal equivalent.
Q Okay.

MS. KLOET: Thank you. I don't have anything further.

THE COURT: Thank you. Any redirect, Mr. Presant?

MR. PRESANT: No, Your Honor.

THE COURT: Thank you. You may step down, Mr. Nye. Thank you for your testimony. Your next witness is going to take how long?

MR. PRESANT: I would have given a different answer this morning but at the pace we are going I would say could be an hour to an hour and a half.

THE COURT: Well, let's do this. Let's get started and see how far you get by 4:15.

MR. PRESANT: That sounds good to me. Your Honor, may I raise just one legal issue now or be better to wait until the end of the day?

THE COURT: Depends on what the issue is.

MR. PRESANT: Your Honor just made some comments at the conclusion of Dr. Buckleton's testimony with respect to the Court's role in Daubert, and I'm making sure that, and I may have misinterpreted the Court's comments, that the evidence was handled properly. And I didn't know if we were just talking about in the lab or chain of custody issues as well. The government hasn't prepared witnesses for this hearing that would address chain of custody, and if the Court needs to hear from them I'll understand that. But it's been treated on page 19 and 20 in our brief and response. The government's view chain of custody is really an issue for trial about whether the officers did it properly. So we prepared our proofs today with respect to the reliability of the technology.

THE COURT: I don't believe the chain of custody or how it was handled has anything to do with what we are about in this hearing. So, no, my answer is no. There is no need for the government to put on any testimony as to that.

MR. PRESANT: Thank you, Your Honor. I just wanted to clarify. With that clarification the government calls

Ms. Amber Smith.

THE COURT: Okay. My purpose in exploring that with Dr. Buckleton was simply to delineate the different areas that are pertinent to the development of this kind of test, of this kind of result, not that we're about examining it here.

MR. PRESANT: That was very well my misunderstanding, Your Honor. I appreciate the clarification.

AMBER SMITH, GOVERNMENT WITNESS, WAS DULY SWORN

THE LAW CLERK: Please be seated. And state your full name for the record, spell your last name.

THE WITNESS: It's amber, A-M-B-E-R, Smith, S-M-I-T-H.

DIRECT EXAMINATION

BY MR. PRESANT:

Q Ms. Smith, where do you work?

A I am a forensic scientist with the Michigan State Police in the Lansing laboratory in the Biology DNA Unit.

Q Can you bring up 3, please? Do you recognize Government's Exhibit 3?

A I do.

Q What is it?

A This is a copy of my curriculum vitae.

MR. PRESANT: Government moves Exhibit 3, Your Honor.

MS. KLOET: No objection, Your Honor.

THE COURT: It's admitted.

BY MR. PRESANT:

Q Can we start at the top with your education.
A Yes. I have a master's degree in biology from Southern Illinois University in Edwardsville, Illinois and I have a bachelor's degree in marine science and biology with minors in chemistry and environmental science from the University of Tampa in Tampa, Florida.
Q Can we back out and go to experience. What has your work experience been, Ms. Smith?
A I was previously employed before I came to Michigan State Police with the St. Louis Metropolitan Police Department. And the formal title there was I was a criminalist for approximately two years.
Q What type of work did you do in St. Louis?
A I did both body fluid identification analysis as well as DNA analysis.
Q And after you left St. Louis?
A When I left St. Louis in March I moved to Michigan and I became a DNA analyst with the Michigan State Police where I also perform body fluid identification as well as DNA analysis.
Q You've been there more than ten years now?
A Yes.
Q Are you trained in using STRmix?
A I am.
Q What trainings have you gone through in STRmix?
A I was a member of the first group that was put through the

STRmix training in March of 2015 where I underwent the four-day class with Dr. Buckleton. And then as part of the validation myself and that original group compiled samples for our technical leader at the time, Mr. Nye, and went through those samples and ran a number of samples for practice with STRmix. And then after we ran through the samples and the validation was complete, we went through and did a competency test involving samples that Mr. Nye had created. And then once that was complete, we then took a written test before we were able to be put online.

Q Page 2, please. And if we zoom in here. Is that the STRmix training you just testified about?

A Yes.

Q That's in the middle of a number of other trainings related to DNA analysis and other forensic topics?

A Yes.

Q And the trainings you have undergone continue to the end of your CV as well, right?

A Yes.

Q Approximately how many samples, DNA samples of any kind do you think you've processed in your career as a forensic analyst?

A Thousands.

Q How many STRmix analyses have you conducted?

A Analyses, well over 200 analyses with the cases being

around 200 as well.

Q How many times have you testified with respect to any type of forensic DNA science, approximately?

A With testimony, this is my 74th time.

Q And how about testimonies including STRmix?

A With STRmix, I have testified 16 times.

Q Those were all in state court here in Michigan?

A Yes.

MR. PRESANT: Your Honor, the government offers Ms. Smith as an expert in forensic DNA analysis.

MS. KLOET: No objection.

BY MR. PRESANT:

Q The Court has already heard Mr. Nye describe the general process for forensic DNA analysis at the Michigan State Police lab. You were in the courtroom for that testimony, correct?

A Yes.

Q Is there anything you would like to add or supplement that you think the Court should know about the general process from the time the evidence comes into the lab until the time you produce your lab report?

A As it pertains to STRmix?

Q As it pertains to the processing of evidence, or to STRmix, just anything you would like to add about the general process before we talk about what you did in this case.

A At this time, not all of the analysts in my unit or across

the state are trained in STRmix. However, they are trained in DNA analysis and mixture interpretation. So because they are not trained in producing statistics on their mixtures, a lot of the samples especially at this time in Lansing there were only two of us that were performing STRmix analysis.

THE COURT: Does the Michigan State Police still do DNA analysis under the pre probability genotyping?

THE WITNESS: Yes. The process is the same and if the analyst that is running their case develops single source samples they will still produce their report that has a random match statistics. It's only when the statistic becomes too complex for mixtures samples that they are then inserted through the STRmix samples.

THE COURT: How is that determined?

THE WITNESS: So once the analyst runs their sample and they see that they have a mixture sample, they then forward their sample to a qualified STRmix analyst and get approval and clarification for this sample if it will fit the guidelines for STRmix analysis.

THE COURT: So when you say a mixture sample, you mean that it appears that there's more than one donor to the DNA that's being tested?

THE WITNESS: Yes.

THE COURT: Okay.

THE WITNESS: So once, in this case anyway, the

analyst developed her profiles and then sent those forward to the two of us at the time to evaluate for potential STRmix analysis. And that still goes on until all of us in the unit will be qualified.

BY MR. PRESANT:

Q So there were in this particular case there are actually three forensic scientists that handled the sample in the lab?

A The actual DNA process there were only two. There was the person that, the person that -- the way our laboratory works is we do perform our own lab work on our own cases and then generate reports. However, the analyst that did her lab work and got, she did issue a report but she could not issue the statistic on the report because the sample needed a mixture statistic. So at that time she issued a report that I believe basically said it's a mixture, a subsequent report will follow with statistics.

Q I missed a question earlier too. You testified you underwent proficiency testing in STRmix, right, and you continue to undergo it as long as you are a STRmix analyst?

A Yes.

Q Have you ever failed a proficiency test?

A Not to my knowledge.

Q Other general background questions. What type of genetic analyzer is in use at the Michigan State Police lab?

A Right now we have a 3500 XL, which is a genetic analyzer

that will inject 24 samples at a time, as well as a regular 3500 which only injects 8 samples at a time.

Q And what sort of software is used to interpret the electropherogram?

A GeneMapper is the name of the software that is used to generate our electropherograms.

Q Then STRmix is used to interpret it.

A STRmix is the statistical tool used to generate the statistics.

Q All right. So let's take a look at Government's Exhibits 6 and 7. Can we do a side-by-side? They will be in the book in front of you if you want to flip back and forth but we will start side-by-side. And would you describe for the Court what 6 and 7 are, if you recognize them?

A 6 are the electropherograms that were generated by the original analyst in the case. Her electropherograms are utilized to help determine number of contributors. And these electropherograms have filters turned on which are the stutter filters which is an artifact that is generated during the process. Our laboratory has thresholds for each location that we test that has stutter filters. So in order for interpretation, if they are below those filters it's automatically filtered out by our GeneMapper software.

And for number 7, I actually have to analyze and insert the electropherogram into STRmix with stutter filters

off. Meaning that all of the artifacts are present on the electropherograms and that is so that STRmix can analyze the sample and determine the probability that this peak is a true artifact or could it possibly be a potential type.

BY MR. PRESANT:

Q Those are your initials at the top of 7, ALS?

A Yes.

Q That's how you know it's yours?

A Yes.

MR. PRESANT: Your Honor, the government moves to admit 6 and 7.

MS. KLOET: No objection.

THE COURT: Admitted.

BY MR. PRESANT:

Q So would you briefly walk through 6 and 7, the process of interpretation as you understand that was undertaken by Ms. Urka was the original analyst, correct?

A Yes.

Q How it was transferred over to you and then we will look through your notes in Government's Exhibit 7.

A So there are different sets of protocols that different analysts follow. So non STRmix qualified analysts follow Section 2.10 in the protocols and they had different rules according to those protocols. And then once the samples come forward to STRmix, I have a different set of protocols I follow

which is 2.11.

So under 2.10, the analyst looks at their data and make sure it's free of any extraneous artifacts which are generally determined to be off ladders at samples, and those samples can generate off ladders if there's oversaturation observed. However --

THE COURT: What does that mean, oversaturation?

THE WITNESS: Oversaturation is when there is too much DNA amplified on the first go around, and that it will generate a bunch of noise that is usually detected below the analytical threshold. So when you have a sample that generates a bunch of excessive plus 4 stutter, or OLs which are called off ladders, means they don't have types, you generally will reinject that sample at a lower time or you will reamplify it with a lesser amount.

THE COURT: Okay.

MR. PRESANT: Can I just interrupt? Can we go to page 5 of Exhibit 7? No need for side-by-side. And zoom in right here. Ms. Smith, would you continue if you would answering the Court's question regarding oversaturation as we look at the D8 locus.

THE WITNESS: Yes. So regarding D8, according to the manual that Ms. Urka followed, there was nothing wrong with the data that was detected in D8. Because Ms. Urka is not yet trained in the STRmix evaluation of the software, she does not

know to look at certain areas more in-depth that I look at when it comes to the analyses. So being that this 13 type at D8 was very high, it was under the parameters for oversaturation regarding STRmix, which is 25,000 RFU, and this is at 23,821, but however there is a potential plus 4 stutter with the 14 present at 910, which I believe was about three and a half percent of the 13 type. And in addition to that, the 12 peak before 13 was in the stutter position for the 13.

And I was unaware if the 11 peak that popped up when the stutter filters were removed could possibly be an N minus 8 stutter which could be artifact generated from the height of the 13 peak or if it was a byproduct of the 12 peak which could have been very high stutter.

So since the version of software with STRmix that we use at 2.307 does not model plus 4 stutter, or N minus 8 stutter, it was my judgment call to remove the locus completely from interpretation as because I could not explain if those were artifacts or if those were real types. Generally when I run my own data I like my types to be around 20,000. When I run my own data I like my types to be around 20,000 at the highest just because that generally does not indicate an overexaggerated plus 4 stutter from that highest peak. And it does not usually generate anything that will go into another color. So that's just a personal preference. That I start to see more artifacts once I see that 20,000 peak height.

BY MR. PRESANT:

Q Now, you talked about what you prefer for your own types when you're running the sample, correct?

A Yes.

Q And so one question that might come to mind is why not just rerun the sample in this case?

A Well, this case was generated under a different protocol which was 2.10, the original data that was run by Ms. Urka. Ms. Urka followed her protocol as she was supposed to. There is no upper bound or upper threshold when it does come to following a 2.10 manual. So when Ms. Urka ran her data it did undergo a technical review as well as an administrative review by two other individuals in the laboratory. Her data was then approved and reported out in her original report following 210. So when that data is supplemented over to me, I am taking data that's already been approved and data that's already been reviewed and passed through. So at that point I do not have the ability to reamplify or rerun or reinject a sample because that's not my data. That's someone else's data. And if on her review the person that reviewed it thought it should have been injected or use a lesser injection time, they would have suggested at that time. But because it did not technically violate the protocol that she was following, this data was acceptable for the purposes of her report.

Q Let's go back to Exhibit 11 previously admitted. You

recognize 11 as the policy manual for MSP, Ms. Smith?

A Yes.

Q And can we go to page 88 of the document? Now, we are in part of the 2.10 series you've testified about, right?

A I cannot -- that's actually regarding the analytical threshold. It says, "All available loci and alleles greater than or equal to the 250 RFU analytical threshold should be analyzed or should be utilized during STR interpretation. It should be noted that peaks not represented by an allele do not need to be included within the interpretation."

Q Can we back out of that please, Ms. Miller, and go to the bottom here. What's the significance of this, Section 2.10?

A This basically says that the single donor reference samples must also surpass the 250 RFU analytical threshold and that a known sample that has a single type present at a location must meet the 900 RFU threshold to be utilized in a statistic.

Q Let's jump ahead to page 106. I'm sorry. We have to start at the bottom, 105. And now we are in the 2.11 series of the manual, is that right?

A I'm not sure.

Q We can go to the next page if you want to see the next.

A Yes.

Q What does the 2.11 series govern exactly?

A 2.11 is a protocol that if I were running my own data through and following forward through the STRmix procedure,

this was the protocol I would follow, as well as if STRmix is to be utilized in a sample, this is also the protocol you follow.

Q The point is these protocols, dozens of pages in this particular area of this hundreds of pages document limit the judgment calls you can make with respect to what you do in processing the DNA sample and in using STRmix, is that right?

A Yes. If I did not generate the data, I can only assess how the data is once I receive it.

Q And the policy allows you to disregard a locus that is deemed to be oversaturated?

A Yes.

Q Is that based in part on the detection of the instrument?

A It is as well as my training and the ability to determine artifacts and actually how well or if the software is able to assess the potential artifact present.

Q Let's go to page 3 of Exhibit 7 where we were before. Zoom in on the note. This note appears on several pages of the report, is that right, Ms. Smith?

A Yes.

Q Or of the electropherograms, rather?

A Yes.

Q Would you just read the note for us?

A So I'm required as part of my protocol to document my number of contributors as well as how I got to that number of

contributors. So this note is me saying that I interpreted this sample, "assuming that there were three donors present due to more than four peaks that were observed at multiple locations, and since D8 exhibited oversaturation and was the only locus to indicate four potential donors, due to the oversaturation and possible artifact observed, D8 was not used in the analysis." So this is my evaluation of the profile as a whole and how I deemed the profile to be three donors.

Q How common is it in your experience to observe a profile where you have to ignore a locus because it's oversaturated?

A At this point it's not uncommon. Again, it's because all of us are not STRmix trained and so I am receiving data that's run by someone else. And because that someone else follows a different protocol, and meets the standards and guidelines in their protocol, it's not uncommon for me to receive an electropherogram that may have demonstrated oversaturation. There also potentially may be a time where I myself may have a sample that does surpass the 25,000 RFU cutoff, however, if you look further out in the profile as a whole I may be losing minor contributors. And so it may be more beneficial for me to sacrifice one locus that exceeds the threshold in order to gain more information from additional contributors.

Q Let's say you were wrong about the oversaturation and it's actually a four-person mixture. Would you be willing to run STRmix again assuming the four-person mixture?

A Yes. My report does offer and say that I will run a sample again if there are different circumstances that people would like me to consider. In fact I have on I believe on at least three occasions that I can recall.

Q So if the defense asked you to rerun STRmix assuming four contributors, that is something you would be willing to do?

A Yes. Although I do feel this is a three contributor sample, I will run it however I am requested to rerun a sample.

Q As you sit on the witness stand today, have you received a request from anyone to rerun STRmix in this case using different assumptions?

A I have not.

Q So let's go to Exhibit 9, please. Do you recognize Exhibit 9?

A 9 is the report that I generated based on my STRmix analysis.

Q It's a minor point, but this Exhibit Number 5 in the bottom right-hand corner, that's not part of your report usually, right?

A Right.

Q I'll represent to you that's just the way it was processed en route to me. That's why it's on there.

But so this is your report, and let's go down to your conclusion section. Mr. Nye testified about the importance of formulating hypotheses and using that to run STRmix; is that

what you've documented here in the report?

A Yes. The first interpretation under number 1 I am required to state how I interpreted the sample, and so I am saying by looking at the evidence I ran the sample assuming there were three individuals. Based on our reporting formats, I also have to state my hypotheses which is generally the person of interest and two unrelated, unknown contributors, or did this sample generate from three unrelated, unknown contributors, not the person of interest.

Q And your conclusion was?

A The conclusion is, "Based on the DNA typing results obtained, it is at least 49 Million times more likely if the observed profile from the swabs of the textured areas of GUN-001 originated from Daniel Gissantaner and two unrelated, unknown contributors than if the data originated from three unrelated, unknown individuals."

MR. PRESANT: Your Honor, the government moves Exhibit 9.

THE COURT: Ms. Kloet.

MS. KLOET: No objection, Your Honor.

THE COURT: It's entered.

MR. PRESANT: Let's go to the second page. We have already heard testimony about the verbal scale table, correct, and you just applied that to the number you reached at the top of the page, is that right?

THE WITNESS: Yes. As part of our reporting format I am required to insert a verbal equivalent to insert the number in lay terms. And this number corresponded with the 10,000 and greater part of the equivalent which was very strong support that Daniel Gissantaner is a contributor to the DNA profile developed from the swabs of the textured area of GUN-001.

BY MR. PRESANT:

Q In this first paragraph here under remarks is the offer you were talking about earlier where you wrote, if other propositions should be considered you would be able to undertake them if instructed with sufficient time, right?

A Yes.

Q Okay. Let's go to Exhibit 8, please. Do you recognize 8, Ms. Smith?

A I do.

Q What is it?

A 8 is the Word Pad file of the comparison report that was run against the deconvolution file for this case. So this is basically an unpretty version of a comparison of the known sample which was 2AX compared to the evidence sample.

Q All right. And let's look at 25, please. What's 25?

A 25 is the original deconvolution file that I ran as it pertained to the gun evidence. And this file gives the summary of my parameters that I ran as well as the diagnostic kickouts and the summary of contributors. It also includes the required

parameters that are used in the software as well as the allele calls that were imported into the software and all the genotype contributor possibility breakdowns.

Q Would it be easier if we talked about what led to the 49 million to 1 number to use this Exhibit 25 as the prettier version of Exhibit 8 or are they different?

A This is not the prettier version. The prettier version is 27 or 26.

Q I'm sorry. I appreciate the correction very much. Let's look at 26 then. And would you just lay a little bit of foundation beyond the prettier version?

A So 26 is the PDF form of the Word Pad version that we previously looked at, and this basically says the same information, however, it's in table format and it's a lot easier to follow.

MR. PRESANT: Your Honor, the government moves to admit 8, 25 and 26.

THE COURT: Ms. Kloet.

MS. KLOET: No objection, Your Honor.

THE COURT: They are admitted.

MR. PRESANT: Let's start where we are with 26. And let's go to page 3, please. Zoom in on this area if we can. So you may recall some testimony from a prior witness where we were looking at these rows where there is no data entered.

THE WITNESS: Yes.

BY MR. PRESANT:

Q Would you explain why there is no data in those rows?

A There is no data at DY because we don't use DY as part of the comparison. Not only with STRmix but with any of our statistics we do not use DY. There's no data at D8 because I removed that locus from the original deconvolution file so because there was no information there for the known sample to be compared to, is why that's blank. Because there is no LR that can be generated because there was no data for that to be compared.

Q Can we back out of that, Ms. Miller? Briefly go to the top, just the column headers. Ms. Smith, would you orient the Court to what each of these three major columns are?

A These columns generate the values that were obtained from the three population data bases that we use, and in Michigan we primarily only report out the African American, Caucasian and Hispanic profiles because those are the races most commonly come across in Michigan.

Q Can we back out of that, please? Now, why did you choose to report -- well let's go to the bottom here, the total likelihood ratios. Let me word it this way. Is the reason you chose to report out the 49 million to 1 number because you had been given information that the defendant in this case, the known sample is African American?

A No. I actually have no idea the race of anyone when I run

the sample. It's the most conservative number out of the three populations that's generated, so that's the number that's reported out.

Q So if you used the Caucasian data set, the number is even bigger, the likelihood ratio is bigger, right?

A Yes. In our previous or my previous reports and in the current random match reports, all three numbers for each of the populations are a given. But because we use STRmix now we just report out the most conservative number.

Q So this E7 means what versus E9 and E8?

A So the E7 is the statistic that was generated or the likelihood ratio that was generated for the African American population, the E to the 9th is for the Caucasian population, and the Hispanic was for E8.

Q That's just a mathematical notation indicating times ten to the power of whatever number is listed?

A Yes.

Q Let me back out, please. Last on this page, just go to the column. So this 49 million number, you've actually rounded down to 49 million from what's reported there, right?

A Yes. We do only report out two significant figures, so that's why it's 49.

Q That's calculated by all these likelihood ratios from the individual loci.

A Yes.

Q Let's go to the next page, please. Ms. Kloet asked Dr. Buckleton some questions earlier about these parameters. Let's look at those briefly. Did you input these settings, the drop-in cap and the drop-in frequency, this 400 number and the .453 number?

A No, these are parameters that were generated during the validation process, and this is what all of the samples run through STRmix are run with. I actually don't have clearance to alter these samples, or these numbers, and if they were to be altered, they would be in bold to signify that they were not the same as the original parameters. However, I do not have that ability to do that.

Q And the next page. Ms. Kloet also highlighted some of these peak heights. In this column right here that were below 400, that drop-in cap, is that right?

A Yes.

Q What's the significance of them being below 400?

A Really there is not much significance that many of them are below 400. 400 just happens to be our drop-in cap. Meaning that if there is a peak that is detected blow 400, the software does consider it to be potential drop-in. However, just because it's 400 doesn't mean that it's drop-in. The software also is considering the possibility that it's stutter of some sort or is it a possible real allele that could be attributed to a contributor in the case. So it does consider every option

and then weighs those options based on the profile as a whole.
Q Is there anything else you would like to highlight for the Court in Exhibit 26 before I turn back to 25?
A Just that 26 also has where on page 7, the last page, is the reference profile from 2AX to show that this is the profile that was compared to the original evidence. And so this is also checked on review to make sure that the right profile is compared to the evidence.
Q So it's just listing here the specific alleles at each of these loci.
A Yes. Known samples don't take peak heights into consideration so that's why none of them are listed; they only take the allele types into consideration.
Q Let's go to 25, please. We have also heard some testimony about 25 today. This is a document you produced in the course of doing the STRmix analysis in this case?
A Yes.
Q Can we zoom in, Ms. Miller, on that area, please? Now, there are three contributors listed, based on your determination that there were three, that STRmix should assume three contributors, is that right?
A Yes. Based on my interpretation of the profile, I believed that these contributor percentages did represent what I was observing in the profile as a whole where there was a more major contributor, as well as a middle contributor, and then

again a minor contributor.

Q And you've listed the measured mixture or proportions here, correct, at 68 percent, 25 percent, and 7 percent.

THE COURT: Does the fact that those three numbers add up to 100 percent confirm that there were three contributors or not?

THE WITNESS: That number, if there were four there, the percentages would also add up to 100 percent. If there were two, the percentages --

THE COURT: Okay.

THE WITNESS: So it's just saying that the percentage of each individual that contributed to the sample. So the sample as a whole is 100 percent, and this breaks down what each contributor approximately contributed.

THE COURT: Okay.

MR. PRESANT: Do you have an opinion, Ms. Smith, on which of these three contributors is the defendant, assuming that he is in fact in the mixture?

THE WITNESS: When I did the comparison first, so before I actually run it through STRmix, I actually look at the electropherograms not only to determine number of contributors but I do my own comparison to look to see if I could visually exclude someone so I don't have to use STRmix. Or I also look through to see is there drop-out in how does this known sample look compared to the evidence. So I make a conclusion just

like I always have before I run the sample. The sample then is entered into STRmix and confirms my conclusion if the person is there or not. And originally when I looked at the electropherogram, I did associate item 2AX as the most minor contributor, and STRmix then confirmed my conclusion.

BY MR. PRESANT:

Q So the sample for comparison, the defendant's sample, you believe to be the 7 percent, correct?

A Yes.

Q What about sex, are you able to make a sex determination with respect to either contributor 1 or contributor 2?

A Generally, again looking at the electropherogram as a whole, I'm able to look at the amelogenin, which is the sex defining chromosome, and generally if the X is larger in the XY combination, that may indicate a potential female contributor is predominantly contributing to this sample, and in this case there was a significant imbalance with the XY. So based on looking at the profile, in my opinion the predominant donor was a female in this case. And I knew at least one of contributor 2 or 3 had to be a male because there was also a Y present as well as there was a type present at DY, which is also another male indicator chromosome. So I knew before comparing anything that there was at least one male present out of this mixture.

Q And that male would be the defendant if he is in the mixture, correct?

A Yes.
Q And your opinion then with respect to contributor 1 is that that's likely a female?
A In my opinion, yes, contributor 1 is predominantly the female.
Q And what contributor 2, can you make a conclusion as for sex for contributor 2?
A I would not make a conclusion on contributor 2.
Q Now, is this 7 percent number acceptable because of the internal validation studies showing STRmix can be used in the Michigan State Police laboratory down to levels where the minor contributor is below 7 percent?
A Yes.
Q There's also been testimony today about this number right here, the Gelman-Rubin conversions diagnostic. Do you recall hearing that testimony? You were in the courtroom, right?
A Yes.
Q Why did you settle on a run of STRmix where this diagnostic was 1.41?
A So according to the 2.11 protocol there are diagnostics that you look at. The primary diagnostic isn't always is the contributor percentages does it appear to make sense as well as the genotype breakdowns. The secondary diagnostics that are to be considered is the Gelman-Rubin convergence, as well as allele and stutter variance, and the average likelihood or

average log likelihood. Those numbers aren't the end all be all and they are to be considered as a whole, and our manual does say that 1.2 is generally an indicative indication that the chains did converge. However, being that is slightly above 1.2, is not a big deal because the ultimate is does the weight of the breakdown of the evidence make sense, and then once the comparison was run, does the individual likelihoods at each location make sense.

So just because one diagnostic is slightly out it doesn't cancel the other 7 or 8 diagnostics that you look at. I have actually had a run where the Gelman-Rubin was at 3.87 but the overall breakdown and likelihood ratios made sense and it was because of the other minor donor that slightly elevated it, but overall the run made sense so it was able to be reported out.

THE COURT: When you say made sense, what, is that a subjective determination by the analyst?

THE WITNESS: Yes. Made sense in a term that it did what I expected it to do. So if in this run say with the contributor proportions there, it would have said that contributors 2 and 3 both generated around 15 percent, and then contributor 1 generated 7 percent or 70 percent. That breakdown would not have made sense to me because I can tell by looking at the electropherogram that there is a major, a mid contributor, as well as a minor contributor. So when I say

made sense, it means that it made sense with what I was observing in the electropherogram. And once I did the comparison, and based on looking at the individual LRs at each individual locus, those numbers weren't anywhere outside of the scope that I would have expected them to be.

BY MR. PRESANT:

Q Is there anything else you would like to highlight for the Court in Government's Exhibit 25?

A Just again that the evidence input file does represent the stutter filters off and that all those numbers are checked by a technical reviewer to make sure they were inputted correctly and that there is data present where there should be data present.

THE COURT: And I think we are going to wrap it up there, Mr. Presant.

MR. PRESANT: Your Honor, if I may I may have one or two more questions then I'm done with my direct.

THE COURT: Okay.

MR. PRESANT: Thank you. Ms. Smith, in your career as a forensic DNA analyst, have you testified of likelihood ratios before?

THE WITNESS: I have.

BY MR. PRESANT:

Q How long have you been testifying to likelihood ratios?

A I have been a forensic scientist for 12 plus years, and I

used likelihood ratios in St. Louis, so I have been testifying and running them since 2006.

MR. PRESANT: That's all I have, Your Honor.

THE COURT: Thank you. You may step down right now, Ms. Smith. We are going to have to invite you back tomorrow for cross-examination.

Counsel, what I would like to do before we adjourn is to go through the, to make sure we have the exhibits correctly, which ones have been offered and admitted, and I'm going to ask the clerk to read her list off and once she's finished with that, would you please indicate whether you have anything different from what she has.

THE LAW CLERK: Do you want the ones, just the ones admitted?

THE COURT: Yes.

THE LAW CLERK: Okay. Those are for the government, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 18, 19, 23, 25, 26, 27, and 28.

THE COURT: Is that what you have, Mr. Presant?

MR. PRESANT: That's what I have, Your Honor.

THE COURT: Thank you.

THE LAW CLERK: And for the defense I have G and CC.

MS. KLOET: Yes, Your Honor. And just to make the record clear, the Defense Exhibit F that Your Honor hesitated to admit I believe has now been admitted through the

Government's Exhibit 8. Government's Exhibit 8. So that issue is resolved.

THE COURT: Thank you. Is there anything else we need to cover before we adjourn for the day?

MR. PRESANT: Not from the government.

MS. KLOET: Give me one second, Your Honor. Just give me a second to -- I think we are all set. Thank you.

THE COURT: Let's be back here ready to start promptly at 9:00 o'clock tomorrow morning.

THE LAW CLERK: All rise. Court is adjourned.

(Proceedings concluded, 4:34 p.m.)

INDEX

Government Exhibits:		Admitted
1	Buckleton CV	15
4	validation studies	53
5	validation studies	53
12	validation studies	53
23	validation studies	53
13	list of labs using STRmix	60
15	Iyer/Lund paper	61
18	Buckleton paper	63
28	accepted Buckleton response	63
17	Dr. Adams report	66
27	Metropolis-Hastings algorithms	73
14	miscodes on STRmix	77
2	Nye CV	144
19	SWGDAM guidelines	148
10	MSP internal validation	174
11	MSP Procedure & Training Manual	188
3	Amber Smith CV	206
6	electropherograms	211
7	electropherograms	211
9	lab report	221
8	Word Pad file	223
25	deconvolution file	223
26	PDF of file	223

Defense Exhibits

REPORTER'S CERTIFICATE

I, Kathy J. Anderson, Official Court Reporter for the United States District Court for the Western District of Michigan, appointed pursuant to the provisions of Title 28, United States Code, Section 753, do hereby certify that the foregoing is a full, true and correct transcript of the proceedings had in the within entitled and numbered cause on the date hereinbefore set forth; and I do further certify that the foregoing transcript has been prepared by me or under my direction.

/s/ Kathy J. Anderson

Kathy J. Anderson, RPR, FCRR

U.S. District Court Reporter

412 Federal Building

Grand Rapids, Michigan 49503